Large Scale Computing for the Modelling of Whole Brain Connectivity

The human brain constitutes an impressive network formed by the structural and functional connectivity patterns between billions of neurons. Modern functional and diffusion magnetic resonance imaging (fMRI and dMRI) provides unprecedented opportunities for exploring the functional and structural organization of the brain in continuously increasing resolution. From these images, networks of structural and functional connectivity can be constructed. A Bayesian stochastic blockmodelling provides a prominent data-driven approach for uncovering the latent organization, by clustering the networks into groups of nodes with a shared connectivity pattern. Modelling the brain in great detail on a whole-brain scale is essential to fully understand the underlying organization of the brain and reveal the relations between structure and function, that allows sophisticated cognitive behaviour to emerge from ensembles of neurons. Relying on Markov Chain Monte Carlo (MCMC) simulations as the workhorse in Bayesian inference however poses significant computational challenges, especially when modelling networks at the scale and complexity supported by high-resolution whole-brain MRI. In this thesis, we present how to overcome these computational limitations and apply Bayesian stochastic block models for unsupervised data-driven clustering of whole-brain connectivity in full image resolution. We implement high-performance software that allows us to efficiently apply stochastic blockmodelling with MCMC sampling on large complex networks. To obtain the necessary computational performance, we find that both hardware and model specific properties must be taken into consideration - to an extend not supported by generic modelling tools. Computational overhead is reduced by an approach, where key values are cached to avoid re-computations, while tablelookups are utilized for frequently computed special functions. The efficient memory-management of C++ is utilized to implement dedicated data-structures, optimized to facilitate performance-critical operations related to the inference procedure. Furthermore, the software is based on a modular design, which allows us to couple and explore different models and sampling procedures in runtime, still being applied to full-sized data. Using the implemented tools, we demonstrate that the models successfully can be applied for clustering whole-brain connectivity networks. Without being informed of spatial information, the data-driven models can discover spatial homogeneous regions that are meaningful and in agreement with existing anatomical atlases. We further demonstrate that structural and functional connectivity share information, allowing us to jointly model both modalities. For limited, noisy fMRI data we find that integrating structural information aids in discovering the functional organization better than using the fMRI data alone. Though structure and function describes very different properties of the brain, we find that probabilistic modelling provides an intuitive data-driven approach for uncovering the latent organization in connectivity networks. We find that the stochastic block models can be computationally scaled to model wholebrain connectivity, and by doing so allows us to better utilize the full potential of high-resolution MRI and advances our understanding of both the functional and structural organization of the entire brain.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Albers, K. J. (Intern), Schmidt, M. N. (Intern), Merup, M. (Intern)
Number of pages: 224
Publication date: 2017

Publication information
Publisher: DTU Compute
Original language: English

Series: DTU Compute PHD-2017
Volume: 450
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
Large Scale Computing for the Modelling of Whole Brain Connectivity
Publication: Research › Ph.D. thesis – Annual report year: 2017

Probabilistic models for structured sparsity
Sparsity has become an increasingly popular choice of regularization in machine learning and statistics. The sparsity assumption for a matrix $X$ means that most of the entries in $X$ are equal to exactly zero. Structured sparsity is generalization of sparsity and assumes that the set of locations of the non-zero coefficients in $X$ contains structure that can be exploited. This thesis deals with probabilistic models for structured sparsity for regularization of ill-posed problems. The aim of the thesis is two-fold: to construct sparsity promoting prior distributions for structured sparsity and to derive efficient inference algorithms for these distributions. The work explores a class of models that uses Gaussian processes (Rasmussen and Williams, 2006) as a latent representation of the structure of sparsity patterns. This representation allows prior knowledge of the structure of the sparsity patterns to be encoded using generic covariance functions through the Gaussian process. This thesis focuses on two specific instances of ill-posed problems: linear inverse problems and time-varying covariance estimation. The first part of the thesis deals with probabilistic methods for finding structured sparse solutions to linear inverse problems. In this part, the sparsity promoting prior known as the spike-and-slab prior (Mitchell and Beauchamp, 1988) is generalized to the structured sparsity setting. An expectation propagation algorithm is derived
for approximate posterior inference. The proposed model and the associated inference algorithm are studied and evaluated using a set of numerical experiments, which include phase transition experiments, compressed sensing, phoneme classification and electroencephalography (EEG) source localization. The second part of the thesis deals with the problem of time-varying covariance estimation. A hierarchical model for a set of non-stationary time series with time-varying covariance matrices is proposed. The model is tailored to address the problem of dynamic functional connectivity in neuroimaging and it assumes that the instantaneous covariance matrix of each time series is decomposed into a non-negative linear combination of elements from a dictionary of shared covariance matrix components. A variational Bayes algorithm is derived for approximate posterior inference. The proposed model is validated using a functional magnetic resonance imaging (fMRI) dataset.

An Investigation of Methods for CT Synthesis in MR-only Radiotherapy
In recent years, the interest in using magnetic resonance (MR) imaging in radiotherapy (RT) has increased. This is because MR has a superior soft tissue contrast compared to computed tomography (CT), which makes it a better modality for delineating the target volume (tumor) and possible organs at risk (OARs). In an MR/CT work-flow, independent MR and CT scans are acquired. The target and possible OARs are delineated on the MR and then transferred to CT by aligning the data using a registration. This introduces the risk of systematic registration errors especially in non-rigid body structures, the consequence being a systematic miss of target or increased dose to healthy tissue.

Radiotherapy based on MR as the only modality removes this uncertainty and simplifies the clinical work-flow. However, the information on electron density which is usually contained in the CT must now be derived from the MR. A way to achieve this is to computationally estimate a so-called synthetic CT (sCT) from the MR data, which can then act as a substitute for the CT. This is a challenging task, since no unique relationship between MR and electron density exists.

The goal of this thesis is to develop and investigate the right combination of MR acquisition protocols and computational models for accurate MR-based CT synthesis for use in RT. We investigate different categories of methods for CT synthesis and validate them using clinically relevant quality measures. Specifically, we implement a patch-based multi-atlas method in the brain, which compares favorably to state-of-the-art methods. In our next effort, we substantially improve the speed of the method and apply it in the pelvis, again with promising results. Our final contribution is a voxel-based method, which is developed to be registration-free and broadly applicable. In initial results, the performance of this method is close to the patch-based.
Formal Analysis of Graphical Security Models
The increasing usage of computer-based systems in almost every aspects of our daily life makes more and more
dangerous the threat posed by potential attackers, and more and more rewarding a successful attack. Moreover, the
complexity of these systems is also increasing, including physical devices, software components and human actors
interacting with each other to form so-called socio-technical systems. The importance of socio-technical systems to
modern societies requires verifying their security properties formally, while their inherent complexity makes manual
analyses impracticable.

Graphical models for security offer an unrivalled opportunity to describe socio-technical systems, for they allow to
represent different aspects like human behaviour, computation and physical phenomena in an abstract yet uniform
manner. Moreover, these models can be assigned a formal semantics, thereby allowing formal verification of their
properties. Finally, their appealing graphical notations enable to communicate security concerns in an understandable way
also to non-experts, often in charge of the decision making.

This dissertation argues that automated techniques can be developed on graphical security models to evaluate qualitative
and quantitative security properties of socio-technical systems and to synthesise optimal attack and defence strategies.

In support to this claim we develop analysis techniques for widely-used graphical security models such as attack trees and
attack-defence trees. Our analyses cope with the optimisation of multiple parameters of an attack and defence scenario.
Improving on the literature, in case of conflicting parameters such as probability and cost we compute the set of optimal
solutions in terms of Pareto efficiency. Moreover, we investigate the relation between attack and attack-defence trees and
stochastic models in a verification-oriented setting, with the aim of leveraging the great many mature tools and analysis
techniques developed for instance in the area of games.

A Diagnostic and Predictive Framework for Wind Turbine Drive Train Monitoring
Vast amount of data are collected minute by minute from wind turbines around the world. This thesis represents a focused
research effort into discovering new ways of processing these data streams in order to gain insights which can be used to
lower the maintenance costs of wind turbines and increase the turbine availability.
First, it is demonstrated how simple sensor data streams can be leveraged based on a combination of non-linear predictive models and unsupervised fault detection to provide warnings of a critical bearing failure more than a month earlier compared to existing alarm systems. Second, early fault identification based on analysis of complex vibration patterns which is a domain previously reserved for human experts, is shown to be solved with high accuracy using deep learning architecture strained in a fully supervised sense from the data collected in a large scale wind turbine monitoring platform. The research shows a way towards a fully automatized data-driven wind turbine diagnostic processing system that is highly scalable and requires little or no feature engineering and system modeling.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Copenhagen Center for Health Technology
Authors: Bach-Andersen, M. (Intern), Winther, O. (Intern), Rømer-Odgaard, B. (Ekstern)
Number of pages: 111
Publication date: 2017

Publication information
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2017
Volume: 449
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
A Diagnostic and Predictive Framework for Wind Turbine Drive Train Monitoring
Publication: Research › Ph.D. thesis – Annual report year: 2017

Enhancing User Experience in Next Generation Mobile Devices Using Eye Tracking as a Biometric Sensor
A good User Experience is not about just “getting the job done” in the most efficient way. It is also about the subjective elements, providing a positive experience to the user while doing so; emotionally and affectively, having the user engage with the service or product.

Knowing when this takes place means we need ways of measuring concepts like attention. The basis for this should preferably be rooted in our understanding of the anatomically based attention networks of the brain.

This thesis looks at biometric markers of cognitive and affective processes; at the overview level Electroencephalography (EEG), Galvanic Skin Conductance (GSR), Heart Rate and Heart Rate Variability as well as Face Expression Detection – and in much more detail Eye Tracking.

A simple framework for relating eye movements and pupil dilations to the visual processing system and to the attentional networks is suggested. It is demonstrated that it is possible to identify components of attention and cognitive load using low cost eye tracking in conventional office settings. It is also shown that aspects of surprise, similar to negativity feedback error coding, is measurable. Behavioural patterns possibly related to time on target, cognitive load, performance or stimuli are inferred. The existence of possibly unique individual gaze patterns related to visual stimuli or to the brain’s Default Mode Network are shown.

A way of synchronizing EEG and Eye Tracking is also suggested, and in addition, a few software assets (a Python interface to The Eye Tribe tracker and an implementation of the Attention Network Test (ANT)) have been created.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Copenhagen Center for Health Technology
Authors: Bækgaard, P. (Intern), Petersen, M. K. (Intern), Larsen, J. E. (Intern)
Number of pages: 260
Publication date: 2017

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2016
Methods and Tools for the Analysis, Verification and Synthesis of Genetic Logic Circuits,

Synthetic biology has emerged as an important discipline in which engineers and biologists are working together to design new and useful biological systems composed of genetic circuits. The purpose of developing genetic circuits is to carry out desired logical functions inside a living cell. This usually requires simulating the mathematical models of these genetic circuits and perceiving whether or not the circuit behaves appropriately. Furthermore, synthetic biology utilizes the concepts from electronic design automation (EDA) of abstraction and automated construction to generate genetic circuits with the aim to reduce the in-vitro (wet-lab) experiments. To address this, several automated tools have been developed to improve the process of genetic design automation (GDA) with different capabilities. This thesis attempts to contribute to the advancement of GDA tools by introducing capabilities which we believe that no other existing GDA tools support. First, we introduce a user-friendly simulation tool, called D-VASim, which allows users to perform virtual laboratory experimentation by dynamically interacting with the model during runtime. This dynamic interaction with the model gives users a feeling of being in the lab performing wet-lab experiments virtually. This tool allows users to perform both deterministic and stochastic simulations. Next, this dissertation introduces a methodology to perform timing analyses of genetic logic circuits, which allows users to analyze the threshold value and propagation delays of genetic logic circuits. In this thesis, it has been demonstrated, through in-silico experimentation, that the threshold value and propagation delay play a vital role in the correct functioning of genetic circuit. It has also been shown how some circuit parameters affect these two important design characteristics. This thesis also introduces an automated approach to analyze the behavior of genetic logic circuits from the simulation data. With this capability, the boolean logic of complex genetic circuits can be analyzed and/or verified automatically. It is also shown in this thesis that the proposed approach is effective to determine the variation in the behavior of genetic circuits when the circuit’s parameters are changed. In addition, the thesis also attempts to propose a synthesis and technology mapping tool, called GeneTech, for genetic circuits. It allows users to construct a genetic circuit by only specifying its behavior in the form of boolean expression. For technology mapping, this tool uses a gates library developed by the collective efforts of the researchers at MIT and Boston universities. It is shown experimentally that the tool is able to provide all feasible solutions, containing different genetic components, to achieve the specified boolean behavior. Finally, it has been shown how D-VASim can be used along with other tools for useful purposes, like model checking. With respect to this, an experimental workflow is proposed for checking genetic circuits using the statistical model checking (SMC) utility of the Uppaal tool and the timing analysis capability of D-VASim. We further demonstrated how the reliability of a simulation can be improved by using the real parameter values. In this regard, the relationship between the simulation parameters and real parameters have been derived.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Embedded Systems Engineering
Authors: Baig, H. (Intern), Madsen, J. (Intern), Pop, P. (Intern)
Number of pages: 253
Publication date: 2017

Publication information
Original language: English
Series: DTU Compute PHD-2017
Volume: 456
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd456_Baig_H.pdf

Relations
Projects:
Methods and Tools for the Analysis, Verification and Synthesis of Genetic Logic Circuits,
Geometric singular perturbation analysis of systems with friction
This thesis is concerned with the application of geometric singular perturbation theory to mechanical systems with friction. The mathematical background on geometric singular perturbation theory, on the blow-up method, on non-smooth dynamical systems and on regularization is presented. Thereafter, two mechanical problems with two different formulations of the friction force are introduced and analysed. The first mechanical problem is a one-dimensional spring-block model describing earthquake faulting. The dynamics of earthquakes is naturally a multiple timescale problem: the timescale of earthquake ruptures is very short, when compared to the time interval between two consecutive ruptures. We identify a small parameter ε that describes the separation between the timescales, so that ε = 0 idealises the complete timescale separation. Earthquake faulting problems also have multiple spatial scales. The action of friction is generally explained as the loss and restoration of linkages between the surface asperities at the molecular scale. However, the consequences of friction are noticeable at much larger scales, like hundreds of kilometers. By using geometric singular perturbation theory and the blow-up method, we provide a detailed description of the periodicity of the earthquake episodes. In particular, we show that attracting limit cycles arise from a degenerate Hopf bifurcation, whose degeneracy is due to an underlying Hamiltonian structure that leads to large amplitude oscillations. We use a Poincaré compactification to study the system near infinity. At infinity, the critical manifold loses hyperbolicity with an exponential rate. We use an adaptation of the blow-up method to recover the hyperbolicity. This enables the identification of a new attracting manifold, that organises the dynamics at infinity for ε = 0. This in turn leads to the formulation of a conjecture on the behaviour of the limit cycles as the timescale separation increases for 0 < ε 1. We illustrate our findings with numerics, and outline the proof of the conjecture. We also discuss how our results can be used to study a similar class of problems. The second mechanical problem is a friction oscillator subject to stiction. The vector field of this discontinuous model does not follow the Filippov convention, and the concept of Filippov solutions cannot be used. Furthermore, some Carathéodory solutions are unphysical. Therefore, we introduce the concept of stiction solutions: these are the Carathéodory solutions that are physically relevant, i.e. the ones that follow the stiction law. However, we find that some of the stiction solutions are forward non-unique in subregions of the slip onset. We call these solutions singular, in contrast to the regular stiction solutions that are forward unique. In order to further the understanding of the non-unique dynamics, we introduce a regularization of the model. This gives a singularly perturbed problem that captures the main features of the original discontinuous problem. We identify a repelling slow manifold that separates the forward slipping to forward sticking solutions, leading to a high sensitivity to the initial conditions. On this slow manifold we find canard trajectories, that have the physical interpretation of delaying the slip onset. We show numerically that the regularized problem has a family of periodic orbits interacting with the canards. We observe that this family is unstable of saddle type and that it connects, in the rigid body limit, the two regular, slip-stick branches of the discontinuous problem, that were otherwise disconnected.
time-to-solution by having a larger basin of attraction, faster initial convergence, data locality and a lower memory footprint. The study is extended to include a hybrid strategy, where FAS is combined with Newton’s method to construct a multilevel nonlinear preconditioner. This method demonstrates high efficiency and robustness.

Second, an improved IMPES formulated reservoir simulator is implemented using a novel variational upscaling approach based on element-based Algebraic Multigrid (AMGe). In particular, an advanced AMGe technique with guaranteed approximation properties is used to construct a coarse multilevel hierarchy of Raviart-Thomas and L2 spaces for the Galerkin coarsening of a mixed formulation of the reservoir simulation equations. By experimentation it is found that the AMGe based upscaling technique provided very accurate results while reducing the computational time proportionally to the reduction in degrees of freedom. Furthermore, it is demonstrated that the AMGe coarse spaces (interpolation operators) can be used for both variational upscaling and the construction of linear solvers. In particular, it is found to be beneficial (even necessary) to apply an AMGe based multigrid solver to solve the upscaled problems. It is found that the AMGe upscaling changes the spectral properties of the matrix, which renders well-known state-of-the-art solvers for this type of system useless.

Third, FAS is combined with AMGe with guaranteed approximation properties to obtain a nonlinear multigrid solver for unstructured meshes. The FAS-AMGe solver is applied to a simplistic but numerically challenging mixed (velocity-pressure) model for porous media flow. In a fair way, FAS-AMGe is compared to Newton’s method and Picard iterations. It is found that FAS-AMGe is faster for the cases considered.

Finally, a number of multigrid linear solvers and preconditioners are implemented for various linear systems. In particular AMGe are used in the construction of multigrid preconditioners. These are compared to two state-of-the-art block diagonal preconditioners based on 1) a Schur complement with an Algebraic Multigrid (AMG) solver and 2) an augmented Lagrangian formulation using the Auxiliary Space AMG solver.

In addition to the research mentioned above, a sequential in-house COmpositional reservoir Simulator (COSI) with many features is parallelized in a distributed setting (MPI) using the PETSc framework. A parallel preconditioner based on the Constrained Pressure Residual method, Algebraic Multigrid and Restricted Additive Overlapping Schwarz with Incomplete LU solves on each subdomain is implemented. It is found that switching the traditionally used method, namely parallel ILU, with Restricted Additive Overlapping Schwarz results in a significant increase in parallel scalability while still maintaining similar robustness and efficiency.

---

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing, Center for Energy Resources Engineering
Authors: Christensen, M. L. C. (Intern), Engsig-Karup, A. P. (Intern)
Number of pages: 213
Publication date: 2017

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 430
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd430_Christensen_MlC.pdf

**Relations**

Projects:
Multilevel techniques for Reservoir Simulation
Publication: Research › Ph.D. thesis – Annual report year: 2017

**Data Mining and Visualization of Large Human Behavior Data Sets**

Traditional methods for studying human behavior such as surveys and manual collection are expensive, time-consuming and therefore cannot be easily applied at large scale. In recent years an explosive amount of digital traces of human activity – for example social network interactions, emails and credit card transactions – have provided us new sources for studying our behavior. In particular smartphones have emerged as new tools for collecting data about human activity, thanks to their sensing capabilities and their ubiquity. This thesis investigates the question of what we can learn about human behavior from this rich and pervasive mobile sensing data. In the first part, we describe a large-scale data collection deployment collecting high-resolution data for over 800 students at the Technical University of Denmark using smartphones, including location, social proximity, calls and SMS. We provide an overview of the technical infrastructure,
the experimental design, and the privacy measures. The second part investigates the usage of this mobile sensing data for understanding personal behavior. We describe two large-scale user studies on the deployment of self-tracking apps, in order to understand the patterns of usage and non-usage. Moreover we provide some design guidelines for facilitating reflection in self-tracking systems. Finally we propose a model for inferring sleep patterns from smartphone interactions.

In the third part, we focus on a specific aspect of collective behavior: human mobility. We perform an experiment to verify the feasibility of inferring places from location traces using mobile sensing data. We develop a hierarchical model for human mobility, which is able to measure mobility properties at multiple scales. We perform a study on the factors influencing the accuracy of nextplace prediction models. Finally we present an open-source tool for creating geographical visualizations.

**General information**

State: Published  
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems  
Authors: Cuttone, A. (Intern), Larsen, J. E. (Intern), Jørgensen, S. L. (Intern)  
Number of pages: 258  
Publication date: 2017

**Publication information**

Place of publication: Kgs. Lyngby  
Publisher: Technical University of Denmark (DTU)  
Original language: English

Series: DTU Compute PHD-2016  
Number: 422  
ISSN: 0909-3192  
Main Research Area: Technical/natural sciences  
Electronic versions:  
phd422_Cuttone_A.pdf

**Relations**

Projects:  
Data Mining and Visualization of Large Human Behavior Data Sets  
Publication: Research › Ph.D. thesis – Annual report year: 2017

**Topological bifurcations of coherent structures and dimension reduction of plasma convection models**

Research in fusion energy seeks to develop a green, safe, and sustainable energy source. Nuclear fusion can be achieved by heating a hydrogen gas to temperatures of millions of kelvin. At fusion temperatures, some or all the electrons leave the atomic nucleus of the hydrogen atom. This results in an overall neutral gaseous state of negatively charged free electrons and positively charged ions. This state of matter is called plasma. To achieve and maintain fusion temperatures, the plasma must avoid direct contact with any solid material. Since the plasma consists of charged particles, it can be confined with an appropriate configuration of strong magnetic fields. Toroidal magnetic confinement devices, such as the tokamak, are the most promising designs for a fusion reactor. A tokamak can operate in two distinct modes of operation. These are the low confinement mode (L-mode) and the high confinement mode (H-mode). H-mode is the preferred operating mode for a fusion reactor. The transition from L-mode to H-mode is called the L–H transition. The confinement properties of a plasma are largely determined by the physics near the edge of the confinement region of the plasma. The edge transport of a magnetically confined plasma is predominantly caused by recurring bursts of coherent plasma structures. These structures are in L-mode called blob filaments (blobs) and in H-mode categorized into edge localized mode (ELM) filaments or inter-ELM filaments. To improve the plasma confinement, it is important to understand the evolution of these structures. We apply a dynamical systems approach to quantitatively describe the time evolution of these structures. Three state variables describe blobs in a plasma convection model. A critical point of a variable defines a feature point where that variable is significant. For a range of Rayleigh and Prandtl numbers, we analyze the bifurcations of the critical points of the three variables with time as the main bifurcation parameter. Plasma simulations can be computationally demanding. We apply a Galerkin method to approximate a plasma convection model with a reduced model. The time evolution of the energies of the pressure profile, the turbulent flow, and the zonal flow capture the dynamic behavior of the convection model. Rayleigh decomposition splits the variables of the model into averaged variables and fluctuation variables. We approximate the fluctuation variables by truncated Fourier series and project the equations onto the Fourier basis functions. This results in a computationally simpler model with the spatial dimension reduced by one. Bifurcation diagrams for the energies show consistency between the bifurcation structures of the full and the reduced model. Finally, we utilize a data-driven modeling approach called SINDy to identify a reduced model from simulation data of a convection model. The reduced model reveals a predator-prey relationship between the zonal flow energy and the turbulent energy. The analytically derived bifurcation diagram for the reduced model has the same structure as the data-based bifurcation diagram for the full model.
Computer Vision for Additive Manufacturing.

Ever since the commercialization of additive manufacturing in the late 80’s, it has been clear what enormous potential the technology could have, potentially disrupting several industries. However, we have yet to see the technology fully adopted by the manufacturing industry. One of the issues that has prevented widespread adoption of 3D printing for use within manufacturing is the apparent lack of quality control during and after the printing process. This thesis demonstrates how computer vision may be applied in beneficial ways within additive manufacturing. The main contributions aim at solving part of the challenges required for the technology to reach its full envisioned potential, and to reach widespread industry adoption as a de-facto manufacturing modality. Quality control has been a major milestone to overcome in this regard. As a result, a core part of the contributions revolves around this central topic. The work is separated into three main categories: The first two concerning process and quality control of appearance and geometry. The third category concerns machine interaction paradigms within additive manufacturing. Here, challenges are addressed within the 3D ecosystem, aiming towards facilitating a fluid integration of additive manufacturing within the factory of tomorrow.

Authentication for E-Government in Developing Countries - With special focus on the North Africa Countries

Recently, many countries include both developed countries as well as developing countries have transformed paper based systems into electronic systems using ICT technologies in order to improve service delivery and reduce cost. Several researches and International Organizations in the field of e-Government reports that many countries over the world have not achieved transaction stages of government e-services and most of those countries are from developing countries. One of the main issues challenge government e-service inclusion is digital divide which barriers achieving principle of equal access and benefit of government e-service. Therefore, This thesis aims to investigate digital divide and IDM issues challenge government e-service in developing countries such as North Africa Countries (NAC) from achieving the priciple of equal access in a secure manner. To achieve this aim we, developed a framework that consists of two
components include digital divide variables and a simple IDM model in order to assess the current state of government e-service in NAC. Moreover, we analyzed the existing IDM protocol’s concept to understand whether those concepts consider disadvantaged user’s needs. Based on the identified challenges in NAC using the developed framework and the analysis of IDM protocol’s concept we identify the requirements to be satisfied in order to allow large portion of citizens access and benefit of government e-service in equal and secure manner. One possible solution to improve e-Government inclusion is to consider vulnerable group needs such as the case in which users (citizens) do not have the ability either to read or write and as a result are excluded from e-services. Thus, a solution should enable such users to benefit from e-services. Introducing vulnerable group such as illiterate individuals might introduce new risks which have not existed in citizens-government face to face interaction. Thus, considering security property include confidentiality, integrity, non-repudiation and accountability for a proposed solution is needed. User authentication based on social relationship protocol is proposed in order to bridge digital divide. We formalized the proposed protocol as well as IDM protocol’s concept using Open Source Fixed Point Model Checker tool (OFMC) To verify security properties include secrecy of exchanged information and authenticity of communication parties of the target protocols. OFMC is an automatic protocol security verification tool to identify the strengths of the verified protocol. Based on the verification result of OFMC tool, an attack is found against the existing IDM protocol’s concept when considering vulnerable users while the proposed protocol has achieved the specified goals without ant attack at least in one session. We also, performed a simple usability comparison between the proposed protocol and public kiosk service delivery channel and the proposed protocol shows its effectiveness as well as efficient.

General information
State: Submitted
Organisations: Cyber Security, Department of Applied Mathematics and Computer Science
Authors: Elaswad, O. (Intern), Jensen, C. D. (Intern)
Number of pages: 130
Publication date: 2017

Publication information
Publisher: DTU Compute
Original language: English

Series: DTU Compute PHD-2017
Volume: 455
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
Authentication for E-Government in Developing Countries - With special focus on the North Africa Countries
Publication: Research › Ph.D. thesis – Annual report year: 2017

The Impact of Parametrization on Randomized Search Heuristics
In this work we present runtime analyses of randomized search heuristics (RSH) in various settings that are determined by parameters of the problems, the algorithms and also exogenous parameters like noise. In the process we provide new techniques for the theoretical analysis of RSH as well as new optimization algorithms. We consider the following topics. Escaping local optima using local search. We analyze memetic algorithms, i.e. evolutionary algorithms equipped with a local search after mutation. To this end we consider the (1+1) EA with mutation rate $c/n$. The differences of the fitness landscape that lead to the (1+1) EA using SLS outperforming the (1+1) EA using VDS with an exponential performance gap. Moreover, we present a new local search operator, Opportunistic Local Search (OLS), that can deal with such features in the landscape and show that the (1+1) EA with OLS can efficiently optimize a discretized Rastergin function. Stochastic fitness functions. We analyze the role of populations in stochastic optimization. We assume that the objective function is subject to noise, introducing stochastic errors in its evaluation. On classical test functions, such noise makes optimization by the simple (1+1) EA hillclimber infeasible even in exponential time. Interestingly, the use of parent and offspring populations of only logarithmic size turns the algorithm into an efficient one. The results are obtained by drift analysis. An asymptotic expansion of the expected runtime of the (1+1) EA on ONE MAX. We consider the (1+1) EA with mutation probability $c/n$, where $c > 0$ is a constant on ONE MAX. We give an asymptotic expansion for the expected runtime depending on both $c$ and $\lambda$. Our results show that $c = 1$ is the optimal mutation rate for $\lambda = o(logloglogn/logloglogn)$ and that $c$ only has an impact on the lower-order terms of the expected runtime, i.e. $c = 1$ is no longer the only optimal mutation rate. Our methods are strongly based on variable drift theorems for upper and lower bounds and a precise analysis of order statistics of the binomial distribution. To the best of our knowledge this is the first tight runtime analysis of a population-based EA, up to lower-order terms. Furthermore, we develop helpful stochastic tools for runtime analyses. Optimal mutation rates for the (1+1) EA on ONE MAX. We consider the (1+1) EA with mutation probability $c/n$ on ONE MAX, where $c > 0$ and $\lambda$ are constant. We present an improved variable drift theorem that weakens the requirement that no large steps towards the optimum may occur in the process to a stochastic one, reducing the analysis of the expected optimization time to finding an exact expression for the drift. We formalize an exact closed-form expression for the drift and provide small error approximations that are very efficient to compute. Self-adjusting mutation rates for the (1+1) EA on ONE MAX. We propose a new mechanism to self-adjust the mutation rate in population-based evolutionary algorithms. It consists of
creating half the offspring with a higher and the rest with a lower mutation rate. The mutation rate is then adjusted, based on the success of the subpopulations. We show that the \((1+\lambda)\) EA optimizes \textsc{Onemax} in an expected optimization time of \(O(n\lambda/\log\lambda + \log n)\) which has been shown to be best-possible among all \(\lambda\)-parallel mutation-based unbiased black-box algorithms.

**General information**
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic
Authors: Gießen, C. (Intern), Witt, C. (Intern)
Number of pages: 156
Publication date: 2017

**Publication information**
Publisher: DTU Compute
Original language: English

Series: DTU Compute PHD-2017
Volume: 460
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

**Relations**
Projects:
The Impact of Parametrization on Randomized Search Heuristics
Publication: Research › Ph.D. thesis – Annual report year: 2017

Generative Temporal Modelling of Neuroimaging - Decomposition and Nonparametric Testing
The goal of this thesis is to explore two improvements for functional magnetic resonance imaging (fMRI) analysis; namely our proposed decomposition method and an extension to the non-parametric testing framework. Analysis of fMRI allows researchers to investigate the functional processes of the brain, and provides insight into neuronal coupling during mental processes or tasks.

The decomposition method is a Gaussian process-based independent components analysis (GPICA), which incorporates a temporal dependency in the sources. A hierarchical model specification is used, featuring both instantaneous and convolutive mixing, and the inferred temporal patterns. Spatial maps are seen to capture smooth and localized stimuli-related components, and often identifiable noise components. The implementation is freely available as a GUI/SPM plugin, and we recommend using GPICA as an additional tool when performing ICA on fMRI data to investigate the effect of the temporal source prior.

In fMRI, statistical tests are used to investigate the significance of activation in specific brain regions. By extending the non-parametric testing framework to incorporate functional prior knowledge, an increase in sensitivity can be achieved, entailing better evaluations and conclusions. The functional prior knowledge is incorporated by use of a proposed Graph-Based Cluster Permutation Test (GBCPT), entailing the possibility to expand the use of cluster permutations to multiple applications, wherever a graph-based setup can be used.
EEG Based Inference of Spatio-Temporal Brain Dynamics
Electroencephalography (EEG) provides a measure of brain activity and has improved our understanding of the brain immensely. However, there is still much to be learned and the full potential of EEG is yet to be realized. In this thesis we suggest to improve the information gain of EEG using three different approaches; 1) by recovery of the EEG sources, 2) by representing and inferring the propagation path of EEG sources, and 3) by combining EEG with functional magnetic resonance imaging (fMRI). The common goal of the methods, and thus of this thesis, is to improve the spatial dimension of EEG.

The main topic of this thesis is the localization of the EEG generators. This entails solving both a forward and an inverse problem. The inverse problem maps the EEG signal recorded on the scalp to its origin in the brain. It is a highly ill-posed problem which we tackle by employing a sparsity promoting 'spike and slab' like method augmented with physiologically relevant source priors. The incorporated temporal and spatial priors exploit coherence between neighboring time samples and between neighboring source locations, respectively. We show that these augmentations effectively increase the source recovery ability.

The forward problem describes the propagation of neuronal activity in the brain to the EEG electrodes on the scalp. The geometry and conductivity of the head layers are normally required to model this path. We propose a framework for inferring forward models which is based on the EEG signal and a low dimensional representation of forward models. The representation is built by principal component analysis of a corpus of forward models. The method can be used to recover subject-specific forward models when structural scans and/or conductivity estimations are not available.

Finally we investigate the extraction of EEG components having bandpower dynamics correlated with fMRI components. We show that adding anatomical information to the inference scheme improves the recovery of correlated components compared to only using functional information. The anatomical information is incorporated through the EEG forward model and assumes that the activity of the fMRI component overlaps spatially with the origin of the coupled EEG component.

General Information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Copenhagen Center for Health Technology
Authors: Hansen, S. T. (Intern), Hansen, L. K. (Intern)
Number of pages: 193
Publication date: 2017

Publication Information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 410
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd410_Hansen_ST.pdf

Relations
Projects:
EEG Based Inference of Spatio-Temporal Brain Dynamics
Publication: Research › Ph.D. thesis – Annual report year: 2017

Analysis of trait-based models in marine ecosystems.
The overarching theme for this thesis is spatial and temporal variations in ecosystems. The focus is on describing mechanisms that are responsible for generating the spatial and temporal patterns. The thesis contains two separate projects, each exploring a possible mechanism for pattern formation. In both projects, the model formulations result in partial integro-differential equations. The first project in the thesis considers temporal patterns in a size structured population. Size structure is relevant for species that goes through significant changes through their lifetime. The population's response to regular temporal variations in the environment is investigated by introducing a periodic forcing in the system. This can for instance represent seasonal changes. The effect of an imposed forcing is explored both when the underlying unforced system has a stable equilibrium and when it has stable oscillatory dynamics. The numerical solutions show regular cycles where the period is equal to, or an integer multiple of, the forcing period and where the population can
have one or more pulses of reproduction in each cycle. Additionally, the numerical results indicate quasi-periodic or chaotic solutions, period doubling bifurcations and coexisting attractors. The bifurcation structure is similar to results for comparable unstructured population models in the literature. This indicates that size structure does not affect the response to periodic forcing. The next project in the thesis considers spatio-temporal pattern formation in a predator–prey system where animals move towards higher fitness. Reaction-diffusion systems have been used extensively to describe spatio-temporal patterns in a variety of systems. However, animals rarely move completely at random, as expressed by diffusion. This has lead to models with taxis terms, describing individuals moving in the direction of an attractant. An example is chemotaxis models, where bacteria are attracted to a chemical substance. From an evolutionary perspective, it is expected that animals act as to optimize their fitness. Based on this principle, a predator–prey system with fitness taxis and diffusion is proposed. Here, fitness taxis refer to animals moving towards higher values of fitness, and the specific growth rates of the populations are used as a measure of the fitness values. To determine the conditions for pattern formation, a linear stability analysis is conducted. The analysis reveals that the fitness taxis leads to mechanisms for pattern formation, which are based on the prey gathering together. It turns out, that in some cases the problem is not well-posed and an ultraviolet catastrophe occurs, i.e., perturbations with infinitely short wavelength grow infinitely fast. To prevent this, the population dynamics are revised with a spatial feeding kernel, that defines a spatial range wherein a predator consumes prey. A linear stability analysis for the revised system reveals the ultraviolet catastrophe is avoided and the basic mechanisms for pattern formation are unchanged. Numerical solutions to the revised system are computed to visualize the patterns. The solutions encompass stationary spatial patterns in addition to traveling waves, standing waves and irregular solutions that might be spatio-temporal chaos. The modeling approach of fitness taxis presents a general way to express movement and it is concluded that the model provides a useful framework for describing generic mechanisms for pattern formation.

General information
State: Published
Organisations: Dynamical Systems, Department of Applied Mathematics and Computer Science, National Institute of Aquatic Resources, Centre for Ocean Life, University of Rostock
Authors: Heilmann, I. L. T. (Intern), Sørensen, M. P. (Intern), Starke, J. (Ekstern), Thygesen, U. H. (Intern), Andersen, K. H. (Intern)
Number of pages: 95
Publication date: 2017

Publication information
Publisher: DTU Compute
Original language: English
Series: DTU Compute PHD-2017
Volume: 453
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd453_Heilmann_ILT.pdf

Relations
Projects:
Analysis of trait-based models in marine ecosystems.
Publication: Research › Ph.D. thesis – Annual report year: 2017

Quantitative evaluation of peptide analogue distribution in mouse tissue using 3D computer modelling
The use of automated image analysis of microscopy images is increasing to enable high throughput approaches and unbiased analysis of the increasingly large data sets produced. This thesis investigates the use of automated image analysis to quantify peptide analogue distribution in mouse brain tissue. The main group of peptides included in this work was glucagon-like peptide 1 receptors agonists (GLP-1RA) used for treatment in diabetes and obesity. Two main image modalities have been applied for image acquisition: Light Sheet Fluorescence Microscopy (LSFM), and slide scanner images of 2D histology sections. The work demonstrates the use of automated image analysis based on image registration to quantify LSFM data of the peptide brain distribution following peripheral administration. The methodology was expanded during the PhD work to also include study of receptor mapping and brain activation. The automated analysis was enabled by integration with a digital multimodality brain atlas from the Allen Institute of Brain Science (AIBS). The work showed that GLP-1RAs accessed multiple brain regions mainly in the hypothalamus and hindbrain and led to increased brain activation in regions related to decreased food intake. The developed integrated brain atlas provides a novel analysis approach for LSFM data to aid researchers understand the complex brain biology related to development of pharmaceuticals with brain mode of action.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Segmentation-Driven Tomographic Reconstruction.

The tomographic reconstruction problem is concerned with creating a model of the interior of an object from some measured data, typically projections of the object. After reconstructing an object it is often desired to segment it, either automatically or manually. For computed tomography (CT), the classical reconstruction methods suffer from their inability to handle limited and/or corrupted data. Form any analysis tasks computationally demanding segmentation methods are used to automatically segment an object, after using a simple reconstruction method as a first step. In the literature, methods that completely combine reconstruction and segmentation have been suggested, but these are often non-convex and have very high computational demand. We propose to move the computational effort from the segmentation process to the reconstruction process, and instead design reconstruction methods such that the segmentation subsequently can be carried out by use of a simple segmentation method, for instance just a thresholding method. We tested the advantages of going from a two-stage reconstruction method to a one stage segmentation-driven reconstruction method for the phase contrast tomography reconstruction problem. The tests showed a clear improvement for realistic materials simulations and that the one-stage method was clearly more robust toward noise. The noise-robustness result could be a step toward making this method more applicable for lab-scale experiments. We have introduced a segmentation-driven reconstruction method which incorporates information about the main texture direction in an object. We proved that this method has mathematically desirable properties such as being convex and lower semicontinuous. We have also demonstrated the practical applicability of the method.

within image denoising, image deblurring and CT reconstruction. In order to use the proposed method we also proposed efficient and robust methods for estimating the main direction in either corrupted images or from limited and corrupted CT projection data. For directional object we also proposed two different reconstruction methods that separates the directional parts in the object from the non-directional parts. These method could for example be used for objects consisting of fibres and cracks. The results can be categorized as either completely combined reconstruction and segmentation of the object, or as highly supporting for the following segmentation process. Computed tomography is used within medical diagnosis, food science, materials science, production inspection, quality assessment, etc. Segmentation-driven reconstruction methods can help to improve both manual and automated segmentation processes that are used to analyze an object after the scanning. The results in this thesis are both of theoretical interest within regularization theory and directly applicable for image denoising, image deblurring and surely within tomographic reconstruction.
**Statistical modelling of space-time processes with application to wind power.**

Short-term wind power forecasts together with a quantification of uncertainties are required for the reliable operation of power systems with significant wind power penetration. A challenge for utilizing wind power as a source of energy is the intermittent and hardly predictable nature of wind. This thesis aims at contributing to the wind power literature by building and evaluating new statistical techniques for producing forecasts at multiple locations and lead times using spatio-temporal information. By exploring the features of a rich portfolio of wind farms in western Denmark, we investigate different types of models and provide several forms of predictions. Starting with spatial prediction, we then extend the methodology to spatio-temporal prediction of individual wind farms and aggregated wind power at monitored locations as well as at locations where recent observations are not available. We propose spatial models for predicting wind power generation at two different time scales: for annual average wind power generation and for a high temporal resolution (typically wind power averages over 15-min time steps). In both cases, we use a spatial hierarchical statistical model in which spatial correlation is captured by a latent Gaussian field. We explore how such models can be handled with stochastic partial differential approximations of Matérn Gaussian fields together with integrated nested Laplace approximations. We show that complex hierarchical spatial models are well suited for wind power data and provide results in reasonable computational time. Moreover, the hierarchical approach for obtaining predictions at a high temporal resolution is found to produce accurate predictions with improved performance compared to a standard geostatistical method at a small additional computational cost. The use of the integrated nested Laplace approximations is motivated by the desire to produce forecasts on large data sets with hundreds of locations, which is critical during periods of high wind penetration. Subsequently, the extension from spatial to spatio-temporal models is given. Three different hierarchical models are developed for obtaining probabilistic wind power forecasts. First, a time series model consisting of an autoregressive process with a location specific intercept is considered. This approach gives satisfactory results for individual forecasts but fails to generate calibrated aggregated forecasts. The second approach has a common intercept for all farms and a spatio-temporal model that varies in time with first order autoregressive dynamics and has spatially correlated innovations given by a zero mean Gaussian process. The third model, which also has a common intercept as well as an autoregressive process to capture the local variability and the spatio-temporal term from the second approach, is able to produce reliable individual and aggregated forecasts for multiple lead times. Finally, very-short-term wind power forecasting is considered. Probabilistic forecasts from 15 minutes up to two hours ahead are produced by using anisotropic spatio-temporal correlation models to account for the propagation of weather fronts and a transformed latent Gaussian field is used to accommodate the probability masses that occur in wind power distribution due to chains of zero measurements. Using what is called kriging equations, even the simplest proposed covariance model is able to produce calibrated spatio-temporal predictions of wind power production.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Electrical Engineering, Center for Electric Power and Energy, Electricity markets and energy analytics
Authors: Lenzi, A. (Intern), Ersbøll, B. K. (Intern), Clemmensen, L. K. H. (Intern), Pinson, P. (Intern)
Number of pages: 192
Publication date: 2017

**Publication information**

Publisher: DTU Compute
Original language: English

Series: DTU Compute PHD-2017
Volume: 451
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd451_Lenzi_A.pdf

**Relations**

Projects:
Statistical modelling of space-time processes with application to wind power.
Publication: Research › Ph.D. thesis – Annual report year: 2017

**Convolutional Neural Networks - Generalizability and Interpretations**

Sufficient data is key when training Machine Learning algorithms in order to obtain models that generalize for operational use. Sometimes sufficient data is infeasible to obtain and this prevents the use of Machine Learning in many applications. The goal of this thesis is to gain insights and learn from data despite it being limited in amount or context representation. Within Machine Learning this thesis focuses on Convolutional Neural Networks for Computer Vision. The research aims to answer how to explore a model's generalizability to the whole population of data samples and how to interpret the model's
function. The thesis presents three overall approaches to gaining insights on generalizability and interpretation. First, one can change the main objective of a problem to study expected insufficiencies and based on this make a better choice of model. For this first approach the thesis presents both a study on translational invariance as well as an example of changing the objective of a problem from classification to segmentation to robustly extract lower level information. The second approach is the use of simulated data which can help by inferring knowledge in our model if real data is scarce. The results show clear advantages both when using rendered Synthetic Aperture Radar images, but also when predictions from physical models are used as target variables which are matched with real data to form a large dataset. The third approach to cope with data insufficiencies is to visualize and understand the internal representations of a model. This approach is explored and concrete examples of learnings that can be obtained are shown. There is no doubt that large quantities of well representing data is the best foundation for training Machine Learning models. On the other hand, there are many tools and techniques available to interpret and understand properties of our models. With these at hand we can still learn about our models and use this knowledge to e.g. collect better datasets or improve on the modeling.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, National Space Institute, Microwaves and Remote Sensing
Authors: Malmgren-Hansen, D. (Intern), Nielsen, A. A. (Intern), Engholm, R. (Ekstern), Skriver, H. (Intern)
Number of pages: 139
Publication date: 2017

Publication information
Publisher: Technical University of Denmark
Original language: English
Series: DTU Compute PHD-2017
Volume: 459
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Optimization on Spaces of Curves
This thesis is concerned with computational and theoretical aspects of Riemannian metrics on spaces of regular curves, and their applications. It was recently proved that second order constant coefficient Sobolev metrics on curves are geodesically complete. We extend this result to the case of Sobolev metrics with coefficient functions depending on the length of the curve. We show how to apply this result to analyse a wide range of metrics on the submanifold of unit and constant speed curves.

We present a numerical discretization of second order Sobolev metrics on the space of regular curves in \(\mathbb{R}^d\), and methods to solve the initial and boundary value problem for geodesics allowing us to compute the Karcher mean and principal components analysis of data of curves. We apply the methods to study shape variation in synthetic data in the Kimia shape database, in HeLa cell nuclei and cycles of cardiac deformations.

Finally we investigate a new application of Riemannian shape analysis in shape optimization. We setup a simple elliptic model problem, and describe how to apply shape calculus to obtain directional derivatives in the manifold of planar curves. We present an implementation based on parametrization of immersions by B-splines, which ties in naturally with Isogeometric Analysis to solve the PDE. We give numerical examples of solutions, and compare the Riemannian optimization algorithms with different choices of metrics to a naive unregularized discretize-first approach.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Mathematics
Authors: Møller-Andersen, J. (Intern), Gravesen, J. (Intern)
Number of pages: 113
Publication date: 2017

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2016
Test-Driven, Model-Based Systems Engineering.

Hearing systems have evolved over many years from simple mechanical devices (horns) to electronic units consisting of microphones, amplifiers, analog filters, loudspeakers, batteries, etc. Digital signal processors replaced analog filters to provide better performance and new features. Central processors were added to provide many functions for monitoring and controlling other parts of the devices. Hearing systems have thus evolved into complex embedded system. Radio systems were added to allow hearing aids to communicate with accessories, auxiliary equipment, third-party products, etc. Many new features are enabled by such radio communication. Monitoring and controlling hearing aids from remote control devices or smart phones have been incorporated into several products. Direct audio streaming between hearing aids and dedicated streaming devices or smart phones is possible with some products. Also emerging are advanced features that are based on interactions with internet services, clouds, etc. Hearing systems are thus evolving into large and complex smart systems. Designing complex embedded systems or large smart systems are notoriously difficult. Many systems are still developed using document-based methods, where requirements and proposed architecture are described textually with the addition of a few figures and tables. Such documents cannot be subjected to testing, so it is impossible to predict the functionality and performance or even feasibility of the intended systems. Replacing documents with models have several advantages. Models can be simulated and analyzed such that functionality and performance can be predicted before any parts have been built. Potential flaws in the specification can therefore be corrected in early phases, which may reduce development effort and costs. This thesis concerns methods for identifying, selecting and implementing tools for various aspects of model-based systems engineering. A comprehensive method was proposed that include several novel steps such as techniques for analyzing the gap between requirements and tool capabilities. The method was verified with good results in two case studies for selection of a traceability tool (single-tool scenario) and a set of modeling tools (multi-tool scenarios). Models must be subjected to testing to allow engineers to predict functionality and performance of systems. Test-first strategies are known to produce good results in software development. This thesis concerns methods for test-driven modeling of hearing systems. A method is proposed for test-driven modeling of embedded systems of medium complexity. It utilizes formal model checking to guarantee functionality and performance. Test-driven design space exploration is enabled by using statistical model checking to obtain estimates that are verified formally at the final stages of the method. The method was applied with good results to a case study, where two solutions to a design problem were developed and verified. Feasible ranges for critical parameters were identified. Both solution conformed to all requirements. Smart systems are typically too large and complex to be verified by formal model checking, and the research showed that statistical model checking in its current form cannot be used for verifying such systems. A new method is therefore proposed for test-driven modeling of smart systems. The method uses formal verification of basic interactions. Simulations are used for verifying the overall system. To predict performance for scenarios that are too large to be simulated, the method uses mathematical forecasting based on simulating series of smaller scenarios, fitting simulation results to estimator functions, and extrapolating beyond the simulated data set. Mathematical forecasting allowed us to predict the performance of system scenarios that were much too large to be simulated. Such performance estimates may be somewhat imprecise but are nevertheless valuable because they provide answers that cannot be obtained otherwise. The research has thus proposed and verified methods for selecting modeling tools and for test-driven systems modeling for the benefit of GN Hearing and other organizations involved in development of complex embedded systems of large smart systems.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Embedded Systems Engineering, Copenhagen Center for Health Technology, GN Hearing Danmark A/S
Authors: Munck, A. (Intern), Madsen, J. (Intern), Lindqvist, L. (Intern), Schmidt, F. (Ekstern), Pop, P. (Intern)
Number of pages: 237
Publication date: 2017

Publication information
Publisher: DTU Compute
Original language: English
Series: DTU Compute PHD-2017
Volume: 444
ISSN: 0909-3192
Rationalization in architecture with surfaces foliated by elastic curves

We develop methods for rationalization of CAD surfaces using elastic curves, aiming at a cost-effective fabrication method for architectural designs of complex shapes. By moving a heated flexible metal rod through a block of expanded polystyrene, it is possible to produce shapes with both positive and negative Gaussian curvature, either for direct use or for use as moulds for concrete casting. If we can control the shape of the rod, while moving, we can produce prescribed shapes.

The flexible rod assumes at all times the shape of an Euler elastica (or elastic curve). The elastica are given in closed analytic form using elliptic functions. We use a gradient-driven optimization to approximate arbitrary planar curves by planar elastic curves. The method depends on an explicit parameterization of the space of elastic curves and on a method for finding a good initial guess for the optimization.

We approximate CAD surfaces by first extracting a collection of planar surface curves and approximating these by elastica. Providing the data for these curves to robots holding the flexible rod, we can produce an elastica-foliated surface that approximates the given CAD surface. Since not all surfaces can be closely approximated by an elastica-foliated surface, an arbitrary CAD surface must first be subdivided into segments that can be approximated. We discuss strategies for subdividing an arbitrary surface into segments that can be closely approximated, taking into account the aesthetics of the segmentation and the production constraints. If the given surface is smooth, we want the approximating surface to be smooth as well, so we must ensure smooth transition between the surface segments of the final result.

As an alternative to rationalization of arbitrary designs, we also present a method for direct generation of design surfaces using foliated Euler elastica. Here we work from a grid of blocks, so the segmentation is given, but we must still ensure smooth transition between segments.
to time-varying parameters and, in addition, the parameters within the regimes and the transition probabilities change over time. Using recursive and adaptive estimation techniques to capture this, we are able to better reproduce the volatility persistence that dynamic asset allocation benefits from. With this approach it is sufficient to distinguish between two regimes in stock returns in order for it to be profitable to change asset allocation based solely on the inferred regimes, both in a single and multi-asset universe. We advocate the use of model predictive control for translating forecasts into a dynamic strategy and controlling drawdowns by solving a multi-period optimization problem. We implement this based on forecasts from a multivariate hidden Markov model with time-varying parameters. Our results show that a substantial amount of value can be added by adjusting the asset allocation to the current market conditions, rather than rebalancing periodically to a static benchmark. By proposing a practical approach to drawdown control, we demonstrate the theoretical link to dynamic asset allocation and the importance of identifying and acting on regime shifts in order to limit losses and build robust portfolios.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems
Authors: Nystrup, P. (Intern), Madsen, H. (Intern), Hansen, B. W. (Ekstern), Larsen, H. O. (Forskerdatabase), Lindstrøm, K. J. E. L. (Ekstern)
Number of pages: 317
Publication date: 2017

Publication information
Publisher: DTU Compute
Original language: English
Series: DTU Compute PHD-2017
Volume: 465
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Risk management, Regime switching, Adaptive estimation, Forecasting, Model predictive control, Portfolio optimization, Drawdown control

Relations
Projects:
Dynamic Asset Allocation - Identifying Regime Shifts in Financial Time Series to Build Robust Portfolios
Publication: Research › Ph.D. thesis – Annual report year: 2017

Roots of the Chromatic Polynomial
The chromatic polynomial of a graph G is a univariate polynomial whose evaluation at any positive integer q enumerates the proper q-colourings of G. It was introduced in connection with the famous four colour theorem but has recently found other applications in the field of statistical physics. In this thesis we study the real roots of the chromatic polynomial, termed chromatic roots, and focus on how certain properties of a graph affect the location of its chromatic roots.

Firstly, we investigate how the presence of a certain spanning tree in a graph affects its chromatic roots. In particular we prove a tight lower bound on the smallest non-trivial chromatic root of a graph admitting a spanning tree with at most three leaves. Here, non-trivial means different from 0 or 1. This extends a theorem of Thomassen on graphs with Hamiltonian paths. We also prove similar lower bounds on the chromatic roots of certain minor-closed families of graphs.

Later, we study the Tutte polynomial of a graph, which contains the chromatic polynomial as a specialisation. We discuss a technique of Thomassen using which it is possible to deduce that the roots of the chromatic polynomial are dense in certain intervals. We extend Thomassen’s technique to the Tutte polynomial and as a consequence, deduce a density result for roots of the Tutte polynomial. This partially answers a conjecture of Jackson and Sokal.

Finally, we refocus our attention on the chromatic polynomial and investigate the density of chromatic roots of several graph families. In particular, we show that the chromatic roots of planar graphs are dense in the interval (3; 4), except for a small interval around 3 + 3 618, where 3 denotes the golden ratio. We also investigate the chromatic roots of related minor-closed classes of graphs and bipartite graphs.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic
Authors: Perrett, T. (Intern), Thomassen, C. (Intern)
Number of pages: 115
Publication date: 2017

Publication information
Place of publication: Kgs. Lyngby
Bayesian Modelling of Functional Whole Brain Connectivity

This thesis deals with parcellation of whole-brain functional magnetic resonance imaging (fMRI) using Bayesian inference with mixture models tailored to the fMRI data. In the three included papers and manuscripts, we analyze two different approaches to modeling fMRI signal; either we accept the prevalent strategy of standardizing of fMRI time series and model data using directional statistics or we model the variability in the signal across the brain and across multiple subjects. In either case, we use Bayesian nonparametric modeling to automatically learn from the fMRI data the number of functional units, i.e. parcels. We benchmark the proposed mixture models against state of the art methods of brain parcellation, both probabilistic and non-probabilistic.

The time series of each voxel are most often standardized using z-scoring which projects the time series data onto a hypersphere. This underlying manifold is often ignored and the data is modeled using Gaussian distributions. In one contribution, we show that using a mixture model based on the directional distribution, the von Mises-Fisher distribution, increase the reliability of inferred parcellations.

We develop a mixture model for modeling time-series using a Gaussian Process as a prior that is informed of the temporal dynamics of the data expected from the blood oxygenation level dependent (BOLD) signal. In two contributions, we explore the potential of this modeling framework. In the first, we show that this mixture model can delineate regions of task activation that can then be identified unsupervised. This forms a promising framework for unsupervised identification of task activated when the task design is unknown. In the final contribution, we evaluate the performance of the mixture model on the problem of clustering whole-brain fMRI. Based on both simulations on synthetic data and analysis of two fMRI datasets, we show that the model provides improved reliability of clustering compared to traditional clustering methods. Furthermore, the inferred parcellations provide the foundation for a method for increasing the reliability and sensitivity in analyses of task activation and for determining the networks of functionally connectivity in fMRI.

The proposed mixture models form promising tools for brain parcellation and we hope the methods can provide a nudge towards using probabilistic models for fMRI parcellation.
Inverse Optimization and Forecasting Techniques Applied to Decision-making in Electricity Markets

This thesis deals with the development of new mathematical models that support the decision-making processes of market players. It addresses the problems of demand-side bidding, price-responsive load forecasting and reserve determination. From a methodological point of view, we investigate a novel approach to model the response of aggregate price-responsive load as a constrained optimization model, whose parameters are estimated from data by using inverse optimization techniques.

The problems tackled in this dissertation are motivated, on one hand, by the increasing penetration of renewable energy production and smart grid technologies in power systems, that is expected to continue growing in the coming years. Non-dispatchable electricity generation cannot ensure a certain production at all times, since it depends on meteorological factors. Also, smart grid technologies are affecting the consumption patterns that the load traditionally exhibited. On the other hand, this thesis is motivated by the decision-making processes of market players. In response to these challenges, this thesis provides mathematical models for decision-making under uncertainty in electricity markets.

Demand-side bidding refers to the participation of consumers, often through a retailer, in energy trading. Under the smart-grid paradigm, the demand bids must reflect the elasticity of the consumers to changes in electricity price. Traditional forecasting models are typically not able to reflect this elasticity, hence we propose two novel approaches to estimate market bids. Both approaches are data-driven and take into account the uncertainty of future factors, as, for example, price. In both cases, demand-side bids that comprise a price-energy term decrease the expected imbalances and also increase the profit of retailers participating in electricity markets.

In the field of load forecasting, this thesis provides a novel approach to model time series and forecast loads under the real-time pricing setup. The relationship between price and aggregate response of the load is characterized by an optimization problem, which is shaped by a set of unknown parameters. Such parameters are estimated from data by using an inverse optimization framework. The usability of the proposed method is studied and we conclude that inverseoptimization-based modeling is a computationally attractive method that outperforms the forecasting capabilities of traditional time series models. Regarding the reserve determination, the special characteristics of the Danish power system do not allow for co-optimizing the unit commitment and reserve requirements. Hence, we propose a probabilistic framework, where the reserve requirements are computed based on scenarios of wind power and load forecast errors and power plant outages. The solution of the stochastic optimization models increases the safety of the overall system while decreases the associated reserve costs, with respect to the method currently used by the Danish TSO.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems
Authors: Saez Gallego, J. (Intern), Madsen, H. (Intern)
Number of pages: 190
Publication date: 2017

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 425
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd425_Gallego_JS.pdf

Relations
Projects:
Inverse Optimization and Forecasting Techniques Applied to Decision-making in Electricity Markets
Publication: Research › Ph.D. thesis – Annual report year: 2017

From Raw Data to Social Systems - Separating the Signal from the Noise in Smartphone Sensor Measurements
Digital tools for communication and information exchange have been ingrained in our lives. We google our information and we skype our parents. We use the Internet to shop for groceries, do banking, and study. We play massively multiplayer online games, belong to online communities, and date online. However, this does not mean that our lives have really moved to the digital domain. Even though the Internet makes it possible to exist without ever leaving the confines of our bedrooms, we still choose to meet our friends in person or to travel through physical, rather than virtual, space. There is a richness to personal contact and direct experience that has not yet been replaced by the digital services. Until this shift happens, we continue to analyze and investigate our offline lives in the pursuit for deepening our understaning of human nature. Digital breadcrumbs, which we leave behind with every online action, are relatively easy to collect. Capturing our offline behaviors, on the other hand, is not trivial. Scientist often rely on data that approximates only one aspect of our
lives. For example, mobile operator logs reveal who we call, but not who we meet. An alternative approach is to derive proxies of certain behaviors from smartphone sensor readings. Copenhagen Networks Study (CNS) employs this method, among others, to build the biggest dataset of the kind available to researchers in academia. The thesis shows a path from collecting raw smartphone data for CNS, through extracting increasingly meaningful information, to gaining novel insights into human behavior. Step by step, I turn a cryptic and seemingly uninteresting collection of hardware identifiers and received signal strengths into a detailed record of people’s lives: where they go, who they encounter, who they become friends with. I compare their offline activities and social ties to their online representations and find a surprisingly small overlap. The methods I propose the thesis constitute a more privacy-aware alternative to currently employed social sensing approaches. I show how to track the mobility and interactions of participants without sharing the results with third parties inadvertently. At the same time, the findings presented in this thesis emphasize the fragility of our privacy: the data we today consider as safe to share today, tomorrow might prove to carry rich information about our lives.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Copenhagen Center for Health Technology
Authors: Sapiezynski, P. (Intern), Jørgensen, S. L. (Intern), Hansen, L. K. (Intern)
Number of pages: 203
Publication date: 2017

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 415
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd415_Sapiezynski_P.pdf

Relations
Projects:
From Raw Data to Social Systems - Separating the Signal from the Noise in Smartphone Sensor Measurements
Publication: Research › Ph.D. thesis – Annual report year: 2017

Image reconstruction under non-Gaussian noise
During acquisition and transmission, images are often blurred and corrupted by noise. One of the fundamental tasks of image processing is to reconstruct the clean image from a degraded version. The process of recovering the original image from the data is an example of inverse problem. Due to the ill-posedness of the problem, the simple inversion of the degradation model does not give any good reconstructions. Therefore, to deal with the ill-posedness it is necessary to use some prior information on the solution or the model and the Bayesian approach.

Additive Gaussian noise has been extensively studied since it produces simple and tractable mathematical models. However, in the real applications, the noise is much more complicated and it cannot be well simulated by additive Gaussian noise, for instance, it may be signal dependent, very impulsive, multiplicative, mixed, etc. This PhD thesis intends to solve some of the many open questions for image restoration under non-Gaussian noise. The two main kinds of noise studied in this PhD project are the impulse noise and the Cauchy noise.

Impulse noise is due to for instance the malfunctioning pixel elements in the camera sensors, errors in analogue-to-digital conversion, faulty memory locations in hardware. Cauchy noise is characterized by a very impulsive behaviour and it is mainly used to simulate atmospheric and underwater acoustic noise, in radar and sonar applications, biomedical images and synthetic aperture radar images. For both noise models we introduce new variational models to recover the clean and sharp images from degraded images. Both methods are verified by using some simulated test problems. The experiments clearly show that the new methods outperform the former ones.

Furthermore, we have carried out a theoretical study on the two most known estimates: maximum a posteriori (MAP) estimate and conditional mean (CM) estimate for non-Gaussian noise. With only the convexity assumption on the data delity term, we introduce some cost functions for which the CM and MAP estimates are proper Bayes estimators and we also prove that the CM estimate outperforms the MAP estimate, when the error depends on Bregman distances.

This PhD project can have many applications in the modern society, in fact the reconstruction of high quality images with less noise and more details enhances the image processing operations, such as edge detection, segmentation, etc.
We present an improved solution to the dynamic partial sums problem in the ultra wide word-RAM model that simulates the RAMBO model to obtain a result by Brodnik et al. [BKMN06]. In this model, we manipulate words of size $w^2$ and access $w$ memory locations. Farzan et al. [FLONS15] additionally gave a solution to the partial sums problem in the ultra wide word-RAM model, recently introduced by Farzan et al. [FLONS15], where we, in constant time, are allowed to access, update and partial sums queries. This classic problem, and its variations, are very well studied in many different computational models [Fre82,FS89,Fen94, HSS11, HR03, HRS96, RRR01, PD04]. We solve the partial sums problem in $\log w$, the query times are faster. Dynamic Partial Sums in Constant Time and Succinct Space with the Ultra Wide Word-RAM Model The dynamic partial sums problem is to dynamically maintain an array of $n$ integers while supporting efficient access, update and partial sums queries. This classic problem, and its variations, are very well studied in many different computational models [Fre82,FS89,Fen94, HSS11, HR03, HRS96, RRR01, PD04]. We solve the partial sums problem in the ultra wide word-RAM model, recently introduced by Farzan et al. [FLONS15], where we, in constant time, are allowed to access words of size $w^2$ and access $w$ memory locations. Farzan et al. [FLONS15] additionally gave a solution to the dynamic partial sums problem by simulating the RAMBO model to obtain a result by Brodnik et al. [BKMN06]. In this paper we present an improved solution to the dynamic partial sums problem in the ultra wide word-RAM model that...
supports all operations in either constant or $O(\log \log n)$ time, depending on whether we allow multiplication, and succinct space. We pose as an open problem whether it is possible in the ultra wide word-RAM model to additionally support the classic select operation in constant time.

**General information**

**State:** Submitted  
**Organisations:** Department of Applied Mathematics and Computer Science, Algorithms and Logic  
**Authors:** Skjoldjensen, F. R. (Intern), Bille, P. (Intern), Gørtz, I. L. (Intern), Thomassen, C. (Intern)  
**Number of pages:** 80  
**Publication date:** 2017

**Publication information**

**Publisher:** DTU Compute  
**Original language:** English

**Series:** DTU Compute PHD-2017  
**Volume:** 446  
**ISSN:** 0909-3192  
**Main Research Area:** Technical/natural sciences

**Relations**

**Projects:**  
Matching and Compression of Strings with Automata and Word Packing  
**Publication:** Research › Ph.D. thesis – Annual report year: 2017

Towards Plug-n-Play robot guidance: Advanced 3D estimation and pose estimation in Robotic applications

Robots are a key technology in the quest for higher productivity in Denmark and Europe. Robots have existed in many years as a part of production lines where they have solved monotonous and repetitive task in mass production industries. Typical the programming of these robots are handled by engineers with special knowledge who have often raised the price for using robots to a given production task. If robots have to be applicable for small and medium sized enterprises where production task often changes and batch sizes are below 50 products it is necessary that the staff is capable of re-programming the robot by themselves.

During the last five years a number of collaborative robots are introduced on the marked e.g. Universal Robot, which enables a production worker to program the robot to solve simple tasks. With the collaborative robot the production worker is able to make the robot grind, mill, weld and move objects, which are physical located at the same positions. In order to place objects in the same position each time, custom-made mechanical fixtures and aligners are constructed to ensure that objects are not moving. It is expensive to design and build these fixtures and it is difficult to quickly change to a novel task. In some cases where objects are placed in bins and boxes it is not possible to position the objects in the same location each time.

To avoid designing expensive mechanical solutions and to be able to pick objects from boxes and bins, a sensor is necessary to guide the robot. Today, primarily 2D vision systems are applied in industrial robotics, which are in-flexible and hard to program for the production workers. Smart cameras, which are easier to re-configure and program to detect objects exist. However, computing the correct position such that a robot can move to this position is still a challenge which requires calibration processes. Moreover, the ability to make the solution robust such that it is running 24/7 in a production is demanding and requires the right skills. Basically, the vision part of a flexible automation solution is difficult to manage for a production worker while the robot motion programming is easily handled with the new collaborative robots. This thesis deals with robot vision technologies and how these are made easier for production workers program in order to get robots to recognize and compute the position of objects in the industry.

This thesis investigates and discusses methods to encapsulate a 2D vision system into a framework in order to make changes in production task easier. The framework is presented in [Contribution B] and [Contribution C] and demonstrates how re-configuration of vision systems is made easier but in the same time reviles some of the fundamental problems that exist by observing a tree dimensional world through a two dimensional vision system. This requires a calibration procedure every time in order to convert 2D to 3D, which still is a cumbersome process for a production worker.

For this reason, the rest of the thesis investigates and discusses how 3D computer vision techniques can ease the problem of recognizing and computing the position of objects. In [Contribution D] a small lightweight 3D sensor is presented. The 3D sensor has a size that makes it suitable for tool mounting at a collaborative robot. It is based on structured light principles and 3D estimation techniques, which enables fast and accurate acquisition of point clouds of low textured and reflective industrial objects.

In [Contribution E] a 3D vision system for easy learning of 3D models is presented. The system creates a 3D model of the object by scanning it from three views. Then the object acts as a reference model in the system when new instances of the object have to be located in the scene. With this approach fast re-configuration is possible. In [Contribution F] a new dataset for 3D object recognition and an evaluation of state-of-the-art local features for object recognition are presented.
The contribution shows as expected that state-of-the-art 3D object recognition algorithms are not good enough to locate industrial objects with few local shape features on the surface.

**Inferring human intentions from the brain data**

The human brain is a massively complex organ composed of approximately a hundred billion densely interconnected, interacting neural cells. The neurons are not wired randomly - instead, they are organized in local functional assemblies. It is believed that the complex patterns of dynamic electric discharges across the neural tissue are responsible for emergence of high cognitive function, conscious perception and voluntary action. The brain’s capacity to exercise free will, or internally generated free choice, has long been investigated by philosophers, psychologists and neuroscientists. Rather than assuming a causal power of conscious will, the neuroscience of volition is based on the premise that “mental states rest on brain processes”, and hence by measuring spatial and temporal correlates of volition in carefully controlled experiments we can infer about their underlying mind processes, including concepts as intriguing as “free will”, “agency” and “consciousness”. Recent developments in electrophysiology and neuroimaging methods allow for increasingly more accurate estimation of spatial and temporal characteristics of decision processes.

The work presented in this thesis is intended to contribute to our understanding of the dynamics of voluntary decision processes about prospective action. In the two presented studies we probe different types of decisions and compare them in terms of behavioral and EEG characteristics. We show that decision processes are manifested by complex, broadband modulation of brain oscillatory patterns, primarily in Alpha(8-12Hz) and Beta (16-30Hz) ranges. Our results suggest that decisions about whether to act or not, what type of action to perform, and about the timing of the action have distinct dynamic representations, and thus are to some extent mediated by different neural components. Furthermore, free action can be partially explained by low level behavioral preferences, especially in contexts where no explicit incentive favors one action over another.

Apart from the investigation of volition, considerable part of the work presented in this thesis is dedicated to experiment design methodology and efficient EEG processing methods. We have developed a dedicated, flexible Virtual Reality Environment (VRE) platform, suitable for investigation of volition and action preparation processes with range of modalities, including electroencephalography (EEG), functional magnetic resonance (fMRI), eye-tracking (ET) and behavioral measures. By providing ecologically valid, semi-realistic experience we aimed at reinforcing the natural decision processes and minimize the problem of random-sequence generation and fatigue in participants undergoing highly repeatable cognitive experiments. Other methodological contributions presented in the thesis are related to efficient, automatized and highly data-preserving methods for processing of EEG data, based on minimal number of arbitrarily selected parameters.
FindZebra - using machine learning to aid diagnosis of rare diseases
FindZebra is a search engine for rare diseases intended to act as a diagnosis decision support system (DDSS) capable of assisting the user both during and after a search. Rare diseases are diseases that affect only a small part of the population (less than one in two thousand). Currently around seven thousand rare diseases are known and it is estimated that 6–8% of the population will be affected by a rare disease during their lifetime. Due to their rarity and large number, diagnosis of rare diseases is difficult and often associated with year long delays and diagnostic errors. These difficulties with diagnosis have a profound human and societal cost. This means that even a small increase in success rate when using a tool such as FindZebra could potentially have a great impact on society. In this dissertation we explore four lines of research for improving FindZebra using machine learning methods. The first line of research is on how to improve the retrieval performance of FindZebra. By using a combination of improved models, medical databases and corpus expansion we show that it is possible to obtain a substantial improvement in retrieval performance compared to current state-of-the-art document retrieval systems. Improving retrieval performance is important, but is not the only way of improving the success rate of a DDSS such as FindZebra. Following an unsuccessful search, the search engine should assist the user by indicating what information is likely to be missing. This idea is called Information Completion (IC) and will be explored in the second line of research. In order to represent words (and other discrete tokens) in a neural network it is necessary to transform each word to a vector form. This is typically accomplished by using a word embedding, which is an essential component in any word based neural network. The third line of research is on how to improve this basic component. Users of FindZebra who do not have English as their primary language often have difficulty expressing complex medical queries in English. Optimally, a user should be able to write a query in his or her native language and the search engine should then give a suggestion for a differential diagnosis based on all the information contained in a multilingual corpus, not only in the native corpus. Methods for performing multilingual search will be the fourth line of research explored in this dissertation.

Secure Block Ciphers - Cryptanalysis and Design
The rapid evolution of computational devices and the widespread adoption of digital communication have deeply transformed the way we conduct both business and everyday life and they continue to do so. The ability to ensure confidentiality and integrity of information sent over digital channels is fundamental to this development and is absolutely
essential for all private and corporate communication, ranging from bank transactions, digital citizen services, and remote computer access, to cell phone calls and instant messaging. The vast majority of secured data sent over all types of networks is encrypted using so-called symmetric ciphers. The security of our digital infrastructure thus rests at its very base on their security.

The central topic of this thesis is the security of block ciphers – the most prominent form of symmetric ciphers. This thesis is separated in two parts. The first part is an introduction to block ciphers and their cryptanalysis, the second part contains publications written and published during the PhD studies. The first publication evaluates the security of a modification of the AES in which the choice of S-box is unknown to the attacker. We find that some of the attacks that can be applied to the AES can be transferred to this block cipher, albeit with a higher attack complexity. The second publication introduces a new block cipher family which is targeted for new applications in fully homomorphic encryption and multi-party computation. We demonstrate the soundness of the design and its superior performance in these applications. The third publication treats the cryptanalysis of Simon, a cipher proposed by the NSA. In particular we discuss how the methods of differential and linear cryptanalysis can correctly be applied to ciphers of this type. The fourth publication introduces a cryptanalytic framework which generalizes differential cryptanalysis. We demonstrate that attacks based on impossible transitions in this framework can competitively break round-reduced block ciphers in the low-data setting.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science
Authors: Tiessen, T. (Intern), Rechberger, C. (Intern), Knudsen, L. R. (Intern)
Number of pages: 151
Publication date: 2017

Publication Information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2016
Number: 412
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd412_Tiessen_T.pdf

Relations
Projects:
Secure Block Ciphers - Cryptanalysis and Design
Publication: Research › Ph.D. thesis – Annual report year: 2017

Characterization of absorption enhancers for orally administered therapeutic peptides in tablet formulations - Applying statistical learning
To develop a successful oral formulation of insulin for treatment of type-2 diabetes patients would be a great mile stone in terms of convenience. Besides protecting insulin from enzymatic cleavage in the small intestine, the formulation must overcome the intestinal epithelia barrier. Absorption enhancers are needed to ensure even a few percent of insulin are taken up. In thesis article 1, various methods to measure the effect of absorption enhancement and enzyme stability of insulin were applied. The major class of absorption enhancers is surfactant-like enhancers and is thought to promote absorption by mildly perturbing the epithelial membranes of the small intestine. The Caco-2 (Carcinoma Colon) cells can grow an artificial epithelial layer, and are used to test the potency of new absorption enhancers. This project was aimed to identify new absorption enhancers, that are both potent and sufficiently soluble. Quantitative structural activity relationship (QSAR) modeling is an empiric approach to learn relationships between molecular formulas and the biochemical properties using statistical models. A public data set testing the potency of absorption enhancers in Caco-2 was used to build a QSAR model to screen for new potent permeation enhancers. Thesis article 2 contains likely the first QSAR model to predict absorption enhancement. The model was verified by predicting molecules not tested before in Caco-2. The Caco-2 model overestimates the clinical effect of lipophilic permeation enhancers. In the Caco-2 model all reagents are pre-dissolved, and therefore the assay cannot predict critical solubility issues and bile salt interactions in the final tablet formulation. A QSAR solubility model was built to foresee and avoid slow tablet dissolution. Due to enzyme kinetics, slow tablet dissolution will allow most insulin to be deactivated by intestinal enzymes. The combined predictions of potency and solubility, will likely provide a more useful in-silico screening of potential permeation enhancers.

Random forest was used to learn relationships between molecular descriptors and potency or solubility. However, unlike multiple linear regression, the explicitly stated random forest model is complex, and therefore difficult to interpret and communicate. Any supervised regression model can be understood as a high dimensional surface connecting any possible combination of molecular properties with a given prediction. This high dimensional surface is also difficult to comprehend, but for random forests, it was discovered that a method, feature contributions, was especially useful to
decompose and visualize model structures. The visualization technique was named forest floor and could replace the otherwise widely use technique partial dependence plots, especially in terms of discovering interactions in the model structure. Thesis article 3 describes the forest floor method. An R package forestFloor was developed to compute feature contributions and visualize these according to the ideas of thesis article 3. Better interpretation of random forest models is an exciting interdisciplinary field, as it allows investigators of many backgrounds to find fairly complicated relationships in data sets without in advance specifying what parameters to estimate. Forest floor was used to explain how potency and solubility were predicted by random forest models.
Time-predictable Stack Caching
Embedded systems are computing systems for controlling and interacting with physical environments. Embedded systems with special timing constraints where the system needs to meet deadlines are referred to as real-time systems. In hard real-time systems, missing a deadline causes the system to fail completely. Thus, in systems with hard deadlines the worst-case execution time (WCET) of the real-time software running on them needs to be bounded.

Modern architectures use features such as pipelining and caches for improving the average performance. These features, however, make the WCET analysis more complicated and less precise. Time-predictable computer architectures provide solutions to this problem. As accesses to the data in caches are one source of timing unpredictability, devising methods for improving the time-predictability of caches are important. Stack data, with statically analyzable addresses, provides an opportunity to predict and tighten the WCET of accesses to data in caches.

In this thesis, we introduce the time-predictable stack cache design and implementation within a time-predictable processor. We introduce several optimizations to our design for tightening the WCET while keeping the time-predictability of the design intact. Moreover, we provide a solution for reducing the cost of context switching in a system using the stack cache. In design of these caches, we use custom hardware and compiler support for delivering time-predictable stack data accesses. Furthermore, for systems where compiler support or hardware changes are not practical, we propose and explore two different alternatives based on only software and only hardware support.

Model Manipulation for End-User Modelers
End-user modelers are domain experts who create and use models as part of their work. They are typically not Software Engineers, and have little or no programming and meta-modeling experience. However, using model manipulation languages developed in the context of Model-Driven Engineering often requires such experience. These languages are therefore only used by a small subset of the modelers that could, in theory, benefit from them.

The goals of this thesis are to substantiate this observation, introduce the concepts and tools required to overcome it, and provide empirical evidence in support of these proposals. To achieve its first goal, the thesis presents the findings of a Systematic Mapping Study showing that human factors topics are scarcely and relatively poorly addressed in model transformation research. Motivated by these findings, the thesis explores the requirements of end-user modelers, and proposes the VM* family of model manipulation languages addressing them. This family consists of the Visual Model Query Language (VMQL), the Visual Model Constraint Language (VMCL), and the Visual Model Transformation Language (VMTL). They allow modelers to specify and execute queries, constraints, and transformations using their modeling notation and editor of choice.

The VM* languages are implemented via a single execution engine, the VM* Runtime, built on top of the Henshin graph-
based transformation engine. This approach combines the benefits of flexibility, maturity, and formality. To simplify model editor integration, the VM* Runtime is deployed as a collection of lightweight Web Services. The claim that VM* languages offer end-user modelers superior learnability compared to existing model manipulation languages is verified empirically via user experiments complemented by qualitative evidence.

**General information**
State: Published  
Organisations: Department of Applied Mathematics and Computer Science, Software Engineering  
Authors: Acretoaie, V. (Intern), Störte, H. (Intern)  
Number of pages: 254  
Publication date: 2016

**Publication information**
Place of publication: Kgs. Lyngby  
Publisher: Technical University of Denmark (DTU)  
Original language: English  
Series: DTU Compute PHD-2016  
Number: 405  
ISSN: 0909-3192  
Main Research Area: Technical/natural sciences

**Computational Modeling of Medical Images of Brain Tumor Patients for Optimized Radiation Therapy Planning**
In brain tumor radiation therapy, the aim is to maximize the delivered radiation dose to the targeted tumor and at the same time minimize the dose to sensitive healthy structures – so-called organs-at-risk (OARs). When planning a radiation therapy session, the tumor and the OARs therefore need to be delineated on medical images of the patient’s head, to be able to optimize a radiation dose plan. In clinical practice, the delineation is performed manually with limited assistance from automatic procedures, which is both time-consuming and typically suffers from poor reproducibility. There is, therefore, a need for automated methods that can segment both brain tumors and OARs. However, there is a noticeable lack in the literature of methods that simultaneously segment both types of structures.

To automatically segment medical images of brain tumor patients is difficult because brain tumors vary greatly in size, shape, appearance and location within the brain. Furthermore, healthy structures surrounding a tumor are pushed and deformed by the so-called mass effect of the tumor. Moreover, medical imaging techniques often result in imaging artifacts and varying intensity across imaging centers.

The goal of this PhD-project was to develop automated segmentation methods that can handle both brain tumors and OARs. In the first part of the project, we developed a model for tumor shape and used it to develop a fully automated generative method specifically for brain tumor segmentation. This method performed favorably compared to other state-of-the-art methods. In the second part of the project, we used a probabilistic atlas-based model capable of detailed modeling of the spatial organization in a healthy brain, and extended it to handle various OARs. We incorporated this model into the previously used modeling framework. In experiments, we showed that the resulting model was capable of simultaneous segmentation of brain tumors and OARs, while also being capable of adapting to varying image sequences and images from different imaging centers.

**General information**
State: Submitted  
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics  
Authors: Agn, M. (Intern), Van Leemput, K. (Intern), Larsen, R. (Intern)  
Number of pages: 60  
Publication date: 2016

**Publication information**
Place of publication: Kgs. Lyngby  
Publisher: Technical University of Denmark (DTU)  
Original language: English  
Series: DTU Compute PHD-2016
Cryptanalysis of Selected Block Ciphers

The focus of this dissertation is to present cryptanalytic results on selected block ciphers. Block ciphers are the mathematical structure that will take a plaintext message and convert it into a ciphertext one block at a time using a secret key. They play an essential role in many cryptographic architectures and frameworks. For a long time they were known as the main building block that will provide confidentiality in an information system. They would also be able to represent a full spectrum of cryptographic services as many block ciphers can be used to construct stream ciphers, hash functions, pseudorandom number generators, and authenticated encryption designs.

For this reason a multitude of initiatives over the years has been established to provide a secure and sound designs for block ciphers as in the calls for Data Encryption Standard (DES) and Advanced Encryption Standard (AES), lightweight ciphers initiatives, and the Competition for Authenticated Encryption: Security, Applicability, and Robustness (CAESAR).

In this thesis, we first present cryptanalytic results on different ciphers. We propose attack named the Invariant Subspace Attack. It is utilized to break the full block cipher PRINTcipher for a significant fraction of its keys. This new attack also gives us new insights into other, more well-established attacks. In addition, we also show that for weak keys, strongly biased linear approximations exists for any number of rounds.

Furthermore, we provide variety of attacks on the family of lightweight block cipher SIMON that was published by the U.S National Security Agency (NSA). The ciphers are developed with optimization towards both hardware and software in mind. While the specification paper discusses design requirements and performance of the presented lightweight ciphers thoroughly, no security assessment is given. We present a series of observations on the presented construction that, in some cases, yield attacks, while in other cases may provide basis of further analysis by the cryptographic community. Specifically, The attacks obtained are using classical- as well as truncated differentials. In addition to that, we also investigate the security of SIMON against different linear cryptanalysis methods, i.e., classic linear, and linear hull attacks. We present a connection between linear characteristic and differential characteristic, multiple linear and differential and linear hull and differential, and employ it to adapt the current known results on differential cryptanalysis of SIMON to linear cryptanalysis results.

Finally, we investigate links between different methods of cryptanalysis and how they can be utilized for block cipher cryptanalysis. We consider the known results on the links among integral, impossible differential and zero-correlation linear hulls in order to prove that constructing a zero-correlation linear hull always implies the existence of an integral distinguisher. Moreover, we show that constructing zero-correlation linear hull on a Feistel structure with SP-type round functions, where P is a binary matrix, is equivalent to constructing impossible differential on the same structure except that P is substituted by the transposed matrix PT . We present an integral distinguishers of 5-round Feistel structure with bijective round functions and 3-round Feistel structure with round functions not necessarily being bijective. In addition to an integral distinguishers of Camellia so far, i.e., 7-round integral distinguishers of Camellia with FL/FL−1 layer and 8-round integral distinguishers of Camellia without FL/FL−1 layer.
Security Protocols: Specification, Verification, Implementation, and Composition

An important aspect of Internet security is the security of cryptographic protocols that it deploys. We need to make sure that such protocols achieve their goals, whether in isolation or in composition, i.e., security protocols must not suffer from any flaw that enables hostile intruders to break their security. Among others, tools like OFMC [MV09b] and Proverif [Bla01] are quite efficient for the automatic formal verification of a large class of protocols. These tools use different approaches such as symbolic model checking or static analysis. Either approach has its own pros and cons, and therefore, we need to combine their strengths. Moreover, we need to ensure that the protocol implementation coincides with the formal model that we verify using such tools.

This thesis shows that we can simplify the formal verification of protocols in several ways. First, we introduce an Alice and Bob style language called SPS (Security Protocol Specification) language, that enables users, without requiring deep expertise in formal models from them, to specify a wide range of real-world protocols in a simple and intuitive way. Thus, SPS allows users to verify their protocols using different tools, and generate robust implementations in different languages. Moreover, SPS has the "ultimate" formal semantics for Alice and Bob notation in the presence of an arbitrary set of cryptographic operators and their algebraic theory. Despite its generality, this semantics is mathematically simpler than any previous attempt.

Second, we introduce two types of relative soundness results that reduce complex verification problems into simpler ones. The first kind is typing results showing that if a security protocol, that fulfills a number of sufficient conditions, has an attack then it has a well-typed attack. The second kind considers the parallel composition of protocols, showing that if the parallel composition of two protocols, that fulfill a number of sufficient conditions, allows for an attack then one of the protocols, at least, has an attack in isolation. In fact, we unify and generalize over prior relative soundness results. The most important generalization is the support for all security properties of the geometric fragment proposed by [Gut14].

Modeling Structural Brain Connectivity

The human brain consists of a gigantic complex network of interconnected neurons. Together all these connections determine who we are, how we react and how we interpret the world. Knowledge about how the brain is connected can further our understanding of the brain’s structural organization, help improve diagnosis, and potentially allow better treatment of a wide range of neurological disorders.

Tractography based on diffusion magnetic resonance imaging is a unique tool to estimate this “structural connectivity” of the brain non-invasively and in vivo. During the last decade, brain connectivity has increasingly been analyzed using graph theoretic measures adopted from network science and this characterization of the brain’s structural connectivity has been shown to be useful for the classification of populations, such as healthy and diseased subjects. The structural connectivity of the brain estimated using tractography is, however, derived by integrating noisy estimates of the local fiber orientation in each voxel, entailing biases and limitations in the estimated connections and resulting in noisy graphs.

In this thesis, the ability of stochastic block models to extract the latent organization of structural brain connectivity graphs is investigated. It is found that both the stochastic block model and its non-parametric extension, the infinite relational model, are able to reliably extract a clustering that better accounts for structural connectivity than cortical atlases based solely upon surface morphology. Furthermore, a statistical prediction framework to quantify the ability of a cortical parcellation to account for structural connectivity is proposed. It is tested on two commonly used cortical atlases that are both based on surface morphology, as well as on a recently proposed cortical parcellation by Glasser et al. (2016) that is based on both task and resting-state functional magnetic resonance imaging, cortical thickness and myelin. It is found that
all three atlases capture the structural connectivity much better than random, but also that the parcellation based on multiple modalities is superior to those solely based on surface morphology.

The generation of structural brain connectivity graphs comprises a comprehensive processing pipeline, with various experimenter-defined parameters. The settings of these parameters are, however, unclear and this subjective aspect complicates the cross-comparison of studies investigating structural brain connectivity derived from tractography. Even though scan acquisition parameters, i.e. spatial resolution, angular resolution and b-value, are often discussed as possible factors influencing the final result, the impact of these factors on the derived structural connectivity graph has not yet been investigated. Herein, structural connectivity graphs, generated using different combinations of the three aforementioned acquisition parameters, are validated by comparison to a connectivity graph derived using invasive tracer injections in monkeys. It is found that the choice of acquisition parameters influences the derived structural connectivity graph and that higher angular resolution is always beneficial. Surprisingly, it is also found that higher spatial resolution does not improve the derived graph, but further investigation is needed to confirm this result.

General information
State: Submitted
Authors: Ambrosen, K. M. S. (Intern), Mørup, M. (Intern), Dyrby, T. B. (Intern), Schmidt, M. N. (Intern)
Number of pages: 70
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2016
Number: 443
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
Modeling Structural Brain Connectivity
Publication: Research › Ph.D. thesis – Annual report year: 2016

Topological Fluid Dynamics For Free and Viscous Surfaces
In an incompressible fluid flow, streamline patterns and their bifurcations are investigated close to wall for two-dimensional system and close to free and viscous surfaces in three-dimensional system. Expanding the velocity field in a Taylor series, we conduct a local analysis at the given expansion point. Applying the boundary conditions, some relations are obtained among the coefficients of the expansions. Series of coordinate transformations, which preserves the boundary conditions, are used to reduce the number of coefficients. Finally, using the normal form and unfolding theory, the velocity field is analysed structurally and bifurcation diagrams are obtained.

First, two-dimensional viscous flow close to wall for non-simple degenerate critical point is considered depending on three-parameter space. Second, threedimensional axisymmetric, viscous and steady flow is analysed close to free and viscous surfaces into three situations: Local analysis close to center axis; away from the axis and close to a stationary wall. Next, in the absence of axisymmetric condition, three-dimensional viscous flow is consider close to a free surface.

As an application of the bifurcation diagrams for three-dimensional axisymmetric viscous flow, three different shaped container driven by a rotating top disk is considered. Using a spectral collocation method, a code is constructed to obtain the meridional and swirl velocities. In a result of this code, all structural changes on the streamline patterns are observed and the occurring bifurcations are determined. These bifurcations are compared with the bifurcations obtained from topologically.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Mathematics
Authors: Balci, A. (Intern), Brøns, M. (Intern)
Number of pages: 194
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Modelling allergenic risk

Up to 20 million Europeans suffer from food allergies. Due to the lack of knowledge about why food allergies developed or how to protect allergic consumers from the offending food, food allergy management is mainly based on food allergens avoidance. The iFAAM project (Integrated approaches to Food Allergen and Allergy Management) aims at developing strategies for food allergies based on evidences.

Especially, food allergen risk assessment helps food producers or authorities to make decisions on withdrawing a food product from the market or adding more information on the label when allergen presence is unintended. The risk assessment method has three different kinds of input. The exposure is calculated from the product consumption and the allergen contamination in the food product. The exposure is then compared to the thresholds to which allergic individuals react in order to calculate the chance of allergic reaction in the population.

In allergen risk assessment, the emphasis was on the threshold data, and no effort was made on consumption data. Moreover, no pan-European consumption data suitable for allergen risk assessment are available. A procedure for grouping food products automatically across countries is proposed. Thus, the allergen risk assessment can be performed cross-nationally and for the correct food group.

Then the two probabilistic risk assessment methods usually used were reviewed and compared. First order Monte-Carlo simulations are used in one method [14], whereas the other one combines second order Monte-Carlo simulations with Bayesian inferences [13]. An alternative method using second order Monte-Carlo simulations was proposed to take into account the uncertainty from the inputs. The uncertainty propagation from the inputs to the risk of allergic reaction was also evaluated for all the methods using uncertainty analysis [11].

The recommended approach for the allergen risk assessment was implemented in a Shiny application with the R software. Thus, allergen risk assessment can be performed easily by non-statisticians with the interactive application.
requiring significant programming effort to exploit efficiently — software developers undertaking such a task will need all the help they can get, in order to keep the programming effort down.

In this thesis we champion using software to improve energy efficiency — in particular we develop guidelines for reasoning and evaluating software performance on modern computers, and a middleware that has been designed for modern computers, improving computational performance both in terms of energy and execution time. Our middleware consists of a new power manager, synchronization libraries using hardware transactional memory (for locks, barriers, and task synchronization), and two concurrent map data structures, which can be deployed in computer systems with little to no effort. At a fundamental level, we are improving computational performance by exploiting modern hardware features, such as dynamic voltage-frequency scaling and transactional memory. Adapting software is an iterative process, requiring that we continually revisit it to meet new requirements or realities; a time consuming process which we hope to simplify by analyzing the realities of modern computers, and providing guidelines explaining how to get the most performance out of them.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Language-Based Technology, Embedded Systems Engineering
Authors: Bonnichsen, L. F. (Intern), Probst, C. W. (Intern), Karlsson, S. (Intern)
Number of pages: 134
Publication date: 2016

Publication Information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 390
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd390_Bonnichsen_L_net.pdf

Relations
Projects:
An Adaptive Middleware for Improved Computational Performance
Publication: Research › Ph.D. thesis – Annual report year: 2016

Analysis of Security Protocols in Embedded Systems
Embedded real-time systems have been adopted in a wide range of safety-critical applications—including automotive, avionics, and train control systems—where the focus has long been on safety (i.e., protecting the external world from the potential damage caused by the system) rather than security (i.e., protecting the system from the external world). With increased connectivity of these systems to external networks the attack surface has grown, and consequently there is a need for securing the system from external attacks. Introducing security protocols in safety critical systems requires careful considerations on the available resources, especially in meeting real-time and resource constraints, as well as cost and reliability requirements. For this reason many proposed security protocols in this domain have peculiar features, not present in traditional security literature.

In this thesis we tackle the problem of analysing security protocols in safety critical embedded systems from multiple perspectives, extending current state-of-the-art analysis techniques where the combination of safety and security hinders our efforts. Examples of protocols in automotive control systems will follow throughout the thesis. We initially take a combined perspective of the safety and security features, by giving a security analysis and a schedulability analysis of the embedded protocols, with intertwined considerations. Then we approach the problem of the expressiveness of the tools used in the analysis, extending saturation-based techniques for formal protocol verification in the symbolic model. Such techniques gain much of their efficiency by coalescing all reachable states into a single set of facts. However, distinguishing different states is a requirement for modelling the protocols that we consider. Our effort in this direction is to extend saturation-based techniques so that enough state information can be modelled and analysed. Finally, we present a methodology for proving the same security properties in the computational model, by means of typing protocol implementations.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Language-Based Technology
Authors: Bruni, A. (Intern), Nielsen, F. (Intern), Nielsen, H. R. (Intern)
Protein Structure Refinement by Optimization

Proteins are the main active elements of life whose chemical activities regulate cellular activities. A protein is characterized by having a sequence of amino acids and a three-dimensional structure. The three-dimensional structure has only been determined experimentally for 50,000 of the seven million sequences that are known. Determining the protein structure from its sequence of amino acids is therefore a major problem in computational structural biology and is referred to as the protein folding problem. The folding problem is solved using de novo methods or comparative methods depending on whether the three-dimensional structure of a homologous sequence is known. Whether or not a protein model can be used for industrial purposes depends on the quality of the predicted structure. A model can be used to design a drug when the quality is high.

The overall goal of this project is to assess and improve the quality of a predicted structure. The starting point of this work is a technique called metric training where a knowledge-based protein potential, for a fixed set of native protein structures and a set of deformed decoys for each native structure, is designed to have native-decoy energy gaps that correlate maximally to a native-decoy distance. The main contribution of this thesis is methods developed for analyzing the performance of metrically trained knowledge-based potentials and for optimizing their performance while making them less dependent on the decoy set used to define them. We focus on using the gradient and the Hessian in the analysis and present a novel smooth solvation potential but otherwise the studied potential is kept close to standard coarse grained potentials.

We analyze the importance of the choice of metric both when used in metric training and when used in the evaluation of the performance of the resulting potential and find a significant improvement by using a metric based on intrinsic geometry. It is well-known that energy minimization of a potential that is efficient in ordering a fixed set of decoys need not bring the decoys closer to the native state. The next part of the work is focused on improving the convergence of decoy structures and we present a method that significantly improves the results of shorter energy minimizations of a metrically trained potential and discuss its limitations. In an ideal potential all near-native decoys will converge toward the native structure being at least a local minimum of the potential. To address how far the current functional form of the potential is from an ideal potential we present two methods for finding the optimal metrically trained potential that simultaneously has a number of native structures as a local minimum. Our results generally indicate that a more fine-grained potential is needed to meet desired model accuracies but even with our coarse-grained model we obtain good results and there is an unexplored possibility to combine it with comparative modeling.

To allow fast energy minimization in Matlab a new set of more sparse formulas to calculate the first and second derivatives of a molecular potential is derived and implemented.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing
Authors: Carlsen, M. (Intern), Røgen, P. (Intern)
Number of pages: 118
Publication date: 2016
Data Analysis of Medical Images: CT, MRI, Phase Contrast X-ray and PET

Data analysis of medical images is an important and growing area, as systems for imaging becomes still more available and complex.

The goal of the thesis is to demonstrate solutions to data analysis problems in a cross disciplinary context. Further, to develop methods for analysis of new imaging modalities and to combine cross disciplinary knowledge from various fields to find new solutions to existing problems.

More specifically the thesis shows segmentation of images, classification and statistics used on a variety of quite different problems. Active Appearance models, Chan-Vese and graph-cut has been used, as well as a variety of statistical tools centred on the General Linear Model.

The point of departure for the thesis is the NanoGuide project, in which gel based x-ray markers for use in radiotherapy has been developed. Two different types of gels has been analysed using segmentation of micro-CT images followed by a statistical analysis of homogeneity, contrast, degradation, and other qualities. By combining knowledge from the different professions in the project, a new application for one of the developed gels - in-vivo dosimetry in radiotherapy - has been studied.

Analysis of differences between groups and of correlations between brain regions and cognitive tests in alzheimers patients is another contribution. Segmentation of fat in abdominal MRI-scans has also been studied and a robust algorithm based on graph-cut is presented.

A relatively new modality phase-contrast x-ray and dark-field has shown promise for diagnosis of a variety of diseases in the lungs. A classification algorithm for differentiation of healthy, emphysematous and fibrotic lung tissue on pixel level is presented.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Image Analysis & Computer Graphics
Authors: Christensen, A. N. (Intern), Conradsen, K. (Intern), Larsen, R. (Intern)
Number of pages: 224
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 386
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd386_Christensen_AN.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Image Analysis for X-ray Imaging of Food

X-ray imaging systems are increasingly used for quality and safety evaluation both within food science and production. They offer non-invasive and nondestructive penetration capabilities to image the inside of food.

This thesis presents applications of a novel grating-based X-ray imaging technique for quality and safety evaluation of food products. In this effort the fields of statistics, image analysis and statistical learning are combined, to provide analytical tools for determining the aforementioned food traits.

The work demonstrated includes a quantitative analysis of heat induced changes in microstructure of meat products. A segmentation framework is presented, from which geometrical parameters are assessed. The grating-based method
embraces the complicated microstructure of the meat products, allowing for an analysis of the full three dimensional structure. The results illustrate that the combination of grating-based X-ray imaging and advanced analysis provides a valuable tool for microstructure analysis. Thus, the method can be considered as an alternative to other existing imaging techniques.

Furthermore, the thesis presents the application of grating-based X-ray imaging for novelty and defect detection in food. Compared to the complex three dimensional analysis of microstructure, here two dimensional images are considered, making the method applicable for an industrial setting. The advantages obtained by grating-based imaging are compared to conventional X-ray imaging, for both foreign object and defect detection. The results further emphasize the applicability of grating-based imaging for evaluation of food quality and food safety.
Decomposition and classification of electroencephalography data

This thesis is about linear and multi-linear analyses of electroencephalography (EEG) data and classification of estimated EEG sources. One contribution consists of an automatic classification method for independent components (ICs) of EEG data and a freely available implementation as an EEGLab plug-in, “IC Classification into Multiple Artefact Classes” (IC_MARC). Four artefact classes (blinks, heart beats, lateral eye movements, and muscle contractions), a neural class, and a mixed class (representing none or a mix of the other classes) were considered. We showed that classification is possible between subjects within studies over all classes. When generalising across studies a high classification rate of neural vs. non-neural ICs was retained but the multi-class performance dropped. In another study, we used IC_MARC to compare the ability to separate artefactual from neural sources of six linear decomposition methods. This study showed that high-pass filtering data at high cut-off frequencies improved artefact removal performances in an Event-Related Desynchronisation setting, providing similar performances of the three included Independent Component Analysis variants. IC_MARC was also used to inspect effects of artefacts on motor imagery based Brain-Computer Interfaces (BCIs) in two studies, where removing artefactual ICs had little performance impact. Finally, we investigated multi-linear classification on single trials of EEG data, proposing a rigorous optimisation approach. To enforce orthonormality of projection matrices, objective functions quantifying class discrimination were optimised on a cross-product of Stiefel (orthonormal matrix) manifolds. Supervised feature extraction outperformed unsupervised methods, but the choice of supervised method mattered less. We suggested completions of methods to include both PARAFAC and Tucker structures. The two structures provided similar performances, making the more interpretable PARAFAC models appealing.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Frølich, L. (Intern), Andersen, T. (Intern), Mørup, M. (Intern)
Number of pages: 208
Publication date: 2016

Prior Information in Inverse Boundary Problems

This thesis gives a threefold perspective on the inverse problem of inclusion detection in electrical impedance tomography: depth dependence, monotonicity-based reconstruction, and sparsity-based reconstruction.

The depth dependence is given in terms of explicit bounds on the datum norm, which shows the change in distinguishability of inclusions (support of an inhomogeneity) as they are placed closer towards the measurement boundary. This is done by determining eigenvalue bounds for differences of pseudodifferential operators on the boundary of the domain. Ultimately, the bounds serves as insight into how much noise that can be allowed in the datum before an inclusion cannot be detected.

The monotonicity method is a direct reconstruction method that utilizes a monotonicity property of the forward problem in order to characterize the inclusions. Here we rigorously prove that the method can be regularized against noise with a uniform regularization parameter, and that the method can be generalized to discrete electrode models. We give examples in 2D and 3D with noisy simulated data as well as real measurements, and give a comparison of reconstructions based on a non-linear and a linear formulation of the method.

Sparsity-based reconstruction is an iterative method, that through an optimization problem with a sparsity prior, approximates the inhomogeneities. Here we make use of prior information, that can cheaply be obtained from the monotonicity method, to improve both the contrast and resolution of the reconstruction. Numerical examples are given in both 2D and 3D for partial data using noisy simulated data as well as real measurements.

General information
Monitoring Animal Well-being

In recent years, animal well-being in industrial slaughterhouses has become a significant concern for consumers, farmers, and meat producers. Different groups have different interpretations of animal well-being. For the majority of consumers, animal well-being is highly influenced by their values and experiences. Meat producers are interested in the stress animals endure because it affects meat quality.

Pigs that arrive at slaughterhouses are more sensitive than usual for several reasons. In some cases, pigs are transported for long distances. Not all animals are used to transportation. Upon their arrival, it is common to mix pigs from different farmers in one area. Such mixing can cause fights between pigs, which can lead to additional stress or the animals being harmed. The unfamiliar environment also increases the animals’ stress levels. In some industrial slaughterhouses, up to 62,000 pigs per week are handled. Ensuring the well-being of such large numbers of pigs using only personnel is a complicated task.

Video surveillance of humans has been widely used to ensure safety and order in multiple situations. Methods have been developed to detect individual actions or abnormal behavior in small groups and dense crowds. In recent years, surveillance has also been used to monitor animals. Research has mainly focused on monitoring laboratory animals and farm animals. In both cases, animals are usually in constrained environment and cameras are used to cover all areas where animals are present. To obtain better results, non-intrusive markers or extracted features are used for tracking. Laboratory environments can be highly controlled; thus, no light and shadow noise are present in videos.

In slaughterhouses, the main focus is on monitoring large groups of animals in locations where additional markers cannot be used and pigs can leave or enter the surveilled area. In addition, pigs have a specific walking pattern; thus, motion analysis is not straightforward. The first aim of this thesis is to monitor the movement of pigs without using any additional markers or feature extraction in an unconstrained environment.

In video surveillance, the behavior of humans and animals is monitored based on extremes: event is present/event is not present, objects behave normally/objects behave abnormally, action 1/action 2/action 3, etc. In nature, the motion of humans and animals is continuous with transitions from one action to another. The second aim of this thesis is to propose a method to monitor motion as a continuous process using common classification methods.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing
Authors: Garde, H. (Intern), Knudsen, K. (Intern)
Number of pages: 56
Publication date: 2016
Modelling Socio-Technical Aspects of Organisational Security

Identification of threats to organisations and risk assessment often take into consideration the pure technical aspects, overlooking the vulnerabilities originating from attacks on a social level, for example social engineering, and abstracting away the physical infrastructure. However, attacks on organisations are far from being purely technical. After all, organisations consist of employees. Often the human factor appears to be the weakest point in the security of organisations. It may be easier to break through a system using a social engineering attack rather than a pure technological one. The StuxNet attack is only one of the many examples showing that vulnerabilities of organisations are increasingly exploited on different levels including the human factor. There is an urgent need for integration between the technical and social aspects of systems in assessing their security. Such an integration would close this gap, however, it would also result in complicating the formal treatment and automatic identification of attacks.

This dissertation shows that applying a system modelling approach to sociotechnical systems can be used for identifying attacks on organisations, which exploit various levels of the vulnerabilities of the systems. In support of this claim we present a modelling framework, which combines many features. Based on a graph, the framework presents the physical infrastructure of an organisation, where actors and data are modelled as nodes in this graph. Based on the semantics of the underlying process calculus, we develop a formal analytical approach that generates attack trees from the model.

The overall goal of the framework is to predict, prioritise and minimise the vulnerabilities in organisations by prohibiting the overall attack or at least increasing the difficulty and cost of fulfilling it. We validate our approach using scenarios from IPTV and Cloud Infrastructure case studies.

Probabilistic Approaches to Energy Systems

Energy generation from wind and sun is increasing rapidly in many parts of the world. This presents new challenges on how to integrate this uncertain, intermittent and non-dispatchable energy source. This thesis deals with forecasting and decision making in energy systems with a large proportion of renewable energy generation. Particularly we focus on producing forecasting models that can predict renewable energy generation, single user demand, and provide advanced forecast products that are needed for an efficient integration of renewable energy into the power generation mix. Such forecasts can be useful on all levels of the energy systems, ranging from the highest level, where the transmission system operator is concerned with minimizing system failures and is aided by wind power forecasts, to the end user of energy where power price forecasts are useful for users with flexible power demand.

The main contributions of this thesis lie in the realm of using gray box models to produce forecasts for energy systems. Gray box models can be defined as a crossover between physical models (or white box models), that base their model on a physical understanding of the system at hand, and data driven models (or black box models) that focus on accurately
describing the data without considering physical limitations of the system. Integrating these physical structures into a data driven approach allows for producing better forecasts with more accurate predictions. In this thesis we have developed and applied methodologies for gray box modeling to produce forecasts for vehicle driving patterns, solar irradiance, wind speeds, wind power, and solar power. The model for driving patterns has subsequently been used as input into an optimization algorithm for charging a single electric vehicle. In a subsequent study the behavior of a fleet of electric vehicles has been studied.

In the thesis we go through various examples of forecasts products and their applications. We emphasize that forecasting can not stand alone and should be complimented by optimization and decision making tools for an efficient integration of renewable energy. Thus forecast products should be developed in unison with the decision making tool as they are two sides of the same overall challenge.

Gabor frames on locally compact abelian groups and related topics
This thesis consists of four papers. The first one introduces generalized translation invariant systems and considers their frame properties, the second and third paper give new results on the theory of Gabor frames, and the fourth is a review paper with proofs and new results on the Feichtinger algebra.

The generalized translation invariant (GTI) systems provide, for the first time, a framework which can describe frame properties of both discrete and continuous systems. The results yield the well-known characterizations of dual frame pairs and Parseval frames of Gabor-, wavelet-, curvelet- and shearlet-type and for (generalized) shift-invariant systems and their continuous formulations.

This thesis advances the theory of both separable and non-separable, discrete, semicontinuous and continuous Gabor systems. In particular, the well established structure theory for separable lattice Gabor frames is extended and generalized significantly to Gabor systems with time-frequency shifts along closed subgroups in the time-frequency plane. This includes density results, the Walnut representation, the Wexler-Raz biorthogonality relations, the Bessel duality and the duality principle between Gabor frames and Gabor Riesz bases.

The theory of GTI systems and Gabor frames in this thesis is developed and presented in the setting of locally compact abelian groups, however, even in the euclidean setting the results given here improve the existing theory.

Finally, the thesis contains a review paper with proofs of all the major results on the Banach space of functions known as the Feichtinger algebra. This includes many of its different characterizations and treatment of its many equivalent norms, its minimality among all time-frequency shift invariant Banach spaces and aspects of its dual space, operators on the space and the kernel theorem for the Feichtinger algebra. The work also includes new findings such as a characterization among all Banach spaces, a forgotten theorem by Reiter on Banach space isomorphisms of the Feichtinger algebra, and new useful inequalities.
Design of Cognitive Interfaces for Personal Informatics Feedback
The emergence of embedded low-cost sensors in mobile devices allows us to capture unprecedented data about human behavior. Hence personal informatics systems are becoming an integrated part of our everyday life: Capturing various aspects from our health, work-life, to economic balance, and utility consumption. All of which are aimed to provide knowledge of oneself, on which we can reflect. Many personal informatics systems are characterized by mainly focusing on collecting and analyzing data, rather than translating the data into meaningful feedback. This dissertation presents challenges related to personal informatics systems, and propose an approach to design cognitive interfaces, which considers both users’ motivations, needs, and goals.

In this thesis I propose a new personal informatics framework, the feedback loop, which incorporates lean agile design principles. Including hierarchical modeling of goals, activities, and tasks to create minimal viable products. While considering how micro-interactions based on an understanding of data, couples with user needs and the context they appear in, can contribute to creating cognitive interfaces. Designing cognitive interfaces requires a focus on translating data into meaningful feedback, which the users can reflect on in order to gain insights. Thus I present tools such as personalized baselines and thresholds to enable reflection, while creating personalized goals, scenarios, trade-offs in order to provide actionable feedback, which can help users to adjust their behavior. Although feedback can be provided in many different ways, it basically consists of audio, visual, and haptic components, which combined may reinforce each other to support the underlying interaction.

The papers included in this thesis cover selected parts of the feedback loop. For instance, examining emotional responses to pleasant and unpleasant media content from brain activity, reveals the large amount of data and extensive analysis required to apply this to future personal informatics systems. In addition we analyse challenges related to temporal aspects of the feedback loop, when users attempt to self-regulate their brain activity based on a real-time feedback. This leads to identification of underlying audio, visual and haptic feedback components, which combined may support the underlying interaction within personal informatics. And with the emerging availability of sensor packed wearable devices, haptic feedback may become an inherent part of personal informatics systems, which could enhance the interaction based visual feedback.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Jensen, C. B. F. (Intern), Larsen, J. E. (Intern), Petersen, M. K. (Intern)
Number of pages: 160
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 383
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd383_Jensen_CBF.pdf

Relations
Projects:
Design of Cognitive Interfaces for Personal Informatics Feedback
Publication: Research › Ph.D. thesis – Annual report year: 2016
High Performance with Prescriptive Optimization and Debugging

Parallel programming is the dominant approach to achieve high performance in computing today. Correctly writing efficient and fast parallel programs is a big challenge mostly carried out by experts. We investigate optimization and debugging of parallel programs.

We argue that automatic parallelization and automatic vectorization is attractive as it transparently optimizes programs. The thesis contributes an improved dependence analysis for explicitly parallel programs. These improvements lead to more loops being vectorized, on average we achieve a speedup of 1.46 over the existing dependence analysis and vectorizer in GCC.

Automatic optimizations often fail for theoretical and practical reasons. When they fail we argue that a hybrid approach can be effective. Using compiler feedback, we propose to use the programmer’s intuition and insight to achieve high performance. Compiler feedback enlightens the programmer why a given optimization was not applied, and suggest how to change the source code to make it more amenable to optimizations. We show how this can yield significant speedups and achieve 2.4 faster execution on a real industrial use case.

To aid in parallel debugging we propose the prescriptive debugging model, which is a user-guided model that allows the programmer to use his intuition to diagnose bugs in parallel programs. The model is scalable, yet capable enough, to be general-purpose. In our evaluation we demonstrate low run time overhead and logarithmic scalability. This enable the model to be used on extremely large parallel systems.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Formal Methods, Embedded Systems Engineering
Authors: Jensen, N. B. (Intern), Probst, C. W. (Intern), Karlsson, S. (Intern)
Number of pages: 186
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 437
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
High Performance with Prescriptive Optimization and Debugging
Publication: Research › Ph.D. thesis – Annual report year: 2016

Influence of Maturation, Pathology and Functional Lateralization on 3D Sulcal Morphology using MRI

The folding of the cortex results in a characteristic pattern of folds called sulci and ridges called gyri. The cortical folding varies greatly both within and between individuals. Despite a century of sustained research, the mechanisms underlying the observed variation in folding is still largely unknown. The shape of cortical sulci and gyri are determined in part by forces exerted by white matter fiber connections between various cortical regions. Studying the shape of the cortical sulci hence contributes to the understanding of the variation in the folding.

This thesis concerns sulcal morphometry using Magnetic Resonance Imaging (MRI) and spatial statistical methods. The sulcal morphology has been studied with respect to: the normal development of a central sulcus; in relation to functional lateralization of the motor hand area in central sulcus and, finally, in relation to a pathological condition, anosmia, in the olfactory sulcus. This thesis describes and uses methods for sulci segmentation, sulci registration, sulci representation, and statistics for modeling sulci shape and testing sulcal morphology.

This thesis describes methods to analyze sulcal morphology and show how sulci variability are influenced under normal development, by a functional ability, and by pathological conditions.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Jensen, B. V. (Intern), Larsen, R. (Intern)
Enhanced Subsea Acoustically Aided Inertial Navigation

This thesis deals with enhancing state-of-the-art underwater acoustic–inertial navigation systems that are necessary for deep water robotic operations. Throughout the project intelligent and simple operational solutions to complex real-world problems was emphasized.

Offshore hydrocarbon, oil and gas, exploration is advancing further into treacherous territories such as deeper waters and arctic region. Deep underwater navigation poses a deluge of challenges; there is no such luxury as Global Navigation Satellite Systems (GNSS) underwater. Many of these challenges have been solved, but vessel time is expensive so lots of effort is put into cutting down on time spent on all tasks. Accuracy demanding tasks such as subsea construction and surveying are subject to strict quality control requirements taking up a lot of time. Offshore equipment is rugged and sturdy as the environmental conditions are harsh, likewise should the use of it be simple and robust to ensure that it actually works.

The contributions of this thesis are all focused on enhancing accuracy and time efficiency while bearing operational reliability and complexity strongly in mind. The basis of inertial navigation, the inertial sensors are treated in a calibration study with three scenarios: factory, in-field and at-sea calibration. Factory calibration compensates for sensor misalignments during the manufacturing process and for intrinsic sensor biases etc. For calibration a precise two-axis turn-table is required. It is shown that long-term effects on inertial sensors can be calibrated and assessed in-field, on land without specialized equipment, or at sea with certain realistic limitations and assumptions.

Automatic calibration of complex multi-sensor acoustic-inertial navigation systems, using parameter estimation, is employed on unprecedented high dynamic trajectories collected from sea-trials. These are needed to increase navigation accuracy to the cm-level and beyond. The same techniques can also be used for regular navigation in order to minimize both time and human error in parameter measurements.

In a unifying litmus test, the entire body of work is applied in a novel and potentially revolutionary methodology for the most challenging of all subsea survey and construction tasks: spool piece and jumper metrology. Two distinct approaches are investigated: One seeks to eliminate acoustic seabed transponders, but keep transponders at desired survey points; the other uses a mapping sensor such as subsea lidar to simply map the area in question. Both approaches are shown to work in practice. Generating high resolution maps, as the latter approach, is how the author anticipates all subsea surveys will be conducted in the near future.
Statistical modelling using CTSM-R

The ability to forecast the future of a system is more important than ever. Countless of applications require precise forecasting models to make smarter decisions in real time. Practitioners need software tools allowing an easy, quick and accurate implementation of their ideas. This thesis shows our software CTSM-R, which means: Continuous Time Stochastic Modelling in R. CTSMR provides the ability to implement statistical models and estimate unknown parameters. Thus allows modellers to model and understand the physical system of interest.

The main contributions of this thesis are the development of a generic software tool for grey box modelling. Grey box models use knowledge about the physical system of interest in combination with data to create models that accurately explain the data. The thesis will demonstrate how to implement linear and nonlinear models. These two model classes then serve as a building block for population models and spatiotemporal models. Population models known from the pharmaceutical industry where nonlinear population modelling has been long used in the industry to analyse trial data from many subjects. Spatiotemporal modelling extends CTSM-R to model complex correlations in space and time. This thesis demonstrates how to use spatio-temporal models for solar power forecasting.

CTSM-R is built in R using fast computations in Fortran when needed. CTSMR provides a simple interface which is quickly learned through our examples on our website http://ctsm.info.

Finally, this thesis demonstrates the importance of identifiability. A model is just a representation of the physical reality and such a representation is not guaranteed to be unique. Identifiability and the use of profile likelihood figures should be a standard tool of any modeller to verify the uniqueness of the solution.
The thesis results in a series of applications relating to CIs. The shape model can be used by CI-manufacturers for virtual product development and testing. At the same time, it can be applied to estimate the detailed inner ear shape from a clinical patient CT scan. This opens up for tools to optimize the programming of the CI, such that the hearing restoration is improved.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Kjer, H. M. (Intern), Paulsen, R. R. (Intern)
Number of pages: 219
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2015
Number: 381
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd381_Kjer_HM.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Design and analysis of cryptographic algorithms
In today’s world computers are ubiquitous. They can be found in virtually any industry and most households own at least one personal computer or have a mobile phone. Apart from these fairly large and complex devices, we also see computers on a much smaller scale appear in everyday objects in the form of micro-controllers and RFID chips.

What truly transformed our society are large scale networks, like the Internet or mobile telephone networks, which can link billions of devices. Our ways of communicating and conducting business have severely changed over the last decades due to this development. However, most of this communication happens over inherently insecure channels requiring methods to protect our communication. A further issue is the vast amount of data generated, which raises serious privacy concerns.

Cryptography provides the key components for protecting our communication. From securing our passwords and personal data to protecting mobile communication from eavesdroppers and our electronic bank transactions from manipulation. These applications would be impossible without cryptography.

The main topic of this thesis is the design and security analysis of the most fundamental algorithms used in cryptography, namely block ciphers and cryptographic hash functions. These algorithms are the building blocks for a vast amount of applications and play a vital role in providing both confidentiality and integrity for our communication.

This work is organized in two parts. First, an introduction to block ciphers and cryptographic hash functions is given to provide an overview over the state-of-the-art, the terminology, and how we can evaluate the security of an algorithm. The second part is a collection of scientific publications that have been written during the PhD studies and published.

In the first publication we analyze the security of cryptographic hash functions based on the AES and demonstrate practical attacks on reduced-round versions of these algorithms. The second publication provides cryptanalysis of the lightweight block cipher SIMON in particular how resistant this type of block ciphers are against differential and linear cryptanalysis. In the fourth publication we present a short-input hash function utilizing AES-specific instructions on modern CPUs in order to improve the performance of hashbased signature schemes. The last publication deals with the design of the tweakable lightweight block cipher Skinny which provides strong security bounds against differential and linear attacks while also competing with the performance of SIMON.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Cyber Security
Authors: Kölbl, S. (Intern), Rechberger, C. (Intern), Knudsen, L. R. (Intern)
Number of pages: 257
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Magnetic resonance imaging (MRI) is the de facto modality in neuroimaging studies, due to its superior image contrast in soft tissue. These studies often employ automated software pipelines that segments the image into structures and tissue. This reduces the time needed for analysis as well as statistical bias that may arise due to disagreements in delineations made by human experts. One such pipeline is Freesurfer.

This thesis presents results from the intervention study 'Preserving cognition, quality of life, physical health and functional ability in Alzheimer’s disease: the effect of physical exercise’ (ADEX), where longitudinal Freesurfer analysis was used to obtain segmentations of the hippocampal subfields and cortical regions in a subgroup of participants before and after a four-month exercise period. The participants performed moderate-to-high aerobic exercise for one hour, three times per week. The study hypothesized that the intervention would lead to reduced loss of tissue in the hippocampus and cortical regions, and that volumetric changes over time would correlate with cognitive performance measures. It was not possible to measure any effects in the hippocampus or cortical regions due to the intervention. However, it was found that exercise load (attendance and training intensity) correlated with changes in the hippocampus and in frontal and cingulate cortical thickness. Furthermore, changes in frontal and cingulate cortical thickness were found to correlate with changes in several cognitive performance measures, including mental speed, attention and verbal uency.

MRI suffers from an image artifact often referred to as the "bias field". This effect complicates automatized analysis of the images. For this reason, bias field correction is typical an early preprocessing step in many pipelines. Freesurfer currently employs the popular N3 bias field correction algorithm early in the pipeline, to solve this problem.

In this thesis, the reader is introduced to generative models for bias field correction. It is further shown how N3, which has traditionally been described as a "histogram sharpening" method, actually employs an underlying generative model, and that the bias field is estimated using an algorithm that is identical to generalized expectation maximization, but relies on heuristic parameter updates.

The thesis progresses to present a new generative model for longitudinal correction of the bias field, as well as a model that does not require brain masking or probabilistic, anatomical atlases in order to perform well. Finally, the thesis presents the realization of these models in the software package "Intensity Inhomogeneity Correction", which will be made publicly available.

Original language: English
Series: DTU Compute PHD-2016
Number: 434
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
Development and Application of Tools for MRI Analysis - A Study on the Effects of Exercise in Patients with Alzheimer's Disease and Generative Models for Bias Field Correction in MR Brain Imaging
Publication: Research › Ph.D. thesis – Annual report year: 2016

Magnetic resonance imaging (MRI) is the de facto modality in neuroimaging studies, due to its superior image contrast in soft tissue. These studies often employ automated software pipelines that segments the image into structures and tissue. This reduces the time needed for analysis as well as statistical bias that may arise due to disagreements in delineations made by human experts. One such pipeline is Freesurfer.

This thesis presents results from the intervention study 'Preserving cognition, quality of life, physical health and functional ability in Alzheimer’s disease: the effect of physical exercise’ (ADEX), where longitudinal Freesurfer analysis was used to obtain segmentations of the hippocampal subfields and cortical regions in a subgroup of participants before and after a four-month exercise period. The participants performed moderate-to-high aerobic exercise for one hour, three times per week. The study hypothesized that the intervention would lead to reduced loss of tissue in the hippocampus and cortical regions, and that volumetric changes over time would correlate with cognitive performance measures. It was not possible to measure any effects in the hippocampus or cortical regions due to the intervention. However, it was found that exercise load (attendance and training intensity) correlated with changes in the hippocampus and in frontal and cingulate cortical thickness. Furthermore, changes in frontal and cingulate cortical thickness were found to correlate with changes in several cognitive performance measures, including mental speed, attention and verbal uency.

MRI suffers from an image artifact often referred to as the "bias field". This effect complicates automatized analysis of the images. For this reason, bias field correction is typical an early preprocessing step in many pipelines. Freesurfer currently employs the popular N3 bias field correction algorithm early in the pipeline, to solve this problem.

In this thesis, the reader is introduced to generative models for bias field correction. It is further shown how N3, which has traditionally been described as a "histogram sharpening" method, actually employs an underlying generative model, and that the bias field is estimated using an algorithm that is identical to generalized expectation maximization, but relies on heuristic parameter updates.

The thesis progresses to present a new generative model for longitudinal correction of the bias field, as well as a model that does not require brain masking or probabilistic, anatomical atlases in order to perform well. Finally, the thesis presents the realization of these models in the software package "Intensity Inhomogeneity Correction", which will be made publicly available.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Larsen, C. T. (Intern), Van Leemput, K. (Intern)
Number of pages: 151
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 378
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd378_Larsen_CT.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016
Effective connectivity and gamma oscillations in a group at risk of psychosis

22q11.2 Deletion Syndrome (22q11.2DS) has been shown to be associated with a markedly increased risk for schizophrenia. Therefore, 22q11.2DS is a homogeneous genetic liability model which enables studies intending to identify functional abnormalities that may precede disease onset of schizophrenia. Being able to define these functional abnormalities could potentially assist in the search of biomarkers for schizophrenia. These are highly desired since early notification as well as early treatment have shown positive effects on everyday functioning in schizophrenia patients.

This thesis aimed at looking for functional abnormalities, known to be found in schizophrenia, in a cohort of 22q11.2 deletion carriers. The search for functional abnormalities in the 22q11.2 deletion syndrome cohort, were carried out measuring EEG while subjects engaged in a roving mismatch negativity (MMN) paradigm as well as an auditory steady state paradigm. Both of these paradigms are known to involve processes that are impaired in schizophrenia. This thesis ties together the three main contributions which are divided into three studies. In the first study, the responses to a roving MMN paradigm were assessed in 22q11.2 deletion carriers and healthy controls. Both conventional analysis of the MMN responses as well as a more sophisticated approach by means of Dynamic Causal Modelling (DCM) were carried out. DCM is a technique to extract effective connectivity between pre-specified brain areas. With this technique we investigated the underlying network of change detection in the two groups. While we found no indication of a reduced MMN response at the scalp level in the 22q11.2 deletion carriers, the underlying network of change detection differed as compared to healthy controls.

Second study extended study number one by employing a parametric DCM to study the underlying network of repetition suppression in 22q11.2 deletion carriers and healthy controls. While results for the control group indicated a processing in accordance with the predictive coding theory, this was not the case for the 22q11.2 deletion carriers.

Finally, in the third study the ability to generate 40 Hz cortical oscillations were assessed in 22q11.2 deletion carriers as well as healthy controls using an auditory steady state paradigm. Here, it was found that both phase and power of the 40 Hz oscillatory activity were reduced in 22q11.2 deletion carriers as compared to healthy controls.

In the three studies, results both similar and dissimilar to what is observed in the schizophrenia literature were found. The studies contribute in understanding the underlying pathology of 22q11.2 deletion syndrome and if results are confirmed by longitudinal follow up studies, the results might contribute in the search of biomarkers for schizophrenia.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Larsen, K. M. (Intern), Mørup, M. (Intern)
Number of pages: 118
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2016
Number: 440
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
Effective connectivity and gamma oscillations in a group at risk of psychosis
Publication: Research › Ph.D. thesis – Annual report year: 2016

Learned image representations for visual recognition
This thesis addresses the problem of extracting image structures for representing images effectively in order to solve visual recognition tasks. Problems from diverse research areas (medical imaging, material science and food processing) have motivated large parts of the methodological development. The solutions are inspired by and extend state-of-the-art techniques for describing and learning image content.

More specifically, the thesis explores two approaches to constructing image representations, namely feature engineering and feature learning. In the feature engineering approach, we devise a new image representation for texture-like patterns based on count statistics of second-order image structure. We demonstrate the discriminative capabilities of this representation on medical images and perform both cell classification and mitosis detection. Moreover, we develop an object identification method based on vector quantized local image descriptors allowing us to distinguish individual meat cuts along a production line and trace them in a non-intrusive manner. In the feature learning approach, we propose to solve the task of segmenting scanning electron microscopy images of calcite crystals by learning a meaningful pixel description to facilitate the actual segmentation. Finally, we present a new unsupervised generative image model
addressing the problem of pixel-based similarity measures for images. We propose a scheme for employing feature-based similarity measures and demonstrate how this improves the ability to learn high-level concepts in images of faces.

The thesis argues in favor of learning features and presents new methods for domains with limited amounts of labeled data allowing feature learning to be applied more broadly.

Design and Analysis of Symmetric Primitives
The subject of this thesis is the study of symmetric cryptographic primitives. We investigate these objects from three different perspectives: cryptanalysis, design and implementation aspects.

The first part deals with cryptanalysis of symmetric primitives, where one tries to leverage a property of the design to achieve some adversarial goal. Two of the most successful types of cryptanalysis are differential- and linear attacks. We apply variants of differential cryptanalysis to the lightweight block cipher SIMON which was proposed by researchers from the National Security Agency (NSA) in 2013. In particular, we present a search heuristic to find differentials of high probability, and we investigate the clustering of characteristics known as the differential effect. Finally, we apply impossible differential attacks using truncated differentials to a number of SIMON variants. Next, we define a theoretical model for key-less linear distinguishers, which captures the meaning of distinguishing a block cipher from an ideal permutation using linear cryptanalysis, when the key is either known or chosen by the adversary. Such models exist using differential properties but were never before defined using linear cryptanalysis. We apply this model to the standardized block cipher PRESENT. Finally, we present very generic attacks on two authenticated encryption schemes, AVALANCHE and RBS, by pointing out severe design flaws that can be leveraged to fully recover the secret key with very low complexity.

In the second part, we delve into the matter of the various aspects of designing a symmetric cryptographic primitive. We start by considering generalizations of the widely acclaimed Advanced Encryption Standard (AES) block cipher. In particular, our focus is on a component operation in the cipher which permutes parts of the input to obtain dependency between the state bits. With this operation in focus, we give a range of theoretical results, reducing the possible choices for the operation in generalized ciphers to a particular set of classes. We then employ a computer-aided optimization technique to determine the best choices for the operation in terms of resistance towards differential- and linear cryptanalysis. Also in the vein of symmetric primitive design we present PRØST, a new and highly secure permutation. Employing existing third-party modes of operation, we present six proposals based on PRØST for the ongoing CAESAR competition for authenticated encryption with associated data. We describe the design criteria, the usage modes and give proofs of security.

Finally, in the third part, we consider implementation aspects of symmetric cryptography, with focus on high-performance software. In more detail, we analyze and implement modes recommended by the National Institute of Standards and Technology (NIST), as well as authenticated encryption modes from the CAESAR competition, when instantiated with the AES. The data processed in our benchmarking has sizes representative to that of typical Internet traffic. Motivated by a significant improvement to special AES instructions in the most recent microarchitecture from Intel, codenamed Haswell, our implementations are tailored for this platform. Finally, we introduce the comb scheduler which is a low-overhead look-ahead strategy for processing multiple messages in parallel. We show that it significantly increases the throughput for sequential modes of operation especially, but also for parallel modes to a lesser extent.
Fine-grained Information Flow for Concurrent Computation
It is essential to protect IT systems against security threats. An example would be the control of aircraft, which uses an internal network that passengers can access. It is important to ensure that malicious code on passenger equipment cannot endanger flight safety.

Information flow control is an important approach to the protection of systems against such threats. Notable examples include tainting analyses in languages such as Javascript, and program transformations on cryptographic algorithms to avoid information leakage through running time. A wide variety of techniques, including type systems and reference monitors, have been proposed in the context of programming languages and process calculi, to enforce such properties. The most widely used definitions of information flow security are noninterference-like properties.

For concurrent systems where processes communicate with each other to accomplish computational tasks, fine-grained security policies can be formulated by distinguishing between whether communication can happen, and what is communicated. As the first contribution of this PhD thesis, we formulate a noninterference-like property that takes all combinations of sensitivity levels for “whether” and “what” into consideration, emphasizing the importance of the integrity case where the former is more sensitive than the latter. This case captures the effect of Message Authentication Codes (MAC) and the consequence of Denial of Service (DoS) attacks. It is also proved that the property degenerates to a classical one when the two dimensions are intentionally blurred.

As the second contribution, we focus on the “what” dimension and further allow the flow policy to vary under different contents stored and communicated. This is the area of content-dependent (or conditional) information flow, which has recently been studied for sequential programs. We generalize the use and enforcement of content-dependent flow policies to concurrent, communicating processes. A security type system is developed, incorporating a Hoare logic component that provides approximations of the memory contents at different program points. Most proofs for the theoretical results on content-dependency are performed in the Coq proof assistant.

The third contribution of this thesis is the obtainment of compositionality results that support modular security analyses of computer systems.

A multiplexer pattern that separates sensitive and non-sensitive network traffic is used as a running example. Whether communications can happen is easily influenced by an attacker — attacking one of the incoming channels would suffice. In any case, the two data paths are still differentiable by the sensitivity levels of what is communicated. In case the destinations of messages are determined by their tagging, content-dependent policies are able to convey the correlation between the sensitivity level of a message and its tagging, and our Hoare-logic equipped type system allows a modular analysis of the overall system.
Analysis of Ant Colony Optimization and Population-Based Evolutionary Algorithms on Dynamic Problems

This thesis presents new running time analyses of nature-inspired algorithms on various dynamic problems. It aims to identify and analyse the features of algorithms and problem classes which allow efficient optimization to occur in the presence of dynamic behaviour. We consider the following settings:

\(\lambda\)-MMAS on Dynamic Shortest Path Problems. We investigate how increasing the number of ants simulated per iteration may help an ACO algorithm to track optimum in a dynamic problem. It is shown that while a constant number of ants per-vertex is sufficient to track some oscillations, there also exist more complex oscillations that cannot be tracked with a polynomial-size colony.

MMAS and \((\mu+1)\) EA on Maze. We analyse the behaviour of a \((\mu + 1)\) EA with genotype diversity on a dynamic fitness function Maze, extended to a finite-alphabet search space. We prove that the \((\mu + 1)\) EA is able to track the dynamic optimum for finite alphabets up to size \(\mu\), while MMAS is able to do so for any finite alphabet size.

Parallel Evolutionary Algorithms on Maze. We prove that while a \((1 + \lambda)\) EA is unable to track the optimum of the dynamic fitness function Maze for offspring population size up to \(\lambda = O(n^{1-\epsilon})\), a simple island model with \(\Omega(\log n)\) islands is able to do so if the migration interval is chosen appropriately.

Migration Topology in Island Models. We investigate the impact of the migration topology on the performance of an island model optimizing a Maze-like dynamic function, demonstrating that in some cases, a less-dense migration topology is preferable to a complete migration topology.

\((1+1)\) EA on Generalized Dynamic OneMax. We analyze the \((1 + 1)\) EA on dynamically changing OneMax, re-proving known results on first hitting times using modern drift analysis, and providing a new anytime analysis showing how closely the EA can track the dynamically moving optimum over time. These results are also extended to a finite-alphabet search space.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic
Authors: Lissovoi, A. (Intern), Witt, C. (Intern)
Number of pages: 130
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark, Transport
Original language: English
Series: DTU Compute PHD-2016
Number: 404
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd404_Lissovoi_A.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Predicting the emotions expressed in music
With the ever-growing popularity and availability of digital music through streaming services and digital download, making sense of the millions of songs, is ever more pertinent. However the traditional approach of creating music systems has treated songs like items in a store, like books and movies. However music is special, having origins in a number of evolutionary adaptations. The fundamental needs and goals of a users use of music, was investigated to create the next generation of music systems. People listen to music to regulate their mood and emotions was found to be the most important fundamental reason. (Mis)matching peoples mood with the emotions expressed in music was found to be an essential underlying mechanism, people use to regulate their emotions. This formed the basis and overall goal of the thesis, to investigate how to create a predictive model of emotions expressed in music. To use in the next generation of music systems.
The thesis was divided into three main topics involved in creating a predictive model 1) Elicitation of emotion, 2) Audio representation and 3) Modelling framework, associating the emotion and audio representation, allowing to predict the emotions expressed in music.

The traditional approach of quantifying musical stimuli on the valence and arousal representation of emotions using continuous or likert scales was questioned. An outline of a number of bias and the so-called confidence effect when using bipolar scales led to the use of relative scales in the form of pairwise comparisons. One issue with pairwise comparisons is the scaling, this was solved using an active learning approach through a Gaussian Process model.

Traditional audio representation disregards all temporal information in audio features used for modelling the emotions expressed in music. Therefore a probabilistic feature representation framework was introduced enabling both temporal and non-temporal aspects to be coded in discrete and continuous features. Generative models are estimated for each feature time-series and used in a discriminative setting using the Probability Product Kernel (PPK) allowing the use of this approach in any kernel machine.

To model the pairwise comparisons directly, a Generalized Linear Model, a kernel extension and a Gaussian Process model were used. These models can predict the ranking of songs on the valence and arousal dimensions directly. Furthermore use of the PPK allowed to find optimal combinations of both feature and feature representation using Multiple Kernel Learning.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Copenhagen Center for Health Technology
Authors: Madsen, J. (Intern), Larsen, J. (Intern), Hansen, L. K. (Intern)
Number of pages: 184
Publication date: 2016

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 369
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd369_Madsen_J.pdf

**Relations**

Projects:

**Predicting the emotions expressed in music**
Publication: Research › Ph.D. thesis – Annual report year: 2016

**Enhancing Security and Privacy in Video Surveillance through Role-Oriented Access Control Mechanism**

Use of video surveillance has significantly increased in the last few decades. Modern video surveillance systems are equipped with techniques that automatically extract information about the objects and events from the video streams and allow traversal of data in an effective and efficient manner. Pervasive usage of such systems gives substantial powers to those monitoring the videos and poses a threat to the privacy of anyone observed by the system. Aside from protecting privacy from the outside attackers, it is equally important to protect the privacy of individuals from the inside personnel involved in monitoring surveillance data to minimize the chances of misuse of the system, e.g. voyeurism. In this context, several techniques to protect the privacy of individuals, called privacy enhancing techniques (PET) have therefore been proposed in the literature which detect and mask the privacy sensitive regions, e.g. faces, from the videos. However, very few research efforts have focused on addressing the security aspects of video surveillance data and on authorizing access to this data. Interestingly, while PETs help protect the privacy of individuals, they may also hinder the usefulness of video surveillance systems resulting in compromising the very purpose of such systems, i.e. public safety. Thus the challenge is to provide sufficient need-specific data to those monitoring the surveillance systems yet preserving the privacy of people as much as possible. This can be achieved through a dynamic access control mechanism that may provide proportionate access to data while allowing reversing the PETs whenever required. In this context, a summary of thesis contributions is given below.

In this thesis, we present an abstract model of video surveillance systems that helps identify the major security and privacy requirements in a video surveillance system. We study existing solutions against these requirements and point out practical challenges in ensuring the security of video surveillance data in all stages (in transit and at storage). Our study
shows a gap, between the security requirements that we identified and the proposed security solutions, where future research efforts may focus in this domain. From the challenges that we outline regarding security in video surveillance, we focus on development of a dynamic access control mechanism.

We develop a general-purpose access control model that is suitable for video surveillance systems as well as other domains sharing similar requirements. As the currently dominant access control models – the role-based access control (RBAC) and the attribute-based access control (ABAC) – suffer from limitations while offering features complementary to each other, their integration has become an important area of research. Our access control model combines the two models in a novel way in order to unify their benefits while avoiding their limitations. Our approach provides a mechanism that not only takes information about the current circumstances into account during access control decision making, but is also suitable for applications where access to resources is controlled by exploiting the contents of resources in the access control policy. We evaluate our model against RBAC and ABAC and demonstrate that our model brings together the benefits offered by RBAC and ABAC while addressing the role- and permission-explosion issues faced in RBAC.

Based on our access control model, we then present an access control mechanism for video surveillance systems. Contrary to the existing approaches, the proposed access control mechanism is role-oriented and retains advantages associated with role-based access control, yet it allows specification of policies using the metadata associated with the objects as well as the attributes of users and environment. In addition to role hierarchies, the content-based permissions in our model allow derivation of several permissions from the explicitly stated ones due to the hierarchical relations between the attributes of different entities. We implement a prototype of the proposed mechanism and demonstrate that the access control policies using our approach may be specified via eXtensible Access Control Markup Language (XACML).

**Graph Decompositions**

The topic of this PhD thesis is graph decompositions. While there exist various kinds of decompositions, this thesis focuses on three problems concerning edgedecompositions. Given a family of graphs $H$ we ask the following question: When can the edge-set of a graph be partitioned so that each part induces a subgraph isomorphic to a member of $H$? Such a decomposition is called an $H$-decomposition. Apart from the existence of an $H$-decomposition, we are also interested in the number of parts needed in an $H$-decomposition.

Firstly, we show that for every tree $T$ there exists a constant $k(T)$ such that every $k(T)$-edge-connected graph whose size is divisible by the size of $T$ admits a $T$-decomposition. This proves a conjecture by Barát and Thomassen from 2006.

Moreover, we introduce a new arboricity notion where we restrict the diameter of the trees in a decomposition into forests. We conjecture that for every natural number $k$ there exists a natural number $d(k)$ such that the following holds: If $G$ can be decomposed into $k$ forests, then $G$ can be decomposed into $k + 1$ forests in which each tree has diameter at most $d(k)$. We verify this conjecture for $k \leq 3$. As an application we show that every 6-edge-connected planar graph contains two edge-disjoint 18/19-thin spanning trees.

Finally, we make progress on a conjecture by Baudon, Bensmail, Przybylo, and Wozniak stating that if a graph can be decomposed into locally irregular graphs, then there exists such a decomposition with at most 3 parts. We show that this conjecture is true if the number 3 is replaced by 328, establishing the first constant upper bound for this problem.
On Practical Sampling of Bidirectional Reflectance

Accurate material models are a key part in producing convincing, photo-realistic, images in computer graphics. Elaborate analytical models exist, allowing graphics designers to manually design material appearance. However, given the complex nature and wide variability of material appearance, measuring this from the real world is an impractical and time-consuming process. Having a practical way of measuring material appearance will not only be of great value to the graphics community, but also open up for a wide range of new application areas, including industrial production quality control, digital prototyping and manufacturing, and interactive real-time product visualization.

In this thesis, the challenge of making material appearance measurements practical is addressed. Specifically, the Bidirectional Reflectance Distribution Function (BRDF), which is the quantity describing material appearance, is thoroughly analysed using both optimisation tools and multivariate statistics, in search of making BRDFs more accessible.

The work demonstrated includes an insight into the challenges of fitting analytical models to measured data and on the compromises one is bound to make when simplifying the real world with a parametric BRDF model. Specifically we identify what error measures work well for obtaining perceptually good results and how a simple BRDF model may be modified to better match real world data. With an offset in this, a linear, data-driven, BRDF model is proposed and a framework for reconstructing full and accurate BRDFs from only a few measurements is presented. It is here demonstrated that with as little as 20 point-samples, a BRDF can accurately be reconstructed. Furthermore utilising the field of view of a camera, this may be reduced to as little as two images. With this, the thesis demonstrates how BRDF measurements can be made practical, and it exemplifies this with a range of datasets intended for various purposes, each including high quality measured BRDFs.

Where the classical approach to BRDF capture may take weeks in measurement time, we here successfully demonstrate that is can in fact be reduced to no more than minutes or even seconds using our framework.
Interactive Topology Optimization

Interactivity is the continuous interaction between the user and the application to solve a task. Topology optimization is the optimization of structures in order to improve stiffness or other objectives. The goal of the thesis is to explore how topology optimization can be used in applications in an interactive and intuitive way. By creating such applications with an intuitive and simple user interface we allow non-engineers like designers and architects to easily experiment with boundary conditions, design domains and other optimization settings. This is in contrast to commercial topology optimization software where the users are assumed to be well-educated both in the finite element method and topology optimization.

This dissertation describes how various topology optimization methods have been used for creating cross-platform applications with high performance. The user interface design is based on theory of from human-computer interaction which is described in Chapter 2. Followed by a description of the foundations of topology optimization in Chapter 3. Our applications for topology optimization in 2D and 3D are described in Chapter 4 and a game which trains the human intuition of topology optimization is presented in Chapter 5. Topology optimization can also be used as an interactive modeling tool with local control which is presented in Chapter 6. Finally, Chapter 7 contains a summary of the findings and concludes the dissertation.

Most of the presented applications of the thesis are available at: http://www.topopt.dtu.dk.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Nobel-Jørgensen, M. (Intern), Bærentzen, J. A. (Intern)
Number of pages: 124
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 375
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd375_Nobel_Joergensen_M.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Approaches for Accommodating Demand Response in Operational Problems and Assessing its Value

This thesis deals with the development of operational models of demand response and the evaluation of this novel resource within existing frameworks for power system dispatch and market clearing.

Increasing shares of power generation from variable renewable sources, and climate change policies that discourage the use of fossil fuel intensive power plants, are among the factors that are currently driving the evolution of power systems towards greater flexibility. Activating the latent flexibility of electricity consumption through demand response can contribute towards facilitating this evolution. However, before the necessary investments can be made to establish and operate this novel resource, its value must be determined.

As with all current power system resources, if distributed demand response is deployed on a large scale it will be required to interface with the power system and market operators through established frameworks. Such frameworks are not suited to interaction with large numbers of individual flexible loads, so it is necessary to establish a representation of their aggregated flexibility that can be effectively communicated to system and market operators. In this thesis we introduce the concept of a saturation curve, which represents the flexibility offered by a broad class of flexible loads capable of providing load shifting demand response: thermal-electric loads such as refrigeration and heating. From this saturation curve we extract dispatch and market offering structures for demand response that respect the physical characteristics and constraints of the individual flexible loads within an aggregate population, while being limited in complexity to that allowable within current operational power system frameworks.

An evaluation of demand response must consider both the social welfare value it generates by reducing overall power system operation costs, and the commercial value it can accrue by participating in competitive electricity markets. Social
welfare value provides an indicator of the viability of any new power system resource, but does not guarantee that the
necessary investments will be made to establish and operate the resource. A positive commercial assessment will signal
to investors that the resource can offer a return on their investment, and that it can thrive in a competitive environment. We
consider both the social welfare and commercial value of demand response in this thesis, by simulating the deployment of
our specialised operational models of demand response within power system dispatch frameworks and by developing
innovative trading strategies for demand response on the day-ahead and intraday markets.

We find through the combined modelling and analysis contained in this thesis that the value offered by demand response
is very low under current power system conditions, and when it is restricted to operating within existing operational
frameworks. Prices and costs on the studied power systems are insufficient to allow demand response to generate
significant value or revenue through energy arbitrage or load curtailment. This does not rule out that there maybe certain
power systems, or sections thereof, that are currently experiencing sufficient resource scarcity to result in a favourable
environment for the successful implementation of demand response. At the current time however, our research finds that
the outlook for the widespread deployment of demand response is poor.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Dynamical Systems
Authors: O'Connell, N. (Intern), Madsen, H. (Intern)
Number of pages: 237
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2015
Number: 401
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd401_OConnel_N.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Aspects of the Tutte polynomial
This thesis studies various aspects of the Tutte polynomial, especially focusing on the Merino-Welsh conjecture.

We write \( T(G;x,y) \) for the Tutte polynomial of a graph \( G \) with variables \( x \) and \( y \). In 1999, Merino and Welsh conjectured that
if \( G \) is a loopless 2-connected graph, then

\[
T(G;1,1) \leq \max\{T(G;2,0), T(G;0,2)\}.
\]

The three numbers, \( T(G;1,1) \), \( T(G;2,0) \) and \( T(G;0,2) \) are respectively the numbers of spanning trees, acyclic orientations
and totally cyclic orientations of \( G \).

First, I extend Negami’s splitting formula to the multivariate Tutte polynomial. Using the splitting formula, Thomassen and I
found a lower bound for the number of spanning trees in a \( k \)-edge-connected graph. Our bound is tight for \( k \) even, but for \( k \)
odd we give a slightly better lower bound which we believe is not tight. We prove that the minimum number of spanning
trees in a 3-edge-connected graph with \( n \) vertices is, not surprisingly, significantly smaller than the minimum number of
spanning trees in a 4-edge-connected graph. However, we conjecture that the minimum number of spanning trees of a 5-
edge-connected graph is actually obtained by a 6-edge-connected graph asymptotically.

Thomassen proved the following partial result for the Merino-Welsh conjecture. Assume the graph \( G \) is loopless,
bridgeless and has \( n \) vertices and \( m \) edges.

If \( m \leq 1.066(n-1) \) then \( T(G;1,1) \leq T(G;2,0) \).
If \( m \geq 4(n-1) \) then \( T(G;1,1) \leq T(G;0,2) \).

I improve in this thesis Thomassen’s result as follows:

If \( m \leq 1.29(n-1) \) then \( T(G;1,1) \leq T(G;2,0) \).
If \( m \geq 3.58(n-1) \) and \( G \) is 3-edge-connected then \( T(G;1,1) \leq T(G;0,2) \).

Strengthening Thomassen’s idea that acyclic orientations dominate spanning trees in sparse graphs, I conjecture that the
ratio \( T(G;2,0)/T(G;1,1) \) increases as \( G \) gets sparser. To support this conjecture, I prove a variant of the conjecture for
series-parallel graphs.

The Merino-Welsh conjecture has a stronger version claiming that the Tutte polynomial is convex on the line segment between (2,0) and (0,2) for loopless 2-connected graphs. Chavez-Lomeli et al. proved that this holds for coloopless paving matroids, and I provide a shorter proof of their theorem. I also prove it for minimally 2-edge-connected graphs. As a general statement for the convexity of the Tutte polynomials, I show that the Tutte polynomial of a sparse paving matroid with fixed rank is almost never convex in the first quadrant. In contrast, I conjecture that the Tutte polynomial of a sparse paving matroid is convex on the line segment between (2,0) and (0,2) for loopless 2-connected graphs. Chavez-Lomeli et al. proved that this holds for coloopless paving matroids, and I provide a shorter proof of their theorem. I also prove it for minimally 2-edge-connected graphs. As a general statement for the convexity of the Tutte polynomials, I show that the Tutte polynomial of a sparse paving matroid with fixed rank is almost never convex in the first quadrant.

The following multiplicative version of the Merino-Welsh conjecture was considered by Noble and Royle:

\[ T(G;1,1)2 \leq T(G;2,0) T(G;0,2). \]

Noble and Royle proved that this multiplicative version holds for series-parallel graphs, using a computer algorithm that they designed. Using a property of the splitting formula which I found, I improve their algorithm so that it is applicable to the class of graphs with bounded treewidth (or pathwidth). As an application, I verify that the multiplicative version holds for graphs with pathwidth at most 3.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic
Authors: Ok, S. (Intern), Thomassen, C. (Intern)
Number of pages: 110
Publication date: 2016

Publication information

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 384
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd384_Ok_S.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Motion Tracking of Infants in Risk of Cerebral Palsy

Every year 2-3 out of 1000 infants are born with cerebral palsy. Among others, the disorder often affects motor, cognitive and perceptual skills. The disorder is usually detected when the infants are old enough the crawl and walk, i.e. when the infant is 1-2 years old. However, studies show that the infant's movements are affected already in the first year of life and methods exist for assessing the movements. The methods often require observation of the movements and qualitative evaluation of these. A more objective measure is desired in order to be able to diagnose cerebral palsy much earlier.

The goal with this thesis is to describe the development of a markerless motion tracking system for infants. Based on data recorded with a low-cost depth sensor, image analysis and mathematical modeling is used to model the infant's body and its movements. Two methods are considered, where the first method is able to do single frame pose estimation, based on simple assumptions on the infant's body. The second method uses an articulated model that incorporates anatomical constraints. Combining the two methods results in a robust motion tracking system for infants.

The results from the motion tracking are used to extract physical features such as velocity and acceleration of the individual body parts. A novel method for estimating scene ow in human motion data is presented, utilizing the results from the motion tracking. A number of examples are given for potential applications for automatic assessment of infant movement. This includes a preliminary study on automatic classification of movements related to cerebral palsy.

The contributions included in this thesis can be divided into two groups. The first two contributions consider the analysis in order to estimate and track the body of the infants. The remaining contributions consider different motion features derived from the motion tracking results. Both pose and motion features are extracted and used for assessing the infants’ motor development.

The presented work is a step closer to automatic motion assessment of infants with focus on early diagnosis of infants with cerebral palsy. Further collaboration with clinicians can result in breakthroughs in the way infants are monitored and assessed during the early years of life.
The main motivation is to be able to assess infants in risk of cerebral palsy based on the previously established connection between infant movement and brain injuries. However, as the data used in this study is recorded simultaneously with the study, the true outcome is not known. Even though some of the included infants were born preterm, none of them have to date been diagnosed with cerebral palsy.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Olsen, M. D. (Intern), Paulsen, R. R. (Intern)
Number of pages: 199
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2015
Number: 393
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd393_Olsen_MD.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Modeling and Control for Price Responsive Electricity Loads
This thesis deals with the development of model-based control architectures to facilitate renewable sources integration in the power system, focusing on residential buildings. Energy use in buildings in developed countries is increasing rapidly, and advanced model based techniques for control of thermal storages are becoming popular due to the high demand for solutions that improve energy efficiency and reduce operating costs.

This presents new challenges on how to integrate uncertain and intermittent energy sources. This work proposes methods for control of price responsive electricity loads in future energy systems and methods for handling stochasticity of, e.g., wind and solar power production. Hierarchies of aggregators and predictive controllers, in flexible demand side response, are implemented to achieve a balance with the non-dispatchable energy production.

Particular focus is given on producing models for control that facilitate better planning for an efficient integration of renewable energy into the power generation. Combining both data and statistical expertise, opens up new possibilities for designing models that describe thermal storages flexibility.

Finally, focus is also put on the problem of managing a virtual power plant equipped with stochastic energy sources and flexible consumers. Two distinct control approaches are described: direct control of the load consumption and indirect control by broadcasting a price signal. The advantages and challenges of these two approaches are discussed providing examples for suggested techniques.

General information
State: Submitted
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Scientific Computing
Authors: Parvizi, J. (Intern), Madsen, H. (Intern), Jørgensen, J. B. (Intern)
Number of pages: 184
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2016
Number: 427
ISSN: 0909-3192
Main Research Area: Technical/natural sciences

Relations
Projects:
Modeling and Control for Price Responsive Electricity Loads
Economic Model Predictive Control for Spray Drying Plants

The main challenge in cost optimal operation of a spray dryer, is to maximize the production rate while minimizing the energy consumption, keeping the residual moisture content of the powder below a maximum limit and avoiding that the powder sticks to the chamber walls. The conventional PI control strategy is simple, but known to be insufficient at providing optimal operation in the presence of variations in the feed and the ambient air humidity. This motivates our investigation of Model Predictive Control (MPC) strategies.

In this thesis, we consider the development and application of new models and MPC strategies to optimize the operation of four-stage spray dryers. The models are first-principle dynamic models with parameters identified from dryer specific experiments and powder properties identified from laboratory tests. A simulation model is used for detailed closed-loop simulations and a complexity reduced control model is used for state estimation and prediction in the controllers. These models facilitate development and comparison of control strategies. We develop two MPC strategies; a linear tracking MPC with a Real-Time Optimization layer (MPC with RTO) and an Economic Nonlinear MPC (E-MPC). We tailor these for the spray drying process to optimize the cost of operation by adjustments to the inputs of the dryer according to the present disturbances and process constraints. Simulations show that MPC strategies improve the profit of operation by up to 9.69%, the production of powder by up to 9.6%, the residual moisture content by up to 0.114 p.p. and the energy efficiency by up to 6.06% while the produced powder is within the given quality specifications and sticky powder on the walls of the chamber is avoided. Thus, we are able to improve the cost of operation significantly compared to the conventional PI control strategy.

The proposed MPC strategies are based on a feedback control algorithm that explicitly handles constrained control inputs and uses a model to predict and optimize the future behavior of the dryer. The solution of the control problem results in a sequence of inputs for a finite horizon, out of which only the first input is applied to the dryer. This procedure is repeated at each sample instant and is solved numerically in real-time. The MPC with RTO tracks a target that optimizes the cost of operation at steady-state. The E-MPC optimizes the cost of operation directly by having this objective directly in the controller. The need for the RTO layer is then eliminated.

We demonstrate the application of the proposed MPC with RTO to control an industrial GEA MSDTM-1250 spray dryer, which produces approximately 7500 kg/hr of enriched milk powder. Compared to the conventional PI controller, our first results shows that the MPC improves the profit of operation by approximately 228,000 €/year, the product rate by 322 kg/hr, the residual moisture content by 0.166 p.p. and the energy efficiency by 1% at comparable ambient air humidity conditions. The demonstrated MPC with RTO is fully integrated in the daily operation of the spray dryer today.

Our primary objectives in the thesis are: 1) Spray dryer modeling of a smallscale four-stage spray dryer. The purpose of the models are to enable simulations of the spray drying process at different operating points, such that the models facilitate development and comparison of control strategies; 2) Development of MPC strategies that automatically adjust the dryer to variations in the feed and the ambient air humidity, such that the energy consumption is minimized, the residual moisture content in the powder is controlled within the specifications and sticky powder is avoided from building up on the dryer walls; 3) Demonstrate the industrial application of an MPC strategy to a full-scale industrial four-stage spray dryer.

The main scientific contributions can be summarized to:

- Modeling of a four-stage spray dryer. We develop new first-principles engineering models for simulation of a four-stage spray dryer. These models enables simulations of the spray dryer at different operating points with high accuracy.
- Development and simulation of control strategies. We develop two control strategies, the MPC with RTO and the E-MPC strategy. The performance of the controllers is studied and evaluated by simulation.
- Industrial application of MPC to a spray dryer. We demonstrate that our proposed MPC with RTO is applicable to an industrial GEA MSDTM-1250 spray dryer, that produces enriched milk powder.

**General information**
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing, Copenhagen Center for Health Technology, Dynamical Systems
Authors: Petersen, L. N. (Intern), Jørgensen, J. B. (Intern), Poulsen, N. K. (Intern)
Number of pages: 229
Publication date: 2016

**Publication Information**
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2016
Computational Analysis of Brain Images: Towards a Useful Tool in Clinical Practice

Due to its excellent soft tissue contrast and versatility, magnetic resonance imaging (MRI) has become arguably the most important tool for studying the structure and disorders of the human brain. Although in recent years tremendous advances have been made in automatic segmentation of brain MRI scans, many of the developed methods are not readily extendible to clinical applications due to the variability of clinical MRI data and the presence of pathologies, such as tumors or lesions. Thus, clinicians are forced to manually analyze the MRI data, which is a time consuming task and introduces rater-dependent variability that reduces the accuracy and sensitivity of the results.

The goal of this PhD-project was to enlarge the scope of the automatic tools into clinical applications. In order to tackle the variability of the data and presence of pathologies, we base our methods on Bayesian generative modeling, which combines detailed prior models of the human neuroanatomy and pathologies with models of the MRI imaging process. This approach allows us to describe the observed MRI data in a principled manner, and to integrate explicit models of different disease effects and imaging artifacts into the framework when needed.

This thesis presents an introduction to the theory behind the generative modeling approach, and an overview of the main results. The first part concentrates on segmenting different neuroanatomical structures in MRI scans of healthy subjects, and the second part describes how this framework can be extended with models of brain lesions. This results in a set of fast, robust and fully automatic tools for segmenting MRI brain scans of both healthy subjects and subjects suffering from brain disorders such as multiple sclerosis. Having access to quantitative measures of both lesions and the surrounding structures opens up avenues for clinicians to study the effect of these type of disorders on the full brain anatomy. This could potentially help in discovering sensitive biomarkers for early diagnosis and tracking of disease development.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Puonti, O. (Intern), Van Leemput, K. (Intern), Larsen, R. (Intern)
Number of pages: 180
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 396
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd396_Puonti_O.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Simultaneous Reconstruction and Segmentation with Class-Specific Priors

Studying the interior of objects using tomography often require an image segmentation, such that different material properties can be quantified. This can for example be volume or surface area. Segmentation is typically done as an image analysis step after the image has been reconstructed. This thesis investigates computing the reconstruction and segmentation simultaneously. The advantage of this is that because the reconstruction and segmentation are computed jointly, reconstruction errors are not propagated to the segmentation step. Furthermore the segmentation procedure can be used for regularizing the reconstruction process. The thesis provides models and algorithms for simultaneous reconstruction and segmentation and their performance is empirically validated.

Two method of simultaneous reconstruction and segmentation are described in the thesis. Also, a method for parameter selection is given. The reconstruction and segmentation are modeled as two parts: the image that is reconstructed and a so-called Hidden Markov Measure Field Model (HMMFM). Pixel values in the image contain material attenuation coefficients and the HMMFM contains pixelwise probabilities for material classes. The number of material classes and their parameters are assumed known a priori. These parameters are the mean value of the class attenuation coefficients and their standard deviations. Given this input together with projection data, the problem is to find the image and HMMFM. The segmentation is obtained from the HMMFM as the most probable class in each pixel.
The solution for the reconstruction and segmentation problem is found using an algorithm that simultaneously minimizes the reprojection error, deviation of the grey levels of pixels from known mean values and the spatial differences in the class probabilities.

In the first Simultaneous Reconstruction and Segmentation (SRS) method data is assumed Gaussian distributed and the minimization is done using standard optimization techniques in two stages. Experimental validation on both phantom and real data shows that modeling the reconstruction and segmentation simultaneously has superior performance, especially when the problem is underdetermined, i.e. when the number of unknowns in the reconstruction exceeds the number of observations.

The second SRS method assumes Poisson distributed data, which is the case for data originating from discrete events like photon counts. The algorithm is again based on solving a minimization problem. In addition a relaxation strategy is employed in order to avoid being stuck in local minimum. This model is also validated on artificial data.

Selecting appropriate regularization parameters can be difficult, so the last thing that we consider is a parameter selection approach. The most promising approach was a modified L-curve algorithm, which was empirically analyzed.

This thesis contributes with methods for simultaneous reconstruction and segmentation and demonstrates the benefits of this approach in situations where only few projections are available and data is noisy. Here a higher precision image as well as segmentation can be computed.

General information
State: Published
Authors: Romanov, M. (Intern), Dahl, A. B. (Intern), Hansen, P. C. (Intern)
Number of pages: 134
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2015
Number: 397
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd397_Romanov_M.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Methods and Algorithms for Economic MPC in Power Production Planning
This thesis concerns methods and algorithms for power production planning in contemporary and future power systems. Power production planning is a task that involves decisions across different time scales and planning horizons. Hours-ahead to days-ahead planning is handled by solving a mixed-integer linear program for unit commitment and economic dispatch of the system power generators. We focus on a minutes-ahead planning horizon, where unit commitment decisions are fixed. Economic model predictive control (EMPC) is employed to determine an optimal dispatch for a portfolio of power generators in real-time. A generator can represent a producer of electricity, a consumer of electricity, or possibly both. Examples of generators are heat pumps, electric vehicles, wind turbines, virtual power plants, solar cells, and conventional fuel-fired thermal power plants. Although this thesis is mainly concerned with EMPC for minutes-ahead production planning, we show that the proposed EMPC scheme can be extended to days-ahead planning (including unit commitment) as well.

The power generation from renewable energy sources such as wind and solar power is inherently uncertain and variable. A portfolio with a high penetration of renewable energy is therefore a stochastic system. To accommodate the need for EMPC of stochastic systems, we generalize certainty-equivalent EMPC (CEEMPC) to mean-variance EMPC (MV-EMPC). In MV-EMPC, the objective function is a trade-off between the expected cost and the cost variance. Simulations show that MV-EMPC reduces cost and risk compared to CE-EMPC. The simulations also show that the economic performance of CE-EMPC can be much improved using a constraint back-off heuristic.

Efficient solution of the optimal control problems (OCPs) that arise in EMPC is important, as the OCPs are solved online. We present special-purpose algorithms for EMPC of linear systems that exploit the high degree of structure in the OCPs. A Riccati-based homogeneous and self-dual interior-point method is developed for the special case, where the OCP objective function is a linear function. We design an algorithm based on the alternating direction method of multipliers
(ADMM) to solve input-constrained OCPs with convex objective functions. The OCPs that occur in EMPC of dynamically decoupled subsystems, e.g. power generators, have a block-angular structure. Subsystem decomposition algorithms based on ADMM and Dantzig-Wolfe decomposition are proposed to solve these OCPs. Subproblems that arise in the decomposition algorithms are solved using structure-exploiting algorithms. To reduce computation time of the EMPC algorithms further, warm-start and early-termination strategies are employed. Benchmarks show that the special-purpose algorithms are significantly faster than current state-of-the-art solvers.

As a potential application area of EMPC, we study power production planning in small isolated power systems. A critical part of power production planning in small isolated power systems is operational reserve planning. The operational reserves are activated to balance production and consumption in real-time. An EMPC scheme is presented for activation of operational reserves. Simulations based on a Faroe Islands case study show that significant cost savings can be achieved using this strategy. For efficient planning of the operational reserves, we present an optimal reserve planning problem (ORPP). The ORPP is a contingency-constrained unit commitment problem that addresses low inertia challenges in small isolated power systems.

In summary, the main contributions of this thesis are:
- A mean-variance optimization strategy for EMPC of linear stochastic systems.
- Tailored algorithms for solution of the OCPs that arise in EMPC of linear stochastic systems.
- Methods for power production planning in small isolated power; the ORPP for unit commitment and economic dispatch, and an EMPC scheme for activation of operational reserves.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing, Center for Energy Resources Engineering, Dynamical Systems
Authors: Sokoler, L. E. (Intern), Jørgensen, J. B. (Intern), Madsen, H. (Intern), Poulsen, N. K. (Intern)
Number of pages: 249
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 377
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd377_Sokoler_LE.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2016

Time-Predictable Communication on a Time-Division Multiplexing Network-on-Chip Multicore
This thesis presents time-predictable inter-core communication on a multicore platform with a time-division multiplexing (TDM) network-on-chip (NoC) for hard real-time systems. The thesis is structured as a collection of papers that contribute within the areas of: reconfigurable TDM NoCs, static TDM scheduling, and time-predictable inter-core communication.

More specifically, the work presented in this thesis investigates the interaction between hardware and software involved in time-predictable inter-core communication on the multicore platform. The thesis presents: a new generation of the Argo NoC network interface (NI) that supports instantaneous reconfiguration, a TDM traffic scheduler that generates virtual circuit (VC) configurations for the Argo NoC, and software functions for two types of intercore communication.

The new generation of the Argo NoC adds the capability of instantaneously reconfiguring VCs and it addresses the identified shortcomings of the previous generation. The VCs provide the guaranteed bandwidth and latency required to implement time-predictable inter-core communication on top of the Argo NoC. This new Argo generation is, in terms of hardware, less than half the size of NoCs that provide similar functionalities and it offers a higher degree of flexibility to the application programmer.

The developed TDM scheduler supports a generic TDM NoC and custom parameterizable communication patterns. These communication patterns allow the application programmer to generate schedules that provide a set of VCs that efficiently uses the hardware resources. The TDM scheduler also shows better results, in terms of TDM period, compared to previous state-of-the-art TDM schedulers. Furthermore, we provide a description of how a communication pattern can be optimized in terms of shortening the TDM period.

The thesis identifies two types of inter-core communication that are commonly used in real-time systems: message passing and state-based communication. We implement message passing as a circular buffer with the data transfer
through the NoC. The worst-case execution time (WCET) of the send and receive functions of our implementation is not
dependent on the message size. We also implement five algorithms for state-based communication and analyze them in
terms of the WCET and worst-case communication delay. The five algorithms each have scenarios where they are better
than the others.

This thesis shows in detail how time-predictable inter-core communication can be implemented in an efficient way, from
the low-level hardware to the high-level software functions.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Embedded Systems Engineering
Authors: Sørensen, R. B. (Intern), Sparsø, J. (Intern), Schoeberl, M. (Intern)
Number of pages: 143
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 423
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd423_Sorensen_RB.pdf

Rationalization with ruled surfaces in architecture
This thesis addresses the problems of rationalizing and segmenting large scale 3D models, and how to handle difficult
production constraints in this area. The design choices when constructing large scale architecture are influenced by the
budget. Therefore I strive to minimize the amount of time and material needed for production. This makes advanced free
form architecture viable for low cost projects, allowing the architects to realize their designs.

By pre-cutting building blocks using hot wire robots, the amount of milling necessary can be reduced drastically. I do this
by rationalizing the intended shape as a piecewise ruled surface; the developed method was able to cut away up to 95%
of the excess material. Methods were developed to minimize the number of blocks necessary to build advanced large
scale 3D shapes. Using stochastic optimization to guide the segmentation, it was possible to remove up to 48% of the
building blocks. Hot blade cutting for constructing models with positive Gauss curvature is an upcoming technology. Three
segmentation algorithms were developed to solve construction constraints that arises when using this technique. One of
the algorithms focuses on creating an aesthetic segmentation.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics,
Mathematics
Authors: Steenstrup, K. H. (Intern), Gravesen, J. (Intern), Bærentzen, J. A. (Intern)
Number of pages: 70
Publication date: 2016

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 413
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd413_Steenstrup_KH_reduced.pdf
Modelling Digital Media Objects
The goal of this thesis is to investigate two relevant issues regarding computational representation and classification of digital multi-media objects. With a special focus on music, a model for representation of objects comprising multiple heterogeneous data types is investigated. Necessary to this work are considerations regarding integration of multiple diverse data modalities and evaluation of the resulting concept representation.

Regarding modelling of data exhibiting certain sequential structure, a number of theoretical and empirical results are presented. These are results related to model parameter estimation and the use of sequence models in a classification scenario. The latter being of importance in various digital multimedia navigation and retrieval tasks.

In the fields of topic modelling and multi-modal integration, we formulate a model to describe entities composed of multiple aspects. The particular aspects considered in the publications are sound, song lyrics, and user-provided metadata. This model integrates the diverse data types comprising the objects and defines concrete unified representations in a joint “semantic” space. Within the context of this model, general measures of similarity between such multi-modal objects are investigated.

In the fields of method of moments and sequence modelling, we increase practical applicability of a certain moment based parameter estimation method for Hidden Markov models by showing how to use full-length sequences in the estimation process. Consequently, this impacts the quality of the estimated model parameters.

Subsequently, we show how to perform time series classification using a composite likelihood formulated from third order moments defined by the Hidden Markov model. Compared to the conventional likelihood based method, our contribution is less computationally expensive, while retaining the level of classification performance.

Game-based verification and synthesis
Infinite-duration games provide a convenient way to model distributed, reactive and open systems in which several entities and an uncontrollable environment interact. Here, each entity as well as the uncontrollable environment are modelled as players.

A strategy for an entity player in the model corresponds directly to a program for the corresponding entity of the system. A strategy for a player which ensures that the player wins no matter how the other players behave then corresponds to a program ensuring that the specification of the entity is satisfied no matter how the other entities and the environment behaves. Synthesis of strategies in games can thus be used for automatic generation of correct-by-construction programs from specifications.

We consider verification and synthesis problems for several well-known game-based models. This includes both model-checking problems and satisfiability problems for logics capable of expressing strategic abilities of players in games with both qualitative and quantitative objectives.
A number of computational complexity results for model-checking and satisfiability problems in this domain are obtained. We also show how the technique of symmetry reduction can be extended to solve finitely-branching turn-based games more efficiently. Further, the novel concept of winning cores in parity games is introduced. We use this to develop a new polynomial-time under-approximation algorithm for solving parity games. Experimental results show that this algorithm performs better than the state-of-the-art algorithms in most benchmark games.

Two new game-based modelling formalisms for distributed systems are presented. The first makes it possible to reason about systems where several identical entities interact. The second provides a game-based modelling formalism for distributed systems with continuous time and probability distributions over the duration of delays. For these new models we provide decidability and undecidability results for problems concerning computation of symmetric Nash equilibria and for deciding existence of strategies that ensure reaching a target with a high probability.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic, Embedded Systems Engineering
Authors: Vester, S. (Intern), Hansen, M. R. (Intern)
Number of pages: 203
Publication date: 2016

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2016
Number: 414
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd414_Vester_S.pdf

**Relations**

Projects:
Game-based verification and synthesis
Publication: Research › Ph.D. thesis – Annual report year: 2016

**Real Time Structured Light and Applications**

Structured light scanning is a versatile method for 3D shape acquisition. While much faster than most competing measurement techniques, most high-end structured light scans still take in the order of seconds to complete.

Low-cost sensors such as Microsoft Kinect and time of flight cameras have made 3D sensor ubiquitous and have resulted in a vast amount of new applications and methods. However, such low-cost sensors are generally limited in their accuracy and precision, making them unsuitable for e.g. accurate tracking and pose estimation.

With recent improvements in projector technology, increased processing power, and methods presented in this thesis, it is possible to perform structured light scans in real time with 20 depth measurements per second. This offers new opportunities for studying dynamic scenes, quality control, human-computer interaction and more.

This thesis discusses several aspects of real time structured light systems and presents contributions within calibration, scene coding and motion correction aspects. The problem of reliable and fast calibration of such systems is addressed with a novel calibration scheme utilising radial basis functions [Contribution B]. A high performance flexible open source software toolkit is presented [Contribution C], which makes real time scanning possible on commodity hardware. Further, an approach is presented to correct for motion artifacts in dynamic scenes [Contribution E].

An application for such systems is presented with a head tracking approach for medical motion correction [Contribution A, F]. This aims to solve the important problem of motion artifacts, which occur due to head movement during long acquisition times in MRI and PET scans. In contrast to existing methods, the one presented here is MRI compatible [Contribution D], not dependent on fiducial markers, and suitable for prospective correction.

Factors contributing to accuracy and precision of structured light systems are investigated with a study of performance factors [Contribution G]. This is also done in the context of biological tissue, which exhibit subsurface effects and other undesirable effects [Contribution H], and it is shown that this error is to a large extent deterministic and can be corrected.
Domain Specific Language for Modeling Waste Management Systems

In order to develop sustainable waste management systems with considering life cycle perspective, scientists and domain experts in environmental science require readily applicable tools for modeling and evaluating the life cycle impacts of the waste management systems. Practice has proved that modeling these systems with general-purpose tools is a cumbersome task. On one hand, the scientists have to spend considerable amount of time to understand these tools in order to develop their models. On another hand, integrated assessments are becoming gradually common in environmental management and therefore scientists are also faced with the problem of integrating models across scales and domains, which is not a straightforward process.

Domain-Specific Languages (DSLs) are languages which are specialized for a specific application domain and they promise to increase developer productivity by raising the level of abstraction. They allow domain experts, who are non-programmers, to directly encode their domain knowledge about what a system under development should do. In this thesis, we utilize domain-specific languages, on the basis of the flow-based programming (FBP) paradigm, to model and evaluate environmental technologies i.e. solid waste management systems. Flow-based programming is used to support concurrent execution of the processes, and provides a model-integration language for composing processes from homogeneous or heterogeneous domains. And a domain-specific language is used to define atomic processes and domain-specific validation rules for composite processes. We call these DSLs, which are based on FBP paradigm, domain-specific flow based languages and we provide a formal framework to develop them. To this end, we advocate aspect-oriented concepts to FBP to separate cross-cutting concerns, by providing an extension called AOFBP. Afterwards, we propose the framework based on this extension, and we use a formal language called ForSpec, which is an extension of FORMULA, to formally specify the structural and behavioral semantics of the sub-languages proposed in this framework. Finally, we propose a domain specific language for modeling of waste-management systems on the basis of our framework. We evaluate the language by providing a set of case studies. The contributions of this thesis are; addressing separation of concerns in Flow-based programming and providing the formal specification of its syntax and semantics; a formal language and framework to specify domain-specific flow based languages; design and develop domain specific languages for waste management modeling; and finally our work also can be considered as another case study for structural and behavioral semantics specifications in ForSpec and FORMULA.
Numerical Modeling of Microelectrochemical Systems

The PhD dissertation is concerned with mathematical modeling and simulation of electrochemical systems. The first three chapters of the thesis consist of the introductory part, the model development chapter and the chapter on the summary of the main results. The remaining three chapters report three independent papers and manuscripts.

As a preliminary to the study, we describe a general model for electrochemical systems and study their underlying mechanisms through electroanalytical techniques. We then extend the model to a more realistic model for microelectrochemical systems which incorporates the finite size of ionic species in the transport equation. The model presents a more appropriate boundary conditions which describe the modified Butler-Volmer reaction kinetics and account for the surface capacitance of the thin electric double layer. We also have found analytical solution for the reactants in the bulk electrolyte that are traveling waves.

The first paper presents the mathematical model which describes an electrochemical system and simulates an electroanalytical technique called cyclic voltammetry. The model is governed by a system of advection–diffusion equations with a nonlinear reaction term at the boundary. We investigate the effect of flow rates, scan rates, and concentration on the cyclic voltammetry. We establish that high flow rates lead to the reduced hysteresis in the cyclic voltammetry curves and increasing scan rates lead to more pronounced current peaks. The final part of the paper shows that the response current in a cyclic voltammetry increases proportionally to the electrolyte concentration.

In the second paper we present an experiment of an electrochemical system in a microfluidic system and compare the result to the numerical solutions. We investigate how the position of the electrodes in the system affects the recorded cyclic voltammetry. The result shows that convection influences the charge transfer dynamics on the electrode surface and hence the cyclic voltammetry recorded. In terms of relative high flow to scan rates, the current response is dominated by the convection due to the fresh supply of reactants towards the electrode surface and quick removal of the products. We also establish that at high scan rates and modest flow rates, peak currents are recorded. Finally, the results show that the position of the electrodes is critical when performing cyclic voltammetry under the flow condition. The numerical results show promising agreement with experimental findings which could be critical in designing highly sensitive electrochemical systems.

The last paper explores the numerical solution which describes the non-linear transient responses to a large applied potential at the electrode in a microelectrochemical system. In our analysis, we account for the finite size properties of ions in the mass and the charge transport of ionic species in an electrochemical system. This term characterizes the saturation of the ionic species close to the electrode surface. We then analyse the responses of the system on the charging of the electric double layer. We consider an arbitrary electrolyte solution that is sandwiched between electrodes and allow for electrochemical reactions at the electrode/electrolyte interface. One of the electrodes is biased with a potential which triggers the reaction and the dynamics of the system. We establish that there is a quick build up of boundary layers in the double layer, but the finite size constraint on the ionic species prevents overcrowding of the ionic species. The result also shows that reactants which undergo charge transfer at the electrode/electrolyte interface crowded the electric double layer and the dynamics of the electric double layer is controlled by the charge transfer.

Towards Theory-of-Mind agents using Automated Planning and Dynamic Epistemic Logic

This thesis is part of a growing body of work in what we call epistemic planning. Epistemic planning is situated at the intersection of automated planning and what can broadly be called dynamic logics. Both are part of the much larger field of Artificial Intelligence.
Automated Planning has been around since at least the 1970s. It is a diverse collection of methods, models, algorithms and specification languages for giving autonomous agents the ability to come up with plans for proactively achieving goals. Autonomous agents can be understood as independent actors, given a purpose by their designer. Whether they are in a software system, connected to the real world with sensors and actuators, or used as a tool for modelling people, for instance in economics, they need to be able to imagine (or predict) outcomes of actions in order to form plans.

The feature that most distinguishes planning from other decision making methods, is that the planner does not know the full system from the beginning. Most of the time it would simply be too big to store in memory! Instead of being given the entire “game”, they use a specification of actions and the initial state to generate only a fraction of the full search space. This means that what an agent can plan for depends crucially on what domains we can describe. This is where logic comes into the picture.

For most of its more than 2500 year long history, logic has been mostly interested in the study of valid reasoning. In later years (in the scheme of things), more attention has been given to studying when reasoning fails in humans. Like using differential equations to analyse and simulate both when a bridge holds and when it collapses, we can use logic to analyse and simulate reasoning both when it is sound and when it isn’t.

The subbranch of logic applied in this work is Dynamic Epistemic Logic. The epistemic part concerns the formalisation of knowledge and belief (mainly) in multi-agent settings. We can describe situations in which many agents are present and have different knowledge and beliefs about the world and each others’ knowledge and belief. Adding the dynamic part of Dynamic Epistemic Logic to our arsenal, we can describe how situations change when, broadly speaking, things happen. In the application to Automated Planning, we let these things be actions of the agents in the system. In doing so we derive new planning formalisms that allow agents to plan under consideration of how what they do changes both the world and knowledge and belief about the world.

In this thesis we give new planning formalisms for single-agent planning and new results for the model theory of multi-agent models. The first of the two fully developed planning formalisms is conditional (single-agent) epistemic planning, allowing an agent to plan with what it knows now and what it knows it will come to know. Though this is nothing new in Automated Planning, it sets the stage for later work.

The second planning formalism extends conditional epistemic planning with beliefs, letting the agent have expectations, without probabilities, of how things will turn out. Our radically different notions of bisimulation for the multi-agent versions of these models are particularly interesting for logicians, as are surprising expressivity results for well known logics on such models.

The final part of the thesis describes ideas on extending the second formalism to a multi-agent setting. With a view towards the practical implementation of agents, we shall also see how an agent can discard the parts of its model that it does not believe to be the case. While this is not necessary for analysing reasoning agents, it does seem a requirement for practical implementations. There are simply too many possibilities for a resource-bounded agent to keep track of. If the agent does discard unlikely possibilities, it must be able to do belief revision if it later turns out to be wrong. Such a procedure is also described.

The long term potential of multi-agent aware planning algorithms is that agents that can predict and understand others in order to plan cooperation, communication, and/or competition. It is the slow edging towards a general framework for multi-agent planning that is the underlying motivation, and some of the main results, of this thesis. While regrettably we haven’t gotten there yet, we’re considerably closer than when we started.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic, Office for Study Programmes and Student Affairs
Authors: Andersen, M. B. (Intern), Bolander, T. (Intern)
Number of pages: 177
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2014
Number: 351
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
Uncertainty Quantification with Applications to Engineering Problems

The systematic quantification of the uncertainties affecting dynamical systems and the characterization of the uncertainty of their outcomes is critical for engineering design and analysis, where risks must be reduced as much as possible. Uncertainties stem naturally from our limitations in measurements, predictions and manufacturing, and we can say that any dynamical system used in engineering is subject to some of these uncertainties.

The first part of this work presents an overview of the mathematical framework used in Uncertainty Quantification (UQ) analysis and introduces the spectral tensor-train (STT) decomposition, a novel high-order method for the effective propagation of uncertainties which aims at providing an exponential convergence rate while tackling the curse of dimensionality. The curse of dimensionality is a problem that afflicts many methods based on meta-models, for which the computational cost increases exponentially with the number of inputs of the approximated function – which we will call dimension in the following.

The STT-decomposition is based on the Polynomial Chaos (PC) approximation and the low-rank decomposition of the function describing the Quantity of Interest of the considered problem. The low-rank decomposition is obtained through the discrete tensor-train decomposition, which is constructed using an optimization algorithm for the selection of the relevant points on which the function needs to be evaluated. The selection of these points is informed by the approximated function and thus it is able to adapt to its features. The number of function evaluations needed for the construction grows only linearly with the dimension and quadratically with the rank.

In this work we will present and use the functional counterpart of this low-rank decomposition and, after proving some auxiliary properties, we will apply PC on it, obtaining the STT-decomposition. This will allow the decoupling of each dimension, leading to a much cheaper construction of the PC surrogate. In the associated paper, the capabilities of the STT-decomposition are checked on commonly used test functions and on an elliptic problem with random inputs.

This work will also present three active research directions aimed at improving the efficiency of the STT-decomposition. In this context, we propose three new strategies for solving the ordering problem suffered by the tensor-train decomposition, for computing better estimates with respect to the norms usually employed in UQ and for the anisotropic adaptivity of the method.

The second part of this work presents engineering applications of the UQ framework. Both the applications are characterized by functions whose evaluation is computationally expensive and thus the UQ analysis of the associated systems will benefit greatly from the application of methods which require few function evaluations.

We first consider the propagation of the uncertainty and the sensitivity analysis of the non-linear dynamics of railway vehicles with suspension components whose characteristics are uncertain. These analysis are carried out using mostly PC methods, and resorting to random sampling methods for comparison and when strictly necessary.

The second application of the UQ framework is on the propagation of the uncertainties entering a fully non-linear and dispersive model of water waves. This computationally challenging task is tackled with the adoption of state-of-the-art software for its numerical solution and of efficient PC methods. The aim of this study is the construction of stochastic benchmarks where to test UQ methodologies before being applied to full-scale problems, where efficient methods are necessary with today's computational resources.

The outcome of this work was also the creation of several freely available Python modules for Uncertainty Quantification, which are listed and described in the appendix.
Algorithms for Electromagnetic Scattering Analysis of Electrically Large Structures

Accurate analysis of electrically large antennas is often done using either Physical Optics (PO) or Method of Moments (MoM), where the former typically requires fewer computational resources but has a limited application regime. This study has focused on fast variants of these two methods, with the goal of reducing the computational complexity while maintaining accuracy.

Regarding MoM, the complexity is reduced by applying the Multi-Level Fast Multipole Method (MLFMM) in combination with an iterative solver. Using MLFMM with a MoM implementation based on Higher-Order (HO) basis functions has, by several authors, been dismissed as being too memory intensive. In the present work, we demonstrate for the first time that by including a range of both novel and previously presented modifications to the standard MLFMM implementation, HO MLFMM can achieve both memory reduction and significant speed increase compared to Lower-Order (e.g., RWG) based MLFMM. Further, issues surrounding an iterative solution, such as the iterative solver and preconditioning, are discussed. Numerical results demonstrate the performance and stability of the algorithm for very large problems, including full satellites at Ku band.

Accelerating PO is an entirely different matter. A few authors have discussed applying the Fast-PO technique to far fields, achieving relative errors of 0.1%–1% for moderately sized scatterers. For near-fields, the state-of-the-art implementation of Fast-PO has several difficulties, in particular low accuracy and limited application regime. For the problems considered in this thesis, the error limit for PO is ≈ 0.01%, and the application limitations of the published Fast-PO are too prohibitive for our use. Therefore, results based on an improved Fast-PO implementation for far-fields, as well as a novel algorithm for near-fields, are presented. These results demonstrate that it is possible to achieve very accurate results, with relative errors around $10^{-5}$, at a much reduced time consumption. The method behind this part of the code is deemed confidential by TICRA.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing
Authors: Borries, O. P. (Intern), Hansen, P. C. (Intern)
Number of pages: 205
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2014
Number: 354
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Publication: Research › Ph.D. thesis – Annual report year: 2015

Combined Shape and Topology Optimization

Shape and topology optimization seeks to compute the optimal shape and topology of a structure such that one or more properties, for example stiffness, balance or volume, are improved. The goal of the thesis is to develop a method for shape and topology optimization which uses the Deformable Simplicial Complex (DSC) method. Consequently, we present a novel method which combines current shape and topology optimization methods. This method represents the surface of the structure explicitly and discretizes the structure into non-overlapping elements, i.e. a simplicial complex. An explicit surface representation usually limits the optimization to minor shape changes. However, the DSC method uses a single explicit representation and still allows for large shape and topology changes. It does so by constantly applying a set of mesh operations during deformations of the structure. Using an explicit instead of an implicit representation gives rise to several advantages including straightforward modeling of the surface, improved scalability and ability to optimize multiple materials.

This dissertation describes the essential parts of the novel method for combined shape and topology optimization. This includes the structural analysis in Chapter 2, the optimization in Chapter 3 and the Deformable Simplicial Complex method in Chapter 4. Finally, four applications of the developed method are presented in the included papers and summarized in Chapter 5.
Textual databases for e.g. biological or web-data are growing rapidly, and it is often only feasible to store the data in compressed form. However, compressing the data comes at a price. Traditional algorithms for e.g. pattern matching requires all data to be decompressed - a computationally demanding task. In this thesis we design data structures for accessing and searching compressed data efficiently.

Our results can be divided into two categories. In the first category we study problems related to pattern matching. In particular, we present new algorithms for counting and comparing substrings, and a new algorithm for finding all occurrences of a pattern in which we may insert gaps. In the other category we deal with accessing and decompressing parts of the compressed string. We show how to quickly access a single character of the compressed string, and present a data structure that supports fast decompression of substrings from prespecified positions.

Given the continuous advancements in the technology of energy harvesting over the last few years, we are now starting to see wireless sensor networks (WSNs) powered by scavenged energy. This change in paradigm has major repercussions not only on the hardware engineering aspects, but also on the software side. The first protocols specifically designed to take advantage of the energy harvesting capabilities of a network have just recently appeared. At the same time, security remains one of the central points of WSNs development, because of their intrinsically unreliable nature that combines a readily accessible communication infrastructure such as wireless data exchange, to an often likewise readily accessible physical deployment. This dissertation provides a comprehensive look at how security can be improved by what energy harvesting has to offer. The main question asked is whether or not it is possible to provide better security in a WSN, by being aware of the fact that the amount of available energy is not going to monotonically decrease over time. The work covers different aspects and components of a WSN and focuses on what is arguably one the most important ones, medium access control (MAC) protocols. An energy-harvesting specific MAC protocol is introduced together with a related security suite. A new attack relevant to a whole class of MAC protocols is
also introduced, along with a scheme that defeats it. A security approach for MAC protocols is discussed to provide an energy-aware solution. In order to address security bootstrapping, a new energy-adaptive key reinforcement scheme is presented. Finally an implementation and some experimental results are provided.
**Statistical learning for predictive targeting in online advertising**

The focus in this thesis is investigation of machine learning methods with applications in computational advertising. Computational advertising is the broad discipline of building systems which can reach audiences browsing the Internet with targeted advertisements. At the core of such systems, algorithms are needed for making decisions. It is in one such particular instance of computational advertising, namely in web banner advertising, that we investigate machine learning methods to assist and make decisions in order to optimize the placements of ads.

The industrial partner in this work is Adform, an international online advertising technology partner. This also means that the analyses and methods in this work are developed with particular use-cases within Adform in mind and thus need also to be applicable in Adform’s technology stack. This implies extra thought on scalability and performance.

The particular use-case which is used as a benchmark for our results, is clickthrough rate prediction. In this task one aims to predict the probability that a user will click on an advertisement, based on attributes about the user, the advertisement the context, and other signals, such as time. This has its main application in real-time bidding ad exchanges, where each advertiser is given a chance to place bids for showing their ad while the page loads, and the winning bid gets to display their banner.

The contributions of this thesis entail application of a hybrid model of explicit and latent features for learning probabilities of clicks, which is a methodological extension of the current model in production at Adform. Our findings confirm that latent features can increase predictive performance in the setup of click-through rate prediction. They also reveal a tedious process for tuning the model for optimal performance.

We also present variations of Bayesian generative models for stochastic blockmodeling for inference of structure based on browsing patterns. Applying this structural information to improve click-through rate prediction becomes a two-step procedure; 1) learn user and URL profiles from browsing patterns, 2) use the profiles as additional features in a click-through rate prediction model. The assumption we implicitly make is reasonable: Users and URLs that are grouped together based on browsing patterns will have similar responses to ads, e.g., can be used as predictors of clicks. We report successful examples of applying this approach in practice.

Finally, we introduce the multiple-networks stochastic blockmodel (MNSBM), a model for efficient overlapping community detection in complex networks which can be assumed to be an aggregation of multiple block-structured subnetworks.

**General information**

State: Published  
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Adform Aps  
Authors: Fruergaard, B. Ø. (Intern), Hansen, L. K. (Intern), Urban, J. (Intern)  
Number of pages: 129  
Publication date: 2015

**Publication information**  
Place of publication: Kgs. Lyngby  
Publisher: Technical University of Denmark (DTU)  
Original language: English

Series: DTU Compute PHD-2014  
Number: 355  
ISSN: 0909-3192  
Main Research Area: Technical/natural sciences  
Electronic versions:  
phd355_Fruergaard_BO.pdf  
Publication: Research › Ph.D. thesis – Annual report year: 2015

**Design and analysis of biomedical studies**

Biomedicine is a field that has great influence on the majority of mankind. The constant development has considerably changed our way of life during the last centuries. This has been achieved through the dedication of biomedical researchers along with the tremendous resources that over time have been allocated this field. It is utterly important to utilize these resources responsibly and efficiently by constantly striving to ensure high-quality biomedical studies. The focus of this project is on statistical aspects that arise within the field of biomedicine.

Two types of errors are frequently accentuated within the framework of statistics, namely type I and type II errors. Type I errors occur when a null hypothesis erroneously is rejected. An acceptable type I error rate is specified prior to conducting the statistical analysis. However, all statistical models make assumptions and if violated the actual type I error rate may deviate from the pre-specified type I error rate. Type II errors occur when we fail to reject a false null hypothesis. On contrary to the type I error rate, the type II error rate is not explicitly specified during the statistical analysis and this entails that assessment of the type II error rate in practice is at risk of being neglected altogether. Concerns regarding type I errors, type II errors and adherence (or lack thereof) to model assumptions for biomedical studies are a recurring theme in this thesis.

Data collected in some biomedical studies are positively skewed; hence methods relying on the normal distribution are not directly applicable. We investigated how data from one of these studies are suitably analyzed. We extracted 23 different summary statistics from data gathered from eleven studies. The degree of adherence to the model assumptions evaluated for each of these summary statistics form basis for our conclusions.
Hierarchically structured data are frequently encountered in biomedical studies. For one type of studies entailing such data we have conducted a literature study strongly indicating that this structure commonly is neglected in the statistical analysis. Based on this closed-form expressions for the approximate type I error rate are formulated. The type I error rates are assessed for a number of factor combinations as they appear in practice and in all cases the type I error rates are demonstrated to be severely inflated.

Prior to conducting a study it is important to perform power and sample size determinations to ensure that reliable conclusions can be drawn from the statistical analysis. We have formulated closed-form expressions for the statistical power of studies with a hierarchical structure to guide biomedical researchers designing future studies of this type. Upon model fitting it is important to examine if the model assumptions are met to avoid that spurious conclusions are drawn. While the range of diagnostic methods is extensive for models assuming a normal response it is generally more limited for non-normal models. An R package providing diagnostic tools suitable for examining the validity of binomial regression models have been developed. The binom Tools package is publicly available at the CRAN repository.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Authors: Hansen, M. K. (Intern), Kulahci, M. (Intern)
Number of pages: 224
Publication date: 2015

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2014
Number: 343
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd343_Hansen_MK.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

**Individualized directional microphone optimization in hearing aids based on reconstructing the 3D geometry of the head and ear from 2D images**

The goal of this thesis is to improve intelligibility for hearing-aid users by individualizing the directional microphone in a hearing aid. The general idea is a three step pipeline for easy acquisition of individually optimized directional filters. The first step is to estimate an individual 3D head model based on 2D images, the second step is to simulate individual head related transfer functions (HRTFs) based on the estimated 3D head model and the final step is to calculate optimal directional filters based on the simulated HRTFs. The pipeline is employed on a Behind-The-Ear (BTE) hearing aid.

We verify the directional filters optimized from simulated HRTFs based on a listener-specific head model against two set of optimal filters. The first set of optimal filters is calculated from HRTFs measured on a 3D printed version of the head model. The second set of optimal filters is calculated from HRTFs measured on the actual human subject.

A verification of the ‘simulated’ directional filters against the optimal filters for the human subject revealed a 0.5 dB reduction in articulation-index weighted directivity index, which corresponds to 5% less speech intelligibility. A comparison against non-individual directional filters revealed equally high Articulation-Index weighted Directivity Index (AI-DI) values for our specific test subject. However, measurements on other individuals indicate that the performance of the non-individual filters vary among subjects, and in particular individuals who deviate from an average of the population could benefit from having individualized filters.

We developed a pipeline for 3D printing of full size human heads. The 3D printed head facilitated the second verification step, which revealed a 0.3 dB reduction from optimal to simulated directional filters. This indicates that the simulation are more similar to measurements on the 3D printed head than measurements on the human subject. We suggest that the larger difference between simulation and human measurements could arise due to small geometrical errors in the head model or due to differences in acoustical properties between human skin and virtual material properties in the simulation.

The BTE hearing aid showed very little room for improvement using individualized directional filters, however the directional filters in an In-The-Ear (ITE) hearing aid revealed an improvement in AI-DI values of up to 3.6 dB between an average filter and an optimal filter. This suggests that hearing-aid users with ITE hearing aids could benefit more from having individualized directional filters than what was shown for a BTE hearing aid.

This thesis is a step towards individualizing the directional microphone in hearing aids, which could contribute with improved sound for a group of hearing-aid users. In particular, we believe that ITE hearing-aid users could have a large benefit from an individualized directional microphone.
Modelling the structure of complex networks

A complex network is a system in which a discrete set of units interact in a quantifiable manner. Representing systems as complex networks has become increasingly popular in a variety of scientific fields including biology, social sciences and economics. Parallel to this development complex networks has been independently studied as mathematical objects in their own right. As such, there has been both an increased demand for statistical methods for complex networks as well as a quickly growing mathematical literature on the subject.

In this dissertation we explore aspects of modelling complex networks from a probabilistic perspective. The first two chapters will be focused on the justification of the use of probabilistic methods for inference problems; we will look at the justification of probabilistic methods from the perspective of consistency and as a general method of updating beliefs. The next chapters will treat some of the various symmetries, representer theorems and probabilistic structures often deployed in the modelling complex networks, the construction of sampling methods and various network models.

The introductory chapters will serve to provide context for the included written work on the topics of (i) updating beliefs (ii) construction of samplers for partition-based problems (iii) applying non-parametric methods for modelling stationary and temporal network data.

Reconstruction Methods for Inverse Problems with Partial Data

This thesis presents a theoretical and numerical analysis of a general mathematical formulation of hybrid inverse problems in impedance tomography. This includes problems from several existing hybrid imaging modalities such as Current Density Impedance Imaging, Magnetic Resonance Electrical Impedance Tomography, and Ultrasound Modulated Electrical Impedance Tomography. After giving an introduction to hybrid inverse problems in impedance tomography and the mathematical tools that facilitate the related analysis, we explain in detail the stability properties associated with the classification of a linearised hybrid inverse problem. This is done using pseudo-differential calculus and theory for overdetermined boundary value problem. Using microlocal analysis we then present novel results on the propagation of
singularities, which give a precise description of the distinct features of solutions in the case of a non-elliptic problem. To conduct a numerical analysis, we develop four iterative reconstruction methods using the Picard and Newton iterative schemes, and the unified approach to the reconstruction problem encompasses several algorithms suggested in the literature. The algorithms are implemented Numerically in two dimensions and the properties of the algorithms and their implementations are investigated theoretically. Novel numerical results are presented for both the full and partial data problem, and they show similarities and differences between the proposed algorithms, which are closely linked to the results of the theoretical analysis. The findings in this thesis justify that the choice of algorithm should be based on a theoretical analysis of the underlying inverse problem.

**General information**

State: Published  
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing  
Authors: Hoffmann, K. (Intern), Knudsen, K. (Intern)  
Number of pages: 130  
Publication date: 2015  

**Publication information**  
Place of publication: Kgs. Lyngby  
Publisher: Technical University of Denmark (DTU)  
Original language: English  
Series: DTU Compute PHD-2014  
Number: 344  
ISSN: 0909-3192  
Main Research Area: Technical/natural sciences  
Electronic versions:  
phd344_Hoffmann_K.pdf  
Publication: Research › Ph.D. thesis – Annual report year: 2015

**The AORTA Reasoning Framework - Adding Organizational Reasoning to Agents**  
Intelligent agents are entities defined by, among other things, autonomy. In systems of many agents, the agents’ individual autonomy can lead to uncertainty since their behavior cannot always be predicted. Usually, this kind of uncertainty is accommodated by imposing an organization upon the system; an organization that defines expected behavior of the agents and attempts to restrict the agents’ behavior to let it match the expectations. Restrictions can lead to a decrease in autonomy, contradicting one of the pillars of intelligent agents.

This thesis presents the AORTA reasoning framework, which is a practical component (founded in logic) that enriches intelligent agents with organizational reasoning capabilities. We take the agent’s perspective by devising a component that integrates with the agent’s usual reasoning capabilities in a non-intrusive way. This results in agents that are both organization-aware and autonomous. The reasoning component makes them organization-aware, and their autonomy is intact because the component does not change the existing reasoning mechanisms. As such, it allows the agents to decide whether to adhere to the system’s expectations.

The ability to reason about organizations has previously been successfully integrated into agent programming languages. However, the operationalization of an organization is usually tailored to a specific language. This makes it hard to apply the same approach to other languages and platforms. The AORTA reasoning framework distinguishes itself by being a generic framework that allows different kinds of agents to reason about different kinds of organizations.

We present our results in three main parts. In the first part, we present the theoretical foundations for the AORTA framework, which consists of semantics of norms, an organizational metamodel, and the AORTA reasoning component. The reasoning component is characterized by being completely decoupled from the cognitive agent, by its automated reasoning about norms and organizational options, and by the reasoning rules specified by the designer to act upon norms and options. We specify the reasoning component using structural operational semantics providing us with a formal, rigid description of the behavior of the component during execution. This enables us to precisely specify each reasoning phases (using transition rules), and it makes the implementation of the system quite straightforward.

The second part moves from theory to practice: we present an implementation of the framework and integrate it into various agent platforms. We show that the same configuration of the component can be used for different agent platforms, providing evidence for its use a general tool for organization-awareness. Furthermore, we use practical verification to show various properties of an implementation of agents and of the system in general.

In the last part, we discuss a potential issue with our framework. The possibility to commit to organizational objectives can affect the agent’s autonomy, which contradicts our main goal. We propose a model that solves this problem by adding a filter to the agent’s decision procedure that takes consequences of fulfilling a goal into account before deciding to commit to it. By considering both the agent’s preferences and the expected outcome of fulfilling the goal, we show that it was possible for the agents to make qualified context-dependent decisions.

We claim that by using the AORTA reasoning framework, agents become organization-aware. The reasoning component provides capabilities to reason about organizations and our decision procedure ensures that the autonomy of the agents is still intact.

**General information**
An Asynchronous Time-Division-Multiplexed Network-on-Chip for Real-Time Systems

Multi-processor architectures using networks-on-chip (NOCs) for communication are becoming the standard approach in the development of embedded systems and general purpose platforms. Typically, multi-processor platforms follow a globally asynchronous locally synchronous (GALS) timing organization. This thesis focuses on the design of Argo, a NOC targeted at hard real-time multi-processor platforms with a GALS timing organization.

To support real-time communication, NOCs establish end-to-end connections and provide latency and throughput guarantees for these connections. Argo uses time division multiplexing (TDM) in combination with a static schedule to implement virtual end-to-end circuits. TDM is a straightforward way to provide guarantees and to share the resources efficiently, and it has an efficient hardware implementation. Argo supports a GALS system organization, and additionally it explores more flexible timing within its structure, to address signal distribution issues, using a network of synchronous routers.

NOCs consist of a switching structure of routers connected by links, with network interfaces (NIs) that connect the processors to the switching structure. Argo uses a novel NI design that supports time-predictability, and asynchronous routers that form a time-elastic network. The NI design integrates the DMA functionality and the TDM schedule, and uses dual-ported local memories. The routers combine the router functionality and asynchronous elastic behavior. They also use a gating mechanism to reduce the energy consumption. The combination of the NI design and the router design supports the formation of end-to-end paths in the NOC, from the local memory of a sending core to the local memory of a receiving core. These end-to-end paths do not require any dynamic arbitration, buffering, flow control, or clock synchronization, in the routers or the NIs.

This thesis explores the implementation of the individual components of Argo, as well as several complete instances of the Argo NOC. The implementations target both FPGA technology and 65 nm CMOS technology. It is shown that (i) the NI design is scalable and four to five times smaller than previously published NIs for similar NOCs, (ii) the router design is power efficient and two to three times smaller than equivalent router designs, and (iii) the overall Argo NOC is around four times smaller than other TDM NOCs. Argo is an important part of the T-CREST platform and used in a number of configurations.

The flexible timing organization of Argo combines asynchronous routers with mesochronous NIs, which are connected to individually clocked cores, supporting a GALS system organization. The mesochronous NIs operate at the same frequency, possibly with some skew, while the network of asynchronous routers absorbs this skew within certain limits. The elasticity of the asynchronous network is explored, answering the question of how much skew the Argo NOC can absorb. A qualitative analysis studies the parameters affecting the elasticity and its limits. A quantitative analysis models the Argo behavior using timed-graph models and worstcase timing separation of events analysis to evaluate the elasticity of Argo. The results show that the skew absorbed by the network of routers can be two or more cycles, depending on the frequency applied at its endpoints, the NIs.

Overall this thesis presents the design and implementation of Argo, and the analysis of its elastic behavior. It shows that Argo provides hard real-time guarantees in a straightforward way, it has an efficient implementation and it is time-elastic.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic
Authors: Jensen, A. S. (Intern), Villadsen, J. (Intern)
Number of pages: 197
Publication date: 2015
Linear mixed models in sensometrics

Today’s companies and researchers gather large amounts of data of different kind. In consumer studies the objective is the collection of the data to better understand consumer acceptance of products. In such studies a number of persons (generally not trained) are selected in order to score products in terms of preferences. In sensory studies the aim is the collection of the data to better describe products and differences of the products according to a number of sensory attributes. Here trained persons, so-called assessors, score the products in terms of different characteristics such as smell, taste, texture, sound - depending on the aim of a study. It is a common approach in both studies to consider persons coming from a larger population, which, from the statistical perspective, leads to the use of mixed effects models, where consumers/assessors enter as random effects (Lawless and Heymann, 1997).

Mixed effects models have been used extensively in analysis of both consumer and sensory studies. However frequently too simplistic models are considered, important effects are not accounted for and as a consequence important information is not gained or analysis leads to improper conclusions. The focus of this project is to propose a methodology for analyzing more complex models together with tools facilitating the methodology. This was accomplished by contributing to the mixed effects ANOVA modelling in general and specifically applied to sensory and consumer studies through a series of papers and software tools facilitating the developed methodologies. The primary advantage of the ANOVA approach is that it gives confidence intervals and significance tests for the various effects including the background variables used in the model and consequently a fast and reliable assessment and ranking of the importance of different factors.

There exists today very little easily available methodology and software which supports consumer studies with both sensory properties and background information related to health benefits, environment and user-friendliness. In close collaboration with the industrial partners an open-source software tool ConsumerCheck was developed in this project and now is available for everyone. This will represent a major step forward when concerns this important problem in modern consumer driven product development. Standard statistical software packages can be used for some of the purposes, but for the specific problems considered here and for the typical users in industry, these programs are far from satisfactory. Therefore, the ConsumerCheck software represents a novel source of information for all quality-oriented industries. The effect is improved procedures for product development and hence improved quality of decision making in Danish as well as international food companies and other companies using the same methods.

The two open-source R packages lmerTest and SensMixed implement and support the methodological developments in the research papers as well as the ANOVA modelling part of the ConsumerCheck software. The SensMixed package is a package for semi-automated analysis of sensory and consumer studies within linear mixed effects framework. The lmerTest package supports tests for linear mixed effects models fitted with the lmer function of the lme4 package (Bates et al., 2013). While SensMixed is closely connected with sensometrics field, the lmerTest package has developed into a generic statistical package.

Reference manuals accompany these R packages.
A Probabilistic Approach for the System-Level Design of Multi-ASIP Platforms

Application Specific Instruction-set Processors (ASIPs) offer a good trade off between performance and flexibility when compared to general purpose processors or ASICs. Additionally, multiple ASIPs can be included in a single platform and they allow the generation of customized heterogeneous MPSoC with a relatively short time-to-market. While there are several commercial tools for the design of a single ASIP, there is still a lack of automation in the design of multi-ASIP platforms.

In this thesis we consider multi-ASIP platforms for real-time applications. Each ASIP is designed to run a specific group of tasks that we identify as a task cluster. With realtime applications, to decide how the tasks should be clustered, we perform a schedulability analysis of the system to verify if the deadlines of the applications can be met. However, to run a schedulability analysis, we need to know the WCET of each task that is available only after an ASIP is designed. Therefore, there is a circular dependency between the definition of the task clusters and the impossibility of defining them without knowing the WCET of the tasks as the ASIPs have not been defined yet.

Many approaches available in the literature break this circular dependency considering pre-defined task clusters or considering a small set of micro-architecture configurations for each ASIP. We propose an alternative approach that uses a probabilistic model to consider the design space of all possible micro-architecture configurations. We introduce a system-level Design Space Exploration (DSE) for the very early phases of the design that automatizes part of the multi-ASIP design flow. Our DSE is responsible for assigning the tasks to the different ASIPs exploring different platform alternatives. We perform a schedulability analysis for each solution to determine which one has the highest chances of meeting the deadlines of the applications and that should be considered in the next stages of the multi-ASIP design flow.

Good towers of function Fields

Algebraic curves are used in many different areas, including error-correcting codes. In such applications, it is important that the algebraic curve C meets some requirements. The curve must be defined over a finite field GF(q) with q elements, and then the curve also should have many points over this field. There are limits on how many points N(C) an algebraic curve C defined over a finite field can have.

An invariant of the curve which is important in this context is the curve’s genus g(C). Hasse and Weil proved that N(C)≤q+1+2g(C) q and this bound can in general not be improved. However if the genus is large compared with q, the bound can be improved. Drinfeld and Vladut showed the asymptotic result:

A(q)=limsup ( N(C) g(C)→∞ g(C)) ≤ √q-1.

The quantity A(q) is called Ihara’s constant. If q is a square, it is known that A(q)=√q-1, while the value of the A(q) is unknown for all other values of q.

In this thesis, we study a construction using Drinfeld modules that produces explicitly defined families of algebraic curves that asymptotically achieve Ihara’s constant. Such families of curves can also be described using towers of function fields.
Restated in this language the aim of the project is to find good and optimal towers. Using the theory of Drinfeld modules and computer algebraic techniques, some new examples of good towers are obtained. We analyse towers of Drinfeld modular curves describing certain equivalence classes of rank 2 Drinfeld modules. Using rank 3 Drinfeld modules further examples of good towers are produced.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Mathematics
Authors: Nguyen, N. (Intern), Beelen, P. (Intern)
Number of pages: 125
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2015
Number: 394
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd394_Nguyen_N.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

Systems for Personalization of Hearing Instruments: A Machine Learning Approach
Today, modern digital devices can be customized significantly to the individual user by adjusting or optimizing multiple parameters affecting the output of the devices. Such personal optimization of devices is referred to as personalization. In the case of hearing aids, personalization is not only a possibility offered to the user, but a requirement that must be performed carefully and precisely in order for the user to utilize the full potential of modern multi-parameter hearing aids. Today though, personalization is still based on a manual time-consuming trial-and-error approach performed by the user himself or, in case of hearing aids, by a hearing-care professional based on typically ambiguous oral feedback from the user. This often results in sub-optimal or even inappropriate settings of multi-parameter devices. This dissertation presents research on a machine-learning based interactive personalization system to improve the personalization of devices and, in particular, of hearing-aid devices. The proposed personalization system iteratively learns a non-parametric probabilistic model of a user’s assumed internal response function over all possible settings of a multi-parameter device based directly on sequential perceptual feedback from the user. A sequential design based on active learning is used to obtain the maximum of the user’s unknown internal response function in as few iterations as possible. Experiments were conducted where the proposed personalization system obtained a significantly preferred setting for individual users within ten to twenty iterations in scenarios with up to four parameters.

Following a short introduction that includes a summary of results and contributions, the first main chapter focuses on the probabilistic modeling framework in which a Gaussian process is used to model the user’s unobserved internal response function. The first main challenge addressed in this context is to account for inconsistent and thus noisy user feedback. The second main challenge addressed is to support feedback which closely reflects the user’s perception while providing maximal information about it without imposing a high cognitive load. In the second main chapter, active learning and sequential design are discussed in relation to the challenge of obtaining the setting that maximizes the user’s unobserved internal response function in as few iterations as possible. For the Gaussian process framework, an active learning criterion is proposed specifically suitable for this type of optimization. The final chapter contains an overall discussion and conclusion of the present work and research based in part on the results from eight scientific paper contributions contained in the appendices.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Cognitive Systems, Widex A/S
Authors: Nielsen, J. B. (Intern), Larsen, J. (Intern), Nielsen, J. (Ekstern)
Number of pages: 176
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: DTU Compute
Original language: English
Series: DTU Compute PHD-2014
An algebraic approach to graph codes
This thesis consists of six chapters. The first chapter, contains a short introduction to coding theory in which we explain the coding theory concepts we use. In the second chapter, we present the required theory for evaluation codes and also give an example of some fundamental codes in coding theory as evaluation codes. Chapter three consists of the introduction to graph based codes, such as Tanner codes and graph codes. In Chapter four, we compute the dimension of some graph based codes with a result combining graph based codes and subfield subcodes. Moreover, some codes in chapter four are optimal or best known for their parameters. In chapter five we study some graph codes with Reed–Solomon component codes. The underlying graph is well known and widely used for its good characteristics. This helps us to compute the dimension of the graph codes. We also introduce a combinatorial concept related to the iterative encoding of graph codes with MDS component code. The last chapter deals with affine Grassmann codes and Grassmann codes. We begin with some previously known codes and prove that they are also Tanner codes of the incidence graph of the point–line partial geometry of the Grassmannian. We expect that the techniques exposed in chapter six are also applicable to other codes as well.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Mathematics
Authors: Pinero, F. (Intern), Beelen, P. (Intern)
Number of pages: 96
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2014
Number: 352
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd352_Pinero_F.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

Modelling and Analysing Access Control Policies in XACML 3.0
XACML (eXtensible Access Control Markup Language) is a prominent access control language that is widely adopted both in industry and academia. XACML is an international standard in the field of information security. The problem with XACML is that its specification is described in natural language (c.f. GM03, Mos05, Ris13) and manual analysis of the overall effect and consequences of a large XACML policy set is a very daunting and time-consuming task.

In this thesis we address the problem of understanding the semantics of access control policy language XACML, in particular XACML version 3.0. The main focus of this thesis is modelling and analysing access control policies in XACML 3.0.

There are two main contributions in this thesis. First, we study and formalise XACML 3.0, in particular the Policy Decision Point (PDP). The concrete syntax of XACML is based on the XML format, while its standard semantics is described normatively using natural language. The use of English text in standardisation leads to the risk of misinterpretation and ambiguity. In order to avoid this drawback, we define an abstract syntax of XACML 3.0 and a formal XACML semantics. Second, we propose a logic-based XACML analysis framework using Answer Set Programming (ASP). With ASP we model an XACML PDP that loads XACML policies and evaluates XACML requests against these policies. The expressivity of ASP and the existence of efficient implementations of the answer set semantics provide the means for declarative specification and verification of properties of XACML policies.

Overall, we focus into two different area. The first part focuses on the access control language. More specifically our focus is on the understanding XACML 3.0. The second part focuses on how we use Logic Programming (LP) to model access control policies. We show that there is a relation between XACML and LP through their semantics. We close the thesis by presenting applications in analysing access control properties and a case study. These applications show that these two approaches (AC paradigm and LP paradigm) can be combined together.
We close the thesis by presenting applications in analysing access control properties and a case study. We present access control security policies in a Smart Grid from Smart Meter perspective.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Language-Based Technology
Authors: Ramli, C. D. P. K. (Intern), Nielson, H. R. (Intern), Nielson, F. (Intern)
Number of pages: 217
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 364
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Access Control Policies, IT Security, Control Systems, XACML, Composition Policies, Logic Programming, Answer Set Programming, Smart Grid, Smart Meter
Electronic versions:
phd364_Ramli_CDPK.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

Dynamics of High-Resolution Networks
NETWORKS are everywhere. From the smallest confines of the cells within our bodies to the webs of social relations across the globe. Networks are not static, they constantly change, adapt, and evolve to suit new conditions. In order to understand the fundamental laws that govern networks we need new, highly detailed maps that uncover the interactions of all constituents, accurately and in a temporal manner. One of the hardest networks to understand, but also one of the most interesting ones, is the human social network. How do humans interact, form friendships, and spread information? And how are we all affected by an ever changing network structure? Answering these questions will enrich our understanding of ourselves, our organizations, and our societies. Yet, mapping the dynamics of social networks has traditionally been an arduous undertaking. Today, however, it is possible to use the unprecedented amounts of information collected by mobile phones to gain detailed insight into the dynamics of social systems. This dissertation presents an unparalleled data collection campaign, collecting highly detailed traces for approximately 1000 people over the course of multiple years. The availability of such dynamic maps allows us to probe the underlying social network and understand how individuals interact and form lasting friendships. More importantly, these highly detailed dynamic maps provide us new perspectives at traditional problems and allow us to quantify and predict human life.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Sekara, V. (Intern), Jørgensen, S. L. (Intern)
Number of pages: 189
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Danmarks Tekniske Universitet (DTU)
Original language: English

Series: DTU Compute PHD-2015
Volume: 367
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd367_Sekara_V.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

Multivariate Analysis Techniques for Optimal Vision System Design
The present thesis considers optimization of the spectral vision systems used for quality inspection of food items. The relationship between food quality, vision based techniques and spectral signature are described. The vision instruments for food analysis as well as datasets of the food items used in this thesis are described. The methodological strategies are
outlined including sparse regression and pre-processing based on feature selection and extraction methods, supervised versus unsupervised analysis and linear versus non-linear approaches.

One supervised feature selection algorithm based on the existing sparse regression methods (EN and lasso) and one unsupervised feature selection strategy based on the local maxima of the spectral 1D/2D signals of food items are proposed. In addition, two novel feature extraction and selection strategies are introduced; sparse supervised PCA (SSPCA) and DCT based characterization of the spectral diffused reflectance images for wavelength selection and discrimination.

These methods together with some other state-of-the-art statistical and mathematical analysis techniques are applied on datasets of different food items; meat, diaries, fruits and vegetables. These datasets are acquired using three different vision systems; a spectral imaging device called VideometerLab, spectroscopy, and diffused reflectance imaging systems called Static Light Scattering (SLS).

These analyses result in significant reduction in the number of required wavelengths and simplification of the design of practical vision systems.
Tomographic Image Reconstruction Using Training Images with Matrix and Tensor Formulations

Reducing X-ray exposure while maintaining the image quality is a major challenge in computed tomography (CT); since the imperfect data produced from the few view and/or low intensity projections results in low-quality images that are suffering from severe artifacts when using conventional reconstruction methods. Incorporating a priori information about the solution is a necessity to improve the reconstruction. For example, Total Variation (TV) regularization method –assuming a piecewise constant image model – has been shown to allow reducing X-ray exposure significantly, while maintaining the image resolution compared to a classical reconstruction method such as Filtered Back Projection (FBP).

Some priors for the tomographic reconstruction take the form of cross-section images of similar objects, providing a set of the so-called training images, that hold the key to the structural information about the solution. The training images must be reliable and application-specific. This PhD project aims at providing a mathematical and computational framework for the use of training sets as non-parametric priors for the solution in tomographic image reconstruction. Through an unsupervised machine learning technique (here, the dictionary learning), prototype elements from the training images are extracted and then incorporated in the tomographic reconstruction problem both with matrix and tensor representations of the training images.

First, an algorithm for the tomographic image reconstruction using training images, where the training images are represented as vectors in a training matrix, is described. The dictionary learning problem is formulated as a regularized non-negative matrix factorization in order to compute a nonnegative dictionary. Then a tomographic solution with a sparse representation in the dictionary is obtained through a convex optimization formulation. Computational experiments clarify the choice and interplay of the model parameters and the regularization parameters. Furthermore, the assumptions in the tomographic problem formulation are analyzed. The sensitivity and robustness of the reconstruction to variations of the scale and rotation in the training images is investigated and algorithms to estimate the correct relative scale and orientation of the unknown image to the training images are suggested.

Then, a third-order tensor representation for the training images images is used. The dictionary and image reconstruction problem are reformulated using the tensor representation. The dictionary learning problem is presented as a nonnegative tensor factorization problem with sparsity constraints and the reconstruction problem is formulated in a convex optimization framework by looking for a solution with a sparse representation in the tensor dictionary. Numerical results show considering a tensor formulation over a matrix formulation significantly reduces the approximation error by the dictionary as well as leads to very sparse representations of both the training images and the reconstructions.

Further computational experiments show that in few-projection and low-dose settings our algorithm is while (not surprisingly) being superior to the classical reconstruction methods, is competitive with (or even better of) the TV regularization and tends to include more texture and sharper edges in the reconstructed images.

The focus of the thesis is the study of mathematical and algorithmic prospectives and thus the training images and tomographic scenarios are mostly simulation based. More studies are however needed for implementing the proposed algorithm in a routine use for clinical applications and materials testing.
Economic Model Predictive Control for Large-Scale and Distributed Energy Systems

In this thesis, we consider control strategies for large and distributed energy systems that are important for the implementation of smart grid technologies. An electrical grid has to ensure reliability and avoid long-term interruptions in the power supply. Moreover, the share of Renewable Energy Sources (RESs) in the smart grids is increasing. These energy sources bring uncertainty to the production due to their fluctuations. Hence, smart grids need suitable control systems that are able to continuously balance power production and consumption. We apply the Economic Model Predictive Control (EMPC) strategy to optimise the economic performances of the energy systems and to balance the power production and consumption. In the case of large-scale energy systems, the electrical grid connects a high number of power units. Because of this, the related control problem involves a high number of variables and constraints and its solution requires high computational times. Energy systems have a hierarchical control framework and the controllers have to work in the time-scale required by their hierarchy level. Dedicated optimisation techniques efficiently solve the control problem and reduce computational time. We implement the Dantzig-Wolfe decomposition technique to efficiently solve the EMPC problem.

The contributions of this thesis are primarily on:

- **Large-scale energy system**: Smart-grids connect a high number of energy units. In such a large-scale scenario the energy units are independent and dynamically decoupled. The mathematical model of the large-scale energy system embodies the decoupled dynamics of each power unit. Moreover, all units of the grid contribute to the overall power production.

- **Economic Model Predictive Control (EMPC)**: This control strategy is an extension of the Model Predictive Control (MPC) strategy. Energy systems often involve stochastic variables due to the share of fluctuating Renewable Energy Sources (RESs). Moreover, the related control problems are multi variables and they are hard, or impossible, to split into single-input-single-output control systems. MPC strategy can handle multi variables control problems and it can embody stochastic variables. The Economic MPC (EMPC) policy optimises the economic performances of the process. In this work, we apply the EMPC to energy systems and it computes the control trajectory for each energy unit. This control policy minimises production costs and ensures that the power production satisfies the customers' demand. The EMPC designs a linear control problem that has a block-angular constraints matrix and it has two sets of constraints. The independent dynamics of the energy units define the decoupling constraints sited on the diagonal. The coupling constraints represent the common goal of all power units in the energy system and this is to satisfy the customers' demand. The Dantzig-Wolfe optimisation technique applies to this structure of the constraints matrix in the view of fastening the control algorithm and increase its applicability.

- **Dantzig-Wolfe decomposition**: The Dantzig-Wolfe decomposition solves the EMPC problem through a distributed optimisation technique. The EMPC problem via Dantzig-Wolfe decomposition algorithm computes the optimal input trajectory for each energy unit and reduces the computation times. Moreover, such a control algorithm applies to large-scale energy systems and the number of energy units does not affect the performances of the controller. In this thesis, we also investigate suboptimal solutions of the EMPC problem via modified versions of the Dantzig-Wolfe decomposition algorithms. The feasibility of the suboptimal solutions suffices for stability. The goal of these modified Dantzig-Wolfe decomposition algorithms is to reduce computation time in the solution of the EMPC problem.
Mobile Phones as Cognitive Systems

Driven by the ubiquitous availability of data and inexpensive data storage, our ability to sense human beings has increased dramatically. Big data has permeated the public discourse and led to surprising insights across the sciences and the humanities. This dissertation presents research on expanding our capabilities in collecting, handling, processing, and using data collected about human beings to create an integrated view of social systems. The goal of the thesis has been threefold.

The first part of the thesis focuses on the need, design, and implementation of large-scale sensor-driven human data collection studies. Social networks can be measured with high resolution and on multiple channels, such as face-to-face meetings, social networks, or phone calls, in order to generate a more comprehensive picture of social systems. The largest study to date measuring large-scale social system—the Copenhagen Networks Study—is described, together with motivation and challenges of the deployment. Preliminary results are presented, indicating how a possibly biased and incomplete picture can be generated when data are collected from a single channel and with a low resolution, thus emphasizing the importance of the proposed approach and deployed implementation.

The second part of the thesis deals with expanding our capabilities to sense the cognitive and emotional state of the users through development of a system for mobile brain imaging—the Smartphone Brain Scanner. A developed framework allows for EEG data collection and processing. It also provides the ability to build end-user applications on top of raw data and extracted features using off-the-shelf and custom-built neuroheadsets and mobile devices, thereby potentially becoming another channel in integrated human sensing. The motivation for creating such system is presented, advanced data processing—3D source reconstruction—is explained, and applications and use-cases are discussed.

In the third part, the privacy issues surrounding the handling of such sensitive behavioral and biomedical data are investigated. A comprehensive review of best privacy practices in sensor-driven human data collection is presented and recommendations for practitioners are made. Based on this review and experiences with the Copenhagen Networks Study and the Smartphone Brain Scanner, the concept of Living Informed Consent is presented, which postulates larger participant control over collected data for the benefit of users, researchers, and society at large. The same privacy principles are applied to a personal neuroinformatics context, resulting in a proposed new approach to sensitive EEG data handling.

Design of Mixed-Criticality Applications on Distributed Real-Time Systems

A mixed-criticality system implements applications of different safety-criticality levels onto the same platform. In such cases, the certification standards require that applications of different criticality levels are protected so they cannot
influence each other. Otherwise, all tasks have to be developed and certified according to the highest criticality level, dramatically increasing the development costs. In this thesis we consider mixed-criticality real-time applications implemented on distributed partitioned architectures.

Partitioned architectures use temporal and spatial separation mechanisms to ensure that applications of different criticality levels do not interfere with each other. With temporal partitioning, each application is allowed to run only within predefined time slots, allocated on each processor. The sequence of time slots for all the applications on a processor are grouped within a Major Frame, which is repeated periodically. Each partition can have its own scheduling policy; we have considered non-preemptive static cyclic scheduling and fixed-priority preemptive scheduling policies. We assume that the communication network implements the TTEthernet protocol, which supports Time-Triggered (TT) messages transmitted based on static schedule tables, Rate Constrained (RC) messages with bounded end-to-end delay, and Best-Effort (BE) messages, for which no timing guarantees are provided. TTEthernet offers spatial separation for mixed-criticality messages through the concept of virtual links, and temporal separation, enforced through schedule tables for TT messages and bandwidth allocation for RC messages.

The objective of this thesis is to develop methods and tools for distributed mixed-criticality real-time systems. At the processor level, we are interested to determine (i) the mapping of tasks to processors, (ii) the assignment of tasks to partitions, (iii) the decomposition of tasks into redundant lower criticality tasks, (iv) the sequence and size of the partition time slots on each processor and (v) the schedule tables, such that all the applications are schedulable and the development and certification costs are minimized. We have proposed Simulated Annealing and Tabu Search metaheuristics to solve these optimization problems. The proposed algorithms have been evaluated using several benchmarks.

At the communication network level, we are interested in the design optimization of TTEthernet networks used to transmit mixed-criticality messages. Given the set of TT and RC messages, and the topology of the network, we are interested to optimize (i) the packing of messages in frames, (ii) the assignment of frames to virtual links, (iii) the routing of virtual links and (iv) the TT static schedules, such that all frames are schedulable and the worst-case end-to-end delay of the RC messages is minimized. We have proposed a Tabu Search-based metaheuristic for this optimization problem.

The proposed algorithm has been evaluated using several benchmarks. The optimization approaches have also been evaluated using realistic aerospace case studies. In this context, we have shown how to extend the proposed optimization frameworks to also take into account quality of service constraints. For TTEthernet networks, we have also proposed a topology selection method to reduce the cost of the architecture.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Embedded Systems Engineering
Authors: Tamas-Selicean, D. (Intern), Pop, P. (Intern), Madsen, J. (Intern)
Number of pages: 160
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2014
Number: 329
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
  phd329_Tamas_Selicean_D.pdf
Publication: Research > Ph.D. thesis – Annual report year: 2015

Multispectral Imaging of Meat Quality - Color and Texture
The use of computer vision systems in food production and development is increasing. Computer vision systems offer fast, reliable, objective and noninvasive methods for assessment of wanted quality traits.

This thesis investigates the applicability of computer vision systems in the assessment of meat quality parameters, especially with regards to meat color and texture. Several image modalities have been applied, all considering multi- or hyper spectral imaging.

The work demonstrates the use of computer vision systems for meat color measurements. The color is assessed by suitable transformations to the CIELAB color space, the common color space within food science. The results show that meat color assessment with a multispectral imaging is a great alternative to the traditional colorimeter, i.e. the vision system meets some of the limitations that the colorimeter possesses. To mention one, it is possible to assess color of very
complicated structures, such as salamis, with a vision system. More importantly though, the vision system embraces the complicated scattering properties of meat.

The images can also lead to other analyses, e.g. image texture analysis relating to the structure of the meat. In the thesis it is presented how simple texture measures can be used for characterizing the texture changes in fermented salamis. Moreover, it was investigated if it was possible to relate structure in images to chemical compounds in lard from boars.

**General information**
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Trinderup, C. H. (Intern), Conradsen, K. (Intern), Dahl, A. B. (Intern)
Number of pages: 159
Publication date: 2015

**Publication information**
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2014
Number: 358
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd358_Trinderup_CH.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

**Interpretation of images from intensity, texture and geometry**
The goal of the thesis is to develop flexible mathematical methods for quantitative interpretation of image content. Problems from research areas as diverse as evolutionary biology, remote sensing and materials science have motivated the methodological development. The solutions are inspired by classical mathematical image analysis techniques, information theory, probabilistic graphical models and manifold learning.

Specifically, the thesis revolves around describing three major components of images, namely intensity, texture and geometry. Intensity distribution modelling is important for obtaining useful global representations of the raw image data. Texture description provides a local representation of the image content, useful for descriptive and discriminative scenarios. Geometrical knowledge of the image content is leveraged within the framework of Markov random fields. Mathematical models are developed around these three topics and constitute building blocks useful for engineering image-based solutions to a wide range of problems.

The contributions include automated quantification of frog patterning from field imagery, statistical methods for estimating the genetic basis of quantified mimicry phenotypes, estimation of the atomic structure of graphene from low-contrast transmission electron microscopy images and patch-based crop classification from synthetic aperture radar data. Further, an information theoretic approach to two-set image decomposition is presented, representing a purely methodological contribution.

This thesis makes statistical image analysis available to fellow researchers with domain specific problems, and provides new methodology relevant for the field itself.

**General information**
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Vestergaard, J. S. (Intern), Larsen, R. (Intern), Nielsen, A. A. (Intern)
Number of pages: 290
Publication date: 2015

**Publication information**
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2014
Number: 346
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
Availability by Design: A Complementary Approach to Denial-of-Service

In computer security, a Denial-of-Service (DoS) attack aims at making a resource unavailable. DoS attacks to systems of public concern occur increasingly and have become infamous on the Internet, where they have targeted major corporations and institutions, thus reaching the general public. There exist various practical techniques to face DoS attacks and mitigate their effects, yet we witness the successfulness of many.

The need for a renewed investigation of availability gains in relevance when considering that our life is more and more dominated by Cyber-Physical Systems (CPSs), large-scale network of sensors that interact with the physical environment. CPSs are increasingly exploited in the realisation of critical infrastructure, from the power grid to healthcare, traffic control, and defence applications. Such systems are particularly prone to DoS attacks: in addition to classic communication-based attacks, their components can be subject to physical capture. Moreover, sensors are often powered by batteries, and time-limited unavailability is usually a stage planned to prolong their life span.

This dissertation argues that techniques rooted in the theory and practice of programming languages, language-based techniques, offer a unifying framework to deal with the consequences of DoS, thereby encompassing inadvertent and malicious sources of unavailability in a uniform manner.

In support to this claim we develop a family of process calculi, the Quality Calculi, where availability considerations are promoted to be first-class object of the language domain. Moreover, these modelling tools are complemented by static analyses that pinpoint where and why unavailability may occur, levering the enhanced expressiveness of the language.

The ultimate aim of the framework is to foster the development of systems resilient to DoS by means of a principled design process, in which formal models allow, and verification tools enforce, the production of such robust code.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science , Language-Based Technology
Authors: Vigo, R. (Intern), Nielson, H. R. (Intern), Nielson, F. (Intern)
Number of pages: 190
Publication date: 2015

Publication information

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2014
Number: 353
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd353_Vigo_R.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

Topics in combinatorial pattern matching

This dissertation studies problems in the general theme of combinatorial pattern matching. More specifically, we study the following topics:

Longest Common Extensions. We revisit the longest common extension (LCE) problem, that is, preprocess a string T into a compact data structure that supports fast LCE queries. An LCE query takes a pair (i, j) of indices in T and returns the length of the longest common prefix of the suffixes of T starting at positions i and j. Such queries are also commonly known as longest common prefix (LCP) queries. We study the time-space trade-offs for the problem, that is, the space used for the data structure vs. the worst-case time for answering an LCE query. Let n be the length of T. Given a parameter τ, 1 ≤ τ ≤ n, we show how to achieve either O(n/√τ) space and O(τ) query time, or O(n/τ) space and O(τ log (|LCE(i, j)|/τ)) query time, where |LCE(i, j)| denotes the length of the LCE returned by the query. These bounds provide the first smooth trade-offs for the LCE problem and almost match the previously known bounds at the extremes when τ = 1 or τ = n. We apply the result to obtain improved bounds for several applications where the LCE problem is the computational bottleneck, including approximate string matching and computing palindromes. We also present an efficient technique to reduce LCE queries on two strings to one string. Finally, we give a lower bound on the time-space product for LCE data structures in the non-uniform cell probe model showing that our second trade-off is nearly optimal.

Fingerprints in Compressed Strings. The Karp-Rabin fingerprint of a string is a type of hash value that due to its strong properties has been used in many string algorithms. We show how to construct a data structure for a string S of size N compressed by a context-free grammar of size n that supports fingerprint queries. That is, given indices i and j, the answer to a query is the fingerprint of the substring S[i, j]. We present the first O(n) space data structures that answer fingerprint queries without decompressing any characters. For Straight Line Programs (SLP) we get O(log N) query time, and
for Linear SLPs (an SLP derivative that captures LZ78 compression and its variations) we get $O(\log \log N)$ query time. Hence, our data structures has the same time and space complexity as for random access in SLPs. We utilize the fingerprint data structures to solve the longest common extension problem in query time $O(\log N \log e)$ and $O(\log e \log \log e + \log \log N)$ for SLPs and Linear SLPs, respectively. Here, $e = |\text{LCE}(i, j)|$ denotes the length of the LCE.

Sparse Text Indexing. We present efficient algorithms for constructing sparse suffix trees, sparse suffix arrays and sparse positions heaps for $b$ arbitrary positions of a text $T$ of length $n$ while using only $O(b)$ words of space during the construction. Our main contribution is to show that the sparse suffix tree (and array) can be constructed in $O(n \log^2 b)$ time. To achieve this we develop a technique, that allows to efficiently answer $b$ longest common prefix queries on suffixes of $T$, using only $O(b)$ space. Our first solution is Monte-Carlo and outputs the correct tree with high probability. We then give a Las-Vegas algorithm which also uses $O(b)$ space and runs in the same time bounds with high probability when $b = O(n)$. Furthermore, additional tradeoffs between the space usage and the construction time for the Monte-Carlo algorithm are given. Finally, we show that at the expense of slower pattern queries, it is possible to construct sparse position heaps in $O(n + b \log b)$ time and $O(b)$ space.

The Longest Common Substring Problem. Given $m$ documents of total length $n$, we consider the problem of finding a longest string common to at least $d \geq 2$ of the documents. This problem is known as the longest common substring (LCS) problem and has a classic $O(n)$ space and $O(n)$ time solution (Weiner [FOCS'73], Hui [CPM'92]). However, the use of linear space is impractical in many applications. We show several time-space trade-offs for this problem. Our main result is that for any trade-off parameter $1 \leq \tau \leq n$, the LCS problem can be solved in $O(\tau)$ space and $O(n^{\tau/\tau'})$ time, thus providing the first smooth deterministic time-space trade-off from constant to linear space. The result uses a new and very simple algorithm, which computes a $\tau$-additive approximation to the LCS in $O(n^{\tau/\tau'})$ time and $O(1)$ space. We also show a time-space trade-off lower bound for deterministic branching programs, which implies that any deterministic RAM algorithm solving the LCS problem on documents from a sufficiently large alphabet in $O(\tau)$ space must use $\Omega(n^{\sqrt{\log(n/(\tau \log n))}})$ time.

Structural Properties of Suffix Trees. We study structural and combinatorial properties of suffix trees. Given an unlabeled tree $T$ on $n$ nodes and suffix links of its internal nodes, we ask the question “Is $T$ a suffix tree?”, i.e., is there a string $S$ whose suffix tree has the same topological structure as $T$? We place no restrictions on $S$, in particular we do not require that $S$ ends with a unique symbol. This corresponds to considering the more general definition of implicit or extended suffix trees. Such general suffix trees have many applications and are for example needed to allow efficient updates when suffix trees are built online. We prove that $T$ is a suffix tree if and only if it is realized by a string $S$ of length $n - 1$, and we give a linear-time algorithm for inferring $S$ when the first letter on each edge is known.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Algorithms and Logic
Authors: Vildhøj, H. W. (Intern), Gørtz, I. L. (Intern), Bille, P. (Intern)
Number of pages: 147
Publication date: 2015

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2014
Number: 348
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:

phd348_Vilhoj_HW.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

Algorithms and Data Structures for Strings, Points and Integers: or, Points about Strings and Strings about Points
This dissertation presents our research in the broad area of algorithms and data structures. More specifically, we show solutions for the following problems related to strings, points and integers. Results hold on the Word RAM and we measure space in $w$-bit words.

Compressed Fingerprints. The Karp-Rabin fingerprint of a string is a useful type of hash value that has multiple applications due to its strong properties. Given a string $S$ of length $N$ compressed into a straight line program (SLP) of size $n$, we show a $O(n)$ space data structure that supports fingerprint queries, retrieving the fingerprint of any substring of $S$. Queries are answered in $O(\lg N)$ time. If the compression is a Linear SLP (capturing LZ78 compression and variations), we get $O(\lg \lg N)$ query time.

Our structure matches the best known query time bound for random access in SLPs, and is the first for general (unbalanced) SLPs that answers fingerprint queries without decompressing any text. We also support longest common
extension queries, returning the length \( r \) that the substrings from two given positions in \( S \) are equal. Answers are correct w.h.p. and take time \( O(lg N \ lg r) \) and \( O(lg N + lg r \ lg lg r) \) for SLPs and Linear SLPs, respectively.

Dynamic Compression. In the dynamic relative compression scheme, we compress a string \( S \) of length \( N \) into \( n \) substrings of a given reference string of length \( r \). We give data structures that maintain an asymptotically optimal compression in the scheme and support access, replace, insert and delete operations on \( S \). Our solutions support each operation in \( O(lg n / lg lg n) \) time and \( O(n + r) \) space, or \( O(lg n / lg lg n) \) time and \( O(n + r \ lg r) \) space. They can be naturally generalized to compress multiple strings.

Our solutions obtains almost-optimal bounds, and are the first to dynamically maintain a string under a compression scheme that can achieve better than entropy compression. We also give improved results for the substring concatenation problem, and an extension of our structure can be used as a black box to get an improved solution to the previously studied dynamic text static pattern problem.

Compressed Pattern Matching. In the streaming model, input data flows past a client one item at a time, but is far too large for the client to store. The annotated streaming model extends the model by introducing a powerful but untrusted annotator (representing "the cloud") that can annotate input elements with additional information, sent as one-way communication to the client. We generalize the annotated streaming model to be able to solve problems on strings and present a data structure that allows us to trade off client space and annotation size. This lets us exploit the power of the annotator.

In compressed pattern matching we must report occurrences of a pattern of length \( m \) in a text compressed into \( n \) phrases (capturing LZ78 compression and variations). In the streaming model, any solution to the problem requires \( O(n) \) space. We show that the problem can be solved in \( O(lg n) \) client space in the annotated streaming model, using \( O(lg n) \) time and \( O(lg n) \) words of annotation per phrase. Our solution shows that the annotator lets us solve previously impossible problems, and it is the first solution to a classic problem from combinatorial pattern matching in the annotated streaming model.

Pattern Extraction. The problem of extracting important patterns from text has many diverse applications such as data mining, intrusion detection and genomic analysis. Consequently, there are many variations of the pattern extraction problem with different notions of patterns and importance measures. We study a natural variation where patterns must 1) contain at most \( k \) don’t cares that each match a single character, and 2) have at least \( q \) occurrences. Both \( k \) and \( q \) are input parameters.

We show how to extract such patterns and their occurrences from a text of length \( n \) in \( O(nk + k3occ) \) time and space, where \( occ \) is the total number of pattern occurrences. Our bound is the first output-sensitive solution for any approximate variation of the pattern extraction problem, with all previous solutions requiring \( O(n2) \) time per reported pattern. Our algorithm is relatively simple, but requires a novel analysis technique that amortizes the cost of creating the index over the number of pattern occurrences.

Compressed Point Sets. Orthogonal range searching on a set of points is a classic geometric data structure problem. Given a query range, solutions must either count or report the points inside the range. Variants of this problem has numerous classic solutions, typically storing the points in a tree.

We show that almost any such classic data structure can be compressed without asymptotically increasing the time spent answering queries. This allows us to reduce the required space use if the point set contains geometric repetitions (copies of equal point set that are translated relative to each other). Our result captures most classic data structures, such as Range Trees, KD-trees, R-trees and Quad Trees. We also show a hierarchical clustering algorithm for ensuring that geometric repetitions are compressed.

Points with Colors. Colored orthogonal range searching is a natural generalization of orthogonal range searching which allows us to perform statistic analysis of a point set. We must store \( n \) points that each have a color (sometimes called a category) and support queries that either count or report the distinct colors of the points inside a query range.

We show data structures that support both types of queries in sublinear time, storing two-dimensional points in linear space and high-dimensional points in almost-linear space. These are the first (almost) linear space solutions with sublinear query time. We also give the first dynamic solution with sublinear query time for any dimensionality. Previous solutions answer queries faster, but require much more space.

Points with Weights in Practice. If points are each assigned a weight, it is natural to consider the threshold range counting problem. A data structure must store the points and be able to count the number of points within a query range with a weight exceeding some threshold. This query appears naturally in a software system built by Milestone Systems, and allows detecting motion in video from surveillance cameras.

We implement a prototype of an index for 3-dimensional points that use little space and answers threshold queries efficiently. In experiments on realistic data sets, our prototype shows a speedup of at least a factor 30 at the expense of 10% additional space use compared to the previous approach. An optimized version of our proposed index is implemented in the latest version of the Milestone Systems software system.

Finger Predecessor. The predecessor problem is to store a set of \( n \) integers from a universe of size \( N \) to support predecessor queries, returning the largest integer in the set smaller than a given integer \( q \). We study a variation where the query additionally receives a finger to an integer \( r \) in the set from which to start the search. We show a linear space data structure that answers such finger predecessor queries in \( O(lg lg |r - q|) \) time. This generalizes and improves the \( O(lg lg N) \) time solutions for the standard predecessor problem. Our data structure is the first with a query time that only depends on
the numerical distance between the finger and the query integer.

Dynamic Partial Sums. The well-studied partial sums problem is to store a sequence of n integers with support for sum and search queries. The sequence is static in the sense that its length cannot change, but the update operation can be used to change the value of an integer in the sequence by a given value. There are matching lower and upper bounds showing that the problem can be solved on the w-bit Word RAM in linear space and \( \Omega(n/\log(w)) \) time per operation, where \( w \) is the maximum number of bits allowed in updates.

As a natural generalization we consider dynamic partial sums, allowing insertions and deletions in the sequence. Our solution requires linear space and supports all operations in optimal worst-case time \( O(n/\log(w)) \), matching lower bounds for all supported operations. Our data structure is the first dynamic partial sums solution that matches the lower bounds, and the first to support storing integers of more than \( \log w \) bits.

Mathematical modelling of membrane separation
This thesis concerns mathematical modelling of membrane separation. The thesis consists of introductory theory on membrane separation, equations of motion, and properties of dextran, which will be the solute species throughout the thesis. Furthermore, the thesis consist of three separate mathematical models, each with a different approach to membrane separation.

The first model is a statistical model investigating the interplay between solute shape and the probability of entering the membrane. More specific the transition of solute particles from being spherical to becoming more elongated as prolate ellipsoids with the same volume. The porous membrane is assumed isotropic such that the model reduces to a two dimensional model. With this assumption ellipsoids with the same volume reduces to ellipses with the same area. The model finds the probability of entering the pore of the membrane. It is found that the probability of entering the pore is highest when the largest of the radii in the ellipse is equal to half the radius of the pore, in case of molecules with circular radius less than the pore radius. The results are directly related to the macroscopic distribution coefficient and the rejection coefficient.

The second model is a stationary model for the flux of solvent and solute in a hollow fibre membrane. In the model we solve the time independent equations for transport of solvent and solute within the hollow fibre. Furthermore, the flux of solute and solvent through the membrane is coupled through the boundary conditions. The model investigates how the true and observed rejection coefficient depends on the transmembrane pressure, the average inlet velocity, and the molecular weight. Furthermore, the effect of concentration dependent viscosity on the rejection coefficients is investigated. The results show that the true rejection coefficient is increasing as a function of increasing transmembrane pressure, increasing inlet velocity, and decreasing molecular weight. Furthermore, it is found that a concentration dependent viscosity decreases the true rejection. The observed rejection is increasing for decreasing molecular weight and increasing inlet velocities. The observed rejection can be either increasing or decreasing as a function of increasing transmembrane pressure. Moreover, the observed rejection is reduced when the viscosity depends on the concentration. The study is a time dependent model of back-shocking. During back-shocking the pressure difference across the membrane is reversed for a given time. This implies that the concentration polarization at the membrane surface is flushed away. When the pressure is reversed back to normal the membrane performs better resulting in an increased average flux. Two models models of the problem was made.

In a two dimensional model, limited to capture the dynamics close to the membrane, a positive effect was observed on both the observed rejection and the average solvent flux. Furthermore, an analytical upper estimate for the optimal back-shock time is given. In a three dimensional model, where the flow within the entire hollow fibre is modelled, the mentioned upper estimate is used to obtain a positive effect on both the observed rejection and the average solvent flux. Moreover,
the effect of a concentration dependent viscosity was investigated. It was found that the average flux compared to the steady-state solution increased when the viscosity depends on the concentration.

**Formal Development and Verification of Railway Control Systems - In the context of ERTMS/ETCS Level 2**

This dissertation presents a holistic, formal method for efficient modelling and verification of safety-critical railway control systems that have product line characteristics, i.e., each individual system is constructed by instantiating common generic applications with concrete configuration data. The proposed method is based on a combination of formal methods and domain-specific approaches. While formal methods offer mathematically rigorous specification, verification and validation, domain-specific approaches encapsulate the use of formal methods with familiar concepts and notions of the domain, hence making the method easy for the railway engineers to use. Furthermore, the method features a 4-step verification and validation approach that can be integrated naturally into different phases of the software development process. This 4-step approach identifies possible errors in generic applications or configuration data as early as possible in the software development cycle, and facilitates debugging/troubleshooting if errors are discovered. The proposed method has successfully been applied to case studies of the forthcoming Danish railway interlocking systems that are compatible with the European standardized railway control systems ERTMS/ETCS Level 2. Experiments showed that the method can be used for specification, verification and validation of systems of industrial size.

**High-resolution imaging methods in array signal processing**

The purpose of this study is to develop methods in array signal processing which achieve accurate signal reconstruction from limited observations resulting in high-resolution imaging. The focus is on underwater acoustic applications and sonar signal processing both in active (transmit and receive) and passive (only receive) mode. The study addresses the limitations of existing methods and shows that, in many cases, the proposed methods overcome these limitations and outperform traditional methods for acoustic imaging.

The project comprises two parts; The first part deals with computational methods in active sonar signal processing for
detection and imaging of submerged oil contamination in sea water from a deep-water oil leak. The submerged oil field is modeled as a uid medium exhibiting spatial perturbations in the acoustic parameters from their mean ambient values which cause weak scattering of the incident acoustic energy. A highfrequency active sonar is selected to insonify the medium and receive the backscattered waves. High-frequency acoustic methods can both overcome the optical opacity of water (unlike methods based on electromagnetic waves) and resolve the small-scale structure of the submerged oil field (unlike low-frequency acoustic methods). The study shows that high-frequency acoustic methods are suitable not only for large-scale localization of the oil contamination in the water column but also for statistical characterization of the submerged oil field through inference of the spatial covariance of its acoustic parameters.

The second part of the project investigates methods that exploit sparsity in order to achieve super-resolution in sound source localization with passive sonars. Sound source localization with sensor arrays involves the estimation of the direction-of-arrival (DOA) of the associated wavefronts from a limited number of observations. Usually, there are only a few sources generating the acoustic wavefield such that DOA estimation is essentially a sparse signal reconstruction problem. Conventional methods for DOA estimation (i.e., beamforming) suffer from resolution limitations related to the physical size and the geometry of the array. DOA estimation methods that are developed up-to-date in order to overcome the resolution limitations of conventional methods involve the estimation or the eigendecomposition of the data cross-spectral matrix. The cross-spectral methods require many snapshots (i.e., observation windows of the recorded wavefield) hence are suitable only for stationary incoherent sources. In this study, the DOA estimation problem is formulated both for single and multiple snapshots in the compressive sensing framework (CS), which achieves sparsity, thus improved resolution, and can be solved efficiently with convex optimization. It is shown that CS has superior performance compared to traditional DOA estimation methods especially under challenging scenarios such as coherent arrivals, single-snapshot data and random array configurations. The high-resolution performance and the robustness of CS in DOA estimation are validated with experimental array data from ocean acoustic measurements.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing, Acoustic Technology
Authors: Xenaki, A. (Intern), Knudsen, K. (Intern)
Number of pages: 121
Publication date: 2015

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2015
Number: 368
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd368_Xenaki_A.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015

**Compilation and Synthesis for Fault-Tolerant Digital Microfluidic Biochips**

Microfluidic-based biochips are replacing the conventional biochemical analyzers, by integrating all the necessary functions for biochemical analysis using microfluidics. The digital microfluidic biochips (DMBs) manipulate discrete amounts of fluids of nanoliter volume, named droplets, on an array of electrodes to perform operations such as dispensing, transport, mixing, split, dilution and detection.

Researchers have proposed compilation approaches, which, starting from a biochemical application and a biochip architecture, determine the allocation, resource binding, scheduling, placement and routing of the operations in the application. During the execution of a bioassay, operations could experience transient faults, thus impacting negatively the correctness of the application. We have proposed both offline (design time) and online (runtime) recovery strategies. The online recovery strategy decides the introduction of the redundancy required for fault-tolerance. We consider both time redundancy, i.e., re-executing erroneous operations, and space redundancy, i.e., creating redundant droplets for fault-tolerance. Error recovery is performed such that the number of transient faults tolerated is maximized and the timing constraints of the biochemical application are satisfied.

Previous work has assumed that the biochip architecture is given, and most approaches consider a rectangular shape for the electrode array, where operations execute on rectangular “modules” formed of electrodes. However, non-regular application-specific architectures are common in practice. Hence, we have proposed an approach to the synthesis of application-specific architectures, such that the cost is minimized and the timing constraints of the application are satisfied.

We propose an algorithm to build a library of non-regular modules for a given applicationspecific architecture, so that the area of a non-regular application-specific biochip can be used effectively. During fabrication, DMBs can be affected by...
permanent faults, which may lead to the failure of the application. Our approach introduces redundant electrodes to synthesize fault-tolerant architectures aiming at increasing the yield of DMBs. We also propose a method to estimate, at design time, the application completion time in case of permanent faults in order to verify if an application can be successfully run on the architecture.

The proposed approaches were evaluated using several real-life case studies and synthetic benchmarks.

**Brain Network Modelling**

Three main topics are presented in this thesis. The first and largest topic concerns network modelling of functional Magnetic Resonance Imaging (fMRI) and Diffusion Weighted Imaging (DWI). In particular nonparametric Bayesian methods are used to model brain networks derived from resting state fMRI data. The models used are the Infinite Relational Model (IRM), Bayesian Community Detection (BCD), and Infinite Diagonal Model (IDM). The models have different constraints on how they cluster nodes. IRM is flexible in the sense that it allows for complex interactions between clusters of nodes. BCD conforms to the definition of community structure in the sense that it forces clusters of nodes to have larger density of internal connections than external connections. IDM models only the linking within a cluster and treats linking between clusters as background noise. The models are evaluated for their ability to reproduce node clustering and predict unseen data. Comparing the models on whole brain networks, BCD and IRM showed better reproducibility and predictability than IDM, suggesting that resting state networks exhibit community structure. This also points to the importance of using models, which allow for complex interactions between all pairs of clusters. In addition, it is demonstrated how the IRM can be used for segmenting brain structures into functionally coherent clusters.

A new nonparametric Bayesian network model is presented. The model builds upon the IRM and can be used to infer shared clustering structure across different types of networks. The model is used to jointly model fMRI and DWI networks. However, results show only a limited amount of sharing across fMRI and DWI networks. Using the model within the same modality can reveal the clustering consistency across scans. A high consistency was found between DWI networks and an intermediate level of consistency was found between fMRI networks. The model is of interest for other applications, for instance in finding dissimilarity between network structure in case-control studies.

The second topic of the thesis concerns local functional connectivity. In particular the local functional connectivity is studied in patients with multiple sclerosis (MS). The functional connectivity in a small neighborhood was estimated using Kendall’s Coefficient of Concordance (KCC). By generating voxelwise KCC maps, MS patients were compared with healthy controls. MS patients had reduced KCC in cerebellum and KCC correlated negatively with disease progression. Lesion load of the left cerebellar peduncles correlated negatively with KCC suggesting that the reduced local connectivity in MS is caused by disrupted inputs to the cerebellum.

The final topic of this thesis concerns model selection for Gaussian Kernel Principal Component Analysis (KPCA) denoising. KPCA can be used for non-linear denoising by mapping data to feature space using a non-linear map. By projecting data onto a subspace in feature space and mapping this projection back to input space noise in data is (hopefully) removed. However, two important parameters must be set, namely the scale of the Gaussian kernel and the subspace dimensionality. A principled method for selecting these two parameters is presented. The method is based on maximizing the signal energy in feature space. When testing on synthetic and real data, the method outperformed a number of other heuristics in terms of signal to noise ratio of the denoised data.
IEA Common Exercise 4: ARX, ARMAX and grey-box models for thermal performance characterization of the test box

In this report results of applying time series models for assessing the thermal performance of the IEA Annex 58 test box based on data given in the Common Exercise 4 (CE4), which was measured in Almeria, Spain. Both ARX, ARMAX and grey-box models are applied. Finally, the same models are fitted for the Common Exercise 3b (CE3) data measured in Belgium and the results are compared.

The focus in this report is on model selection and validation enabling a stable and reliable performance assessment. Basically, the challenge is to find a procedure for each type of model, which can give un-biased and accurate estimates of the essential performance parameters, including reliable uncertainties of the estimates. Important is also the development of methodologies for analyzing the quality of data, for example correlated inputs and lack of information in data (e.g. if no clearsky days with direct solar radiation is present data), these aspects are discussed. Furthermore, new models for enhancing the description of the effect of solar radiation on the test box is presented.

Mathematical Modeling and Dimension Reduction in Dynamical Systems

Processes that change in time are in mathematics typically described by differential equations. These may be applied to model everything from weather forecasting, brain patterns, reaction kinetics, water waves, finance, social dynamics, structural dynamics and electrodynamics to name only a few. These systems are generically nonlinear and the studies of them often become enormously complex. The framework in which such systems are best understood is via the theory of dynamical systems, where the critical behavior is systematically analyzed by performing bifurcation theory. In that context the current thesis is attacking two problems.

The first is concerned with the mathematical modelling and analysis of an experiment of a vibro-impacting beam. This type of dynamical system has received much attention in the recent years and they occur frequently in mechanical applications, where they induce noise and wear which decrease the life time of machines. From the modelling point of view these systems are often particularly rich in nonlinear dynamics. In the present study a mathematical model is derived. Amongst other outcomes the model was successfully applied to predict a nonlinear phenomenon, namely the existence of isolas of subharmonic orbits. These were then verified in the practical experiment in the lab. The second problem that is addressed in the current thesis is a problem that has developed as a consequence of the increasing power of computers which has
created the demand for analysis of ever more advanced and complex systems. These complex systems are computationally very demanding and proper analysis of the qualitative behavior of the systems becomes difficult. In order to overcome this obstacle much research is going into the direction of development of robust methods to perform dimension and model reduction such as to pave the way for a qualitative analysis of the high-dimensional problems by analyzing the low-dimensional models.

In this thesis we demonstrate how to reduce the dimension of a certain class of dynamical systems by construction of k-dimensional submanifolds using the so-called graph transform. The method is suitable for a specific class of problems with spectral gaps, these are often observed. In particular the method is applied to a mechanical system. Furthermore the method has some unique and promising properties compared to other methods.

**General information**
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Scientific Computing
Authors: Elmegård, M. (Intern), Starke, J. (Intern), Evgrafov, A. (Intern), Thomsen, J. J. (Intern)
Number of pages: 158
Publication date: 2014

**Publication information**
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2014
Number: 321
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd321_Elmegaard_M.pdf
Publication: Research > Ph.D. thesis – Annual report year: 2014

**Electrophysiological assessment of audiovisual integration in speech perception**
Speech perception integrates signal from ear and eye. This is witnessed by a wide range of audiovisual integration effects, such as ventriloquism and the McGurk illusion. Some behavioral evidence suggest that audiovisual integration of specific aspects is special for speech perception. However, our knowledge of such bimodal integration would be strengthened if the phenomena could be investigated by objective, neutrally based methods. One key question of the present work is if perceptual processing of audiovisual speech can be gauged with a specific signature of neurophysiological activity, the mismatch negativity response (MMN).

MMN has the property of being evoked when an acoustic stimulus deviates from a learned pattern of stimuli. In three experimental studies, this effect is utilized to track when a coinciding visual signal alters auditory speech perception. Visual speech emanates from the face of the talker. Perception of faces and of speech shares the trait, that they are learned from infancy and seemingly specialized behaviorally and neurally. Due to this, speech and face encoding functions quasi-automatically and with high efficiency. However, perhaps owing to our long experience with human faces, which all are variations on a relatively constrained space of features, face perception is sensitive to manipulations of the structure of the face, the relation between its segments, and the properties of the segments. Does this sensitivity alter the influence of visual speech on the auditory speech percept? In two experiments, which both combine behavioral and neurophysiological measures, an uncovering of the relation between perception of faces and of audiovisual integration is attempted. Behavioral findings suggest a strong effect of face perception, whereas the MMN results are less clear. Another interesting property of speech perception is that it is relatively tolerant towards temporal shifts between acoustic and visual speech signals. Here, behavioral studies report that perception of speech exhibits far greater temporal tolerance than towards non-speech stimuli. Current findings on neural correlates of this tolerance, however, are few and limited. Here, a novel experimental MMN paradigm is used in effort to shed light on integration asynchronous audiovisual speech. Based on individual behavioral estimates of temporal windows of tolerance, we ask if the MMN signal can be evoked at different points within and outside this window. Behavioral findings match earlier behavioral studies, whereas the MMN findings are ambiguous.

In conclusion, the work presented here sheds light onto two important aspects of speech perception. It also presents important methodological conclusions on the use of MMN as a neural marker of audiovisual integration.

**General information**
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems, Hearing Systems, Department of Electrical Engineering
Authors: Eskelund, K. (Intern), Andersen, T. (Intern), MacDonald, E. (Intern), Dau, T. (Intern)
Number of pages: 148
Publication date: 2014
Medium Access Control in Energy Harvesting - Wireless Sensor Networks

Focusing on Wireless Sensor Networks (WSN) that are powered by energy harvesting, this dissertation focuses on energy-efficient communication links between senders and receivers that are alternating between active and sleeping states of operation. In particular, the focus lies on Medium Access Control (MAC) protocols that are following the receiver-initiated paradigm of asynchronous communication. According to the receiver-initiated paradigm the communication is initiated by the receiver that states its availability to receive data through beacons. The sender is passively listening to the channel until it receives the beacon of interest.

In this context, the dissertation begins with an in-depth survey of all the receiver-initiated MAC protocols and presents their unique optimization features, which deal with several challenges of the link layer such as mitigation of the energy consumption, collision avoidance, provision of Quality of Service (QoS) and security. Focusing on the particular requirements of an energy harvesting application, the dissertation continues with the presentation of a MAC protocol, named ODMAC, which extends the receiver-initiated paradigm with several energy-efficient features that aim to adapt the consumed energy to match the harvested energy, distribute the load with respect to the harvested energy, decrease the overhead of the communication, address the requirements for collision avoidance, prioritize urgent traffic and secure the system against beacon replay attacks.

The performance and behavior of ODMAC and its features are compared to the state-of-the-art and evaluated using mathematical models, simulations and testbed experiments that are based on eZ430-rf2500 wireless development platform. The results validate the efficient use of the harvested energy and demonstrate sustainable operation.

Interior Point Methods on GPU with application to Model Predictive Control

The goal of this thesis is to investigate the application of interior point methods to solve dynamical optimization problems, using a graphical processing unit (GPU) with a focus on problems arising in Model Predictive Control (MPC). Multi-core processors have been available for over ten years now, and manycore processors, such as GPUs, have also become a standard component in any consumer computer. The GPU offers faster floating point operations and higher memory bandwidth than the CPU, but requires algorithms to be redesigned and implemented, to match the underlying architecture. A large number of different optimization algorithms are available for solving optimization problems. Some of the most common method are the simplex method and interior point methods. We focus on interior point methods in this thesis, due to its polynomial complexity, and since the use of the simplex method with GPUs have been investigated by several other authors already. The main computational task in interior point methods is the solution of a linear system to compute the Newton direction in each iteration. Direct interior point methods use a direct method such as Cholesky factorization to
factorize the normal equations of the Hessian matrix. The use of a GPU has been shown to be very efficient in the factorization of dense matrices, and several numeric libraries, which utilize the GPU, have become available during the course of this thesis. We have developed a direct interior point method, which utilizes the GPU, and demonstrate that our implementation can reduce the solution time substantially.

There are multiple software packages available for solving optimization problems with interior point methods, such as GLPK, IPOPT, MOSEK and many more. However, none of these support the GPU yet. With this thesis, we include a new software package called GPUOPT, available under the non-restrictive MIT license. GPUOPT includes a primal-dual interior-point method, which supports both the CPU and the GPU. It is implemented as multiple components, where the matrix operations and solver for the Newton directions is separated from the core interior point method. This makes it possible to replace the matrix operations and solver with alternative, and potentially problem-specific, implementations.

In this thesis, we include different implementations of the matrix operations, including general dense, general sparse and problem-specific implementation of a test problem from model predictive control. Multiple solvers are implemented as well, including a direct solver based on CHOLMOD, and an iterative solver which uses preconditioned conjugate gradient. The iterative solver is based on the matrix-free iterative interior point method.

Tradeoff analysis for Dependable Real-Time Embedded Systems during the Early Design Phases

Embedded systems are becoming increasingly complex and have tight competing constraints in terms of performance, cost, energy consumption, dependability, flexibility, security, etc. The objective of this thesis is to propose design methods and tools for supporting the tradeoff analysis of competing design objectives during the early design phases, which are characterized by uncertainties. We consider safety-critical real-time applications modeled as task graphs, to be implemented on distributed heterogeneous architectures consisting of processing elements (PEs), interconnected by a shared communication channel. Tasks are scheduled using fixed-priority preemptive scheduling, and we use non-preemptive scheduling for messages.

As a first step, we address the problem of function-to-task decomposition. In this context we have assumed that the application functionality is captured by a set of functional blocks, with different safety requirements. We propose a Genetic Algorithm-based metaheuristic to solve the function-to-task decomposition problem. Our algorithm also decides the mapping of tasks to the PEs of a distributed architecture and the reliability of each PE in the architecture, such that the safety and integrity constraints are satisfied, the schedulability of the real-time application is guaranteed and the overall development and product unit costs are minimized.

Next, we investigate tradeoffs between performance, energy and reliability. Addressing energy and reliability simultaneously is especially challenging, since lowering the voltage to reduce the energy consumption has been shown to increase the transient fault rate. We are interested to tolerate transient faults and we use task replication for recovery. We propose a Tabu Search-based approach, which decides the mapping of tasks to processing elements, as well as the processor voltage and frequency levels for executing each task, such that transient faults are tolerated, the real-time constraints of the application are satisfied, and the energy consumed is minimized.

In this thesis, we target the early design phases, when decisions have a high impact on the subsequent implementation choices. However, due to a lack of information, the early design phases are characterized by uncertainties, e.g., in the worst-case execution times (WCETs), in the functionality requirements, or in the hardware component costs. In this context, we select the hardware components for the architecture and derive a mapping of tasks in the application, such that the resulted implementation is both robust and flexible. The architecture also has a high chance to have its unit cost within the cost budget. Robust means that the application has a high chance of being schedulable, considering the WCET uncertainties, whereas a flexible mapping has a high chance to successfully accommodate future functionality changes. We propose a Genetic Algorithm-based approach to solve this optimization problem. The proposed tradeoff analysis
methods have been evaluated using several synthetic and real-life benchmarks.

**Model Predictive Control for Smart Energy Systems**

In this thesis, we consider control strategies for flexible distributed energy resources in the future intelligent energy system – the Smart Grid. The energy system is a large-scale complex network with many actors and objectives in different hierarchical layers. Specifically, the power system must supply electricity reliably to both residential and industrial consumers around the clock. More and more fluctuating renewable energy sources, like wind and solar, are integrated in the power system. Consequently, uncertainty in production starts to affect an otherwise controllable power production significantly. A Smart Grid calls for flexible consumers that can adjust their consumption based on the amount of green energy in the grid. This requires coordination through new large-scale control and optimization algorithms. Trading of flexibility is key to drive power consumption in a sustainable direction. In Denmark, we expect that distributed energy resources such as heat pumps, and batteries in electric vehicles will mobilize part of the needed flexibility.

Our primary objectives in the thesis were threefold:

1. Simulate the components in the power system based on simple models from literature (e.g. heat pumps, heat tanks, electrical vehicle battery charging/discharging, wind farms, power plants).
2. Embed forecasting methodologies for the weather (e.g. temperature, solar radiation), the electricity consumption, and the electricity price in a predictive control system.
3. Develop optimization algorithms for large-scale dynamic systems. This includes decentralized optimization and simulation on realistic large-scale dynamic systems.

Chapter 1 introduces the power system, the markets, and the main actors. The objectives and control hierarchy is outlined while Aggregators are introduced as new actors.

Chapter 2 provides linear dynamical models of Smart Grid units: Electric Vehicles, buildings with heat pumps, refrigeration systems, solar collectors, heat storage tanks, power plants, and wind farms. The models can be realized as discrete time state space models that fit into a predictive control system.

Chapter 3 introduces Model Predictive Control (MPC) including state estimation, filtering and prediction for linear models.

Chapter 4 simulates the models from Chapter 2 with the certainty equivalent MPC from Chapter 3. An economic MPC minimizes the costs of consumption based on real electricity prices that determined the flexibility of the units. A predictive control system easily handles constraints, e.g. limitations in power consumption, and predicts the future behavior of a unit by integrating predictions of electricity prices, consumption, and weather variables. The simulations demonstrate the expected load shifting capabilities of the units that adapts to the given price predictions. We furthermore evaluated control performance in terms of economic savings for different control strategies and forecasts.

Chapter 5 describes and compares the proposed large-scale Aggregator control strategies. Aggregators are assumed to play an important role in the future Smart Grid and coordinate a large portfolio of units. The developed economic MPC controllers interfaces each unit directly to an Aggregator. We developed several MPC-based aggregation strategies that coordinates the global behavior of a portfolio of units by solving a large-scale optimization and control problem. We applied decomposition methods based on convex optimization, such as dual decomposition and operator splitting, and developed price-based aggregator strategies.
Chapter 6 provides conclusions, contributions and future work.

The main scientific contributions can be summarized to:

• Linear dynamical models of flexible Smart Grid units: heat pumps in buildings, heat storage tanks, and electric vehicle batteries.
• Economic MPC that integrates forecasts in the control of these flexible units.
• Large-scale distributed control strategies based on economic MPC, convex optimization, and decomposition methods.
• A Matlab toolbox including the modeled units for simulating a Smart Energy System with MPC.
Resilient Infrastructure and Building Security

Traditional authentication systems are considered persistent as they rarely limit the time the authentication is valid. Conversely, sensor-based authentication systems are considered transient as they allow continuous authentication of the users.

In this thesis we present a new approach to authentication that combines traditional access control systems with the sensing technologies and tracking capabilities offered by smart environments. Our approach is called Persistent Authentication for Location-based Services. Persistent authentication enables the secure provision of location-based services through non-intrusive authentication of mobile users in a smart environment. The objective is to shift the current authentication paradigm from a single discrete event to a continuous session. This is accomplished by utilising the contextual awareness provided by the smart environment to track principals from the point of initial authentication to the point where authorisation is requested by location-based services.

Facial recognition and appearance analysis are integrated in the persistent authentication system as remote biometric experts that operate at a distance and require no interaction from the users. The experts perform continuous authentication by processing samples of the biometric modalities as they become available.

Combining scores from multiple biometric experts is known as sensor fusion. A common challenge in this field is that the results from evaluating different biometric characteristic are usually incompatible, as they have different score ranges as well as different probability distributions. Error-rate-based fusion is presented as a novel fusion technique that transforms individual scores from different biometric systems into objective evidences and combine them using Bayesian inference.

Persistent authentication offers an effective integrated protection measure that is distributed directly in the facility and is non-intrusive to the public and affordable to the facility owners. Persistent authentication is suitable for security sensitive applications and can help protect the facility against terrorism and other types of crime.

Epistemic and Doxastic Planning

This thesis is concerned with planning and logic, which are both core areas of Artificial Intelligence (AI). A wide range of research disciplines deal with AI, including philosophy, economy, psychology, neuroscience, mathematics and computer science. The approach of this thesis is based on mathematics and computer science. Planning is the mental capacity that allow us to predict the outcome of our actions, thereby enabling us to exhibit goal-directed behaviour. We often make use of planning when facing new situations, where we cannot rely on entrenched habits, and the capacity to plan is therefore closely related to the reflective system of humans. Logic is the study of reasoning. From certain fixed principles logic enables us to make sound and rational inferences, and as such the discipline is virtually impossible to get around when working with AI.
The basis of automated planning, the term for planning in computer science, is essentially that of propositional logic, one of the most basic logical systems used in formal logic. Our approach is to expand this basis so that it is based on richer and more expressive logical systems. To this end we work with logics for describing knowledge, beliefs and dynamics, that is, systems that allow us to formally reason about these aspects. By letting these elements be used in a planning context, we obtain a system that extends the degree to which goal-directed behaviour can, at present, be captured by automated planning.

In this thesis we concretely apply dynamic epistemic logic to capture knowledge, and dynamic doxastic logic for capturing belief. We highlight two results of this thesis. The first pertains to how dynamic epistemic logic can be used to describe the (lack of) knowledge of an agent in the mid of planning. This perspective is already incorporated in automated planning, and seen in isolation this result appears mainly as an alternative to existing theory. Our second result underscores the strengths of the first. Here we show how the kinship between the aforementioned logics enable us to extend automated planning with doxastic elements. An upshot of expanding the basis of automated planning is therefore that it allows for a modularity, which facilitates the introduction of new aspects into automated planning. We round things o_ by describe what we consider to be the absolutely most fascinating perspective of this work, namely situations involving multiple agents. Reasoning about the knowledge and beliefs of others are essentialy to acting rationally. It enables cooperation, and additionally forms the basis for engaging in a social context. Both logics mentioned above are formalized to deal with multiple agents, and the first steps have been taken towards extending automated planning with this aspect.

Unfortunately, the first results in this line of research have shown that planning with multiple agents is computationally intractable, and additional work is therefore necessary in order to identify meaningful and tractable fragments.
Vitamin D status and effects of food fortification in families

Background and aims: The importance of vitamin D in bone health is recognised and low concentrations have been associated with increased risk of disease. Cutaneous synthesis is considered the major source of vitamin D, but during winter when sufficient sun exposure is restricted at Northern latitudes, intake from food and dietary supplements become essential. Vitamin D intakes are lower than dietary recommendations in most populations and low vitamin D status is common. The PhD thesis is based on the VitmaD study in which a realistic and model derived fortification strategy was investigated in a real-life setting. The aim was to investigate the effect of increasing vitamin D intake by fortification of milk and bread to the amount recommended in the Nordic Nutrition recommendations (NNR) on serum 25(OH)D concentration in families during winter in Denmark (paper 1). Secondly, the aim was to assess vitamin D status and its determinants at baseline of the study (paper 2). Further, to model the relationship between total vitamin D intake and serum 25(OH)D taking into account potential effect modifiers and estimate required vitamin D intake during winter (paper 3).

Methods: The VitmaD study was a randomized controlled trial in 782 children and adults (4-60 years) recruited as 201 families. Families were randomly assigned to vitamin D fortified or nonfortified milk and bread for 6 months starting from September. The milk and bread replaced the subjects’ usual consumptions of products. Information on dietary intake, supplement use, health and lifestyle was obtained by self-administered web-based questionnaires. Serum 25(OH)D was analysed by liquid chromatography-tandem mass spectrometry (LC/MS-MS). Mixed models with family as a random factor were applied in all the statistical analyses.

Results: At baseline of the study (late summer) the geometric mean (IQR) serum 25(OH)D concentration was 72.1 (61.5-86.7) nmol/L with no overall differences between age (P=0.190), gender (P=0.332) or age and gender groups (P=0.223) (paper 2). The prevalence of serum 25(OH)D <50 nmol/L was 9 %. In the multiple analysis of all subjects, vitamin D status was negatively associated with BMI (P<0.001) and positively associated with dietary vitamin D (P=0.008), multivitamin use (P=0.019), solarium use (P=0.006), outdoor stay in light clothes (P=0.001), sun preference (P=0.002) and sun vacation (P<0.001). The intra-family correlation was stronger in children (0.42) compared with adults (0.24). Thus children within a family seemed to be more alike than adults within a family with respect to vitamin D status. The planned fortification strategy was to increase the vitamin D intake to 7.5 µg/day. This succeeded in 66 % of the subjects in the fortification group with a median vitamin D intake (habitual diet plus fortified milk and bread) of 9.4 µg/day compared with 2.2 µg/day in the control group (paper 1). During winter the serum 25(OH)D concentration decreased from 73.1 to 67.6 nmol/L (Δ-5.5 nmol/L) in the fortification group (P<0.001) and from 71.1 to 41.7 nmol/L (Δ-29.4 nmol/L) in the control group (P<0.001). The final serum 25(OH)D concentration was significantly higher in the fortification group compared with in the control group (P<0.001, interpreted estimate 1.59) and the treatment effect was not affected by BMI, multivitamin use and sun vacation. The prevalence of serum 25(OH)D <50 nmol/L remained low in the fortification group (16 %) whereas it increased to 65 % in the control group. The relationship between total vitamin D intake from natural foods, fortified milk and bread and dietary supplements and serum 25(OH)D concentration in winter was best fitted by a non-linear curve (paper 3). The effect of total vitamin D intake on serum 25(OH)D concentration was 4 % higher in men compared with women (P=0.014) and 10 % higher in the group with lowest initial 25(OH)D concentration (<61.5 nmol/L) compared with the group with highest initial 25(OH)D concentration (>86.9 nmol/L) (P<0.001). It was not modified by age (P=0.132) or BMI (P=0.884). Estimated required vitamin D intake was 5, 11, 23 and 39 µg/day for 50, 75, 90 and 95 % of the population to maintain vitamin D status >50 nmol/L during winter. These figures were higher for the group with lowest initial 25(OH)D concentration (11, 18, 34 and >34 µg/day) and lower for the group with highest initial 25(OH)D concentration (<1, 3, 8 and 17 µg/day). Conclusions: In the population of Danish families, serum 25(OH)D concentration was above 50 nmol/L in late summer and it was associated with both dietary and sun related factors. Children within a family seemed to be more alike than adults within a family with respect to vitamin D status. Vitamin D fortification of milk and bread reduced the decrease in serum 25(OH)D concentration during winter and ensured concentrations above 50 nmol/L. The relationship between total vitamin D intake and vitamin D status was non-linear. Estimated total vitamin D intake to maintain serum 25(OH)D above 50 nmol/L was largely dependent on the initial vitamin D status.
Coarse Analysis of Microscopic Models using Equation-Free Methods

Mathematical models of real-world problems from physics, biology and chemistry have become very complex over the last three decades. Although increasing computational power allows to solve even larger systems of differential equations, the number of differential equations is still a main limiting factor for the complexity of models, e.g., in real-time applications. With the increasing amount of data generated by computer simulations a challenge is to extract valuable information from the models in order to help scientists and managers in a decision-making process. Although the dynamics of these models might be high-dimensional, the properties of interest are usually macroscopic and low-dimensional in nature. Examples are numerous and not necessarily restricted to computer models. For instance, the power output, energy consumption and temperature of engines are interesting quantities for engineers, although the models they base their design on are described for the gas mixture (a system with many degrees-of-freedom) inside a combustion engine. Since good models are often not available on the macroscopic scale the necessary information has to be extracted from the microscopic, high-dimensional models.

The goal of this thesis is to investigate such high-dimensional multiscale models and extract relevant low-dimensional information from them. Recently developed mathematical tools allow to reach this goal: a combination of so-called equation-free methods with numerical bifurcation analysis is used and further developed to gain insight into high-dimensional systems on a macroscopic level of interest. Based on a switching-procedure between a detailed microscopic and a coarse macroscopic level during simulations it is possible to obtain a closure-on-demand for the macroscopic dynamics by only using short simulation bursts of computationally-expensive complex models. Those information is subsequently used to construct bifurcation diagrams that show the parameter dependence of solutions of the system. The methods developed for this thesis have been applied to a wide range of relevant problems. Applications include the learning behavior in the barn owl's auditory system, traffic jam formation in an optimal velocity model for circular car traffic and oscillating behavior of pedestrian groups in a counter-flow through a corridor with narrow door. The methods do not only quantify interesting properties in these models (learning outcome, traffic jam density, oscillation period), but also allow to investigate unstable solutions, which are important information to determine basins of attraction of stable solutions and thereby reveal information on the long-term behavior of an initial state.

Multivariate phase type distributions - Applications and parameter estimation

The best known univariate probability distribution is the normal distribution. It is used throughout the literature in a broad field of applications. In cases where it is not sensible to use the normal distribution alternative distributions are at hand and well understood, many of these belonging to the class of phase type distributions. Phase type distributions have several advantages. They are versatile in the sense that they can be used to approximate any given probability distribution on the positive reals. There exist general probabilistic results for the entire class of phase type distributions, allowing for different
estimation methods for the whole class or subclasses of phase type distributions. These attributes make this class of distributions an interesting alternative to the normal distribution.

When facing multivariate problems, the only general distribution that allows for estimation and statistical inference, is the multivariate normal distribution. Unfortunately only little is known about the general class of multivariate phase type distribution. Considering the results concerning parameter estimation and inference theory of univariate phase type distributions, the class of multivariate phase type distributions shows potential for similar great results.

My PhD studies were part of the the work package 3 of the UNITE project. The overall goal of the UNITE project is to improve the decision support prior to deciding on a project by reducing systematic model bias and by quantifying and reducing model uncertainties.

Research has shown that the errors on cost estimates for infrastructure projects clearly do not follow a normal distribution but is skewed towards cost overruns. This skewness can be described using phase type distributions. Cost benefit analysis assesses potential future projects and depend on reliable cost estimates. The Successive Principle is a group analysis method primarily used for analyzing medium to large projects in relation to cost or duration. We believe that the mathematical modeling used in the Successive Principle can be improved. We suggested a novel approach for modeling the total duration of a project using a univariate phase type distribution. The model is then extended to catch the correlation between duration and cost estimates using a bivariate phase type distribution. The use of our model can improve estimates for duration and costs and therefore help project management to make the optimal decisions.

The work conducted during my PhD studies aimed at shedding light on the class of multivariate phase type distributions. This thesis contains analytical and numerical results for parameter estimations and inference theory for a family of multivariate phase type distributions. The results can be used as a stepping stone towards understanding multivariate phase type distributions better. However, we are far from uncovering the full potential of general multivariate phase type distributions. Deeper understanding of multivariate phase type distributions will open up a broad field of research areas they can be applied to.

This thesis consists of a summary report and two research papers. The work was carried out in the period 2010 - 2014.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Department of Transport, Transport policy and behaviour
Authors: Meisch, D. (Intern), Nielsen, B. F. (Intern), Leleur, S. (Intern)
Number of pages: 143
Publication date: 2014

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2014
Number: 331
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd331_Meisch_D.pdf

Broadband optical characterization of material properties
Optical inspection of material properties is of great interest to industry because it can perform objective and non-invasive characterisation of large sample quantities. This may be used in various ways to lower production costs and improve product quality. In this thesis the objective has been to develop and investigate the applicability of optical broadband characterization techniques in industrially relevant production process. Both combined broad and high resolution techniques have the potential to provide important information on scattering properties related to particle size distributions, as well as details of the absorption spectrum which relate to chemical composition.

The thesis focuses on two production process from the food industry. The first process is from the dairy industry where discrimination between chemical and structural properties is of importance. To explore the applicability of optical techniques for this purpose, the fermentation of milk into yogurt has been used as a model system. Studies have been conducted on commercially available products, but also of on-line measurement of the fermentation process. The second process is from the aquaculture industry, quantification of the fish feed additive astaxanthin has been investigated. A measurement campaign has been carried out on a series of pellets specially produced for the purpose.
To investigate these processes, the following three measurement techniques have been developed and applied. (I) A camera based inspection system for spectrally resolved Static Light Scattering (SLS). (II) Photon Time-of-Flight (PToF) spectroscopy, which is a state of the art technique for characterization of turbid media. (III) A new hyperspectral imaging system based on full-field illumination by diffuse laser light. This thesis reports on the design and operation of the different measurement techniques together with the necessary theoretical background for the industrial applications.

For the purpose of milk fermentation this work has demonstrated that the reduced scattering properties of milk change significantly throughout the fermentation process. It has also been shown that the optical inspection methods sense changes to structural properties before any are detected by traditional mechanical rheology. Finally, the developed hyperspectral imaging system was used to quantify the content of astaxanthin in fish feed, and performed at an equal level to a state of the art multi-spectral vision system.

Model Identification for Control of Display Units in Supermarket Refrigeration Systems
In this paper we propose a method for identifying and validating a model of the heat dynamics of a supermarket refrigeration display case for the purpose of advanced control. The model is established to facilitate the development of novel model-based control techniques for individual display units in a supermarket refrigeration system. The grey-box modelling approach is adopted, using stochastic differential equations to define the dynamics of the model, combining prior knowledge of the physical system with data-driven modelling. Model identification is performed using the forward selection method, and the performance of candidate models is evaluated through cross-validation. The model developed in this work uses operational data from a small Danish supermarket. A three-state model is determined to be most appropriate for describing the dynamics of this system. Advanced local control employing the identified model can contribute to the extension of the control capabilities of the entire supermarket refrigeration system.
Predictive Food Microbiology: new tools for risk assessment and dairy product development

Listeria monocytogenes is a well-known food borne pathogen that potentially causes listeriosis. No outbreaks or cases of listeriosis have been associated with cottage cheese, but several confirmed cases and outbreaks in the EU and the US have been related to dairy products made from raw or pasteurised milk. This, in combination with the fact that cottage cheese support growth of Listeria monocytogenes, induces a documentation requirement on the food producer. In the EU regulatory framework, mathematical models are recognised as a suitable supplement to traditional microbiological methods. The models can be used for documentation of compliance with microbiological criteria for Listeria monocytogenes under reasonably foreseeable conditions. Cottage cheese is a fresh, fermented dairy product. It consists of a fermented cheese curd mixed with a fresh or cultured cream dressing. The product contains considerable concentrations of lactic acid bacteria from the added starter or aroma cultures. The presence of these microorganisms induces some complexity to the product, since the lactic acid bacteria metabolites and e.g. bacteriocins exhibit an inhibitory effect towards co-culture microorganisms such as Listeria monocytogenes. During storage at temperatures allowing the mesophilic lactic acid bacteria to grow (> 8-10°C), a pronounced inter-bacterial interaction and growth inhibition of co-culture Listeria monocytogenes was observed. These observations emphasised the need for inter-bacterial interaction models when predicting the growth response of Listeria monocytogenes in fermented dairy products. The objective of the PhD-project was to develop new, or extend existing mathematical models to be used for risk assessment and product development. When the project was initiated, none of the existing predictive models were found to appropriately describe the simultaneous growth of lactic acid bacteria from the added starter or aroma culture and Listeria monocytogenes in cottage cheese. New, deterministic growth models were developed for Listeria monocytogenes, starter lactic acid bacteria and aroma lactic acid bacteria. The new cardinal parameter type growth models included the effect of temperature, pH, NaCl, lactic and sorbic acid. The models were developed based on growth data obtained from absorbance measurements in liquid laboratory media and growth data obtained in cottage cheese with fresh or cultured cream dressing. An important step in the modelling procedure was the calibration of the reference growth rate ($\mu_{ref}$, h$^{-1}$ at 25°C) which was strongly affected by the dominating lactic acid bacteria culture. By combining the developed secondary growth models with the empirical Jameson approach, good predictions of the simultaneous growth of Listeria monocytogenes and lactic acid bacteria were obtained. Both growth rate and maximum population densities of Listeria monocytogenes were accurately described under constant and dynamic storage temperatures (between 5°C and 15°C). The inter-bacterial interaction was clearly important to include when predicting growth response of Listeria monocytogenes in fermented dairy products. Alternative, semi-mechanistic, imodelling approaches were evaluated based on methods applied in the fermentation technology. The dynamics of lactic acid concentration and product pH was related to growth of lactic acid bacteria by the yield factor concept. The ability to predict the maximum population density of Listeria monocytogenes in cottage cheese based on dynamic lactic acid and pH was evaluated. For cottage cheese with fresh cream dressing, the semi-mechanistic interaction model successfully predicted the maximum population density. Lactic acid and pH was, however, insufficient to describe the growth inhibition of Listeria monocytogenes observed in cottage cheese with cultured cream dressing. Improved, mechanistic, prediction of Listeria monocytogenes in cottage cheese with cultured cream dressing would require that additional mechanisms were included in the model, such as other metabolites or bacteriocins. Finally, the semi-mechanistic and the empirical Jameson approach to inter-bacterial interaction modelling were compared. The empirical Jameson model consistently performed equally well or better than the more complex semi-mechanistic model. In order to evaluate the growth response of more realistic concentrations of Listeria monocytogenes and to take variability into account, a stochastic approach was applied. The deterministic growth models were used in combination with stochastic input values for bacterial concentration; lag time duration and product characteristics. Good agreement between predicted and observed growth was obtained, when applying broth based lag time distributions for Listeria monocytogenes single cells in combination with the relative lag time concept. Furthermore, application of relative lag time distributions from Listeria monocytogenes population data provided good predictions of the growth response of only a few Listeria monocytogenes cells in cottage cheese at chilled temperatures. From the results of the present PhD-project it was found that once solid, deterministic, secondary growth models have been developed and validated, they can be modified and/or extended to a range of other modelling procedures. Furthermore, inclusion of inter-bacterial interaction was considered to be an inevitable part when modelling and predicting growth of L. monocytogenes in fermented dairy products. In general, simple approaches to describe interaction and growth inhibition (empirical approach), lag time prediction of individual cells (variability in population RLT-values) and representation of e.g. variable product characteristics (bootstrapping from empirical distributions) were advocated. It is believed that it is necessary to define some applicable methodologies for the development of growth models for complex products such as fermented dairy products. Model development is a comprehensive process with an almost infinite data requirement and the findings of the present PhD-project is thought to be important in relation to the development of predictive models that are valuable for, and readily applicable in the food industry.
Electric vehicle integration in a real-time market

This project is rooted in the EDISON project, which dealt with Electrical Vehicle (EV) integration into the existing power grid, as well as with the infrastructure needed to facilitate the ever increasing penetration of fluctuating renewable energy resources like e.g. wind turbines. In the EDISON project, the EV is introduced as an energy buffer used to store excess energy produced at off-peak hours, while at the same time potentially benefiting the consumer by offering cheaper charging. This role as a buffer, predominantly used for delayed charging, also known as “smart charging”, can also be used for ancillary services to help stabilize the grid at critical periods, e.g. by providing near instant up- or down regulation. The initial goal of this project is to develop the components for a simulation platform for large scale EV integration studies. By interfacing the EV simulation with an externally simulated model of the power grid, it is be possible, in real-time, to simulate the impact of EV charging and help to identify bottlenecks in the system. In EDISON the vehicles are aggregated using an entity called a Virtual Power Plant (VPP); a central server monitoring and controlling the distributed energy resources registered with it, in order to make them appear as a single producer in the eyes of the market. Although the concept of a VPP is used within the EcoGrid EU project, the idea of more individual control is introduced through a new proposed real-time electricity market, where the consumers will have direct access to the current price. As opposed to the hourly spot-price market of today, the real-time market see price updates as often as every couple of minutes. To allow the individual resources to react to these changes, independent of each other, so called “smart controllers” are needed at the device level. In order for this market to work, however, the proper ICT network- and server-infrastructure has to be developed. The primary goal of this PhD project, has been to investigate the scope of this ICT infrastructure, required to realise price-signal based charging of electric vehicles, in accordance with the EcoGrid EU market.

Safety-Critical Java for Embedded Systems

Safety-critical systems are real-time systems whose failure can have severe or catastrophic consequences, possibly endangering human life. Many safety-critical systems incorporate embedded computers used to control different tasks. Software running on safety-critical systems needs to be certified before its deployment and the most time-consuming step of this process is the testing and verification phase. Due to the increasing complexity in safety-critical systems there is a need for new technologies that can facilitate testing and verification activities. The safety-critical specification for Java aims at providing a reduced set of the Java programming language that can be used for systems that need to be certified at the highest levels of criticality. Safety-critical Java (SCJ) restricts how a developer can structure an application by
providing a specific programming model and by restricting the set of methods and libraries that can be used. Furthermore, its memory model do not use a garbage-collected heap but scoped memories.

In this thesis we examine the use of the SCJ specification through an implementation in a time-predictable, FPGA-based Java processor. The specification is now in a mature state and with our implementation we have proved its feasibility in an embedded platform. Moreover, we have explored how simple hardware extensions can reduce the execution time of time-critical operations required by the SCJ specification.

The scoped memory model used in SCJ is perhaps one of its most difficult features to use correctly. Therefore, in this work we have also studied practical aspects of its usage by developing scoped memory use patterns and reusable libraries aiming at facilitating the development of complex software systems.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Embedded Systems Engineering
Authors: Rios Rivas, J. R. (Intern), Schoeberl, M. (Intern)
Number of pages: 162
Publication date: 2014

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PhD-2014
Number: 340
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd340_Rivas_JRR.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2014

Performance Aspects of Synthesizable Computing Systems
Embedded systems are used in a broad range of applications that demand high performance within severely constrained mechanical, power, and cost requirements. Embedded systems implemented in ASIC technology tend to provide the highest performance, lowest power consumption and lowest unit cost. However, high setup and design costs make ASICs economically viable only for high volume production. Therefore, FPGAs are increasingly being used in low and medium volume markets. The evolution of FPGAs has reached a point where multiple processor cores, dedicated accelerators, and a large number of interfaces can be integrated on a single device.

This thesis consists of five parts that address performance aspects of synthesizable computing systems on FPGAs. First, it is evaluated how synthesizable processor cores can exploit current state-of-the-art FPGA architectures. This evaluation results in a processor architecture optimized for a high throughput on modern FPGA architectures. The current hardware implementation, the Tinuso I core, can be clocked as high as 376MHz on a Xilinx Virtex 6 device and consumes fewer hardware resources than similar commercial processor configurations. The Tinuso architecture leverages predicated execution to circumvent costly pipeline stalls due to branches and exposes hazards to the compiler to keep the hardware simple. Second, it is investigated if a production compiler, GCC, is able to successfully leverage predicated execution and schedule instructions so as to mitigate the hazards. The third part of this thesis describes the design and implementation of communication structures for Tinuso multicore configurations and evaluates the scalability of these systems. Forth, a case study shows how to map a high performance synthetic aperture radar application to a synthesizable multicore system. The proposed system includes 64 processor cores and a 2D mesh interconnect on a single FPGA device and consumes about 10 watt only. Finally, a task based programming model is proposed that allows for easily expressing parallelism and simplifying memory management.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Embedded Systems Engineering
Authors: Schleuniger, P. (Intern), Karlsson, S. (Intern), Madsen, J. (Intern)
Number of pages: 213
Publication date: 2014

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2014
Number: 337
2D Static Light Scattering for Dairy Based Applications

Throughout this thesis we investigate a recently introduced optical technique denoted 2D static light scattering (2DSLS). The technique is remote sensing, non-invasive, highly flexible, and appears to be well suited for in-line process control. Moreover, the output signal contains contributions from several different optical phenomena, which can be utilised to provide information on chemical composition and underlying microstructure of an investigated sample.

The main goal of this thesis is to provide an exploratory study of the 2DSLS technique in relation to dairy based applications. This includes getting an understanding of the various parameters in the setup as well as understanding the output signal in terms of potential and limitations. Furthermore, suitable ways of quantifying the signal are investigated. Here, both established physical models and statistical descriptions of the signal are evaluated and discussed.

There is a major emphasis on using 2DSLS to discriminate between different protein microstructures in yogurt products. This potentially allows for process control, in relation to microstructure, during yogurt manufacture. As microstructure is critical for consumer acceptability, this specific process control can be highly beneficial. To provide suitable reference measures on the actual microstructure, we investigate how to quantify micrographs of yogurts objectively. We provide a comparative study, that includes a broad range of different image texture descriptors.
Drops on hydrophobic surfaces & vibrated fluid surfaces

The first part of this thesis deals with a droplet on a hydrophobic surface. We first present a basic introduction to fluid dynamics, including a description of relevant dimensionless numbers and a derivation of the Young-Laplace equation. An analytic approach to describing the oscillations of a droplet is then given, after which we set up a 2D computational Finite-Element Method (FEM) model for a neutrally buoyant drop immersed in another liquid. The model is validated by considering the volume loss over time.

Subsequent to an introduction to the physics of wetting, the developed FEM model is then extended to include drop-surface interactions, and we describe a) the initial descent of a droplet down an inclined hydrophobic substrate, and b) the motion of the droplet in a potential well created through spatial contact angle variations. We solve the full Navier-Stokes equations inside the drop domain, and use the Arbitrary Lagrangian-Eulerian method to keep track of the droplet surface; the contact angle is included by using the Frennet-Serret equations.

In situation a), we investigate the behavior of the drop velocity as a function of the slip length and compare with experimental results found in the literature. Furthermore, we quantify the energy associated with center of mass translation and internal fluid motion. The model predicts trajectories for tracer particles deposited inside the drop, and satisfactorily describes the sliding motion of steadily accelerating droplets. The model can be used for determining a characteristic slip parameter, associated with slip lengths and drag reduction for hydrophobic surfaces.

In situation b), we observe that the droplet oscillations (frequency, amplitude and decay time) in the potential is not linear with respect to the forcing, i.e. the strength of the potential, and contribute this to preferred eigenmodes of the droplet oscillation.

The second part of this thesis deals with a droplet bouncing on a vertically vibrated fluid bath of the same liquid, a system which is the first known macroscopic example of pilot-wave dynamics. An introduction to the experimental set-up is given, followed by a description of the mathematical models governing the vertical and horizontal motion of the drop. Two in-depth studies are then presented.

In the first, the results of a comprehensive series of experiments are presented. The most detailed characterisation to date of the system’s dependence on fluid properties, droplet size, and vibrational forcing is provided. A number of new bouncing and walking states are reported, including complex periodic and aperiodic motions. Specific attention is given to the first characterisation of the different gaits arising within the walking regime. In addition to complex periodic walkers and limping droplets, we highlight a previously unreported mixed state, in which the droplet switches periodically between two distinct walking modes. The experimental results are compared to previously developed theoretical predictions.

In the second study, we consider the case where the fluid bath is also rotated around its center-line. The drop then experiences an effective Coriolis force, and previous studies have made a comparison between emerging unstable radii in this system and Landau levels for a charged particle in a magnetic field. The system is treated numerically, and the results are compared to experiments. We provide, again, the most detailed regime diagram of the possible orbits depending on the forcing and the rotation rate of the fluid bath. We highlight each class of orbit, and analyze in depth the wobbling state, precessing orbits, wobble-leap dynamics, exotic trajectories and the emergence of statistical behavior when the forcing is near the Faraday threshold.
Kernel Methods for Machine Learning with Life Science Applications

Kernel methods refer to a family of widely used nonlinear algorithms for machine learning tasks like classification, regression, and feature extraction. By exploiting the so-called kernel trick straightforward extensions of classical linear algorithms are enabled as long as the data only appear as innerproducts in the model formulation. This dissertation presents research on improving the performance of standard kernel methods like kernel Principal Component Analysis and the Support Vector Machine. Moreover, the goal of the thesis has been two-fold.

The first part focuses on the use of kernel Principal Component Analysis for nonlinear denoising. In this context stable solution of the inverse and inherently ill-posed pre-image problem constitutes the main challenge. It is proposed to stabilize the estimation by augmenting the cost function with either an \(1\)-or \(2\)-norm penalty, and solution schemes are derived for both approaches. The methods are experimentally validated on several biomedical data sets. Furthermore, frameworks for exploiting label information for improved denoising in the semisupervised case are proposed.

The second part of the thesis examines the effect of variance inflation in kernel methods. Variance inflation occurs in high-dimensional problems when the training data are insufficient to describe the entire signal manifold. Thereby leading to a potential mismatch between the subspaces spanned by the training and test data, respectively. It is shown how this effect extends from linear models to kernel learning, and means for restoring the generalizability in both kernel Principal Component Analysis and the Support Vector Machine are proposed. Viability is proved on a wide range of benchmark machine learning data sets.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Abrahamsen, T. J. (Intern), Hansen, L. K. (Intern), Winther, O. (Intern)
Number of pages: 168
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: PHD-2013
Number: 299
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd299_Abrahamsen_TJ.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2013

Assessing Miniaturized Sensor Performance using Supervised Learning, with Application to Drug and Explosive Detection

This Ph.D. thesis titled “Assessing Miniaturized Sensor Performance using Supervised Learning, with Application to Drug and Explosive Detection” is a part of the strategic research project “Miniaturized sensors for explosives detection in air” funded by the Danish Agency for Science and Technology’s, Program Commission on Nanoscience Biotechnology and IT (NABIIT), case number: 2106-07-0031. The project, baptized “Xsense” was led by professor Anja Boisen, DTU Nanotech. DTU Informatics participate in the project as data analysis partner.

This thesis presents advances in the area of detection of vapor emanated by explosives and drugs, similar to an electronic nose. To evaluate sensor responses a data processing and evaluation pipeline is required. The work presented herein focuses on the feature extraction, feature representation and sensor accuracy. Thus the primary aim of this thesis is twofold; firstly, present methods suitable for assessing sensor accuracy, and secondly improve sensor performance by enhancing the preprocessing and feature extraction.

Five different miniaturized sensors are presented. Naturally, each sensor require its own special preprocessing and feature extraction techniques before the sensor responses can be applied to supervised learning algorithms. The technologies used for sensing consist of Calorimetry, Cantilevers, Chemoselective compounds, Quartz Crystal Microbalance and Surface Enhanced Raman Scattering. Each of the sensors have their own strength and weaknesses. The reasoning for using multiple sensors was the desire to investigate the feasibility for an integrated multisensor solution. A unique setup of multiple independent detectors is able to vastly enhance accuracy compared to what a single sensor can deliver. As we are detecting hazardous compounds this enables the need for sensors to deliver not only decisions but also certainty about decisions. This requirement is handled by introducing classifiers that offer posterior probabilities and not only decisions. The three probabilistic classification models utilized are Artificial Neural Networks, Logistic Regression and Gaussian Processes. Often, there is no tradition for using these methods in the communities of the prescribed sensors. Here, a method of too much complexity is often undesired so it is a balance when to utilize more sophisticated methods. For this reason, an array of methods that only discriminate between samples are used as baseline. The methods used vary from sensor to sensor, as these methods serve as baseline performance when introducing new methods. The most widely used baseline method in this thesis is the k-nearest-neighbor algorithm. This method is of particular
interest in the application of sensors, as the sensors are designed to provide robust and reliable measurements. That means, the sensors are designed to have repeated measurement clusters. Sensor fusion is presented for the sensor based on chemoselective compounds.

An array of color changing compounds are handled and in unity they make up an colorimetric sensor array. In this setting it is valuable to qualify which compounds in the colorimetric sensor array are important. That knowledge enables the ability to either reduce the size of the sensor or replace less sensitive and unimportant compounds with more selective and responsive compounds. A framework based on forward selection Gaussian Process classification is demonstrated to successfully identify a set of important compounds.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Alstrøm, T. S. (Intern), Larsen, J. (Intern)
Number of pages: 262
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2012
Number: 292
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd292_tsal.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2013

Low-energy house in Sisimiut - Data overview
Experiments with persistently exciting heat inputs are a fundamental tool in identification of heat dynamics in buildings. The Low-energy house in Sisimiut, Greenland, provides an advanced experimental setup with frequent measurements of temperatures, heat inputs, and much more. This paper presents an overview of data collected since the installation of the new measurement and control system. Focus is on heat dynamics so only data related to that will be shown. 5 experiments have been conducted. They are described, and resulting data is shown.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Department of Civil Engineering, Section for Building Physics and Services
Authors: Andersen, P. H. D. (Intern), Rode, C. (Intern), Madsen, H. (Intern)
Number of pages: 19
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute-Technical Report-2013
Number: 18
ISSN: 1601-2321
Main Research Area: Technical/natural sciences
Electronic versions: tr13_18_Delff_P_1.pdf
Publication: Research › Report – Annual report year: 2013

Low-energy house in Sisimiut - Measurement equipment
This paper documents the measurement equipment in a low-energy house in Sisimiut, Greenland. Detailed measurements are being taken on energy consumption, indoor temperatures, floor heating, ventilation, open/closed state of doors and windows, and indoors climate. Equipped with a central control unit, experiments can be designed in order to study heat dynamics of the building. It is described how to plan and execute such experiments in one apartment in the building. The building also features both a solar thermal system and extra buffer tank facilitating testing of storage strategies on the power generated by the solar thermal system. A weather station equipped with thermometer, pyranometer and anemometer is installed on the building as well. Finally, it is described how to retrieve data from an SQL server which is
configured to take monthly backups. R functions have been implemented to fetch and prepare the data for time series analysis. Examples are given on the use of these.

**General information**

State: Published  
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Department of Civil Engineering, Section for Building Physics and Services  
Authors: Andersen, P. H. D. (Intern), Rode, C. (Intern), Madsen, H. (Intern)  
Number of pages: 54  
Publication date: 2013

**Publication information**

Place of publication: Kgs. Lyngby  
Publisher: Technical University of Denmark (DTU)  
Original language: English  
Series: DTU Compute-Technical Report-2013  
Number: 17  
ISSN: 1601-2321  
Main Research Area: Technical/natural sciences  
Electronic versions:  
tr13_17_Delff_P_1.pdf  
Publication: Research › Report – Annual report year: 2013

**Models for the energy performance of low-energy houses**

The aim of this thesis is data-driven modeling of heat dynamics of buildings. Traditionally, thermal modeling of buildings is done using simulation tools which take information about the construction, weather data, occupancy etc. as inputs and generate deterministic energy profiles of the buildings. This approach often fails in predicting the actual heat consumption of buildings once they are constructed. The approach taken in this work is deriving models from observations collected after the construction, aiming at describing the actual characteristics of the building.

Identification of heat dynamics of buildings is needed both in order to assess energy-efficiency and to operate modern buildings economically. Energy signatures are a central tool in both energy performance assessment and decision making related to refurbishment of buildings. Also for operation of modern buildings with installations such as mechanical ventilation, floor heating, and control of the lighting effect, the heat dynamics must be taken into account. Hence, this thesis provides methods for data-driven modeling of heat dynamics of modern buildings.

While most of the work in this thesis is related to characterization of heat dynamics of buildings, the first topic analyzed is the variation of presence of occupants. As buildings get more energy-efficient, internal loads and user-behavior increasingly influence the energy consumption. Most simulation tools use deterministic occupancy profiles to simulate internal loads. However, such occupancy patterns will largely depend on the specific use of the building, and hence the profiles must be empirically based. A probabilistic method for modeling timedependence and dynamics of presence of occupants is developed and applied by estimation and model validation on data from an office building. The approach to modeling occupants’ presence provides a flexible method where no assumptions in the application.

The rest of the thesis deals with statistical modeling of heat dynamics of buildings. First, discrete-time models are applied. Discrete-time models are computationally relatively simple and provide a flexible framework for dynamical modeling as a natural extension of the often-used static energy-balance models. The importance of applying dynamical models, even for deriving thermostatic or steady-state properties, is stressed, and methods for doing so are outlined.

Since heat transfer is fundamentally described by partial differential equations, modeling of heat dynamics using differential equations is an obvious approach. A quasi-Gaussian maximum likelihood estimation technique, where the likelihood function is evaluated using the extended Kalman filter on state-space models, is used. In this framework referred to as "grey-box" modeling - one-step predictions can be generated and used for model validation by testing statistically whether the model describes all variation and dynamics observed in the data. The possibility of validating the model dynamics is a great advantage from the use of stochastic differential equations compared to ordinary differential equations.

The strengths of the discrete-time and the continuous-time approach are discussed. Besides the parametrization, which is directly physically interpretable, grey-box models intrinsically provide variable prediction uncertainty, which is important in relation to design of controllers and decision making for comfort requirements. In the framework of stochastic differential equations, there are normally more parameters related to noise processes than in discrete-time models which increases the complexity of the estimation. Here, the state space formulation is often used. Since there is normally infinitely many state space representations corresponding to a transfer function model, structural identifiability is important in relation to state space modeling.
A low-energy building in Sisimiut, Greenland is used as a test-building. The building is well-insulated and features large modern energy-efficient windows and floor heating. These features lead to increased non-linear responses to solar radiation and longer time constants. The building is equipped with advanced control and measuring equipment. Experiments are designed and performed in order to identify important dynamical properties of the building, and the collected data is used for modeling.

The thesis emphasizes the statistical model building and validation needed to identify dynamical systems. It distinguishes from earlier work by focusing on modern low-energy construction and going further into studying and characterizing the dynamical properties of the fitted models.

**General information**

State: Published  
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Department of Civil Engineering, Section for Building Physics and Services  
Authors: Andersen, P. H. D. (Intern), Madsen, H. (Intern), Rode, C. (Intern)  
Number of pages: 267  
Publication date: 2013

**Publication information**

Place of publication: Kgs. Lyngby  
Publisher: Technical University of Denmark (DTU)  
Original language: English

Series: PHD-2013  
Number: 312  
ISSN: 0909-3192  
Main Research Area: Technical/natural sciences  
Electronic versions:  
phd_thesis_philip_delff_net.pdf  
Publication: Research › Ph.D. thesis – Annual report year: 2013

**Topology of streamlines and vorticity contours for two-dimensional flows**

Considering a coordinate-free formulation of helical symmetry rather than more traditional definitions based on coordinates, we discuss basic properties of helical vector fields and compare results from the literature. For inviscid flow where a velocity field is generated by a sum of helical vortex filaments with same pitch we use the established results to prove briefly that the velocity field is helical. We discuss the role of the stream function for the topology of the streamlines in incompressible, helical flows. On this basis, we perform a comprehensive study of the topology of the flow field generated by a helical vortex filament in an ideal fluid. The classical expression for the stream function obtained by Hardin (Phys. Fluids 25, 1982) contains an infinite sum of modified Bessel functions. Using the approach by Okulov (Russ. J. Eng. Thermophys. 5, 1995) we obtain a closed-form approximation which is considerably easier to analyse. Critical points of the stream function can be found from the zeroes of a single real function of one variable, and we show that three different flow topologies can occur, depending on a single dimensionless parameter. Including the self-induced velocity on the vortex filament by the localised induction approximation the stream function is slightly modified and an extra parameter is introduced. In this setting two new flow topologies arise, but not more than two critical points occur for any combination of the parameters. The analysis of the closed form show promise for analysing more complex flow with helical symmetry e.g. multiple helical vortex filaments inside a cylinder which has industrial relevance.

We then change focus and study creation, destruction and interaction of vortices in two dimensional flow. A vortex is advected above a wall causing a viscous response near the wall which generates a new vortex structure. The problem is studied numerically relying on the code developed by Prof. M. Thompson and his group at Monash University, Australia. We also investigate the problem analytically using normal form theory. It is not a simple task to define a vortex in a proper way that allow the study of creation and destruction of vortices. We investigate three sound choices: the vorticity extrema, the streamline centers in a coordinate system with zero wall speed and the streamline centers in a frame moving with constant velocity as predicted by a point vortex above a wall in inviscid fluid. There is no reason to a priori expect equivalent results of the three vortex definitions. However, the study is mainly motivated by the findings of Kudela & Malecha (Fluid Dyn. Res. 41, 2009) who find good agreement between the vorticity and streamlines in the fixed wall system. For small Re no new vortices are observed. Creation of a vortex occurs for sufficiently large Re for all the applied vortex definitions. The new vortex alters the generating vortex motion by slowing its horizontal motion and lifting it further from the wall. In the fixed wall system vortex eruption happens through a characteristic ‘figure 8’ bifurcation. Considering the other coordinate system there is no topological change indicating when a vortex has left the boundary layer. However, here there is remarkable good agreement between streamlines and the vorticity contours even for short-lived vortices close to the wall.

The normal form approach does not reveal simple connections between the streamline topology and the vorticity contour topology. Only for a non simple degenerate on wall critical point may a bifurcation occur in both the streamlines and the vorticity contours. The streamline bifurcations in this normal form contain the lower part of the ‘figure 8’ bifurcation observed in numerics. The similarities and differences of the streamlines in the two different coordinate systems are well described by normal form theory.
We derive the criterion, \( \mathbf{u} \cdot \nabla \omega = 0 \), for exactly matching contours of the vorticity contours and streamlines. This is fulfilled when the Navier-Stokes equations and the heat equation have identical solutions.

Finally, we focus on the superposition of two rotational invariant vortices in R2. The topology of the streamlines and the topology of the vorticity contours are determined by the zeros of a single real function. For the canonical example of two Gaussian vortices, three parameters exist. Three structurally stable topologies are observed. For the streamlines, two of the topologies are well known for the corresponding situation of two point vortices when the singularities are treated as centers. The last topology is a single center, which is consistent with the powerful result on the long time behavior proved by Gallay & Wayne (Comm. in Math. Phys. 255, 2005). The case of three critical points of the streamlines is a subset of three critical points of the vorticity. This explains an observation in the simulations of vortex generation near a wall. Here, a long living erupted vortex disappears due to viscosity. This happens first considering the streamlines while being more robust when considering the vorticity formulation.
Adaptive Trailing Edge Flaps for Active Load Alleviation in a Smart Rotor Configuration

The work investigates the development of an active smart rotor concept from an aero-servo-elastic perspective. An active smart rotor is a wind turbine rotor that, through a combination of sensors, control units and actuators, is able to alleviate the fluctuating part of the aerodynamic loads it has to withstand. The investigation focuses on a specific actuator type: the Adaptive Trailing Edge Flap (ATEF), which introduces a continuous deformation of the aft part of the airfoil camber-line. An aerodynamic model that accounts for the steady and unsteady effects of the flap deflection on a 2D airfoil section is developed, and, considering both attached and separated flow conditions, is validated by comparison against Computational Fluid Dynamic solutions and a panel code method. The aerodynamic model is integrated in the BEM-based aeroelastic simulation code HAWC2, thus providing a tool able to simulate the response of a wind turbine equipped with ATEF.

A load analysis of the NREL 5 MW reference turbine in its baseline configuration reveals that the highest contribution to the blade flapwise fatigue damage originates from normal operation above rated wind speed, and from loads characterized by frequencies below 1 Hz. The analysis also reports that periodic load variations on the turbine blade account for nearly 11 % of the blade flapwise lifetime fatigue damage, while the rest is ascribed to load variations from disturbances of stochastic nature.

The study proposes a smart rotor configuration with flaps laid out on the outer 20 % of the blade span, from 77 % to 97% of the blade length. The configuration is first tested with a simplified cyclic control approach, which gives a preliminary indication of the load alleviation potential, and also reveals the possibility to enhance the rotor energy capture below rated conditions by using the flaps.

Two model based control algorithms are developed to actively alleviate the fatigue loads on the smart rotor with ATEF. The first algorithm features a linear quadratic regulator with periodic disturbance rejection, and controls the deflection of the flap on each blade based on measurements of the root flapwise bending moment; each blade is considered as an independent Single Input-Single Output system. The second algorithm is a Multiple Input-Multiple Output Model Predictive Control (MIMO-MPC), which monitors the whole turbine response, and controls all the available actuators: ATEF, individual blade pitch, and generator. Both algorithms include frequency-dependent weighting of the control actions in order to limit high frequency control activity, and thus effectively reduce actuators use and wear.

The smart rotor performances are evaluated from HAWC2 simulations reproducing the response to standard turbulent wind fields. Both algorithms reduce the lifetime fatigue damage on the blade root flapwise bending moment by 15 % using the ATEF actuators. Whereas, by combining pitch and flap actions, the MIMO-MPC reports alleviation results close to 30 %. The MIMO-MPC requires lower flap activity, and also achieves higher reductions of the tower fatigue loads, thus indicating that a combined control approach that coordinates and integrates all available sensors and actuators has the potential for overall better results than achieved by a series of independent control systems.
Automated Image-Based Procedures for Adaptive Radiotherapy

Fractionated radiotherapy for cancer treatment is a field of constant innovation. Developments in dose delivery techniques have made it possible to precisely direct ionizing radiation at complicated targets. In order to further increase tumour control probability (TCP) and decrease normal-tissue complication probability (NTCP), margins used to account for interfraction and intrafraction anatomical changes and motion need to be reduced. This can only be achieved through proper treatment plan adaptations and intrafraction motion management.

This thesis describes methods in support of image-based treatment replanning and real-time intrafraction guidance techniques. The selected contributions detail a number of findings and techniques, in particular:
- For ten head & neck cancer patients, changes in tumour density were well described by linear functions with patient-specific slope and intercept. This is of particular interest for proton therapy as delivered dose to a tissue and calculated dose distributions rely on density. Furthermore, tumour density changes might be indicative of treatment response.
- It is demonstrated how spatially varying elasticity parameters can be employed in image registration to encourage bone rigidity and local tissue volume change only in the gross tumour volume and the lungs. This is highly relevant in adaptive radiotherapy when modelling significant tumour volume changes.
- It is described how cone beam CT reconstruction can be modelled as a deformation of a planning CT scan of the same patient, using a non parametric diffusion based deformation model, opening the door to the use of a number of advanced non-parametric algorithms. An advantage of reconstruction by deformation is that no subsequent image registration is needed in order to obtain the deformation which can be employed for contour propagation in adaptive radiotherapy.
- MRI-radiotherapy devices have the potential to offer near real-time intrafraction imaging without any additional ionising radiation. It is detailed how the use of multiple, orthogonal slices can form the basis for reliable 3D soft tissue tracking.

General information

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Bjerre, T. (Intern), Larsen, R. (Intern)
Number of pages: 115
Publication date: 2013

Publication information

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: PHD-2013
Number: 305
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd305_Bjerre_T.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2013

Analysis of the Indented Cylinder by the use of Computer Vision

The research summarised in this PhD thesis took advantage of methods from computer vision to experimentally analyse the sorting/separation ability of a specific type of seed sorting device – known as an “indented cylinder”. The indented cylinder basically separates incoming seeds into two sub-groups: (1) “long” seeds and (2) “short” seeds (known as length-separation). The motion of seeds being physically manipulated inside an active indented cylinder was analysed using various computer vision methods. The data from such analyses were used to create an overview of the machine’s ability to separate certain species of seed from each other. Seeds are processed in order to achieve a high-quality end product: a batch of a single species of crop seed. Naturally, farmers need processed clean crop seeds that are free from non-seed impurities, weed seeds, and non-viable or dead crop seeds. Since the processing is based on physical manipulation of the seeds themselves, their individual shape and size becomes very relevant. The problem of modelling such physical parameters for various species of seed, grown under various environmental circumstances, is a very complex one. The general problem of modelling and controlling seed processing equipment can be expected to be complex. Due to the involvement of seeds, the problem will naturally inherit all their biological complexities. In addition to this, because of the very large number of individual seeds, the problem also involves a granular media and thus inherits all the complexities related to that as well.

The project arrived at a number of results of high scientific and practical value to the area of applied computer vision and seed processing and agricultural technology in general. The results and methodologies were summarised in one conference paper and two journal papers. These three papers, referred to as Paper I, Paper II, and Paper III can be found in Appendix A, B, and C, respectively. These three papers represent the very first examples of published/submitted work that thoroughly analyse and verify the separation ability of the indented cylinder by the use of computer vision (or image analysis). Moreover, the imagery data sets, generated as a result of actual recordings of sorting experiments using the
indented cylinder, are novel by their high dimensionality and size. Paper II in Appendix B makes one of these data sets available online as a cite-aware imagery data set. The work summarised in this thesis is very much related to the task of constructing models from observed data. This field is known as empirical model development or more specifically as “system identification”. System identification deals specifically with estimating mathematical models from observed dynamic states (time series) of inputs and outputs to and from some physical system under investigation. The contribution of the work is to be found primarily within the problem domain of experimentation for system identification. Computer vision techniques were used to acquire observations of a measure of separation efficiency of the indented cylinder. Such techniques for observation are likely to be very relevant for experimentation in a laboratory for system identification purposes. This work should therefore be seen as an important step towards future research in system identification of the indented cylinder. The technical solutions developed are currently novel and represent an ideal platform for future applied research into empirical model development. Finally, this work should also be considered as an early step toward a paradigm shift where the best parameters for the indented cylinder are not mainly determined by “rule of thumb” and other forms of heuristics, but are instead optimized parameters tied to an actual theory of seed separation in the indented cylinder.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Aarhus University
Authors: Buus, O. T. (Intern), Jørgensen, J. R. (Ekstern), Carstensen, J. M. (Intern)
Number of pages: 169
Publication date: 2013

Publication information
Publisher: University of Aarhus, Denmark
Original language: English
Main Research Area: Technical/natural sciences
Electronic versions: olebuus_dissertation_pub.pdf
Source: PublicationPreSubmission
Source-ID: 97259267
Publication: Research › Ph.D. thesis – Annual report year: 2014

Nonlinear Model Predictive Control for Oil Reservoirs Management
The current world average recovery factor from oil fields is widely agreed to be about 30-35%. An increase of 10% point of this recovery factor would bring about 500 billion of oil barrels, sufficient to meet 16 years of current global production. To realize this potential production increase, the research community is working on improving current feedback model-based optimal control technologies. The topic of this thesis is production optimization for water flooding in the secondary phase of oil recovery. We developed numerical methods for nonlinear model predictive control (NMPC) of an oil field. The controller consists of

- A model based optimizer for maximizing some predicted financial measure of the reservoir (e.g. the net present value).
- A parameter and state estimator.
- Use of the moving horizon principle for data assimilation and implementation of the computed control input.

The optimizer uses gradient-based optimization and the required gradients are computed by the adjoint method. We propose the use of efficient high order implicit time integration methods for the solution of the forward and the adjoint equations of the dynamical model. The Ensemble Kalman filter is used for data assimilation. Further, we studied the use of robust control strategies in both open-loop, i.e. without measurement feedback, and closed-loop, i.e. with measurement feedback, configurations.

This thesis has three main original contributions:

The first contribution in this thesis is to improve the computationally expensive gradient computation by using high-order ESDIRK (Explicit Singly Diagonally Implicit Runge-Kutta) temporal integration methods and continuous adjoints. The high order integration scheme allows larger time steps and therefore faster solution times. We compare gradient computation by the continuous adjoint method to the discrete adjoint method and the finite-difference method. We demonstrate that the optimization algorithm can be accelerated by using the continuous time adjoint equations. This is the first time in the literature that the higher order continuous adjoint and higher order discrete adjoint methods have been investigated for oil production optimization.

The second contribution of this thesis is the application of the Robust Optimization strategy in both open-loop (i.e. without measurement feedback) and closed-loop (i.e. with measurement feedback). In the oil industry, Robust Optimization has been suggested to compensate for inherent geological uncertainties in an oil field. In robust optimization of an oil reservoir,
the water injection and production borehole pressures are computed such that the predicted net present value of an ensemble of permeability field realizations is maximized. In our study, the permeability field is the uncertain parameters. We compare the performance of the RO strategy to a certainty equivalent optimization strategy, based on the ensemble mean of the permeability field realizations as its permeability field, and to a reactive strategy. In open-loop, for the case studied, the reactive strategy performed better than the open-loop RO strategy. These observations are nontrivial, as previous literature suggests that the open-loop RO strategy performs better than the reactive strategy. Simulations indicate that the inferior performance of the open-loop RO strategy compared to the reactive strategy is due to the inability of the RO strategy to efficiently encompass ensembles with very different and conflicting optimal control trajectories. Hence, we propose a modified RO strategy that allow shut in of uneconomical wells. The modified RO strategy performs significantly better than the other open-loop strategies and the reactive strategy. Finally, this is the first time in literature that the RO optimization has been investigated in closed-loop. Surprisingly, for the case studied, the closed-loop certainty equivalent strategy yields a higher NPV than the closed-loop RO strategy. The uncertainty reduction of the permeability field estimate due to data assimilation explains the good performance of the closed-loop certainty equivalent optimization strategy. Consequently, in closed-loop, the increased computational effort of the RO strategy compared to the certainty equivalent strategy is not justified for the particular case studied in this paper.

The third contribution of this thesis is a mean-variance method for risk mitigation in production optimization of oil reservoirs. We introduce a return-risk bicriterion objective function for the profit-risk tradeoff. With this objective function we link the optimization problem in production optimization to the Markowitz portfolio optimization problem in finance or to the robust design problem in topology optimization. In this study we focus on open-loop configuration, i.e. without measurement feedback. We demonstrate that a return-risk bi-criterion objective function is a valuable tool for the profit-risk tradeoff. If combined with the previous contribution, this result trigger the necessity of comparing the closed-loop CE strategy with the closed-loop MV strategy.

The thesis consists of a summary report and a collection of five research papers written during the period May 2010 to August 2013. Three papers are published in conference proceedings, one paper is published in Computational Geosciences journal and another paper is submitted to Journal of Petroleum Science and Engineering.

---

**General information**

**State:** Published  
**Organisations:** Department of Applied Mathematics and Computer Science, Center for Energy Resources Engineering, Scientific Computing  
**Authors:** Capolei, A. (Intern), Jørgensen, J. B. (Intern)  
**Number of pages:** 199  
**Publication date:** 2013

**Publication information**

**Place of publication:** Kgs. Lyngby  
**Publisher:** Technical University of Denmark (DTU)  
**Original language:** English

**Series:** PHD-2013  
**Number:** 315  
**ISSN:** 0909-3192  
**Main Research Area:** Technical/natural sciences  
**Electronic versions:** phd315_Capolei_A.pdf  
**Publication:** Research › Ph.D. thesis – Annual report year: 2014

**On Building Secure Communication Systems**

This thesis presents the Guided System Development (GSD) framework, which aims at supporting the development of secure communication systems.

A communication system is specified in a language similar to the Alice and Bob notation, a simple and intuitive language used to describe the global perspective of the communications between different principals. The notation used in the GSD framework extends that notation with constructs that allow the security requirements of the messages to be described.

From that specification, the developer is guided through a semi-automatic translation that enables the verification and implementation of the system. The translation is semi-automatic because the developer has the option of choosing which implementation to use in order to achieve the specified security requirements. The implementation options are given by plugins defined in the framework. The framework’s flexibility allows for the addition of constructs that model new security properties as well as new plugins that implement the security properties.

In order to provide higher security assurances, the system specification can be verified by formal methods tools such as the Beliefs and Knowledge (BAK) tool — developed specifically for the GSD framework —, LySatool and OFMC. The framework’s flexibility and the existence of the system model in different perspectives — an overall global perspective and
an endpoint perspective —allow the connection to new formal methods tools.

The modeled system is also translated into code that implements the communication skeleton of the system and can then be used by the system designer. New output languages can also easily be added to the GSD framework.

Additionally, a prototype of the GSD framework was implemented and an example of using the GSD framework in a real world system is presented.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Language-Based Technology
Authors: Carvalho Quaresma, J. N. (Intern), Probst, C. W. (Intern), Nielsen, F. (Intern)
Number of pages: 139
Publication date: 2013

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: PHD-2013
Number: 313
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
On_Building_Secure.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2014

**Stochastic Differential Equations in Artificial Pancreas Modelling**

Type 1 diabetes accounts for approximately 5% of the total diabetes population. It is caused by the destruction of insulin producing β-cells in the pancreas. Various treatment strategies are available today, some of which include advanced technological devices such as an insulin pump and a continuous glucose monitor (CGM). Despite these technological advances in the treatment of type 1 diabetes, the disease still poses an enormous and constant challenge for the patients. To obtain tight glucose control the patients are required to assess how much they will eat prior to the meal. They have to assess the timing, intensity and duration of physical exercise in advance, to adjust the insulin dose accordingly. Additionally, several uncontrollable and unpredictable factors such as stress, hormonal cycles and sickness changing the metabolic state make this task even more difficult.

The development of the insulin pump and the CGM has paved the way for a fully automatic treatment regime, the artificial pancreas. The idea is to connect the CGM with the insulin pump via a control algorithm running on e.g. the patients smart phone. The CGM observations are sent to the smart phone and based on this information, the control algorithm computes the optimal dose adjustment and sends instructions to the insulin pump.

To develop control algorithms, mathematical models of the physiological dynamics are needed. They attempt to describe the significant dynamics of the system and hence they approximate the system behavior. However, uncertainty in the model occurs due to the nature of physiological systems and due to the presence of unknown disturbances. An attractive approach to deal with this uncertainty is to use stochastic differential equations (SDEs). In a model based on SDEs, the noise is separated into two terms: 1) a diffusion term occurring from model misspecifications, effects of unknown disturbances, or just true stochastic behavior of the system and 2) a measurement noise term representing the serially uncorrelated error occurring due to the imperfect analysing equipment. The diffusion term affects the evolution of the system directly.

The purpose of this PhD-project was to investigate the potential of SDEs in the artificial pancreas development. Especially, the emerging continuous monitoring of glucose levels makes SDEs highly applicable to this field. The current thesis aims at demonstrating and discussing the benefits and challenges by using SDEs compared to traditional methods on the basis of the results of the project.

First of all, we designed a clinical study to obtain high quality data from type 1 diabetes patients to identify the models from. The study included the main factors influencing the glucose level: insulin boluses, meals, and exercise. A modelling study showed that using SDEs in model development can be advantageous in several ways. We were able to pinpoint model deficiencies in a well-known model and to track parameter variation probably caused by a differences in meal type. This information could be added to the model to improve the fit. The study was limited by the lack of a software capable of handling SDE models of population effects instead of single-subject effects. A prototype of this type of software was developed parallel to the end of the project. Thus, we could finally identify a population model of the effect of exercise on the insulin absorption rate. The small amount of observations made it impossible to use SDEs to track parameter variation. Instead, we formulated a model structure with showed to be significantly better than the base model with a constant rate.
Two studies specifically related to the CGM observations were performed during the project. In the first study, we showed that SDEs could be used to tune a control algorithm for overnight glucose control on the basis of CGM observations. The tuned algorithm improved the controller performance in a subsequent clinical study. Further attempts to deal with the problems related to the CGM included a Bayesian estimation scheme. By incorporating prior knowledge about the uncertainty in the CGM observations into the estimation method, we succeeded in predicting the plasma glucose level with acceptable confidence from the CGM observations only.

Overall, the project confirms that SDEs have a large potential within this field. However, future modeling requires a robust software capable of handling the nonlinear population SDE models. When this is available, larger modeling studies can be initiated and the impact of SDEs would be expected to increase.

Facial Analysis: Looking at Biometric Recognition and Genome-Wide Association
The goal of this Ph.D. project is to present selected challenges regarding facial analysis within the fields of Human Biometrics and Human Genetics. In the course of the Ph.D. nine papers have been produced, eight of which have been included in this thesis.

Three of the papers focus on face and gender recognition, where in the gender recognition papers the process of human perception of gender is analyzed and used to improve machine learning algorithms.

One paper addresses the issues of variability in human annotation of facial landmarks, which most papers regard as a static “gold standard”. However, we document intra- and inter-operator variability associated with annotating these landmarks, which is a valuable result for applications that are sensitive to such variability.

One paper presents a comprehensive proof-of-concept study of the prediction of facial characteristics based solely on genetic information, a new area that holds great potential.

Two papers explore the connection between minor physical anomalies in the face and schizophrenic disorders. Schizophrenia is a life long disease, but early discovery and treatment can have a significant impact on the course of the disease.

Finally, one paper presents a novel appearance model that is a fusion of the active appearance models and the Riemannian elasticity framework.
The Improved Relevance Voxel Machine

The concept of sparse Bayesian learning has received much attention in the machine learning literature as a means of achieving parsimonious representations of features used in regression and classification. It is an important family of algorithms for sparse signal recovery and compressed sensing and enables basis selection from overcomplete dictionaries.

One of the trailblazers of Bayesian learning is MacKay who already worked on the topic in his PhD thesis in 1992 [1]. Later on Tipping and Bishop developed the concept of sparse Bayesian learning [2, 3] and Tipping published the Relevance Vector Machine (RVM) [4] in 2001. While the concept of RVM was intriguing, problems with the approach were the run time which is approximately cubic in the number of basis functions as well as the greedy optimization. Hence, different approaches to overcome these shortcomings were developed, e.g. [5] or [6] as well as Tipping himself in [7] (FastRVM).

Recently, Sabuncu and Van Leemput [8, 9] extended the relevance vector machine by incorporating an additional spatial regularization term in the Gaussian prior on the regression weights or classification features (RVoxM). RVoxM encourages spatial clustering of the relevance voxels and computes predictions as linear combinations of their content. While the model of RVoxM produced nice results on age regression data [8, 9], the algorithm used a simple fixed point optimization scheme, which is not guaranteed to decrease the cost function at every step and is computationally expensive. In addition, RVoxM prunes voxels from the regression model by applying an artificial numerical threshold to the weight hyperparameters and hence has a free parameter that influences model sparsity. Finally, RVoxM can only remove voxels from the model, but not reintroduce them later on. Hence in its current form it is reminiscent of a greedy forward feature selection algorithm.

In this report, we aim to solve the problems of the original RVoxM algorithm in the spirit of [7] (FastRVM). We call the new algorithm Improved Relevance Voxel Machine (IRVoxM). Our contributions are an improvement of the greedy optimization algorithm employed in RVoxM by exploiting the form of the marginal likelihood function and deriving an analytic expression for the optimal hyperparameter of each voxel, given the current hyperparameter of all other voxels. This enables us to maximize the marginal likelihood function in a principled and efficient manner. As a result IRVoxM optimizes the objective function better during training and the resulting models predict better on unseen cases. Finally, IRVoxM enables us to flexibly add and/or remove voxels during the optimization procedure.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Massachusetts General Hospital
Authors: Ganz, M. (Ekstern), Sabuncu, M. (Ekstern), Van Leemput, K. (Intern)
Number of pages: 19
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute-Technical Report-2013
Number: 10
ISSN: 1601-2321
Main Research Area: Technical/natural sciences
Electronic versions:
tr13_10_Ganz_M_updated.pdf
Publication: Research › Report – Annual report year: 2013

Designing Scientific Software for Heterogeneous Computing: With application to large-scale water wave simulations
The main objective with the present study has been to investigate parallel numerical algorithms with the purpose of running efficiently and scalably on modern many-core heterogeneous hardware. In order to obtain good efficiency and scalability on modern multi- and many-core architectures, algorithms and data structures must be designed to utilize the underlying parallel architecture. The architectural changes in hardware design within the last decade, from single to multi
and many-core architectures, require software developers to identify and properly implement methods that both exploit concurrency and maintain numerical efficiency.

Graphical Processing Units (GPUs) have proven to be very effective units for computing the solution of scientific problems described by partial differential equations (PDEs). GPUs have today become standard devices in portable, desktop, and supercomputers, which makes parallel software design applicable, but also a challenge for scientific software developers at all levels. We have developed a generic C++ library for fast prototyping of large-scale PDEs solvers based on flexible-order finite difference approximations on structured regular grids. The library is designed with a high abstraction interface to improve developer productivity. The library is based on modern template-based design concepts as described in Glimberg, Engsig-Karup, Nielsen & Dammann (2013). The library utilizes heterogeneous CPU/GPU environments in order to maximize computational throughput by favoring data locality and low-storage algorithms, which are becoming more and more important as the number of concurrent cores per processor increases.

We demonstrate in a proof-of-concept the advantages of the library by assembling a generic nonlinear free surface water wave solver based on unified potential flow theory, for fast simulation of large-scale phenomena, such as long distance wave propagation over varying depths or within large coastal regions. Simulations that are valuable within maritime engineering because of the adjustable properties that follow from the flexible-order implementation. We extend the novel work on an efficient and robust iterative parallel solution strategy proposed by Engsig-Karup, Madsen & Glimberg (2011), for the bottleneck problem of solving a _-transformed Laplace problem in three dimensions at every time integration step. A geometric multigrid preconditioned defect correction scheme is used to attain high-order accurate solutions with fast convergence and scalable work effort. To minimize data storage and enhance performance, the numerical method is based on matrix-free finite difference approximations, implemented to run efficiently on many-core GPUs. Also, single-precision calculations are found to be attractive for reducing transfers and enhancing performance for both pure single and mixed-precision calculations without compromising robustness. A structured multi-block approach is introduced that decomposes the problem into several subdomains, supporting flexible block structures to match the physical domain. For data communication across processor nodes, messages are sent using MPI to repeatedly update boundary information between adjacent coupled subdomains. The impact on convergence and performance scalability using the proposed hybrid CUDA-MPI strategy will be presented. A survey of the convergence and performance properties of the preconditioned defect correction method is carried out with special focus on large-scale multi-GPU simulations. Results indicate that a limited number of multigrid restrictions are required, and that it is strongly coupled to the wave resolutions. These results are encouraging for the heterogeneous multi-GPU systems as they reduce the communication overhead significantly and prevent both global coarse grid corrections and inefficient processor utilization at the coarsest levels.

We find that spatial domain decomposition scales well for large problems sizes, but for problems of limited sizes, the maximum attainable speedup is reached for a low number of processors, as it leads to an unfavorable communication to compute ratio. To circumvent this, we have considered a recently proposed parallel-in-time algorithm referred to as Parareal, in an attempt to introduce algorithmic concurrency in the time discretization. Parareal may be perceived as a two level multigrid method in time, where the numerical solution is first sequentially advanced via course integration and then updated simultaneously on multiple GPUs in a predictor-corrector fashion. A parameter study is performed to establish proper choices for maximizing speedup and parallel efficiency. The Parareal algorithm is found to be sensitive to a number of numerical and physical parameters, making practical speedup a matter of parameter tuning. Results are presented to confirm that it is possible to attain reasonable speedups, independently of the spatial problem size.

To improve application range, curvilinear grid transformations are introduced to allow representation of complex boundary geometries. The curvilinear transformations increase the complexity of the implementation of the model equations. A number of free surface water wave cases have been demonstrated with boundary-fitted geometries, where the combination of a flexible geometry representation and a fast numerical solver can be a valuable engineering tool for large-scale simulation of real maritime scenarios.

The present study touches some of the many possibilities that modern heterogeneous computing can bring if careful and parallel-aware design decisions are made. Though several free surface examples are outlined, we are yet to demonstrate results from a real large-scale engineering case.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Scientific Computing
Authors: Glimberg, S. L. (Intern), Engsig-Karup, A. P. (Intern), Dammann, B. (Intern)
Number of pages: 153
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2013
Number: 317
Large-scale Machine Learning in High-dimensional Datasets

Over the last few decades computers have gotten to play an essential role in our daily life, and data is now being collected in various domains at a faster pace than ever before. This dissertation presents research advances in four machine learning fields that all relate to the challenges imposed by the analysis of big data.

In the field of kernel methods, we present an information-based denoising technique based on semi-supervised kernel Principal Component Analysis (PCA), that incorporates label information into the kernel PCA objective. Effectively, this guides the low-rank representation towards relevant components, while exploiting intrinsic manifold structures exposed by the data. In the same field, we also introduce a scalable randomized heuristic for optimizing kernel hyperparameters, that is based on maximizing the Minimum Enclosing Ball (MEB) of the class means in the associated Reproducing Kernel Hilbert Space (RKHS).

In the field of spectral methods, we introduce semi-supervised eigenvectors of a graph Laplacian, that inherit many of the properties that characterize the global eigenvectors, but by using side-information in the form of a seed set, the semi-supervised eigenvectors are better at modeling local heterogeneities.

In the field of machine learning for neuroimaging, we introduce learning protocols for real-time functional Magnetic Resonance Imaging (fMRI) that allow for dynamic intervention in the human decision process. Specifically, the model exploits the structure of fMRI data by incorporating a temporal Gaussian Process (GP) smoothness prior, which reduces model degeneracy caused by mislabeled data samples.

Finally, in the field of topic modeling, we introduce a Graphics Processing Unit (GPU) accelerated framework for co-clustering in large-scale sparse bipartite networks. By implementing the Infinite Relational Model (IRM) in this framework we achieve speedups of two orders of magnitude compared to estimation based on conventional processors.

Power Management for Energy Systems

In this thesis, we consider the control of two different industrial applications that belong at either end of the electricity grid; a power consumer in the form of a commercial refrigeration system, and wind turbines for power production. Our primary studies deal with economic model predictive control of a commercial multi-zone refrigeration system, consisting of several cooling units that share a common compressor, and is used to cool multiple areas or rooms, e.g., in supermarkets. Substantial amounts of energy are consumed in refrigeration systems worldwide and there is a strong motivation for introducing more energy efficient as well as cost reducing control techniques. At the same time, the power grid is evolving from a centralized system with rather controllable production in the conventional power plants to a much more decentralized network of many independent power generators and a large penetration of renewable, fossil-free energy sources such as solar and wind power. To facilitate such intermittent power producers, we must not only control the production of electricity, but also the consumption, in an efficient and exible manner. By enabling the use of thermal energy storage in supermarkets, we open up for exible power consumption schemes with the possibility of reducing operational costs and we develop and demonstrate prototype control technology that creates completely new business opportunities for selling regulating power to the grid. Moreover, this enables a larger penetration of wind energy in the power production
and increases the potential market size for wind power generators and other renewable energy sources. Thus, we aim at promoting the use of environmentally sustainable power production technologies while creating new business opportunities for both power consumers and producers of renewable energy.

The second application, wind turbines, takes us to the production side of the power grid. The key concern here is to improve the quality and integrability of power delivered to the grid from large parks of wind turbines. Our goal is to reduce the fluctuating nature of the power output and to meet tightened demands from the grid by enabling a more intelligent control at both the individual turbine level, at the park controller level, and in cooperation with exible power consumers or other means of energy storage. The possible interaction and synergies of the two applications are obvious reasons to consider both in this thesis, and as we will see, the similarities in our formulations of the different control problems allow us to apply almost identical techniques despite the lack of immediate similarity.

For control of the commercial refrigeration application as well as the wind turbine application, we propose an economic optimizing model predictive controller, economic MPC. MPC is a feedback control technique that is characterized by its explicit handling of constrained control problems in which a model is used to predict the future behavior of a system along with forecasts of future disturbances. At each time step the values of the control inputs are computed by solving an open-loop nite time optimal control problem over a dened prediction horizon. Only the rst step in this optimal open-loop sequence is implemented as a control command. Feedback is obtained by solving the open-loop problem repeatedly, in a receding horizon fashion, as new predictions become available.

Our investigations are primarily concerned with: 1) modeling of the applications to suit the chosen control framework; 2) formulating the MPC controller laws to overcome challenges introduced by the industrial applications, and de ning economic objectives that reect the real physics of the systems as well as our control objectives; 3) solving the involved, non-trivial optimization problems eciently in real-time; 4) demonstrating the feasibility and potential of the proposed methods by extensive simulation and comparison with existing control methods and evaluation of data from systems in actual operation. We present contributions on:

Economic MPC for commercial refrigeration systems, including
-Linear economic MPC formulations that utilize the exibility in refrigeration systems to counteract uctuations in the balance between power consumption and production.
-Economic MPC with probabilistic constraints, ensuring a robust performance and constraint satisfaction in spite of inaccurate system models and forecasts.
-Nonlinear economic MPC, re ecting the nonconvexity in the realistic description of temperature dependent e ciencies in the refrigeration cycle.
-Nonlinear economic MPC with uncertain predictions and the implementation of very simple predictors that use entirely historical data of, e.g., electricity prices and outdoor temperatures.

Economic MPC for wind turbines, including
-Optimal steady-state calculation for wind farms.
-Nonlinear economic MPC for individual turbines.

-Change of variables and convex formulations of economic MPC for individual turbines.

Tractable optimization methods for the MPC problems, including
-Sequential convex programming (SCP) for specic nonconvex problems originating from our studies of commercial refrigeration as well as from our studies concerning wind power.
-Successful demonstration of the SCP approach on three different problems the commercial refrigeration system with linear dynamics and constraints and a nonconvex objective, the individual wind turbine with nonlinear dynamics and constraints, and the static optimization of the wind farm with a black-box model.

The major contribution is the formulation of these problems and the demonstrations to show that the SCP method can be used for their solution. We demonstrate, i.a., substantial cost savings, on the order of 30 %, compared to a standard thermostat-based supermarket refrigeration system and show how our methods exhibit sophisticated demand response to real-time variations in electricity prices. Violations of the temperature ranges can be kept at a very low frequency of occurrence inspite of the presence of uncertainty. For the power output from wind turbines, ramp rates, as low a 3 % of the rated power per minute, can be efciently ensured with the use of energy storage and we show how the active use of rotor inertia as an additional energy storage can reduce the needed storage capacity by up to 30 % without reducing the power output.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing, Center for Energy Resources Engineering, Department of Electrical Engineering, Automation and Control, VESTAS Wind Systems A/S
Authors: Hovgaard, T. G. (Intern), Jørgensen, J. B. (Intern), Blanke, M. (Intern), Larsen, L. F. S. (Ekstern)
Number of pages: 255
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
ISBN (Print): 978-87-643-1147-1
Original language: English
Formal methods for design and simulation of embedded systems

Cyber physical systems (CPSs) are present in many variants in our daily life. The complexity of developing a CPS is quickly increasing and the interaction between different CPSs is increasingly important. The interaction of the systems is becoming more and more fluent and seamless.

This thesis presents the development of a formal systems modelling (ForSyDe) framework for modelling CPSs. The formalism of the framework makes computer aided design (CAD) a possibility for developing CPSs. The framework consists of four models of computation (MoCs): synchronous (SY), synchronous data flow (SDF), discrete event (DE), and continuous time (CT).

Usage of the framework is demonstrated with two use cases. A company use case featuring a hearing aid calibration device and the distributed energy harvesting aware routing (DEHAR) algorithm for wireless sensor networks (WSNs). These two use cases illustrate different design challenges. With the ForSyDe framework, the use cases are expressed as homogeneous and heterogeneous models.

The company use case illustrates that the ForSyDe framework handles systems with well defined interactions very well. The WSN use case illustrates that networked systems with complex interaction are more challenging to express naturally, yet the ForSyDe framework is able to express such systems.

Challenges in 3D scanning: Focusing on Ears and Multiple View Stereopsis

It is the goal of this thesis to address some of the challenges in 3D scanning. This has been done with focus on direct in-ear scanning and on Multiple View Stereopsis. Seven papers have been produced over the course of the Ph.D., out of which, six have been included. Two papers concern volumetric segmentation based on Markov Random Fields. These have been formulated to address problems relating to noise filtering in direct in-ear scanning and Intracranial Volume estimation. Another two papers have been produced on the topic of recovering surface data based on a strong statistical prior. This was done in particular on scans of ear canals, but the methods are general. Finally, an experimental setup has been constructed, which has produced a large versatile data set. The data set has been used as the foundation for two papers on the evaluation of Multiple View Stereopsis. The data have a great potential to be used for advances in Multiple View Stereopsis, robust surface reconstruction and photorealistic modelling.
Sparse Image Reconstruction in Computed Tomography

In recent years, increased focus on the potentially harmful effects of x-ray computed tomography (CT) scans, such as radiation-induced cancer, has motivated research on new low-dose imaging techniques. Sparse image reconstruction methods, as studied for instance in the field of compressed sensing (CS), have shown significant empirical potential for this purpose. For example, total variation regularized image reconstruction has been shown in some cases to allow reducing x-ray exposure by a factor of 10 or more, while maintaining or even improving image quality compared to conventional reconstruction methods.

However, the potential in CT has mainly been demonstrated in individual proof-of-concept studies, from which it is hard to distill general conditions for when sparse reconstruction methods perform well. As a result, there is a fundamental lack of understanding of the effectiveness and limitations of sparse reconstruction methods in CT, in particular in a quantitative sense. For example, relations between image properties such as contrast, structure and sparsity, tolerable noise levels, sufficient sampling levels, the choice of sparse reconstruction formulation and the achievable image quality remain unclear. This is a problem of high practical concern, because the large scale of CT problems makes detailed exploration of the parameter space very time-consuming. Due to the limited quantitative understanding, sparse reconstruction has not yet become the method of choice in practical CT applications.

This thesis takes a systematic approach toward establishing quantitative understanding of conditions for sparse reconstruction to work well in CT. A general framework for analyzing sparse reconstruction methods in CT is introduced and two sets of computational tools are proposed:

1. An optimization algorithm framework enabling easy derivation of algorithms for sparse reconstruction problems, and
2. Tools for characterizing sparse reconstruction in CT, i.e., establishing relations between parameters governing reconstruction quality.

The flexibility of the optimization algorithm framework is demonstrated by constructing convergent optimization algorithms for a range of sparse reconstruction problems of interest to CT. The practical usefulness of the framework is shown through case studies of the efficacy of specific sparse reconstruction problems in tomographic reconstruction.

The characterization methods proposed in the thesis focus on the role of image sparsity for the level of sampling required for accurate CT reconstruction. While a relation between sparsity and sampling is motivated by CS, no theoretical guarantees of accurate sparse reconstruction are known for CT. In simulation studies, a sparsity-sampling relation is established in CT. This enables quantification of the undersampling allowed by sparse reconstruction methods.

Both the prototyping framework and the characterization methods add to the understanding of sparse reconstruction methods in CT and serve as initial contributions to a general set of computational characterization tools. Thus, the thesis contributions help advance sparse reconstruction methods toward routine use in...
Spatio-temporal abundance and dispersal of *Culicoides*

This PhD project comprises studies of biting midges (*Culicoides*) in Denmark with regards to vector-borne diseases such as bluetongue virus (BTV) and Schmallenberg virus (SBV). Both diseases are new in northern Europe. In Denmark there was an outbreak of BTV in 2007 and 2008. BTV infects ruminants, and especially infected sheep and cattle are constitute a problem for farmers. The symptoms of BTV include fever, cyanotic tongue, oedemas and decreased milk production. The last symptom affects the economy and animal welfare in the farming industry. In 2011 and 2012, outbreaks of SBV were also recorded in Denmark. The symptoms of SBV are similar to BTV but also include a high proportion of malformations and stillbirths in lambs. Models of vector-borne diseases can be used to predict an outbreak and evaluate e.g. the optimal control strategy, the economic impact and the number of infected animals. These models need to have proper input regarding the abundance and behavior of the vectors. If no vectors are present in an area, the disease will not spread. Thus the vector abundance is a very important factor for models of vector-borne diseases. This PhD project investigates different key factors important for the abundance and behavior of vectors.

Inverse Problems in Geosciences: Modelling the Rock Properties of an Oil Reservoir

Even the most optimistic forecasts predict that Danish oil production will decrease by 80% in the period between 2006 and 2040, and only a strong innovative technological effort can change that. Due to the geological structures of the subsurface in the Danish part of the North Sea, Denmark is currently missing out on approximately 70% of the oil, which is left behind, trapped in unreachable parts of the reservoirs.

An increase in the oil recovery rate can be achieved by better planning and optimisation of oil production. Both require an improved description of the rock properties of the subsurface of the reservoirs. Hence the focus of this work has been on acquiring models of spatial parameters describing rock properties of the subsurface using geostatistical a priori knowledge and available geophysical data. Such models are solutions to often severely under-determined, inverse problems.

The focus of the study has been on the computational aspects of inferring such models. Reservoir modelling is a large-scale problem with great computational complexity and the work should be seen as a first part of a foundation for one day,
when the computational resources are available, being able to handle the large scale problems of the petroleum industry. But for now most of the study is based on simplified and idealised models.

We have proposed a method for efficient and accurate interpolation of rock properties from seismic data. It is based on a recently published paper on interpolation of rock properties that breaks with the dominating influence of spatial coordinates in traditional interpolation methods. The thesis contains work involving a test case study of the method demonstrating how the interpolation in attribute space ensures the geological structures of the computed models and how the method can be further improved by an orthogonal transformation of the attribute space.

We have formulated a closed form expression of an a priori probability density function that quantifies the statistical probability of models describing the rock properties of a reservoir. This can be used to evaluate the probability that a model adhere to prior knowledge by having specific multiple-point statistics, for instance, learned from a training image. Existing methods efficiently sample an a priori probability density function to create a set of acceptable models; but they cannot evaluate the probability of a model.

We have developed and implemented the Frequency Matching method that uses the closed form expression of the a priori probability density function to formulate an inverse problem and compute the maximum a posteriori solution to it. Other methods for computing models that simultaneously fit data observations and honour a priori knowledge are not capable of computing the maximum a posteriori solution. Instead they either sample the posterior probability density function or they sample the a priori probability density function to optimise the likelihood function.

This thesis consists of a summary report and seven research papers submitted, reviewed and/or published in the period 2010 - 2013.

**General information**
- **State:** Published
- **Organisations:** Center for Energy Resources Engineering, Department of Applied Mathematics and Computer Science, Scientific Computing, Department of Chemistry
- **Authors:** Lange, K. (Intern), Mosegaard, K. (Intern), Hansen, P. C. (Intern), Stenby, E. H. (Intern)
- **Number of pages:** 248
- **Publication date:** 2013

**Publication information**
- **Place of publication:** Kgs. Lyngby
- **Publisher:** Technical University of Denmark (DTU)
- **Original language:** English

**Series:** IMM-PHD-2013
**Number:** 296
**ISSN:** 0909-3192
**Main Research Area:** Technical/natural sciences

**Electronic versions:**
- [phd296_Lange_K.pdf](phd296_Lange_K.pdf)

**Publication:** Research › Ph.D. thesis – Annual report year: 2013

**Determination of magnetic resonance imaging biomarkers for multiple sclerosis treatment effects**
This thesis describes methods for deriving multiple sclerosis (MS) biomarkers from Magnetic resonance images (MRI).

MS results in a neurodegenerative disease course to which MRI has proven sensitive. In particular diffusion MRI (dMRI), a modality reflecting microstructural properties of brain tissue has shown sensitivity towards the disease pathology of MS. We introduce three different methods for analysing MRI/dMRI in the white matter (WM) tracts, of an MS population. One method detects groupwise, tract-oriented differences based on features of the local diffusion tensor model. The next method, anatomical connectivity mapping (ACM) reflects voxel-wise whole-brain connectivity and is used to investigate cross sectional disease-related connectivity alterations. The third method presented is a voxelbased segmentation method able to detect WM abnormalities (WM lesions), with the potential of being used as lesion load markers often reported in clinical studies.

The main result of the first method is statistical differences between healthy controls and MS patients in 11 WM tracts. The ability to distinguish the clinically defined subtypes of relapse remitting and secondary progressive MS patients is found based on the ACM method. Using ACM, localized statistical differences were detected in the bilateral motor tracts. The most interesting result of the lesion segmentation method study, was that it achieved a segmentation performance which was better than two competing methods relative to the manual segmentations of the radiographers.

The methods presented in the thesis are useful in studies of MS and are expected to have widespread applications in neuroscience.
RF subsystem power consumption and induced radiation emulation
The thesis introduces a novel approach towards the emulation of the RF subsystem power consumption when transmitting a LTE signal. The RF subsystem which is made up of analog components has not been covered by the status quo emulation methodologies which are compatible with digital circuits. Though the study of the RF subsystem architectures revealed numerous architectures with different impacts on power consumption, we have decided to consider the RF subsystem as a black box.

The RF subsystem power emulation has been studied for the telecommunication technology Long Term Evolution (LTE). Given the fact that major power consumptions of wireless devices are largely functions of sequences of protocol/logical activities, it is this technology that provided the inputs to the RF subsystem as a black box which are Tx power, carrier frequency and signal bandwidth. The physical environmental variable temperature has also proven to be very influential on power consumption. These inputs also do constitute to the input parameters of the emulation methodology.

The emulation methodology has been proven to be a mathematical mapping between the input parameters and a predefined mathematical model. For the mathematical model, multivariate modeling approaches were analyzed for an approach with the least modeling error and complexity. Herein, the homotopy continuation numerical approach proved to have the least modeling error of 3%. The RF subsystem power consumption has been emulated with accuracies of 84 ±2.25% and 94.3% ±2.25% on different devices.

List Decoding of Algebraic Codes
We investigate three paradigms for polynomial-time decoding of Reed–Solomon codes beyond half the minimum distance: the Guruswami–Sudan algorithm, Power decoding and the Wu algorithm. The main results concern shaping the computational core of all three methods to a problem solvable by module minimisation; by applying the fastest known algorithms for this general problem, we then obtain realisations of each paradigm which are as fast or faster than all previously known methods. An element of this is the “2D key equation”, a heavily generalised form of the classical key
equation, and we show how to solve such using module minimisation, or using our new Demand–Driven algorithm which is also based on module minimisation.

The decoding paradigms are all derived and analysed in a self-contained manner, often in new ways or examined in greater depth than previously. Among a number of new results, we give: a fast maximum-likelihood list decoder based on the Guruswami–Sudan algorithm; a new variant of Power decoding, Power Gao, along with some new insights into Power decoding; and a new, module based method for performing rational interpolation for the Wu algorithm. We also show how to decode Hermitian codes using Guruswami–Sudan or Power decoding faster than previously known, and we show how to Wu list decode binary Goppa codes.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Mathematics
Authors: Nielsen, J. S. R. (Intern), Beelen, P. (Intern), Høholdt, T. (Intern)
Number of pages: 178
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: DTU Compute PHD-2013
Number: 309
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
jsrn_thesis.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2013

Viscoelastic Modelling of Road Deflections for use with the Traffic Speed Deflectometer
This Ph.D. study is at its core about how asphalt and road structures responds to dynamic loads. Existing models for the deflections under a moving load using beam equations are revisited and it is concluded they leave room for improvement for the particular setup and problem at hand. Then a different approach is set up to model visco-elastic deflections starting from the physically based framework of continuum mechanics by using Finite Element Methods (FEM) combined with the Laplace transform. It is shown that this approach coincides with a more standard time-stepping FEM setup in the case of a generalized Maxwell model. Validations by comparison to ViscoRoute simulations are also made. This justifies the use of the Laplace FEM for generating simulated data using a Huet-Sayegh model for the visco-elastic behaviour of asphalt. These simulated data, along with measured data, are then used to suggest an approach for a computationally simpler synthetic model capturing essential behaviour of deflection basins under a moving wheel.

Additionally the setup allows for simulated comparisons of the cases of loadings emulating the use of a Falling Weight Deflectometer with loadings emulating a moving wheel as in the case of using a Traffic Speed Deflectometer. The flexibility of the method also allows for looking into cases excluded by imposing simplifying assumptions such as the structure imagined to be an infinite halfspace.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Scientific Computing
Authors: Pedersen, L. (Intern), Hjorth, P. G. (Intern), Knudsen, K. (Intern)
Number of pages: 97
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2013
Number: 310
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
The Homogeneous Interior-Point Algorithm: Nonsymmetric Cones, Warmstarting, and Applications

The overall topic of this thesis is convex conic optimization, a sub-field of mathematical optimization that attacks optimization problem with a certain geometric structure. These problems allow for modelling of an extremely wide range of real-world problems, but the availability of solution algorithms for these problems is still limited.

The goal of this thesis is to investigate and shed light on two computational aspects of homogeneous interior-point algorithms for convex conic optimization:

The first part studies the possibility of devising a homogeneous interior-point method aimed at solving problems involving constraints that require nonsymmetric cones in their formulation. The second part studies the possibility of warmstarting the homogeneous interior-point algorithm for conic problems. The main outcome of the first part is the introduction of a completely new homogeneous interior-point algorithm designed to solve nonsymmetric convex conic optimization problems. The algorithm is presented in detail and then analyzed. We prove its convergence and complexity. From a theoretical viewpoint, it is fully competitive with other algorithms and from a practical viewpoint, we show that it holds lots of potential, in several cases being superior to other solution methods.

The main outcome of the second part of the thesis is two new warmstarting schemes for the homogeneous interior-point algorithm for conic problems. Again, we first motivate and present the schemes and then analyze them. It is proved that they, under certain circumstances, result in an improved worst-case complexity as compared to a normal coldstart. We then move on to present an extensive series of computational results substantiating the practical usefulness of these warmstarting schemes. These experiments include standard benchmarking problem test sets as well as an application from smart energy systems.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Scientific Computing, Center for Energy Resources Engineering
Authors: Skajaa, A. (Intern), Hansen, P. C. (Intern), Jørgensen, J. B. (Intern)
Number of pages: 149
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2013
Number: 311
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd311_Skajaa_A.pdf Publication: Research › Ph.D. thesis – Annual report year: 2014

Quantitative assessment of course evaluations
Student evaluation of teaching has been used in educational institutions around the world as a means of providing feedback on the quality of teaching. Nowadays, it is one of the most widespread tools used to inform teachers and administration about the instruction given in an institution.

The goal of the thesis is to develop efficient tools to analyze the data from student evaluations of teaching and courses at the Technical University of Denmark.

The thesis explores both classical and modern methods of multivariate statistical data analysis to address different issues of student evaluation of teaching (SET). In particular, the thesis includes results on the investigation of the association between the student evaluations of the course and the student evaluations of the teachers, the investigation of the effects of the mid-term evaluation on the end-of-term evaluations and the investigation of the student non-response on SETs. In order to utilize information from open-ended qualitative student answers, text-mining methods were applied in order to extract points of students praise and complaints.

The methods proposed contribute to the knowledge about student evaluation at the Technical University of Denmark. The results provided some new information that will help teachers and university managers to better understand results of course evaluations.
Mid-term course evaluation was found to be able to capture both types of course issues: issues that can be addressed during the semester and also issues that can only be addressed at the next semester. Therefore, it seems to be preferable to conduct general mid-term evaluations instead of end-of-term evaluation, so the current course students can benefit. Additionally, it might be beneficial to conduct a short end-of-term evaluation with very limited number of questions that focus on general course issues after the final exams in order to obtain student feedback on the entire teaching and learning process, including the alignment of assessment of students’ learning with course objectives and teaching activities.

Student-specific and course-specific characteristics was found to be related with whether students participate in SETs and with how students evaluate courses and teachers. The DTU administrations should be aware that high achievers are more likely to participate in course evaluation survey and are more likely to give higher scores to courses. Students diversity on the course should be taken into account while making comparisons of evaluation results between courses.

In the student written feedback was found be able to provide additional knowledge of student point of satisfaction or dissatisfaction. However, in order to build an automated tool that can help to extract patterns from student comments higher quality of the collected data is needed.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Statistics and Data Analysis
Authors: Sliusarenko, T. (Intern), Ersbøll, B. K. (Intern), Clemmensen, L. K. H. (Intern)
Number of pages: 196
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: DTU Compute PHD-2013
Number: 318
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd318_Sliusarenko_T_web.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2014

Short-term wind power forecasting: probabilistic and space-time aspects
Optimal integration of wind energy into power systems calls for high quality wind power predictions. State-of-the-art forecasting systems typically provide forecasts for every location individually, without taking into account information coming from the neighbouring territories. It is however intuitively expected that owing to the inertia in meteorological systems such local approach to power forecasting is sub-optimal. Indeed, errors in meteorological forecasts might translate to fronts of imbalances, i.e. taking the form of a band of forecast errors propagating across entire regions.

My research work deals with the proposal and evaluation of new mathematical models and forecasting methods for short-term wind power forecasting, accounting for space-time dynamics based on geographically distributed information. Different forms of power predictions are considered, starting from traditional point forecasts, then extending to marginal predictive densities and, finally, considering multivariate space-time trajectories.

Point predictions is the most classical approach to wind power forecasting, only providing single-valued estimates of the expected future power generation. The objective is to introduce a statistical model which would improve the quality of state-of-the-art prediction methods by accounting for the fact that forecasts errors made by such locally-optimized forecasting methods propagate in space and in time under the influence of prevailing weather conditions.

Subsequently, the extension from point to probabilistic forecasts is dealt with, hence requiring to describe the uncertainty associated with the point predictions previously generated. Both parametric and non-parametric approaches to forming predictive densities are considered, while ways to include space-time effects into the corresponding models are analysed.

As a final step, emphasis is placed on generating space-time trajectories: this calls for the prediction of joint multivariate predictive densities describing wind power generation at a number of distributed locations and for a number of successive lead times.

In addition to new improved approaches to wind power forecasting, a part of the research is devoted to problems related to the assessment of high-dimensional (multivariate) probabilistic forecasts.

General information
An algebraic approach to analysis of recursive and concurrent programs

This thesis focuses on formal techniques based on static program analysis, model checking and abstract interpretation that offer means for reasoning about software, verification of its properties and discovering potential bugs.

First, we investigate an algebraic approach to static analysis and explore its connections to abstract interpretation framework. We introduce the notion of a flow algebra, which is an algebraic structure similar to semirings, but closer to the classical monotone frameworks. We also generalize Galois connections to flow algebras and discuss when a flow algebra is an upper-approximation of (or induced from) another flow algebra.

Furthermore, we show how flow algebras can be used in communicating or weighted pushdown systems. To achieve that, we show that it is possible to relax some of the requirements imposed by original formulation of those techniques without compromising the soundness or completeness results.

Moreover, we present a new application of pushdown systems in the context of an aspect-oriented process calculus. The addition of aspect-oriented features makes it possible for a process to exhibit a recursive structure. We show how one can faithfully model and analyze such a language.

We also introduce an abstract domain that symbolically represents the messages sent between the concurrently executing processes. It stores prefixes or suffixes of communication traces including various constraints imposed on the messages. Since the problem has exponential complexity, we also present a compact data structure as well as efficient algorithms for the semiring operations.

Apart from that, we discuss an improvement to Pre* and Post* algorithms for pushdown systems, making it possible to directly use program representations such as program graphs. We present a modular library implementing those algorithms, which also provides a lot of flexibility with respect to, e.g., various constraints solvers.

Finally, we describe one such experimental solver based on Newton’s method. It allows solving equation systems over abstract domains that were not accommodated by other solving techniques, e.g., Kleene iteration. We present such a domain and provide a preliminary evaluation of our implementation.

To conclude, we believe the thesis presents a number of contributions interesting both from the theoretical point of view as well as from an implementation one.
Modeling of Craniofacial Anatomy, Variation, and Growth
The topic of this thesis is automatic analysis of craniofacial images with respect to changes due to growth and surgery, inter-subject variation and intracranial volume estimation. The methods proposed contribute to the knowledge about specific craniofacial anomalies, as well as provide a tool for detailed analyses for clinical and research purposes.

Most of the applications in this thesis rely on non-rigid image registration by the means of warping one image into the coordinate system of another image. This warping results in a deformation field that describes the anatomical correspondence between the two images. To elaborate further: a computational atlas of the average anatomy was constructed. Using non-rigid registration, image data from a subject is automatically transformed into the coordinate space of the atlas. In this process, all knowledge built into the atlas is transferred to the subject, thus creating a personalized atlas. The knowledge built into the atlas is e.g. location of anatomical regions and landmarks of importance to surgery planning and evaluation or population studies. With these correspondences, various analyses could be carried out e.g. quantification of growth, inter-subject variation etc. Besides image registration, a volumetric segmentation method using graph cuts was developed and applied for intracranial volume estimation. Graph cut is a fast method for segmentation utilizing a suitable graph.

Three different craniofacial anomalies were examined in this thesis: Cleft lip and palate, unicoronal synostosis, and Crouzon syndrome. Using the proposed methods, highly detailed variation was assessed for cleft lip and palate, correspondence between images obtained before and after lip repair was established for cleft lip and palate, the intracranial volume was estimated for infants with unicoronal synostosis, and finally, craniofacial growth patterns were quantified for Crouzon syndrome in a mouse model.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Image Analysis & Computer Graphics
Authors: Thorup, S. S. (Intern), Larsen, R. (Intern), Darvann, T. A. (Intern), Paulsen, R. R. (Intern)
Number of pages: 174
Publication date: 2013

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2013
Number: 290
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd290_Thorup_SS.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2013

Modeling and forecasting of wind power generation - Regime-switching approaches
The present thesis addresses a number of challenges emerging from the increasing penetration of renewable energy sources into power systems. Focus is placed on wind energy and large-scale offshore wind farms. Indeed, offshore wind power variability is becoming a serious obstacle to the integration of more renewable energy into power systems since these systems are subjected to maintain a strict balance between electricity consumption and production, at any time. For this purpose, wind power forecasts offer an essential support to power system operators. In particular, there is a growing demand for improved forecasts over very short lead times, from a few minutes up to a few hours, because these forecasts, when generated with traditional approaches, are characterized by large uncertainty. In this thesis, this issue is considered from a statistical perspective, with time series models. The primary case study is the Horns Rev wind farm located in the North Sea. Regime-switching aspects of offshore wind power fluctuations are investigated. Several formulations of Markov-Switching models are proposed in order to better characterize the stochastic behavior of the underlying process and improve its predictability. These models assume the existence of a hidden or unobservable regime sequence. Estimation methods are presented in both Bayesian and Frequentist frameworks. Markov-Switching models enable to highlight structural breaks in the dynamics of offshore wind power generation, with alternating periods of high and low variability. They also yield substantial gains in probabilistic forecast accuracy for lead times of a few minutes. However, these models only integrate historical and local measurements of wind power and thus have a limited ability for notifying
regime changes for larger lead times. For that purpose, there is a long tradition in using meteorological forecasts of wind speed and direction that are converted into wind power forecasts. Nevertheless, meteorological forecasts are not informative on the intra-hour wind variability and thus cannot be used in the present context focusing on temporal resolutions of a few minutes. Instead, this thesis investigates the use of weather radar observations for monitoring weather conditions in the vicinity of offshore wind farms, with the ambition of establishing a link between the passage of precipitation systems and high wind variability. The underlying motivation of this approach is twofold. First, it aims at providing a meteorological interpretation of the hidden regimes as estimated by regime-switching models. Second, it aims at determining an observed sequence of regimes based on the information extracted from the observations supplied weather radar observations. This approach, combining both meteorological and statistical expertise, opens up new possibilities for designing prediction systems in wind energy.

**General information**

State: Published
Organisations: Department of Informatics and Mathematical Modeling, Mathematical Statistics
Authors: Trombe, P. (Intern), Madsen, H. (Intern), Pinson, P. (Intern)
Number of pages: 166
Publication date: 2013

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: PHD-2013
Number: 287
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd287_Trombe_PJ.pdf
Publication: Research > Ph.D. thesis – Annual report year: 2014

**Optimization Under Uncertainty for Management of Renewables in Electricity Markets**

This thesis deals with the development and application of models for decision-making under uncertainty to support the participation of renewables in electricity markets.

The output of most renewable sources, e.g., wind, is intermittent and, furthermore, it can only be predicted with a limited accuracy. As a result of their non-dispatchable and stochastic nature, the management of renewables poses new challenges as compared to conventional sources of electricity. Focusing in particular on short-term electricity markets, both the trading activities of market participants (producers, retailers and consumers) and the decision-making processes of system and market operators are challenged.

As far as producers are concerned, participation in electricity markets imposes them to make their trading decisions with a certain advance in time as compared to energy delivery. Since their actual output is uncertain at the time of bidding, the trading problem for a renewable power producer translates into a stochastic optimization problem, whose objective is the maximization of the expected revenues. In this thesis, we consider the trading problem for a wind power producer both in markets with low penetration of renewables, where the producer is a price-taker, and in markets where the producer acts as a price-maker.

Owing to the demand response initiatives to be undertaken in future power systems, the operation of electricity retailers and the behavior of consumers are also going to be influenced by renewable power production. Another focus of this thesis is on time-varying price mechanisms to make the most of end consumers' flexibility. In particular, the problem of managing optimally a virtual power plant equipped with renewable production facilities and flexible consumers is addressed through control-by-price. In a similar setup, the optimal trading (and pricing) problem for a retailer connected to flexible consumers is considered.

Finally, market and system operators are challenged by the increasing penetration of renewables, which put stress on markets that were designed to accommodate a generation mix largely dominated by conventional sources. Indeed, the traditional market design, based on the sequential clearing of successive market floors and on deterministic rules and criteria, is characterized by higher and higher degrees of suboptimality and lower reliability as the penetration of renewables increases. This work contributes to the state-of-the-art by proposing new mechanisms for day-ahead dispatch and reserve determination in markets with high penetration of renewables, on the basis of stochastic criteria.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Dynamical Systems, Department of Electrical Engineering
Structured Intuition: A Methodology to Analyse Entity Authentication

Entity authentication is a process of verifying a claimed identity of a network party. It may appear to be a simple goal, but, depending on the application and context, it entails a number of modalities, such as whether the party is currently active on the network, whether the party is willing to communicate, and whether the party knows that it has been authenticated. Combining such goals in different ways leads to different flavours of entity authentication.

On an unauthenticated channel, an adversary can present a false claim of identity. Clearly, if the adversary succeeds, it may have serious consequences for the security of the system, e.g., private information of legitimate parties may be leaked or the security policy of a trusted system may be violated. At a corporate level, such a failure of authentication may result in loss of proprietary technology or customers' credit card information. Sometimes, a single failure of authentication affects the system for a long time, e.g., if an adversary is able to install a malicious program, such as a root kit, back door, key logger, bot, or other malware. Therefore, security protocols, which can resist a resourceful adversary, are used to authenticate network parties.

Verification of an authentication protocol to show that it is secure is a hard problem. Most of the reported flaws in authentication protocols are not due to some weakness in the cryptographic primitives used in these protocols. The usual problems lie in improper use of cryptographic primitives, and failure to unambiguously specify protocol assumptions and goals. Therefore, it is important that an authentication protocol is analysed with clear goals and explicitly stated assumptions.

There are many different formal definitions of authentication goals, and the decision of which definition is most appropriate depends on the requirements and constraints imposed by the larger system. Whether a reported flaw in a protocol is exploitable depends on the protocol goals and the environment in which the protocol is deployed. Whether a 'secure' protocol is indeed secure depends on the security model and the level of abstraction used in the analysis. Thus, the goal of developing a high level methodology that can be used with different notions of security, authentication, and abstraction is worth considering.

In this thesis, we propose a new methodology, called the structured intuition (SI), which addresses the issues mentioned above. In the SI, we divide entity authentication into fine grained properties, which we call FLAGs (fine level authentication goals). FLAGs are protocol independent goals and represent one's expectations in an authentication-as-a-service paradigm. There is a single notion of security in our methodology, which is called canonicity, which is a weaker form of message authenticity. As compared to many contemporary analysis techniques, an SI based analysis provides detailed results regarding the design rationales and entity authentication goals of a protocol.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Embedded Systems Engineering, Computer Science and Engineering
Authors: Ahmed, N. (Intern), Jensen, C. D. (Intern)
Number of pages: 191
Publication date: 2012
Models for efficient integration of solar energy

Efficient operation of energy systems with substantial amount of renewable energy production is becoming increasingly important. Renewables are dependent on the weather conditions and are therefore by nature volatile and uncontrollable, opposed to traditional energy production based on combustion. The "smart grid" is a broad term for the technology for addressing the challenge of operating the grid with a large share of renewables. The "smart" part is formed by technologies, which models the properties of the systems and efficiently adapt the load to the volatile energy production, by using the available flexibility in the system.

In the present thesis methods related to operation of solar energy systems and for optimal energy use in buildings are presented. Two approaches for forecasting of solar power based on numerical weather predictions (NWPs) are presented, they are applied to forecast the power output from PV and solar thermal collector systems. The first approach is based on a developed statistical clear-sky model, which is used for estimating the clear-sky output solely based on observations of the output. This enables local effects such as shading from trees to be taken into account. The second approach to solar power forecasting is based on conditional parametric modelling. It is well suited for forecasting of solar thermal power, since it can be make non-linear in the inputs. The approach is also extended to a probabilistic solar power forecasting model.

The statistical clear-sky model is furthermore used as basis for a method for correction of global radiation observations. This method can use for correction of typical errors, for example from shading trees or buildings.

Two methods for efficient energy use in buildings are presented in the last part of the thesis. First a method for forecasting of the heat load in single-family houses based on weather forecasts is presented. A model is identified, which works well when applied to forecast the heat load for sixteen single-family houses. The model adapts to the individual houses and needs only no specific information about the buildings. Finally a procedure for identification of a suitable model for the heat dynamics of a building is presented. The applied models are greybox model based on stochastic differential equations and the identification is carried out with likelihood ratio tests. The models can be used for providing detailed information of the thermal characteristics of buildings and as basis for optimal control for exible heating of buildings.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling, Department of Electrical Engineering
Authors: Bacher, P. (Intern), Madsen, H. (Intern), Nielsen, H. A. (Intern)
Number of pages: 188
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PhD-2012
Number: 272
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd272_Peder_Bacher.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2012

Model Predictive Control Algorithms for Pen and Pump Insulin Administration

Despite recent developments within diabetes management such as rapidacting insulin, continuous glucose monitors (CGM) and insulin pumps, tight blood glucose control still remains a challenge. A fully automated closedloop controller, also known as an artificial pancreas (AP), has the potential to ease the life and reduce the risk of acute and chronic diabetic complications. However, the noise associated to CGMs, the long insulin action time for continuous subcutaneous infusion of insulin (CSII) pumps, and the high intra- and inter-patient variability significantly limits the performance of current closed-loop controllers.

In this thesis, we present different control strategies based on Model Predictive Control (MPC) for an artificial pancreas. We use Nonlinear Model Predictive Control (NMPC) in order to determine the optimal insulin and blood glucose profiles. The optimal control problem (OCP) is solved using a multiple-shooting based algorithm. We use an explicit Runge-Kutta method (DOPRI45) with an adaptive stepsize for numerical integration and sensitivity computation. The OCP is solved using a Quasi-Newton sequential quadratic programming (SQP) with a linesearch and a BFGS update for the Hessian of
In addition, we apply a Continuous-Discrete Extended Kalman Filter (CDEKF) in order to simulate cases where the meal size is uncertain, or even unannounced.

We also propose a novel control strategy based on linear MPC for overnight stabilization of blood glucose. The model parameters are personalized using a priori available patient information. We consider an autoregressive integrated moving average with exogenous input (ARIMAX) model. We summarize and discuss the results of the overnight clinical studies conducted at Hvidovre Hospital. Based on these results, we propose improvements for the stochastic part of our controller model. We state and compare three different stochastic model structures. The first one is the ARIMAX structure that has been used for the clinical studies. The second one is an autoregressive moving average with exogenous input (ARMAX) model. The third one is an adaptive ARMAX model in which we estimate the parameters of the stochastic part using a Recursive Least Square (RLS) method. We test the controller in a virtual clinic of 100 patients. This virtual clinic is based on the Hovorka model. We consider the case where only half of the bolus is administrated at mealtime, and the case where the insulin sensitivity increases during the night.

This thesis consists of a summary report, glucose and insulin profiles of the clinical studies and research papers submitted, peer-reviewed and/or published in the period September 2009 - September 2012.

**General information**
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Center for Energy Resources Engineering, Scientific Computing, Dynamical Systems
Authors: Boiroux, D. (Intern), Jørgensen, J. B. (Intern), Poulsen, N. K. (Intern), Madsen, H. (Intern)
Number of pages: 239
Publication date: 2012

**Publication information**
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2012
Number: 283
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd283_Boiroux_D.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2013

**Sensometrics: Thurstonian and Statistical Models**
This thesis is concerned with the development and bridging of Thurstonian and statistical models for sensory discrimination testing as applied in the scientific discipline of sensometrics. In sensory discrimination testing sensory differences between products are detected and quantified by the use of human senses. Thurstonian models provide a stochastic model for the data-generating mechanism through a psychophysical model for the cognitive processes and in addition provides an independent measure for quantification of sensory differences.

In the interest of cost-reduction and health-initiative purposes, much attention is currently given to ingredient substitution. Food and beverage producing companies are consequently applying discrimination testing to control and monitor the sensory properties of evolving products and consumer response to product changes. Discrimination testing is as relevant as ever because it enables more informed decision making in quantifying the degree to which an ingredient substitution is successful and the degree to which the perceptual properties of the product remain unchanged from end user perspectives.

This thesis contributes to the field of sensometrics in general and sensory discrimination testing in particular in a series of papers by advancing Thurstonian models for a range of sensory discrimination protocols in addition to facilitating their application by providing software for fitting these models. The main focus is on identifying Thurstonian models for discrimination methods as versions of well-known statistical models.

The Thurstonian models for a group of discrimination methods leading to binomial responses are shown to be versions of a statistical class of models known as generalized linear models. Thurstonian models for A-not A with sureness and 2-Alternative Choice (2-AC) protocols have been identified as versions of a class of statistical models known as cumulative link models. A theme throughout the contributions has been the development of likelihood methods for computing improved confidence intervals in a range of discrimination methods including the above mentioned methods as well as the same-different test.

A particular analysis with 2-AC data involves comparison with an identifiability norm. For such tests we propose a new test statistic that improves on previously proposed methods of analysis.

In a contribution to the scientific area of computational statistics, it is described how the Laplace approximation can be implemented on a case-by-case basis for flexible estimation of nonlinear mixed effects models with normally distributed response. The two R packages sensR and ordinal implement and support the methodological developments in the
research papers. sensR is a package for sensory discrimination testing with Thurstonian models and ordinal supports analysis of ordinal data with cumulative link (mixed) models. While sensR is closely connected to the sensometrics field, the ordinal package has developed into a generic statistical package applicable to statistical problems far beyond sensometrics. A series of tutorials, user guides and reference manuals accompany these R packages. Finally, a number of chapters provide background theory on the development and computation of Thurstonian models for a range of binomial discrimination protocols, the estimation of generalized linear mixed models, cumulative link models and cumulative link mixed models. The relation between the Wald, likelihood and score statistics is expanded upon using the shape of the (profile) likelihood function as common reference.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Christensen, R. H. B. (Intern), Brockhoff, P. B. (Intern)
Number of pages: 432
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2012
Number: 271
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd271_Rune_Haubo_net.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2012

A Succinct Approach to Static Analysis and Model Checking
In a number of areas software correctness is crucial, therefore it is often desirable to formally verify the presence of various properties or the absence of errors. This thesis presents a framework for concisely expressing static analysis and model checking problems. The framework facilitates rapid prototyping of new analyses and consists of variants of ALFP logic and associated solvers.

First, we present a Lattice based Least Fixed Point Logic (LLFP) that allows interpretations over complete lattices satisfying Ascending Chain Condition. We establish a Moore Family result for LLFP that guarantees that there always is a single best solution for a problem under consideration. We also develop a solving algorithm, based on a dierential worklist, that computes the least solution guaranteed by the Moore Family result.

Furthermore, we present a logic for specifying analysis problems called Layered Fixed Point Logic. Its most prominent feature is the direct support for both inductive computations of behaviors as well as co-inductive specications of properties. Two main theoretical contributions are a Moore Family result and a parametrized worst-case time complexity result. We develop a BDD-based solving algorithm, which computes the least solution guaranteed by the Moore Family result with worst-case time complexity as given by the complexity result.

We also present magic set transformation for ALFP, known from deductive databases, which is a clause-rewriting strategy for optimizing query evaluation. In order to compute the answer to a query, the original ALFP clauses are rewritten at compile time, and then the rewritten clauses are evaluated bottom-up. It is usually more eciently than computing entire solution followed by selection ii of the tuples of interest, which was the case in the classical formulation of ALFP logic.

Finally, we show that the logics and the associated solvers can be used for rapid prototyping. We illustrate that by a variety of case studies from static analysis and model checking.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Language-Based Technology, Computer Science and Engineering
Authors: Filipiuk, P. (Intern), Nielson, H. R. (Intern), Nielson, F. (Intern)
Number of pages: 187
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Modelling spread of Bluetongue and other vector borne diseases in Denmark and evaluation of intervention strategies

The main outcome of this PhD project is a generic model for non-contagious infectious vector-borne disease spread by one vector species between up to two species of hosts distributed on farms and pasture. The model features a within-herd model of disease, combined with a triple movement kernel that describes spread of disease using vectors or hosts as agents of the spread.

The model is run with bluetongue as the primary case study, and it is demonstrated how an epidemic outbreak of bluetongue 8 in Denmark is sensitive to the use of pasture, climate, vaccination, vector abundance, and flying parameters.

In constructing a more process oriented agent-based approach to spread modeling new parameters describing vector behavior were introduced. When these vector flying parameters have been quantified by experiments, this model can be implemented on areas naïve to the modeled disease with a high predictive power.

Furthermore this PhD has provided a new method of estimating the effect of light traps, which can estimate the additive effect of closely placed traps, and determine trap range of individual traps based on emitted light intensity. Moreover there has been devised a method to sample in time which maximizes information about time dependence and is robust to changes.
Distributed security in closed distributed systems

The goal of the present thesis is to discuss, argue and conclude about ways to provide security to the information travelling around computer systems consisting of several known locations.

When developing software systems, security of the information managed by these plays an important role in their design. There should always exist techniques for ensuring that the required security properties are met. This has been thoroughly investigated through the years, and many varied methodologies have come through.

In the case of distributed systems, there are even harder issues to deal with. Many approaches have been taken towards solving security problems, yet many questions remain unanswered. Most of these problems are related to some of the following facts: distributed systems do not usually have any central controller providing security to the entire system; the system heterogeneity is usually reflected in heterogeneous security aims; the software life cycle entails evolution and this includes security expectations; the distribution is useful if the entire system is “open” to new (a priori unknown) interactions; the distribution itself poses intrinsically more complex security-related problems, such as communication, cryptography, performance and reliability. We do not expect to solve all of these, but we shall approach the first three.

In this dissertation, we take the view of a distributed system from a high-level of abstraction. We then focus on the interactions that can take place between the locations, and aim at providing security to each of these individually. The approach taken is by means of access control enforcement mechanisms, providing security to the locations they are related to. We provide a framework for modelling so. All this follows techniques borrowed from the aspect-orientation community.

As this needs to be scaled up to the entire distributed system, we then focus on ways of reasoning about the resulting composition of these individual access control mechanisms. We show how, by means of relying on the semantics of our framework, we can syntactically guarantee some limited set of global security properties. This is also restricted to distributed systems in which the set of locations is known a priori. All this follows techniques borrowed from both the model checking and the static analysis communities.

In the end, we reach a step towards solving the problem of enforcing security in distributed systems. We achieve the goal of showing how this can be done, though we restrict ourselves to closed systems and with a limited set of enforceable security policies. In this setting, our approach proves to be efficient.

Finally, we achieve all this by bringing together several fields of Computer Science. These include aspect orientation, model checking and static analysis, and of course some ingredients of logics and formal methods as well. All this is in an attempt to approach a software engineering problem, such as security in distributed systems. This shows how the full field of Computer Science can benefit from combining its subfields.
Planning and Evaluation of Radio-Therapeutic Treatment of Head-and-Neck Cancer Using PET/CT scanning

Radiation therapy relies in great extent on delineations of tumour and organs on medical images. These delineations are essential for the entire treatment. Unfortunately manual delineations are both prone to variation. At the same time the manual delineation process is time-consuming. This thesis represent a work within the automatic definition of organs and tumours. The thesis includes a summary of the prior methods employed for automatic segmentation and 3 articles describing segmentation algorithms of different areas of application for radiation therapy. Variation within and between manual and automatic segmentation methods is documented in the thesis. The last article of the thesis analyses treatment outcome difference due to manual delineation variation.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics, Department of Electrical Engineering, Biomedical Engineering
Authors: Hollensen, C. (Intern), Larsen, R. (Intern), Højgaard, L. (Intern)
Number of pages: 123
Publication date: 2012

Integration of top-down and bottom-up information for audio organization and retrieval

The increasing availability of digital audio and music calls for methods and systems to analyse and organize these digital objects. This thesis investigates three elements related to such systems focusing on the ability to represent and elicit the user's view on the multimedia object and the system output. The aim is to provide organization and processing, which aligns with the understanding and needs of the users.

Audio and music is often characterized by the large amount of heterogenous information. The rst aspect investigated is the integration of such multi-variate and multi-modal information sources based on latent Dirichlet allocation (LDA). The model is used to integrate bottom-up features (reflecting timbre, loudness, tempo and chroma), meta-data aspects (lyrics) and top-down aspects, namely user generated open vocabulary tags. The model and representation is evaluated on the auxiliary task of genre and style classification.

Eliciting the subjective representation and opinion of users is an important aspect in building personalized systems. The thesis contributes with a setup for modelling and elicitation of preference and other cognitive aspects with focus on audio applications. The setup is based on classical regression and choice models placed in the framework of Gaussian processes, which provides flexible non-parametric Bayesian models. The setup consist of a number of likelihood functions suitable for modelling both absolute ratings (direct scaling) and comparative judgements (indirect scaling). Inference is performed by analytical and simulation based methods, including the Laplace approximation and expectation propagation. In order to minimize the cost of the often expensive and lengthy experimentation, sequential experiment design or active learning is supported. The setup is applied in the eld of music emotion modelling and optimization of a parametric audio system with high-dimensional input spaces.

The final aspect, considered in the thesis, concerns the general context of users, such as location and social context. This is important in understanding user behavior and in determining the users current information needs. The thesis investigates the predictability of the user context, in particular location, based on information theoretic bounds and a particular experimental approach based on context sensing using the ubiquitous mobile phone.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Cognitive Systems
Authors: Jensen, B. S. (Intern), Larsen, J. (Intern), Hansen, L. K. (Intern)
Number of pages: 199
Publication date: 2012
Forecasting and decision-making in electricity markets with focus on wind energy

This thesis deals with analysis, forecasting and decision making in liberalised electricity markets. Particular focus is on wind power, its interaction with the market and the daily decision making of wind power generators. Among recently emerged renewable energy generation technologies, wind power has become the global leader in terms of installed capacity and advancement. This makes wind power an ideal candidate to analyse the impact of growing renewable energy generation capacity on the electricity markets. Furthermore, its present status of a significant supplier of electricity makes derivation of practically applicable tools for decision making highly relevant.

The main characteristics of wind power differ fundamentally from those of conventional thermal power. Its effective generation capacity varies over time and is directly dependent on the weather. This dependency makes future production uncertain and difficult to contract even on a day-to-day basis. Consequently decisions about market bids for next-day delivery are based on production forecasts which are bound to come with some uncertainty. Naturally markets that experience large scale integration of wind power are affected by these different characteristics. The thesis presents analyses of how this impact is realised in markets significantly penetrated by wind power. Due to its representation by forecasts in the supply curve, such predictions are used to describe their non-linear influence on the market prices.

Methods adequately accounting for this effect in models for day-ahead forecasting of the prices are also presented in the thesis. Prompted by the volatile behaviour of electricity markets, considerable focus has been on time-varying and robust parameter estimates. The models derived are all based on well know methods from the statistical literature.

The stochastic production of wind turbines prompts the need for alternative methods for optimally bidding wind power to day-ahead markets. Such bidding strategies are formulated in this thesis, which utilise the information provided by the market models. Bids that maximise expected revenues are found and the possibility of risk averse behaviour is discussed.
This thesis focuses on three subjects within cognitive science related to hearing. Initially, a novel method for automatic speech recognition using binary features from binary masks, is discussed. The performance of binary features in terms of robustness to noise is compared with the ASR state of the art features, mel frequency cepstral coefficients. Secondly, human top-down auditory attention is studied. A computational top-down attention model is presented and behavioral experiments are carried out to investigate the role of top-down task driven attention in the cocktail party problem. Finally, automatic emotion recognition from speech is studied using a dimensional approach and with a focus of integrating semantic and acoustic features. An emotional speech corpus that consists of short movie clips with audio and text parts, rated by human subjects in two affective dimensions (arousal and valence), is prepared to evaluate the method proposed.
Realistic Virtual Cuts

Pigs and pig meat are major sources of income for Denmark. As one of the country's primary exports, it is no wonder that Denmark strives to maintain its competitive edge in the meat market. As part of an on-going effort to lower costs and maintain high standards, X-ray computed tomography (CT), along with image analysis, is being deployed in Danish abattoirs. The data made available from scanning pig carcasses paves the way for new means to optimize the production process.

This thesis is concerned with the development of a communication tool intended to make use of the aforementioned technology in the product prototyping process. In broad terms, the focus can be divided into two areas of focus: visualization and interaction.

Visualizing volume data, obtained via CT-scanning, is a common area of research within other areas of research, e.g. for medical applications. The availability of graphics processing units, and the subsequent programmability of the unit, has allowed for computationally heavy visualization algorithms to execute in real-time. Despite the flexibility of modern GPUs, their architecture still poses problems that require further study. The thesis presents research within the area of texture synthesis and data interpolation in an effort to create even more realistic volume data visualization.

The potential advantages provided by volume data, is exponentially expanded when we are free to interact with it. The food industry sees a significant benefit in volume interaction when concerned with product development. Product earnings projection, product specifications, and interactive training are just a few of the applicable areas. In this thesis we present an interaction method intended for the commercial development of meat product prototypes. The interaction method is evaluated in a thorough usability study with eight volunteer participants from the target user group.

This thesis presents technology and research which, combined with the advent of using CT in the abattoir, paves the way for new possibilities and advantages when designing meat product prototypes. I have no doubt that this is just the tip of the iceberg in regards to modernizing and optimizing the way animal carcasses are processed and handled before becoming consumer goods.

Multivariate image analysis for quality inspection in fish feed production

Aquaculture is today one of the fastest growing food producing sectors in the world. Access to good and effective fish feed is a condition for optimised and sustainable aquaculture activity. In the aquaculture industry it is of utmost importance that...
the fish get feed of proper size and nutrition. The colour appearance of fish products is important for customers. Salmonid fish get their red colour from a natural pigment called astaxanthin. To ensure a similar red colour of fish in aquaculture astaxanthin is used as an additive coated on the feed pellets. Astaxanthin can either be of natural origin, or synthesised chemically. Common for both types is that they are relatively expensive in comparison to the other feed ingredients. This thesis investigates multi-variate data collection for visual inspection and optimisation of industrial production in the fish feed industry. Quality parameters focused on here are: pellet size, type and concentration level of astaxanthin in pellet coating, as well as astaxanthin type detected in salmonid fish. Methods used are three different devices for multi- and hyper-spectral imaging, together with shape analysis and multi-variate statistical analysis. The results of the work demonstrate a high potential of image analysis and spectral imaging for assessing the product quality of fish feed pellets, astaxanthin and fish meat. We show how image analysis can be used to inspect the pellet size, and how spectral imaging can be used to inspect the surface quality of biological materials. This technology and method can be a useful tool for optimising the industrial process, e.g. the utilisation of the expensive astaxanthin. The development of automatic quality inspection methods by machine vision can improve the industry’s position in the competition for high quality products and efficient processes.

General information
State: Published
Organisations: Division of Industrial Food Research, Department of Informatics and Mathematical Modeling, DTU Data Analysis, National Food Institute
Authors: Ljungqvist, M. G. (Intern), Ersbøll, B. K. (Intern), Frosch, S. (Intern)
Number of pages: 250
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2012
Number: 273
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
Multivariate_image_analysis.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2012

System-Level Modeling and Synthesis Techniques for Flow-Based Microfluidic Very Large Scale Integration Biochips
Microfluidic biochips integrate different biochemical analysis functionalities on-chip and offer several advantages over the conventional biochemical laboratories. In this thesis, we focus on the flow-based biochips. The basic building block of such a chip is a valve which can be fabricated at very high densities, e.g., 1 million valves per cm2. By combining these valves, more complex units such as mixers, switches, multiplexers can be built up and the technology is therefore referred to as microfluidic Very Large Scale Integration (mVLSI).

The manufacturing technology for the mVLSI biochips has advanced faster than Moore’s law. However, the design methodologies are still manual and bottom-up. Designers use drawing tools, e.g., AutoCAD, to manually design the chip. In order to run the experiments, applications are manually mapped onto the valves of the chips (analogous to exposure of gate-level details in electronic integrated circuits). Since mVLSI chips can easily have thousands of valves, the manual process can be very time-consuming, error-prone and result in inefficient designs and mappings.

We propose, for the first time to our knowledge, a top-down modeling and synthesis methodology for the mVLSI biochips. We propose a modeling frame-work for the components and the biochip architecture. Using these models, we present an architectural synthesis methodology (covering steps from the schematic design to the physical synthesis), generating an application-specific mVLSI biochip. We also propose a framework for mapping the biochemical applications onto the mVLSI biochips, binding and scheduling the operations and performing fluid routing. A control synthesis framework for determining the exact valve activation sequence required to execute the application is also proposed. In order to reduce the macro-assembly around the chip and enhance chip scalability, we propose an approach for the biochip pin count minimization. We also propose a throughput maximization scheme for the cell culture mVLSI biochips, saving time and reducing costs. We have extensively evaluated the proposed approaches using real-life case studies and synthetic benchmarks. The proposed framework is expected to facilitate programmability and automation, enabling the emergence of a large biochip market.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Embedded Systems Engineering
Authors: Minhass, W. H. (Intern), Pop, P. (Intern), Madsen, J. (Intern)
Number of pages: 115
Wind Turbine Control: Robust Model Based Approach

In the 1970s the oil price crisis encouraged investigation of non-petroleum energy sources of which wind energy was the most promising one. Lately global warming concerns have even intensified the demand for green and sustainable energy resources and opened up several lines of research in this area. Wind turbines are the most common wind energy conversion systems and are hoped to be able to compete economically with fossil fuel power plants in near future. However this demands better technology to reduce the price of electricity production. Control can play an essential part in this context. This is because, on the one hand, control methods can decrease the cost of energy by keeping the turbine close to its maximum efficiency. On the other hand, they can reduce structural fatigue and therefore increase the lifetime of the wind turbine.

The power produced by a wind turbine is proportional to the square of its rotor radius, therefore it seems reasonable to increase the size of the wind turbine in order to capture more power. However as the size increases, the mass of the blades increases by cube of the rotor size. This means in order to keep structural feasibility and mass of the whole structure reasonable, the ratio of mass to size should be reduced. This trend results in more flexible structures.

Control of the flexible structure of a wind turbine in a wind field with stochastic nature is very challenging. In this thesis we are examining a number of robust model based methods for wind turbine control. Firstly we examine potentials of synthesis methods and use tools to analyze robustness of the resulting controllers both in terms of robust stability and robust performance. Afterwards we employ model predictive control (MPC) and show that the way MPC solves control problems suits wind turbine control problems very well, especially when we have preview measurements of wind speed using LIDARs. For the control problem with LIDAR measurements we have proposed a new MPC approach which gives better results than linear MPC while it has almost the same computational complexity. We have also tackled wind turbine control using robust MPC. In general, robust MPC problems are very computationally demanding, however we have shown that with some approximations the resulting robust MPC problem can be specialized with reduced computational complexity.

After a short introduction on wind energy and wind turbines in chapter 1, we briefly explain wind turbine modeling in chapter 2. Introductions to different control design methods are given in chapter 3. The goal of this chapter is to show how different control methods are chosen. The next eight chapters comprise the body of the thesis and are scientific papers that are published or going to be published. Control methods which were briefly introduced in chapter 3 are explained in these chapters in details.
Imaging Food Quality

Imaging and spectroscopy have long been established methods for food quality control both in the laboratories and online. An ever increasing number of analytical techniques are being developed into imaging methods and existing imaging methods to contain spectral information. Images and especially spectral images contain large amounts of data which should be analysed appropriately by techniques combining structure and spectral information.

This dissertation deals with how different types of food quality can be measured by imaging techniques, analysed with appropriate image analysis techniques and finally use the image data to predict or visualise food quality.

A range of different food quality parameters was addressed, i.e. water distribution in bread throughout storage, time series analysis of chocolate milk stability, yoghurt glossiness, graininess and dullness and finally structure and meat colour of dry fermented sausages. The imaging techniques ranged from single wavelength images, multispectral to hyperspectral images. The effect of different light geometries were utilised in measuring the light reflection of yoghurt surfaces.

What the best imaging technique for a given problem is, should be addressed by visually evaluation of a detectable difference between known samples. While doing image analysis, it was found to be advantageous to combine several small models. The combined model was used for extraction of object relevant information, i.e. spectral, texture or size. The data extracted was used for explorative or predictive data analysis.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics
Authors: Møller, F. (Intern), Larsen, R. (Intern), Carstensen, J. M. (Intern)
Number of pages: 136
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2012
Number: 288
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:

Model Based Analysis of Ethnic Differences in Type 2 Diabetes
The present thesis deals with different aspects of population pharmacokinetic/pharmacodynamic (PK/PD) modelling of the glucose homeostatic system. The thesis consist of a summary report and four scientific research papers.

A description of the main topics covered in the thesis is given in the summary report. This includes a short introduction to the mathematical methods applied in the thesis, followed by an outline of the physiological and pathological aspects of the glucose homeostatic system and how to obtain diagnostic indices for characterising the condition of the system. Finally an overview of ethnic differences in type 2 diabetes (T2D) is given, which relates to the subject of the last 2 papers included in the thesis.

One of the main objectives of the thesis was to investigate possible ethnic differences between development of T2D in Caucasian and Japanese and investigate the applicability of stochastic differential equations (SDEs) and non-linear mixed effects (NLME) models for such an assessment. One way to perform such an investigation is to characterise the pathophysiology of the two groups at different stages of disease progression. For T2D this involves a characterisation of the glucose homeostatic system, which is a complex feedback system mainly involving mainly organs such as the liver and the pancreas, the hormones insulin and glucagon, and the carbohydrate glucose.

As for any other dynamical system, a proper characterisation at non-steady state, requires a proper input to the system. This input must reflect the circumstances in which one wants to draw conclusions. In this thesis the intake of oral glucose, which closely resembles the intake of food under daily living has been applied.

Mathematical modelling of such complex physiological phenomena as the glucose homeostatic system will usually be based on both insight into the system and experimental data. Through estimation techniques, free parameters in the models are estimated and can be related directly to behaviour of the system. These semi-physical (grey box) models are well suited for understanding the system, although in many cases they are not able to fully describe the systematic behaviour observed in the applied data sets. This issue can be addressed through an inspection of the autocorrelation function (ACF) of residuals and the description can be improved by switching to the use of stochastic differential equations.
(SDEs) or another improved description of residuals. For characterising disease progression in Caucasian and Japanese, established models that include parameters for insulin sensitivity and beta-cell function were implemented in a non-linear mixed-effects setting with ODEs. Based on the ACF of residuals it was clear that the two models provide a good, although not perfect, description of the systematic variation in the analysed data sets. Based on this the models were extended to SDE models for improved description of residuals. Using the SDE models it was not possible to obtain convergence with the full covariate models so the results presented in the thesis mainly originate from the ODE models. This also caused a more fair comparison with the well-established single-subject models implemented using ODEs.

Previous research have stated the importance of the gut hormone glucagonlikepeptide-1 (GLP-1) as determinant for normal beta-cell function. Based on this a population PK/PD model for secretion of (GLP-1) following an oral glucose tolerance test (OGTT) was developed. This model can be used as a tool to analyse potential differences in the secretion capabilities of GLP-1 between subjects. ACF of residuals did not show any signs of strong serial correlation, and the model was thus not implemented using SDEs.

Assessment of simple and model-based measures for insulin sensitivity and beta-cell function in Japanese and Caucasian subjects stratified according to normal glucose tolerance (NGT), impaired glucose tolerance (IGT), and T2D showed that Japanese in general have higher insulin sensitivity and lower beta-cell function compared to Caucasians. In spite of this, the pattern going from NGT to T2D appeared similar in the two cohorts and the majority of the difference in insulin sensitivity and beta-cell function, measured by simple insulin based measures, could be explained by difference in body size (BMI). This was supported by Forest plots of covariate effects obtained from population models, in general indicating that race had no clinical relevant effect on either the insulin sensitivity or the beta-cell function when measures for obesity (android fat mass or BMI) was taken into account.

### General information

**State:** Published  
**Organisations:** Mathematical Statistics, Department of Informatics and Mathematical Modeling  
**Authors:** Möller, J. B. (Intern), Madsen, H. (Intern)  
**Number of pages:** 82  
**Publication date:** 2012

### Publication information

**Place of publication:** Kgs. Lyngby, Denmark  
**Publisher:** Technical University of Denmark (DTU)  
**Original language:** English  
**Series:** IMM-PHD-2012  
**Number:** 268  
**ISSN:** 0909-3192  
**Main Research Area:** Technical/natural sciences  
**Pharmacokinetic/pharmacodynamic (PK/PD), Type 2 diabetes (T2D), Autocorrelation function (ACF), Stochastic dieriential equations (SDEs), Oral glucose tolerance test (OGTT), Glucagon-like-peptide 1 (GLP-1), Disease progression, Ethnic dierences**  
**Electronic versions:** phd268_Jonas_Bech_Moller.pdf  
**Source-ID:** 318951  
**Publication:** Research › Ph.D. thesis – Annual report year: 2012

### Temporal constraints on visual perception: A psychophysical investigation of the relation between attention capture and the attentional blink

While the richness of our visual perceptions is nearly boundless, the rate with which we can perceive information is limited. For instance when we are required to perceive two consecutive target objects following briefly after each other, the accuracy with which we can report the second target is often reduced in the first half second. This phenomenon is known as the attentional blink (Raymond, Shapiro & Arnell, 1992) and as suggests by the name is assumed to pertain to how fast attention can be reallocated. Bottleneck models suggest that the attentional blink is caused by limited capacity in processing targets, which effectively causes a perceptual bottleneck (Chun & Potter, 1995). According to bottleneck models, making the first target easier to perceive should improve processing in the bottleneck and reduce the attentional blink. However, recent studies suggest that an attentional blink may be triggered by attention capture to the first object (Folk, Leber & Egeth, 2008) and that if making the first target easier to perceive increase its saliency this may increase the attentional blink (Chua, 2005).

This thesis examines the attention capture hypothesis with focus on empirical investigations and a theoretical review. Specifically this work presents studies in which first target contrast is varied in two different attentional blink paradigms, while potential influences from bottleneck effects are controlled. Publication 1 describes findings using the two-target paradigm (Duncan, Ward & Shapiro, 1994) where two masked targets are presented in different locations. Here we find that the attentional blink increases with first target contrast, however, only when no mask follows the first target. To further examine the effect of first target contrast, we disentangle the potential influence of bottleneck effects and vary first target contrast while maintaining target difficulty constant. Again we find that first target contrast increases the attention blink.
Publication describes finding using the rapid serial visual presentation paradigm (Potter & Levy, 1969), in which two targets are presented centrally in the same location embedded in a stream of distractor objects. These findings replicate those from Publication 1, and suggest that the effect is not entirely spatial, since the rapid serial visual presentation paradigm does not require a spatial shift of attention to a new location. In addition to the findings in Publication 1, Publication 2 shows that the effect of first target contrast can be cancelled by the opposing effect of second target contrast.

Thus the results presented here are consistent with an attention capture hypothesis and suggest that the first target can trigger an attentional blink, and that the size of the blink increases with first target contrast.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Nielsen, S. (Intern), Andersen, T. (Intern), Hansen, L. K. (Intern)
Number of pages: 44
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2012
Number: 279
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd279_SimonNielsen.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2012

Attacker Modelling in Ubiquitous Computing Systems
Within the last five to ten years we have experienced an incredible growth of ubiquitous technologies which has allowed for improvements in several areas, including energy distribution and management, health care services, border surveillance, secure monitoring and management of buildings, localisation services and many others. These technologies can be classified under the name of ubiquitous systems.

The term Ubiquitous System dates back to 1991 when Mark Weiser at Xerox PARC Lab first referred to it in writing. He envisioned a future where computing technologies would have been melted in with our everyday life. This future is visible to everyone nowadays: terms like smartphone, cloud, sensor, network etc. are widely known and used in our everyday life.

But what about the security of such systems. Ubiquitous computing devices can be limited in terms of energy, computing power and memory. The implementation of cryptographic mechanisms that comes from classical communication systems could be too heavy for the resources of such devices, thus forcing the use of lighter security measures if any at all. The same goes for the implementation of security protocols. The protocols employed in classical communication systems were not designed for the ubiquitous environment, hence their security has to be proven again, leading to the definition of new protocols designed specifically to address new vulnerabilities introduced by the ubiquitous nature of the system.

Throughout the network security community this problem has been investigated for some time now and this has resulted in some lightweight cryptographic standards and protocols, as well as tools that make it possible for security properties of communication protocols which are typical of ubiquitous systems. However the abilities of the ubiquitous attacker remain somehow undefined and still under extensive investigation.

This Thesis explores the nature of the ubiquitous attacker with a focus on how she interacts with the physical world and it defines a model that captures the abilities of the attacker. Furthermore a quantitative implementation of the model is presented. This can be used by a security analyst as a supporting tool to analyse the security of an ubiquitous system and identify its weak parts. Most importantly this work is also useful for system designers, who wish to implement an effective secure solution while developing their system.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science , Embedded Systems Engineering
Authors: Papini, D. (Intern), Sharp, R. (Intern), Jensen, C. D. (Intern)
Number of pages: 182
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2012
Functional Brain Imaging by EEG: A Window to the Human Mind

This thesis presents electroencephalography (EEG) brain imaging by covering topics as empirical evaluation of source confusion, probabilistic inverse methods, and source analysis performed on infant EEG data. In terms of source confusion we inspect how current sources within the brain may be confused with each other as noise is present in the EEG recordings. Moreover, we examine how errors in the forward model affect the source confusion.

The primary aim of this thesis is to provide sharper EEG brain images by improving current inverse methods. In this relation we focus the attention on two topics in EEG source reconstruction, namely, the forward propagation model (describing the mapping from the current sources within the brain to the sensors at the scalp) and the temporal patterns present in the EEG.

As forward models may suffer from a number of errors including the geometrical representation of the human head, the tissue conductivity distribution, and electrode positions, we propose an algorithm which consider forward model uncertainties. Bayesian graphical models provide a powerful means of incorporating prior assumptions that narrow the solution space and lead to tractable posterior distributions over the unknown sources given the observed data. Here, we propose a hierarchical Bayesian model that attempts to minimize the influence of uncertainties associated with the forward model on the source estimates.

Similarly, we develop a hierarchical spatio-temporal Bayesian model that accommodates the principled computation of sparse spatial and smooth temporal EEG source reconstructions consistent with neurophysiological assumptions in a variety of event-related imaging paradigms.

Logics and Models for Stochastic Analysis Beyond Markov Chains

Within the last twenty years, logics and models for stochastic analysis of information systems have been widely studied in both theory and practice. The quantitative properties, such as performance and reliability, are evaluated over discrete–time and continuous–time Markov chains. This thesis lifts the stochastic analysis techniques from the class of Markov chains to the more general classes of stochastic processes having PHase--type (PH) distributions and Matrix–Exponential (ME) distributions, such that a Markov renewal process with ME kernels that cannot be formulated as a Markov process with finite or countable state space.

PH distributions are known for many explicit analytic properties, such that systems having PH distributed components can still be formulated as Markov chains. This thesis presents several results related to PH distributions. We first show how to use the explicit analytic form of discrete PH distributions as computational vehicle on measuring the performance of concurrent wireless sensor networks. Secondly, choosing stochastic process algebras as a widely accepted formalism, we study the compositionality of continuous PH distributions in order to support modelling concurrent stochastic systems having PH representations as building blocks. At last, we consider discrete–time point processes having PH distributed
interarrival times with multiple marks, we propose time-lapse bisimulation, a state-based characterisation of the equivalence relation between the point processes. We clarify that time-lapse bisimulation is a new contribution to the existing bisimulation family, which captures probabilistic behaviour over time for labelled discrete–time Markov chains.

ME distributions is a strictly larger class than PH distributions, such that many results from PH distributions also are valid for ME distributions. ME distributions have a very appealing property, called minimality property: generally a ME representation of a PH distribution will be of lower dimension than PH representations, and one can always find a ME representation with the minimal dimension. However, because of the generality of ME distributions, we have to leave the world of Markov chains. To support ME distributions with multiple exits, we introduce a multi-exits ME distribution together with a process algebra MEME to express the systems having the semantics as Markov renewal processes with ME kernels. The most appealing feature is that all the components before and after compositions are secured to have a minimal state space representation. To support quantitative verification, we also propose stochastic model checking algorithms to our problem.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Language-Based Technology, Statistics and Data Analysis
Authors: Zeng, K. (Intern), Nielsen, B. F. (Intern), Nielsen, F. (Intern)
Number of pages: 132
Publication date: 2012

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2012
Number: 282
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd282_Kebin_Zeng.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2013

Model Checking as Static Analysis
Both model checking and static analysis are prominent approaches to detecting software errors. Model Checking is a successful formal method for verifying properties specified in temporal logics with respect to transition systems. Static analysis is also a powerful method for validating program properties which can predict safe approximations to program behaviors. In this thesis, we have developed several static analysis based techniques to solve model checking problems, aiming at showing the link between static analysis and model checking.

We focus on logical approaches to static analysis. Alternation-free Least Fixed Point Logic (ALFP), an extension of Datalog, has been used as the specification language in most of our research results.

We have first considered the CTL model checking and developed an ALFP-based technique to solve the CTL model checking problem. We have shown that the set of states satisfying a CTL formula can be characterized as the least model of ALFP clauses specifying this CTL formula. The existence of the least model of ALFP clauses is ensured by the Moore Family property of ALFP. Then, we take fairness assumptions in CTL into consideration and have shown that CTL fairness problems can be encoded into ALFP as well.

To deal with multi-valued model checking problems, we have proposed multivalued ALFP. A Moore Family result for multi-valued ALFP is also established, which ensures the existence and uniqueness of the least model. When the truth values in multi-valued ALFP constitute a finite distributive complete lattice, multi-valued ALFP can be reduced to two-valued ALFP. This result enables to implement a solver for multi-valued ALFP by reusing existing solvers for two-valued ALFP. Our ALFP-based technique developed for the two-valued CTL naturally generalizes to a multi-valued setting, and we therefore obtain a multivalued analysis for temporal properties specified by CTL formulas. In particular, we have shown that the three-valued CTL model checking problem over Kripke modal transition systems can be exactly encoded in three-valued ALFP.

Last, we come back to two-valued settings and have considered the model checking for the modal μ-calculus. Our results have shown that ALFP suces to deal with the model checking problem for the alternation-free μ-calculus. However, to deal with the full fragment of the μ-calculus, we need to go beyond ALFP. Therefore, we proposed Succinct Fixed Point Logic (SFP), as an extension of ALFP. We have established a Moore Family result for SFP, which ensures the existence and uniqueness of the intended model of SFP. We have shown that SFP is well suited to specify nested xed points in the μ-
calculus and the model checking problem for the $\mu$-calculus can be encoded as the intended model of SFP.

Our research results have strengthened the link between model checking and static analysis. This provides a theoretical foundation for developing a unified tool for both model checking and static analysis techniques.

**General information**

State: Published
Organisations: Department of Informatics and Mathematical Modeling, Language-Based Technology, Computer Science and Engineering
Authors: Zhang, F. (Intern), Nielson, F. (Intern), Nielson, H. R. (Intern)
Number of pages: 174
Publication date: 2012

**Publication information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2012
Number: 280
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:

phd280_Fuyuan_Zhang.pdf

Publication: Research › Ph.D. thesis – Annual report year: 2012

**3D Shape Modeling Using High Level Descriptors**

The goal of this Ph.D. project is to investigate and improve the methods for describing the surface of 3D objects, with focus on modeling geometric texture on surfaces. Surface modeling being a large field of research, the work done during this project concentrated around a few smaller areas corresponding to the research papers presented here. One of those areas is formulating surface priors by utilizing local surface properties. A well defined prior can, in a Bayesian framework, assist many common task in geometry processing, like denoising, object recovery, object matching and classification. Some of the priors described here are defined on the main entities of the triangular mesh, vertices, edges and faces. Other priors are defined on small planar patches, denoted surfels. Another area of research deals with textures which cannot be described by height fields, for example biological features like thorns, bark and scales. Presented here is a simple method for easy modeling, transferring and editing that kind of texture. The method is an extension of the height-field texture, but incorporates an additional tilt of the height field. Related to modeling non-heightfield textures, a part of my work involved developing feature-aware resizing of models with complex surfaces consisting of underlying shape and a distinctive texture detail. The aim was to deform an object while preserving the shape and size of the features.

**General information**

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Andersen, V. (Intern), Aanæs, H. (Intern), Bærentzen, J. A. (Intern)
Number of pages: 139
Publication date: 2011

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark, DTU Informatics, Building 321
Original language: English

Series: IMM-PHD-2011
Number: 233
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:

phd233_va-m.pdf

Source: orbit
Source-ID: 263603
Publication: Research › Ph.D. thesis – Annual report year: 2012
Hyperspectral Image Analysis of Food Quality
Assessing the quality of food is a vital step in any food processing line to ensure the best food quality and maximum profit for the farmer and food manufacturer. Traditional quality evaluation methods are often destructive and labour-intensive procedures relying on wet chemistry or subjective human inspection. Near-infrared spectroscopy can address these issues by offering a fast and objective analysis of the food quality. A natural extension to these single spectrum NIR systems is to include image information such that each pixel holds a NIR spectrum.

This augmented image information offers several extensions to the analysis of food quality. This dissertation is concerned with hyperspectral image analysis used to assess the quality of single grain kernels. The focus is to highlight the benefits and challenges of using hyperspectral imaging for food quality presented in two research directions. Initially, the visualisation and interpretation of hyperspectral images are discussed. A Bayesian based unmixing method is presented as a novel approach to decompose a hyperspectral image into interpretable components. Secondly, hyperspectral imaging is applied to a dedicated application of predicting the degree of pre-germination in single barley kernels using a customised classification framework. Both contributions serve to illustrate the improvement of adding image information to NIR systems in food quality assessment applications.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics
Authors: Arngren, M. (Intern), Larsen, J. (Intern), Larsen, R. (Intern)
Number of pages: 156
Publication date: 2011

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
ISBN (Print): 978-87-643-0821-1
Original language: English
Series: IMM-PHD-2011
Number: 255
Main Research Area: Technical/natural sciences
Electronic versions: phd255_ma.pdf
Source: orbit
Source-ID: 276974
Publication: Research › Ph.D. thesis – Annual report year: 2012

A Bio-Inspired Self-Healing Reconfigurable Hardware Architecture: Concept, design, prototype, and evaluation

General information
State: Published
Organisations: Embedded Systems Engineering, Department of Informatics and Mathematical Modeling
Authors: Boesen, M. R. (Intern), Madsen, J. (Intern), Pop, P. (Intern)
Number of pages: 260
Publication date: 2011

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD
Number: 259
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 277783
Publication: Research › Ph.D. thesis – Annual report year: 2011

New vision technology for multidimensional quality monitoring of food processes
Spectroscopy and spectral imaging in combination with multivariate data analysis and machine learning techniques have proven to be an outstanding tool for rapid analysis of different products. This may be utilized in various industries, but especially rapid assessment of food products in food research and industry is of importance in this thesis. The non-invasive spectroscopic imaging techniques are able to measure individual food components simultaneously in situ in the food matrix while pattern recognition techniques effectively are able to extract the quantitative information from the vast
data amounts collected. Underlying qualitative features (latent structures) are extracted from multivariate spectral data in order to quantify desired quality parameters properly. Specifically multispectral imaging which has been explored to a lesser extent than ordinary spectroscopy, having the possibility to exploit the inherent heterogeneity that exists in foodstuffs have been investigated here. An extra feature obtained by combining spectroscopy, imaging and chemometrics is exploratory analysis. This is central in food research, since novel hypotheses about the food systems under observation may be generated using this inductive analytical approach. For the food industry it is an additional advantage that the fast, non-invasive, remote sensing nature of the spectroscopic imaging methods allows on-line measurements. In this way spectroscopic imaging in combination with advanced data analysis meets the high throughput needs for quality control, process control and monitoring. In this Ph.D. project the possibilities provided by spectroscopic imaging and chemometrics have been utilized to improve the analysis and understanding of different food products. The work is presented in seven papers and two additional technical reports which make up the core of the thesis. Furthermore an introduction together with a linking of the contributions is presented in this thesis. The papers puts an emphasis on the use of multispectral imaging in the baking industry where especially the non-enzymatic browning appearance and features related to this are highlighted. These are features such as colour, water content and internal structure of bread. A paper presenting enzymatic browning in pre stir fried and thawed vegetables is also presented showing that imaging techniques such as the one investigated in this thesis is able to detect even subtle colour changes. The possibility for quantifying early as well as late spoilage in raw pork meat is investigated where use of the heterogenetic structure is utilized to obtain good results on predicting sensory evaluations as well on laboratory analysis. Colour in other settings such as in the shery industry is equally important, and a paper describing detection of carotenoid pigment in trouts using spectral images shows promising results. Finally, two technical papers present possible ways of mapping multispectral images to a visible colour space, as well as how an alternative multispectral imaging system, making use of filters, may be used to design new more broad ranged filters. Fewer filters will increase the speed of such systems. Methods for solving such problems is to the knowledge of the authors rarely covered in the literature.

**General information**
State: Published
Organisations: DTU Data Analysis, Department of Informatics and Mathematical Modeling, National Food Institute
Authors: Dissing, B. S. (Intern), Ersbøll, B. K. (Intern), Adler-Nissen, J. (Intern)
Number of pages: 194
Publication date: 2011

**Publication information**
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

**Series**: IMM-PHD-2011
**Number**: 256
**Main Research Area**: Technical/natural sciences

Electronic versions:
phd256_bdi.pdf
Source: orbit
Source-ID: 276639
Publication: Research › Ph.D. thesis – Annual report year: 2011

**Maximum likelihood estimation of phase-type distributions**
This work is concerned with the statistical inference of phase-type distributions and the analysis of distributions with rational Laplace transform, known as matrix-exponential distributions. The thesis is focused on the estimation of the maximum likelihood parameters of phase-type distributions for both univariate and multivariate cases. Methods like the EM algorithm and Markov chain Monte Carlo are applied for this purpose. Furthermore, this thesis provides explicit formulae for computing the Fisher information matrix for discrete and continuous phase-type distributions, which is needed to find confidence regions for their estimated parameters. Finally, a new general class of distributions, called bilateral matrix-exponential distributions, is defined. These distributions have the entire real line as domain and can be used, for instance, for modelling. In addition, this class of distributions represents a generalization of the class of matrix-exponential distributions.

**General information**
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Esparza, L. J. R. (Intern), Nielsen, B. F. (Intern), Bladt, M. (Intern)
Publication date: 2011

**Publication information**
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Sparse Multivariate Modeling: Priors and Applications

This thesis presents a collection of statistical models that attempt to take advantage of every piece of prior knowledge available to provide the models with as much structure as possible. The main motivation for introducing these models is interpretability since in practice we want to be able to use them as hypothesis generating tools. All of our models start from a family of structures, for instance factor models, directed acyclic graphs, classifiers, etc. Then we let them be selectively sparse as a way to provide them with structural flexibility and interpretability. Finally, we complement them with different prior assumptions in order to make them appropriate at handling different domain specific situations as time series, non-linearities, batch effects, missing values, etc. In particular, we present a framework for linear Bayesian networks we call sparse identifiable multivariate modeling, a model for peptide-protein/protein-protein interactions called latent protein tree, a framework for sparse Gaussian process classification based on active set selection and a linear multi-category sparse classifier specially targeted to gene expression data. The thesis is organized to provide a general yet self-contained description of every model in terms of generative assumptions, interpretability goals, probabilistic formulation and target applications. Case studies, benchmark results and practical details are also provided as appendices published elsewhere, containing reprints of peer reviewed material.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Henao, R. (Intern), Winther, O. (Intern)
Publication date: 2011

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Model Predictive Control of Wind Turbines

Wind turbines play a major role in the transformation from a fossil fuel based energy production to a more sustainable production of energy. Total-cost-of-ownership is an important parameter when investors decide in which energy technology they should place their capital. Modern wind turbines are controlled by pitching the blades and by controlling the electro-magnetic torque of the generator, thus slowing the rotation of the blades. Improved control of wind turbines, leading to reduced fatigue loads, can be exploited by using less materials in the construction of the wind turbine or by reducing the need for maintenance of the wind turbine. Either way, better total-cost-of-ownership for wind turbine operators can be achieved by improved control of the wind turbines. Wind turbine control can be improved in two ways, by improving the model on which the controller bases its design or by improving the actual control algorithm. Both possibilities have been investigated in this thesis. The level of modeling detail has been expanded as dynamic infl ow has been incorporated into the control design model where state-of-the-art controllers usually assume quasi-steady aerodynamics. Floating wind turbines have been suggested as an alternative to ground-fixed wind turbines as they can be placed at water depths usually thought outside the realm of wind turbine placement. The special challenges posed by controlling a floating wind turbine have been addressed in this thesis. Model predictive control (MPC) has been the foundation on which the control algorithms have been build. Three controllers are presented in the thesis. The first is based on four different linear model predictive controllers where appropriate switching conditions determine which controller is active. Constraint handling of actuator states such as pitch angle, pitch rate and pitch acceleration is the primary focus of this controller. The wind turbine is a highly nonlinear plant and a gain scheduling or relinearizing model predictive controller forms the next step to improve performance compared to a linear controller. Finally, a nonlinear model predictive controller has been devised and tested under simplified conditions. At present, the nonlinear model predictive controller is however not expected to be an realistic option for real world application as the computation burden is to heavy to achieve real-time performance. This thesis is comprised of a collection scientific papers dealing with the various topics presented in this
Integrating Design Decision Management with Model-based Software Development

Design decisions are continuously made during the development of software systems and are important artifacts for
design documentation. Dedicated decision management systems are often used to capture such design knowledge. Most
such systems are, however, separated from the design artifacts of the system. In model-based software development,
where design models are used to develop a software system, outcomes of many design decisions have big impact on
design models. The realization of design decisions is often manual and tedious work on design models. Moreover,
keeping design models consistent with all made decisions is a time-consuming manual task that is often performed in peer
reviews. In this thesis, a generic technology has been developed for extracting model differences from models and
transferring them to other models. These concepts, called model-independent differences, can be used to specify
realizations of decisions in design models. This way, recurring realization work of design decisions can be automated.
Since the concepts are generic and not bound to design decisions, other recurring work on models can be automated as
well, for instance, design patterns and refactorings. With such a technology at hand, design decision realizations can
easily be specified and parts of the realization work can be automated. A binding is produced as a by-product that links
documented decision outcomes to design model elements which are affected by the respective decisions. With a set of
constraints, such a binding can be used to validate the consistency between the design and made design decisions.
Whenever the evolving design models become inconsistent with realized decisions, developers are notified about the
violations. The violations can be fixed by correcting the design, adjusting the binding, or by ignoring the causes. This
substitutes manual reviews to some extent. The concepts, implemented in a tool, have been validated with design
patterns, refactorings, and domain level tests that comprise a replay of a real project. This proves the applicability of the
solution to realistic examples. The implementation of model-independent differences, called MPatch, is further contributed to the Eclipse open source project.

**General information**

State: Published
Organisations: Software Engineering, Department of Informatics and Mathematical Modeling
Authors: Könemann, P. (Intern), Kindler, E. (Intern)
Publication date: 2011

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2011-249
Main Research Area: Technical/natural sciences
Electronic versions:

phd249_pk-net.pdf
Source: orbit
Source-ID: 274577
Publication: Research › Ph.D. thesis – Annual report year: 2011

---

**Efficient Unbiased Rendering using Enlightened Local Path Sampling**

Most global illumination algorithms today solve the light transport problem using Monte Carlo ray tracing. These algorithms are capable of producing photo-realistic imagery and in addition have few limitations with respect to the kind of input (geometry, reflection models, etc.) they support. The downside to using these algorithms is that they can be slow to converge. Due to the nature of Monte Carlo methods, the results are random variables subject to variance. This manifests itself as noise in the images, which can only be reduced by generating more samples. The reason these methods are slow is because of a lack of effective methods of importance sampling. Most global illumination algorithms are based on local path sampling, which is essentially a recipe for constructing random walks. Using this procedure paths are built based on information given explicitly as part of scene description, such as the location of the light sources or cameras, or the reflection models at each point. In this work we explore new methods of importance sampling. Our idea is to analyze the scene before rendering and compute various statistics that we use to improve importance sampling. The first of these are adjoint measurements, which are the solution to the adjoint light transport problem. The second is a representation of the distribution of radiance and importance in the scene. We also derive a new method of particle sampling, which is advantageous compared to existing methods. Together we call the resulting algorithm enlightened local path sampling and demonstrate how the algorithm improves efficiency in some hard scenes.

**General information**

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Kristensen, A. W. (Intern), Christensen, N. J. (Intern)
Number of pages: 188
Publication date: 2011

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2010-240
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:

phd240_awk_vers2.pdf
Source: orbit
Source-ID: 268337
Publication: Research › Ph.D. thesis – Annual report year: 2010

---

**Feedback Driven Annotation and Refactoring of Parallel Programs**

This thesis combines programmer knowledge and feedback to improve modeling and optimization of software. The research is motivated by two observations. First, there is a great need for automatic analysis of software for embedded systems - to expose and model parallelism inherent in programs. Second, some program properties are beyond reach of such analysis for theoretical and practical reasons - but can be described by programmers. Three aspects are explored. The first is annotation of the source code. Two annotations are introduced. These allow more accurate modeling of
parallelism and communication in embedded programs. Runtime checks are developed to ensure that annotations correctly describe observable program behavior. The performance impact of runtime checking is evaluated on several benchmark kernels and is negligible in all cases. The second aspect is compilation feedback. Annotations are not effective unless programmers are told how and when they are beneficial. A prototype compilation feedback system was developed in collaboration with IBM Haifa Research Labs. It reports issues that prevent further analysis to the programmer. Performance evaluation shows that three programs perform significantly faster - up to 12.5x - after modification directed by the compilation feedback system. The last aspect is refinement of compilation feedback. Out of numerous issues reported, few are important to solve. Different compilers and compilation flags are used to estimate whether an issue can be resolved or not. On average, 43% of the issues reported can be categorized as potentially resolvable (27%) or unresolvable (15%).

General information
State: Published
Organisations: Embedded Systems Engineering, Department of Informatics and Mathematical Modeling
Authors: Larsen, P. (Intern), Karlsson, S. (Intern), Madsen, J. (Intern)
Number of pages: 163
Publication date: 2011

Power and Thermal Management of System-on-Chip
With greater integration of VLSI circuits, power consumption and power density have increased dramatically resulting in high chip temperatures and presenting a heat removal challenge. To effectively limit the high temperature inside a chip, thermal specific approaches, besides low power techniques, are necessary at the chip design level. In this work, we investigate the power and thermal management of System-on-Chips (SoCs). Thermal analysis is performed in a SPICE simulation approach based on the electrical-thermal analogy. We investigate the impact of inter-connects on heat distribution in the substrate and present a way to consider temperature dependent signal delay in global wires at early design stages. With the aim of reducing high local power density in hotspots, we propose two placement techniques to spread hot cells over a larger area. The proposed methods are compared in terms of temperature reduction, timing and area overhead to the general method, which enlarges the circuit area uniformly. A case study analyzes the design of Floating Point Units (FPU) from an energy and a thermal perspective. For the division operation, we compare different implementations and illustrate the impact of power efficient dividers on the energy consumption and thermal distribution within the FPU and the on-chip cache. We also characterize the temperature dependent static dissipation to evaluate the reduction in leakage obtained from the decrease in temperature.

General information
State: Published
Organisations: Embedded Systems Engineering, Department of Informatics and Mathematical Modeling
Authors: Liu, W. (Intern), Nannarelli, A. (Intern)
Publication date: 2011
Synthesis of Digital Microfluidic Biochips with Reconfigurable Operation Execution

Microfluidic biochips are an alternative to conventional biochemical laboratories, and are able to integrate on-chip all the necessary functions for biochemical analysis. The "digital" biochips are manipulating liquids not as a continuous flow, but as discrete droplets on a two-dimensional array of electrodes. The main objective of this thesis is to develop top-down synthesis techniques for digital microfluidic biochips. So far, researchers have assumed that operations are executing on virtual modules of rectangular shape, formed by grouping adjacent electrodes, and which have a fixed placement on the microfluidic array. However, operations can actually execute by routing the droplets on any sequence of electrodes on the biochip. Thus, we have proposed a routing-based model of operation execution, and we have developed several associated synthesis approaches, which progressively relax the assumption that operations execute inside fixed rectangular modules. The proposed synthesis approaches consider that i) modules can dynamically move during their execution and ii) can have non-rectangular shapes. iii) We have relaxed the assumption that all electrodes are occupied during the operation execution, by taking into account the position of droplets inside modules. Finally, iv) we have eliminated the concept of virtual modules and have allowed the droplets to move on the chip on any route. In this context, we have also shown how contamination can be avoided. We have extensively evaluated the proposed approaches using several real-life case studies and synthetic benchmarks. The experiments show that by considering the dynamically reconfigurable nature of microfluidic operations, significant improvements can be obtained, decreasing the biochemical application completion times, reducing thus the biochip area and implementation costs.

General information
State: Published
Organisations: Embedded Systems Engineering, Department of Informatics and Mathematical Modeling
Authors: Maftei, E. (Intern), Pop, P. (Intern), Madsen, J. (Intern)
Number of pages: 123
Publication date: 2011

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2011-257
Main Research Area: Technical/natural sciences
Electronic versions:
phd257_Maftei_E.pdf
Source: orbit
Source-ID: 276602
Publication: Research › Ph.D. thesis – Annual report year: 2011

Stochastic State Space Modelling of Nonlinear systems - With application to Marine Ecosystems

This thesis deals with stochastic dynamical systems in discrete and continuous time. Traditionally dynamical systems in continuous time are modelled using Ordinary Differential Equations (ODEs). Even the most complex system of ODEs will not be able to capture every detail of a complex system like a natural ecosystem, and hence residual variation between the model and observations will always remain. In stochastic state-space models the residual variation is separated into observation and system noise and a main theme of the thesis is a proper description of the system noise. Additive Gaussian noise is the standard approach to introduce system noise, but this may lead to undesirable consequences for the state variables. In biological models, where the statespace generally contains positive real numbers only, modelling in the log-domain ensures positive state variables, however, this transformation is likely to conflict with the concept of mass balances. One of the central conclusions of the thesis is that the stochastic formulations should be an integral part of the model formulation. As discrete-time stochastic processes are simpler to handle numerically than continuous-time stochastic processes, I start by considering discrete-time processes. An novel approach combining multiplicative and additive log-normal noise has been developed in discrete time, and used to demonstrate the effect of stochastic forcing in simple discrete-time regime shift models. An approximate maximum likelihood estimation procedure based on the second order moment representation of the multiplicative and additive log-normal noise model was developed and tested in simulation studies. The transition to continuous-time stochastic models (here Stochastic Differential Equations (SDEs)) offers the opportunity of embedding parts of the ODE processes into the stochastic part of the model (the diffusion term). The estimation method we use here (maximum likelihood and the Extended Kalman Filter (EKF)) rely on state-independent diffusion, but for a wide class of SDEs there exist an alternative description (given by the Lamperti transform) of the input-output relation, where the diffusion term is independent of the state. This alternative description is used to develop better parametric descriptions of the diffusion term, while maintaining the opportunity of estimation by standard software. Additionally, the state-space formulation facilitates estimation of unobserved states. Based on estimation of random walk hidden states and examination of simulated distributions and stationarily characteristics, a methodological framework for structural identification based on information embedded in the observations of the system has been developed. The applicability of the methodology is demonstrated using phytoplankton and nitrogen data from a Danish estuary as well as bacterial growth data from a controlled experiment. In summary, the novelty of the work presented here
is the introduction of more appropriate stochastic descriptions in non-linear state-space models, which can include combinations of additive and multiplicative noise components under various distributional assumptions. A model identification and estimation framework for working with such models has been developed and tested using data from biological and ecological systems typically characterised by non-linear and non-Gaussian responses.

**General information**

State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling, Department of Environmental Science and Engineering
Authors: Møller, J. K. (Intern), Carstensen, N. J. (Intern), Madsen, H. (Intern)
Publication date: 2011

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2010
Number: 246
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd246-jkm.pdf
Source: orbit
Source-ID: 271398
Publication: Research › Ph.D. thesis – Annual report year: 2012

**Computed Tomography in the Modern Slaughterhouse**

The Danish pig meat industry has been seeing a growing international competition in the past years. In the quest to maintain both competitive prices and high product standards in spite of the higher Danish factor costs, a substantial effort is being put into innovation, research and development of technology. Recently, the use of X-ray computed tomography (CT) coupled with methods from image analysis has been introduced as a powerful means to optimise production, by providing detailed information on the raw materials. This thesis covers two aspects of the application of CT in the modern abattoir. In the first aspect we use CT to analyse the biological diversity of carcass populations. The images form the basis for a data-driven tissue deformation model. The results provide valuable input to assist the development of an automated robotic tool for trimming the rind off pig backs. The second aspect concerns measurements of each single carcass, to improve the raw material utilisation by individually adapted processing. Measurements performed online in the abattoir demand fast, robust and cost-effective imaging. We propose a tomographic reconstruction algorithm, enabling a substantial reduction of the subject-specific X-ray data needed to produce high quality images for accurate measurements. This is very beneficial for the abattoirs, as a reduction in acquired data translates directly into higher speed and a lower cost. The thesis demonstrates the great potential of CT as a technology for improving the yield of the Danish pig meat industry. An introduction of efficient online CT will especially open a vast number of possibilities for optimising the production.

**General information**

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, DTU Data Analysis
Authors: Mosbech, T. H. (Intern), Ersbøll, B. K. (Intern), Larsen, R. (Intern)
Number of pages: 147
Publication date: 2011

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2011
Number: 258
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions: phd258_tm.pdf
Source: orbit
Source-ID: 276481
Markerless 3D Head Tracking for Motion Correction in High Resolution PET Brain Imaging

This thesis concerns application specific 3D head tracking. The purpose is to improve motion correction in position emission tomography (PET) brain imaging through development of markerless tracking. Currently, motion correction strategies are based on either the PET data itself or tracking devices relying on markers. Data-driven motion correction is problematic due to the physiological dynamics. Marker-based tracking is potentially unreliable, and it is extremely hard to validate when the tracking information is correct. The motion estimation is essential for proper motion correction of the PET images. Incorrect motion correction can in the worst cases result in wrong diagnosis or treatment. The evolution of a markerless custom-made structured light 3D surface tracking system is presented. The system is targeted at state-of-the-art high resolution dedicated brain PET scanners with a resolution of a few millimeters. State-of-the-art hardware and software solutions are integrated into an operational device. This novel system is tested against a commercial tracking system popular in PET brain imaging. Testing and demonstrations are carried out in clinical settings. A compact markerless tracking system was developed with an accuracy sufficient for PET imaging (<0.1 degrees and <0.3 mm). Furthermore, the first non-visible structured light system using Pico DLP technology was used. In a proof-of-principle study with two human PET scans, the system was demonstrated to improve PET image quality significantly. The results were similar to motion correction using an integrated commercial marker-based system. Furthermore, phantom studies were performed supporting the system's abilities for PET motion correction.

General information

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Olesen, O. V. (Intern), Larsen, R. (Intern), Paulsen, R. R. (Intern)
Number of pages: 174
Publication date: 2011

Publication information

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
ISBN (Print): 9788764309300
Original language: English
Series: IMM-PHD-2011
Number: 266
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd266_ovol.pdf
Source: orbit
Source-ID: 312476
Publication: Research › Ph.D. thesis – Annual report year: 2011

Mining of Ship Operation Data for Energy Conservation

This thesis presents two state-of-the-art systems approaches to statistical modelling of fuel efficiency in ship propulsion: a regression model and a dynamical model. Three statistical regression model approaches are investigated and compared: Artificial Neural Networks (ANN), Gaussian processes (GP), and Gaussian Mixture Models (GMM). A dynamical modelling approach is introduced. This modelling approach has not been used before in the context of ship propulsion modelling, and solves problems encountered with the regression model in an onboard trim optimization application. The dynamical model is introduces through a study of the wellknown sunspot time series, and then on ship data. The dynamical modelling approach is investigated using both the Artificial Neural Network and the Gaussian mixture model. The thesis also presents a novel and publicly available data set of high quality sensory data on which all the models are based and tested. No other similar publicly available data set exists. The data presented is a publicly available full-scale data set, with a whole range of features sampled over a period of 2 months. The data is online with an accompanying homepage, where all the results are also presented.

General information

State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Petersen, J. P. (Intern), Winther, O. (Intern)
Number of pages: 104
Publication date: 2011

Publication information

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
A Framework for Constraint-Programming based Configuration

Product configuration systems play an important role in the development of Mass Customisation, allowing the companies to reduce their costs while offering highly customised products. Such systems are often based on a configuration model, representing the product knowledge necessary to perform the configuration task. Several challenges arise when dealing with product configuration. One of those issues concerns how to model a configurable product family, i.e. how to represent the different types of configuration knowledge and their interactions. Another challenge is to provide adequate formalisms and efficient algorithms to solve the dependencies of the models at runtime. In this dissertation, we present a constraint-based framework for configuration. The design of this framework is partly based on a study of product configuration requirements as well as a comparison of several general modelling languages. We then develop ProCoLa, a configuration-specific modelling language based on a conceptual framework that synthesizes, unifies and extends several approaches to modelling configuration in different design disciplines, e.g. physical products, software or services. A rigorous formalisation of the ProCoLa language is given and used to verify and analyse the configuration models. Another goal of this dissertation is to describe the semantics of ProCoLa by providing a translation to a Constraint Satisfaction Problem (CSP) representation. For that purpose, several CSP formalisms are discussed and a new algorithm DnSTR is developed in order to solve the dynamic addition and retraction of table constraints at runtime. Finally, we present and evaluate a prototype implementation of ProCoLa and the configuration framework, including the integration in a development environment, tool support and interaction with UML, databases and spreadsheet applications.
accuracy. Our results instead suggest that model regularization parameters must be carefully selected, so that the model and its visualization enhance our ability to interpret the brain. The second part concerns interpretation of nonlinear models and procedures for extraction of ‘brain maps’ from nonlinear kernel models. We assess the performance of the sensitivity map as means for extracting a global summary map from a trained model. Such summary maps provides the investigator with an overview of brain locations of importance to the model’s predictions. The sensitivity map proves as a versatile technique for model visualization. Furthermore, we perform a preliminary investigation of the use of pre-image estimation for lo- calized interpretation of nonlinear models. In the context of image denoising the pre-image analysis proves to enhance the reliability of brain patterns extracted from multivariate models of the neuroimaging data.

Verification of Stochastic Process Calculi

Stochastic process calculi represent widely accepted formalisms within Computer Science for modelling nondeterministic stochastic systems in a compositional way. Similar to process calculi in general, they are suited for modelling systems in a hierarchical manner, by explicitly specifying subsystems as well as their interdependencies and communication channels. Stochastic process calculi incorporate both the quantified uncertainty on probabilities or durations of events and nondeterministic choices between several possible continuations of the system behaviour. Modelling of a system is often performed with the purpose to verify the system. In this dissertation it is argued that the verification techniques that have their origin in the analysis of programming code with the purpose to deduce the properties of the code's execution, i.e. Static Analysis techniques, are transferable to stochastic process calculi. The description of a system in the syntax of a particular stochastic process calculus can be analysed in a compositional way, without expanding the state space by explicitly resolving all the interdependencies between the subsystems which may lead to the state space explosion problem. In support of this claim we have developed analysis methods that belong to a particular type of Static Analysis { Data Flow / Pathway Analysis. These methods have previously been applied to a number of non-stochastic process calculi. In this thesis we are lifting them to the stochastic calculus of Interactive Markov Chains (IMC). We have devised the Pathway Analysis of IMC that is not only correct in the sense of overapproximating all possible behaviour scenarios, as is usual for Static Analysis methods, but is also precise. This gives us the possibility to explicitly decide on the trade-off between precision and complexity while post-processing the analysis results. Another novelty of our methods consists in the kind of properties that we can verify using the results of the Pathway Analysis. We can check both qualitative and quantitative properties of IMC systems. In particular, we have developed algorithms for constructing bisimulation relations, computing (overapproximations of) sets of reachable states and computing the expected time reachability, the last for a linear fragment of IMC. In all the cases we have the complexities of algorithms which are low polynomial in the size of the syntactic description of a system. The presented methods have a clear application in the areas of embedded systems, (randomised) protocols run between a fixed number of parties etc.
Grey Box Modelling of Hydrological Systems: With Focus on Uncertainties

The main topic of the thesis is grey box modelling of hydrologic systems, as well as formulation and assessment of their embedded uncertainties. Grey box model is a combination of a white box model, a physically-based model, and a black box model, which relates to models that are obtained statistically from input-output relations. Grey box model consists of a system description, defined by a finite set of stochastic differential equations, and an observation equation. Together, system and observation equations represent a stochastic state space model. In the grey box model the total noise is divided into a measurement noise and a process noise. The process noise is due to model approximations, undiscovered input and uncertainties in the input series. Estimates of the process noise can be used to highlight the lack of fit in state space formulation, and further support decisions for a model expansion. By using stochastic differential equations to formulate the dynamics of the hydrological system, either the complexity of the model can be increased by including the necessary hydrological processes in the model, or formulation of process noise can be considered so that it meets the physical limits of the hydrological system and give an adequate description of the embedded uncertainty in model structure. The thesis consists of two parts: a summary report and a part which contains six scientific papers. The summary report is divided into three distinct parts that introduce the main concepts and methods used in the following papers. The first part contains the basic concepts in hydrology and related hydrological models. The second part explains the grey box model by presenting stochastic differential equations and show how the equations can be linked to the available measurements. Moreover, impulse response function models are introduced as an alternative to stochastic differential equation based models, but by exploiting known hydrological models as the impulse response function in this model makes this model framework partly physically-based. For estimating the parameters in the grey box models maximum likelihood method is used. The third important part of the summary report is predictions, and with focus on uncertainty of prediction intervals the corresponding performance measures have to include the intervals. The thesis illustrates three performance measures for this performance evaluation: reliability, sharpness and resolution. For decision making, a performance criterion is preferred that quantifies all of these measures in a single number, and for that the quantile skill score criterion is discussed in this thesis. The first half of part of the thesis, which contains the papers, is divided into two different subjects. First are four papers, which consider the grey box model approach to a well field with several operating pumps. The model foundation is the governing equation for groundwater flow, which can be simplified and represented a state space form that resembles the methods used in numerical methods for well field modelling. The objective in the first two papers is to demonstrate how a simple grey box model is formulated and, subsequently, extended in terms of parameter estimation using statistical methods. The simple models in these papers consider only part of the well field, but data analysis reveals that the wells in the well field are highly correlated. In the third paper, all wells pumping from the same aquifer are included in the state space formulation of the model, but instead of extending the physical description of the system, the uncertainty is formulated to handle the spatio-temporal variation in the output. The uncertainty in the model are then evaluated by using the quantile skill score criterion. In the fourth paper, the well field is formulated by considering the impulse response function models to describe water level variation in the wells, as a function of available pumping rates in the well field. The paper illustrates, through a case study, how the model can be used to define and solve the well field management problem. The second half of part II consists of two papers where the stochastic differential equation based model is used for sewer runoff from a drainage system. A simple model is used to describe a complex rainfall-runoff process in a catchment, but the stochastic part of the system is formulated to include the increasing uncertainty when rainwater flows through the system, as well as describe the lower limit of the uncertainty when the flow approaches zero. The first paper demonstrates in detail the grey box model and all related transformations required to obtain a feasible model for the sewer runoff. In the last paper this model is used to predict the runoff, and the performances of the prediction intervals are evaluated by the quantile skill score criterion.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Thordarson, F. Ø. (Intern), Madsen, H. (Intern)
Number of pages: 208
Publication date: 2011

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2011
The advances seen in the semiconductor industry within the last decade have brought the possibility of integrating evermore functionality onto a single chip forming functionally highly advanced embedded systems. These integration possibilities also imply that as the design complexity increases, so does the design time and effort. This challenge is widely recognized throughout academia and the industry and in order to address this, novel frameworks and methods, which will automate design steps as well as raise the level of abstraction used to design systems, are being called upon. To support an efficient system level design methodology, a modelling framework for performance estimation and design space exploration at the system level is required. This thesis presents a novel component based modelling framework for system level modelling and performance estimation of embedded systems. The framework is simulation based and allows performance estimation to be carried out throughout all design phases ranging from early functional to cycle accurate and bit true descriptions of the system, modelling both hardware and software components in a unified way. Design space exploration and performance estimation is performed by having the framework produce detailed quantitative information about the system model under investigation. The project is part of the national Danish research project, Danish Network of Embedded Systems (DaNES), which is funded by the Danish National Advanced Technology Foundation. The project is carried out in collaboration with the Danish company and DaNES partner, Bang & Olufsen ICEpower. Bang & Olufsen ICEpower provides industrial case studies which will allow the proposed modelling framework to be exercised and assessed in terms of ease of use, production speed, accuracy and efficiency. The framework allows a given embedded system to be constructed and explored before a physical realization is present and it can be used in the design of completely new systems or for modification of legacy systems. The primary benefits of the framework are the possibilities of exploring a large number of candidate systems within a short time frame leading to better designs, easier design verification through an iterative refinement of the executable system description, and finally the possibility of a reduction of the time-to-market of the design and implementation of the system under consideration. In practice, however, additional time spent on software development in order to provide commercial quality tools supporting the method is required.

**System Level Modelling and Performance Estimation of Embedded Systems**

With an increasing demand for oil and difficulties in finding new major oil fields, research on methods to improve oil recovery from existing fields is more necessary now than ever. The subject of this thesis is to construct efficient numerical methods for simulation and optimization of oil recovery with emphasis on optimal control of water flooding with the use of smartwell technology. We have implemented immiscible oil injection in water and in isothermal reservoirs with isotropic heterogeneous permeability fields. We use the method of lines for solution of the partial differential equation (PDE) system that governs the fluid flow. We discretize the two-phase flow model spatially using the finite volume method (FVM), and we use the two point upstream (TPFA) and the single-point upstream (SPU) scheme for computing the fluxes. We propose a new formulation of the differential equation system that arise as a consequence of the spatial discretization of the two-phase flow model. Upon discretization in time, the proposed equation system ensures the mass-conserving property of the two-phase flow model. For the solution of the spatially discretized two-phase flow model, we develop mass conserving explicit singly diagonally implicit Runge-Kutta (ESDIRK) methods with embedded error estimators for adaptive step size control.
We demonstrate that high order ESDIRK methods are more efficient than the low-order methods most commonly used in reservoir simulators. Most commercial reservoir simulation tools use step size control, which is based on heuristics. These can neither deliver solutions with predetermined accuracy or guarantee the convergence in the modified Newton iterations. We have established predictive step size control based on error estimates, which can be calculated from the embedded ESDIRK methods. We change the step size control in order to minimize the computational cost per simulation. We implement a numerical method for nonlinear model predictive control (NMPC) along with smart-well technology to maximize the net present value (NPV) of an oil reservoir. The optimization is based on quasi-Newton sequential quadratic programming (SQP) with line-search and BFGS approximations of the Hessian, and the adjoint method for efficient computation of the gradients. We demonstrate that the application of NMPC for optimal control of smart-wells has the potential to increase the economic value of an oil reservoir.

**General information**

State: Published
Organisations: Scientific Computing, Department of Informatics and Mathematical Modeling, Center for Energy Resources Engineering
Authors: Völcker, C. (Intern), Jørgensen, J. B. (Intern), Thomsen, P. G. (Intern)
Number of pages: 186
Publication date: 2011

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2011
Number: 265
Main Research Area: Technical/natural sciences
Electronic versions: phd265_Carsten_Volcker.pdf
Source: orbit
Source-ID: 312479
Publication: Research › Ph.D. thesis – Annual report year: 2012

**Descriptor Based Analysis of Digital 3D Shapes**

Analysis and processing of 3D digital shapes has become an important research area with numerous medical, industrial, and entertainment applications which has gained enormously in importance as optical scanning modalities have started to make acquired 3D geometry commonplace. The area holds many challenges. One such challenge, which is addressed in this thesis, is to develop computational methods for classifying shapes which are in agreement with the human way of understanding and classifying shapes. In this dissertation we first present a shape descriptor based on the process of diffusion on the surface of the shape – the auto diffusion function. When all heat is inserted at a single point, the function describes how much of that heat will remain at the same point after a period of time. This method allows for finding shape features at different scales related to time parameter. For instance, in conjunction with the method of Reeb graphs for skeletonization, it is an effective tool for generating scale dependent skeletons of shapes represented as 3D triangle meshes. The second part of the thesis aims at capturing the style phenomenon. The style of an object is easily recognized by humans but a computational method for finding the style of an object is elusive. Instead of codifying the style explicitly, which can be only done within a specific context, we develop a general method for dealing with both style and function which uses the supervision provided by a set of training examples and can be evaluated using any shape descriptor, that produces dissimilarity measures between different shapes. Our methods decouple the effect of style from the effect of function and assess how suitable a descriptor is to a specific problem.

**General information**

State: Published
Organisations: Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics
Authors: Welnicka, K. (Intern), Bærentzen, J. A. (Intern)
Number of pages: 157
Publication date: 2011

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2011
Number: 262
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Semantic Approaches for Knowledge Discovery and Retrieval in Biomedicine

This thesis discusses potential applications of semantics to the recent literature-based informatics systems to facilitate knowledge discovery, hypothesis generation, and literature retrieval in the domain of biomedicine. The approaches presented herein make use of semantic information extracted from biomedical texts by natural language processing systems supported by biomedical ontologies. The thesis is divided into two main parts: First, a field of literature-based discovery is introduced, with a review of recent approaches of the field; second, literature retrieval in the domain of neuroimaging (neuroscience) is discussed with the emphasis put on the coordinate-based searching of related publications. My own contribution to the first part is a novel literature-based ‘discovery browsing’ methodology incorporating semantic predications, graph theory and path analysis for guiding researchers through the relevant literature on a user-specied biomedical phenomenon. Moreover, the additional analyses of the methodology show its potential application as a support for the recent probabilistic retrieval methods. In the second part of the thesis, I present the BredeQuery plugin which integrates a coordinate-based literature retrieval system with the common in neuroimaging statistical analysis environment. It is followed by the detailed description of a prototype of context-dependent neuroscientic literature retrieval methodology, which thanks to the employment of ontologies, allows the user to define context of interest for a search. The peer reviewed research articles, included in the appendices, discuss further the details of the presented methods, case studies, and provide other related information.

Qualitative and Quantitative Security Analyses for ZigBee Wireless Sensor Networks

Wireless sensor networking is a challenging and emerging technology that will soon become an inevitable part of our modern society. Today wireless sensor networks are broadly used in industrial and civilian application areas including environmental monitoring, surveillance tasks, healthcare applications, home automation, and traffic control. The challenges for research in this area are due to the unique features of wireless sensor devices such as low processing power and associated low energy. On top of this, wireless sensor networks need secure communication as they operate in open fields or unprotected environments and communicate on broadcasting technology. As a result, such systems have to meet a multitude of quantitative constraints (e.g. timing, power consumption, memory usage, communication bandwidth) as well as security requirements (e.g. authenticity, confidentiality, integrity). One of the main challenges arise in dealing with the security needs of such systems where it is less likely that absolute security guarantees can be sustained - because of the need to balance security against energy consumption in wireless sensor network standards like ZigBee.

This dissertation builds on existing methods and techniques in different areas and brings them together to create an efficient verification system. The overall ambition is to provide a wide range of powerful techniques for analyzing models with quantitative and qualitative security information. We stated a new approach that first verifies low level security protocols in a qualitative manner and guarantees absolute security, and then takes these verified protocols as actions of scenarios to be verified in a quantitative manner. Working on the emerging ZigBee wireless sensor networks, we used probabilistic verification that can return probabilistic results with respect to the trade-off between security and performance. In this sense, we have extended various existing ideas and also proposed new ideas to improve verification. Especially in the problem of key update, we believe we have contributed to the solution for not only wireless sensor networks but also many other types of systems that require key updates. Besides we produced automated tools that were intended to
demonstrate what kind of tools can be developed on different purposes and application domains.

**General information**
State: Published
Organisations: Language-Based Technology, Department of Informatics and Mathematical Modeling
Authors: Yuksel, E. (Intern), Nielson, H. R. (Intern), Nielson, F. (Intern)
Number of pages: 306
Publication date: 2011

**Publication information**
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2011
Number: 247
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
B85E2d01.pdf
Source: orbit
Source-ID: 272804
Publication: Research › Ph.D. thesis – Annual report year: 2011

**Advanced Load Alleviation for Wind Turbines using Adaptive Trailing Edge Flaps: Sensoring and Control**
The purpose of wind turbines and their predecessors the windmill, is to convert the energy in the wind to usable energy forms. Whereas windmills of the past focused on the conversion of wind power to torque for grinding, pumping and winching, modern wind turbines convert the wind energy into electric power. They do so through incorporation of generators, which convert mechanical torque into electricity. Wind turbines are designed to keep the overall cost per produced Kilo Watt hour as low as possible. One way of improving the performance and lifetime of the wind turbine is through active flow control. Active control is often considered costly but if the lifespan of the components can be increased it could be justifiable. This thesis covers various aspects of 'smart control' such as control theory, sensoring, optimization, experiments and numerical modeling.

**General information**
State: Published
Organisations: Aeroelastic Design, Wind Energy Division, Risø National Laboratory for Sustainable Energy, Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Andersen, P. B. (Intern), Gaunaa, M. (Intern), Bak, C. (Intern), Buhl, T. (Intern), Poulsen, N. K. (Intern)
Number of pages: 132
Publication date: Jun 2010

**Publication information**
Place of publication: Roskilde
Publisher: Risø National Laboratory for Sustainable Energy
ISBN (Print): 978-87-550-3824-0
Original language: English
Series: Risø-PhD
Number: 61(EN)
Main Research Area: Technical/natural sciences
Wind energy, Aeroelastic design methods, Risø-PhD-61(EN), Risø-PhD-61, Risø-PhD-0061
Electronic versions:
ris-phd-61.pdf
Source: orbit
Source-ID: 262258
Publication: Research › Ph.D. thesis – Annual report year: 2010

**Modelling and Analyses of Embedded Systems Design**
We present the MoVES languages: a language with which embedded systems can be specified at a stage in the development process where an application is identified and should be mapped to an execution platform (potentially multicore). We give a formal model for MoVES that captures and gives semantics to the elements of specifications in the MoVES language. We show that even for seemingly simple systems, the complexity of verifying real-time constraints can be overwhelming - but we give an upper limit to the size of the search-space that needs examining. Furthermore, the formal model exposes important scheduling situations that become central in establishing timed-automata models that can
be used for analysis of MoVES specifications effectively. Finally we present the MoVES tool, which can conduct automatic verification of interesting properties of MoVES specifications. In several examples, we use the MoVES tool to conduct analysis that identifies timing anomalies. We also conduct design space exploration in an example using the MoVES tool. And we show that it can be used for analysis of systems that, in size, resemble industrially-interesting systems. We find that semantically-based verification is a promising approach for assisting developers of embedded systems. We provide examples of system verifications that, in size and complexity, point in the direction of industrially-interesting systems.

General information
State: Published
Organisations: Embedded Systems Engineering, Department of Informatics and Mathematical Modeling
Authors: Brekling, A. W. (Intern), Hansen, M. R. (Intern), Madsen, J. (Intern)
Publication date: 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2011-236
Main Research Area: Technical/natural sciences
Electronic versions:
phd236_aske-foreloebig.pdf
Source: orbit
Source-ID: 265894
Publication: Research › Ph.D. thesis – Annual report year: 2011

Data analysis in high-dimensional sparse spaces: Large p, small n problems
The present thesis considers data analysis of problems with many features in relation to the number of observations (large p, small n problems). The theoretical considerations for such problems are outlined including the curses and blessings of dimensionality, and the importance of dimension reduction. In this context the trade off between a rich solution which answers the questions at hand and a simple solution which generalizes to unseen data is described. For all of the given data examples labelled output exists and the analyses are therefore limited to supervised settings. Three novel classification techniques for high-dimensional problems are presented: Sparse discriminant analysis, sparse mixture discriminant analysis and orthogonality constrained support vector machines. The first two introduces sparseness to the well known linear and mixture discriminant analysis and thereby provide low-dimensional projections of data with few non-zero loadings which give improvements in classification. The latter adds a priori information of pairing between observations to the support vector machine and thereby give solutions with less variation and slight improvements in classification. The classification methods are applied to classifications of fish species, ear canal impressions used in the hearing aid industry, microbiological fungi species, and various cancerous tissues and healthy tissues. In addition, novel applications of sparse regressions (also called the elastic net) to the medical, concrete, and food industries via multi-spectral images for objective and automated systems are presented.

General information
State: Published
Organisations: DTU Data Analysis, Department of Informatics and Mathematical Modeling
Authors: Clemmensen, L. K. H. (Intern), Ersbøll, B. K. (Intern)
Publication date: Mar 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2009-228
Main Research Area: Technical/natural sciences
Electronic versions:
phd228_lhc.pdf
Source: orbit
Source-ID: 254894
Publication: Research › Ph.D. thesis – Annual report year: 2010

Design of Computer Experiments
The main topic of this thesis is design and analysis of computer and simulation experiments and is dealt with in six papers and a summary report. Simulation and computer models have in recent years received increasingly more attention due to their increasing complexity and usability. Software packages make the development of rather complicated computer
models using predefined building blocks possible. This implies that the range of phenomena that are analyzed by means of a computer model has expanded significantly. As the complexity grows so does the need for efficient experimental designs and analysis methods, since the complex computer models often are expensive to use in terms of computer time. The choice of performance parameter is an important part of the analysis of computer and simulation models and Paper I introduces a new statistic for waiting times in health care units. The statistic is a measure of the extent of long waiting times, which are known both to be the most bothersome and to have the greatest impact on patient satisfaction. A simulation model for an orthopedic surgical unit at a hospital illustrates the benefits of using the measure. Another important consideration in connection to simulation models is the design of experiments, which is the decision of which of the possible configurations of the simulation model that should be tested. Since the possible configurations are numerous and the time to test a single configuration may take minutes or hours of computer time, the number of configurations that can be tested is limited. Papers B and C introduce a novel experimental plan for simulation models having two types of input factors. The plan differentiates between factors that can be controlled in both the simulation model and the physical system and factors that are only controllable in the simulation model but simply observed in the physical system. Factors that only are controllable in the simulation model are called uncontrollable factors and they correspond to the environmental factors in fluencing the physical system. Applying the experimental framework on the simulation model in Paper A shows that the effects of changes in the uncontrollable factors are better understood with the proposed design compared to the alternative and commonly used methods. In papers D and E a modeling framework for analyzing simulation models with multiple noise sources is presented. It is shown that the sources of variation of the simulation model can be divided in two components corresponding to changes in the environmental factors (the uncontrollable factor settings) and to random variation. Moreover, the structure of the environmental effects can be estimated, which can be used to put the system in a more robust operating mode. The interpolation technique called Kriging is the topic of Paper F, which is a widely applied technique for building so called models-for-the-model (metamodels). We propose a method that handles both qualitative and quantitative factors, which is not covered by the standard model. Fitting the final Kriging model is done in two stages each based on fitting regular Kriging models. It is shown that this method works well on a realistic example such as a simulation model for a surgical unit.

**General information**

**State:** Published  
**Organizations:** Mathematical Statistics, Department of Informatics and Mathematical Modeling  
**Authors:** Dehlendorff, C. (Intern), Andersen, K. K. (Intern), Kulahci, M. (Intern)  
**Publication date:** Dec 2010

**Publication information**

**Place of publication:** Kgs. Lyngby, Denmark  
**Publisher:** Technical University of Denmark (DTU)  
**Original language:** English  
**Series:** IMM-PHD-2010-237  
**Main Research Area:** Technical/natural sciences  
**Electronic versions:**  
phd237_cd.pdf  
**Source:** orbit  
**Source-ID:** 266323  
**Publication:** Research › Ph.D. thesis – Annual report year: 2010

**Corrections in clinical Magnetic Resonance Spectroscopy and SPECT:** Motion correction in MR spectroscopy Downscatter correction in SPECT

The quality of medical scanner data is often compromised by several mechanisms. This can be caused by both the subject to be measured and the scanning principles themselves. In this PhD project the problem of subject motion was addressed for Single Voxel MR Spectroscopy in a cohort study of preterm infants. In Iodine-123 SPECT the problem of downscatter was addressed. This thesis is based on two papers. Paper I deals with the problem of motion in Single Voxel Spectroscopy. Two novel methods for the identification of outliers in the set of repeated measurements were implemented and compared to the known mean and median filtering. The data comes from non-anesthetized preterm infants, where motion during scanning is a common problem. Both the novel outlier identification and the independent component analysis (ICA) perform satisfactory and better than the common mean and median filtering. ICA performed best in the sense that it recovered most of the lost peak height in the spectra. The ICA motion correction algorithm described in paper I and in this thesis was applied to a quantitative analysis of the Single Voxel Spectroscopy data from the cohort study of preterm infants. This analysis revealed that differences between term and preterm infants are not to be found in the concentrations of Lactate (caused by inflammation or hypoxia-ischemia) and/or NAA (caused by hypoxia-ischemia) as hypothesized before the cohort study. Instead choline levels were decreased in the preterm infants, which might indicate a detrimental effect of the extra-uterine environment on brain development. Paper II describes a method to correct for downscatter in low count Iodine-123 SPECT with a broad energy window above the normal imaging window. Both spatial dependency and weight factors were measured. As expected, the implicitly assumed weight factor of one for energy windows with equal width is slightly too low, due the presence of a backscatter peak in the energy spectrum coming from high-energy photons. The effect on the contrast was tested in 10 subjects and revealed a 20% increase in the specific binding ratio of the striatum due to downscatter correction. This makes the difference between healthy subjects and
patients more profound. Downscatter in Iodine-123 SPECT is not the only deteriorating mechanism. Normal scatter compromises the images quality as well. Since scatter correction of SPECT-images also can be performed by the subtraction of an energy window, a method was developed to perform scatter and downscatter correction simultaneously. A phantom study has been performed, where the in paper II described downscatter correction was extended with scatter correction. This new combined correction was compared to the known Triple Energy Window (TEW) correction method. Results were satisfying and indicate that TEW is more correct from the physics point of view, while the in paper II described method extended with scatter correction gives reasonable results, but is far less noise sensitive than TEW.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling, Biomedical Engineering, Department of Electrical Engineering
Authors: de Nijs, R. (Intern), Hansen, L. K. (Intern), Hanson, L. G. (Intern)
Publication date: Mar 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2009-221
Main Research Area: Technical/natural sciences
Electronic versions:
phd221_rdn.pdf
Source: orbit
Source-ID: 251055
Publication: Research › Ph.D. thesis – Annual report year: 2010

Approximate Inference for Wireless Communications
This thesis investigates signal processing techniques for wireless communication receivers. The aim is to improve the performance or reduce the computationally complexity of these, where the primary focus area is cellular systems such as Global System for Mobile communications (GSM) (and extensions thereof), but also general Multiple-Input Multiple-Output (MIMO) systems are considered. The motivation for a performance improvement is that this is needed to achieve higher capacity in the systems, which can ensure increased bit-rates at the same or lower prices. A reduction in the computationally complexity can potentially lead to limited power consumption, which translates into longer battery life-time in the handsets. The scope of the thesis is more specifically to investigate approximate (nearoptimal) detection methods that can reduce the computationally complexity significantly compared to the optimal one, which usually requires an unacceptable high complexity. Some of the treated approximate methods are based on QL-factorization of the channel matrix. In the work presented in this thesis it is proven how the QL-factorization of frequency-selective channels asymptotically provides the minimum-phase and all-pass filters. This enables us to view Sphere Detection (SD) as an adaptive variant of minimum-phase prefiltered reduced-state sequence estimation. Thus, a novel way of computing the minimum-phase filter and its associated all-pass filter using the numerically stable QL-factorization is suggested. Alternatively, fast QL-factorization methods can be applied which provides a computationally efficient way of obtaining these filers. Additionally, Markov Chain Monte Carlo (MCMC) sampling has been investigated for near-optimal Maximum Likelihood Sequence Detection in MIMO systems. The MCMC method considered in the thesis is the Gibbs sampler, which is proposed as an alternative to the SD in scenarios where the latter type of detector requires an unacceptable high complexity.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Hansen, M. (Intern), Christensen, L. P. (Intern), Winther, O. (Intern)
Publication date: Mar 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2009-227
Main Research Area: Technical/natural sciences
Electronic versions:
phd227_mha.pdf
Source: orbit
Source-ID: 253947
Characterization and optimized control by means of multi-parameter controllers

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling, Section for Indoor Environment, Department of Civil Engineering, Teknologisk Institut, Businessminds
Authors: Jensen, S. Ø. (Ekstern), Andersen, P. H. D. (Intern), Heerup, C. (Ekstern), Larsen, S. (Ekstern), Olsen, L. (Ekstern), Toftum, J. (Intern), Trombe, P. (Intern)
Number of pages: 276
Publication date: 2010

Publication information
Place of publication: Danish Technological Institute
ISBN (Print): 87-7756-772-2
Original language: English
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 272341
Publication: Research › Report – Annual report year: 2010

Quantitative data analysis methods for 3D microstructure characterization of Solid Oxide Cells
The performance of electrochemical ceramic devices such as solid oxide fuel and electrolyser cells depends on the distribution of constituent phases on the micro or nano scale, also known as the microstructure. The microstructure governs key properties such as ion, electron and gas transport through percolating networks and reaction rates at the triple phase boundaries. Quantitative analysis of microstructure is thus important both in research and development of optimal microstructure design and fabrication. Three dimensional microstructure characterization in particular holds great promise for gaining further fundamental understanding of how microstructure affects performance. In this work, methods for automatic 3D characterization of microstructure are studied: from the acquisition of 3D image data by focused ion beam tomography to the extraction of quantitative measures that characterize the microstructure. The methods are exemplified by the analysis of Ni-YSZ and LSC-CGO electrode samples. Automatic methods for preprocessing the raw 3D image data are developed. The preprocessing steps correct for errors introduced by the image acquisition by the focused ion beam serial sectioning. Alignment of the individual image slices is performed by automatic detection of ducial marks. Uneven illumination is corrected by fitting hypersurfaces to the spatial intensity variation in the 3D image data. Routine use of quantitative three dimensional analysis of microstructure is generally restricted by the time consuming task of manually delineating structures within each image slice or the quality of manual and automatic segmentation schemes. To solve this, a framework for the automatic segmentation of 3D image data is developed. The technique is based on a level set method and uses numerical approximations to partial differential equations to evolve a 3D surface to capture the phase boundaries. Vector fields derived from the experimentally acquired data are used as the driving forces. The framework performs the segmentation in 3D rather than on a slice by slice basis. It naturally supplies sub-voxel accuracy of segmented surfaces and allows constraints on the surface curvature to enforce a smooth surface in the segmentation. A high accuracy method is developed for calculating two phase boundary surface areas and triple phase boundary length of triple phase systems. The calculations are based on sub-voxel accuracy segmentations of the constituent phases. The method performs a three phase polygonization of the interface boundaries which results in a non-manifold mesh of connected faces. The triple phase boundaries can be extracted from the mesh as connected curve loops without branches. The accuracy of the method is analyzed by calculations on geometrical primitives. A suite of methods is developed for characterizing the shape and connectivity of phase networks. The methods utilize the fast marching method to compute distance maps and optimal paths in the microstructure network. The extracted measurements are suited for the quantitative comparison and evaluation of microstructures. The quantitative measures characterize properties of network path tortuosity, network thickness, transport path width and dead ends.

General information
State: Published
Organisations: Microstructures and Interfaces, Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, Fuel Cells and Solid State Chemistry Division, Risø National Laboratory for Sustainable Energy
Authors: Jørgensen, P. S. (Intern), Larsen, R. (Intern), Bowen, J. R. (Intern), Hansen, K. V. (Intern), Lassen, N. C. K. (Intern), Wallenberg, R. (Ekstern)
Publication date: Sep 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
**Deformable Simplicial Complexes**

In this dissertation we present a novel method for deformable interface tracking in 2D and 3D (deformable simplicial complexes (DSC)). Deformable interfaces are used in several applications, such as fluid simulation, image analysis, reconstruction or structural optimization. In the DSC method, the interface (curve in 2D; surface in 3D) is represented explicitly as a piecewise linear curve or surface. However, the domain is also subject to discretization: triangulation in 2D; tetrahedralization in 3D. This way, the interface can be alternatively represented as a set of edges/triangles separating triangles/tetrahedra marked as outside from those marked as inside. Such an approach allows for robust topological adaptivity. Among other advantages of the deformable simplicial complexes there are: space adaptivity, ability to handle and preserve sharp features, possibility for topology control. We demonstrate those strengths in several applications. In particular, a novel, DSC-based fluid dynamics solver has been developed during the PhD project. A special feature of this solver is that due to the fact that DSC maintains an explicit interface representation, surface tension is more easily dealt with. One particular advantage of DSC is the fact that as an alternative to topology adaptivity, topology control is also possible. This is exploited in the construction of cut loci on tori where a front expands from a single point on a torus and stops when it self-intersects.

**General information**

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Misztal, M. K. (Intern), Anton, F. (Intern), Bærentzen, J. A. (Intern)
Publication date: 2010

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

**Markov and mixed models with applications**

This thesis deals with mathematical and statistical models with focus on applications in pharmacokinetic and pharmacodynamic (PK/PD) modelling. These models are today an important aspect of the drug development in the pharmaceutical industry and continued research in statistical methodology within these areas are thus important. PK models are concerned with describing the concentration profile of a drug in both humans and animals after drug intake whereas PD models are used to describe the effect of a drug in relation to the drug concentration. PK models for an individual are usually described as a deterministic mean value using ordinary differential equations to which a random error is added. This thesis explores methods based on stochastic differential equations (SDEs) to extend the models to more adequately describe both true random biological variations and also variations due to unknown or uncontrollable factors in an individual. Modelling using SDEs also provides new tools for estimation of unknown inputs to a system and is illustrated with an application to estimation of insulin secretion rates in diabetic patients. Models for the eect of a drug is a broader area since drugs may affect the individual in almost any thinkable way. This project focuses on measuring the eects on sleep in both humans and animals. The sleep process is usually analyzed by categorizing small time segments into a number of sleep states and this can be modelled using a Markov process. For this purpose new methods for non-parametric estimation of Markov processes are proposed to give a detailed description of the sleep process during the night. Statistically the Markov models considered for sleep states are closely related to the PK models based on SDEs as both models share the Markov property. When the models are applied to clinical data there will often be a large variation between individuals and this can be included in both types of models using the mixed modelling approach. Estimation in these models is discussed with emphasis on data with a more complex grouping structure.

**General information**
Call Center Capacity Planning
The main topics of the thesis are theoretical and applied queueing theory within a call center setting. Call centers have in recent years become the main means of communication between customers and companies, and between citizens and public institutions. The extensively computerized infrastructure in modern call centers allows for a high level of customization, but also induces complicated operational processes. The size of the industry together with the complex and labor intensive nature of large call centers motivates the research carried out to understand the underlying processes. The customizable infrastructure allows customers to be divided into classes depending on their requests or their value to the call center operator. The agents working in call centers can in the same way be split into groups based on their skills. The discipline of matching calls from different customer classes to agent groups is known as skills-based routing. It involves designing the routing policies in a way that results in customers receiving a desired service level such as the waiting time they experience. The emphasis of this thesis is on the design of these policies. The first paper, Queues with waiting time dependent service, introduces a novel approach to analyzing queueing systems. This involves using the waiting time of the first customer in line as the primary variable on which the analysis is based. The legacy approach has been to use the number of customers in queue. The new approach facilitates exact analysis of systems where service depends on the waiting time. Two such systems are analyzed, one where a server can adapt its service speed according to the waiting time of the first customer in line. The other deals with a two-server setup where one of the servers is only allowed to take customers who have waited a certain fixed amount of time. The latter case is based on a commonly used rule in call centers to control overflow between agent groups. Realistic call center models require multi-server setups to be analyzed. For this reason, an approximation based on the waiting time of the first in line approach is developed in the paper Waiting time dependent multi-server priority queues, which is able to deal with multi-server setups. It is used to analyze a setup with two customer classes and two agent groups, with overflow between them controlled by a fixed threshold. Waiting time distributions are obtained in order to relate the results to the service levels used in call centers. Furthermore, the generic nature of the approximation is demonstrated by applying it to a system incorporating a dynamic priority scheme. In the last paper Optimization of overflow policies in call centers, overflows between agent groups are further analyzed. The design of the overflow policies is optimized using Markov Decision Processes and a gain with regard to service levels is obtained. Also, the fixed threshold policy is investigated and found to be appropriate when one class is given high priority and when it is desired that calls are answered by the designated agent class and not by other groups through overflow.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling, Networks Technology and Service Platforms, Department of Photonics Engineering
Authors: Nielsen, T. B. (Intern), Nielsen, B. F. (Intern), Iversen, V. B. (Intern)
Publication date: Mar 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2009-223
Main Research Area: Technical/natural sciences
Electronic versions:
phd223_tbn.pdf
Source: orbit
Source-ID: 251790
Publication: Research › Ph.D. thesis – Annual report year: 2010
Hidden Markov modelling of movement data from fish

Movement data from marine animals tagged with electronic tags are becoming increasingly diverse and plentiful. This trend entails a need for statistical methods that are able to filter the observations to extract the ecologically relevant content. This dissertation focuses on the development and application of hidden Markov models (HMMs) for analysis of movement data from fish. The main contributions are represented by six scientific publications. Estimation of animal location from uncertain and possibly indirect observations is the starting point of most movement data analyses. In this work a discrete state HMM is employed to deal with this task. Specifically, the continuous horizontal plane is discretised into grid cells, which enables a state-space model for the geographical location to be estimated on this grid. The estimation model for location is extended with an additional state representing the behaviour of the animal. With the extended model can migratory and resident movement behaviour be related to geographical regions. For population inference multiple individual state-space analyses can be interconnected using mixed effects modelling. This framework provides parameter estimates at the population level and allows ecologists to identify individuals that deviate from the rest of the tagged population. The thesis also deals with geolocation on state-spaces with complicated geometries. Using an unstructured discretisation and the finite element method tortuous shore line geometries are closely approximated. This furthermore enables accurate probability densities of location to be computed. Finally, the performance of the HMM approach in analysing nonlinear state space models is compared with two alternatives: the AD Model Builder framework and BUGS, which relies on Markov chain Monte Carlo estimation.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling, Section for Population Ecology and Genetics, National Institute of Aquatic Resources
Publication date: 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2010-243
Main Research Area: Technical/natural sciences
Electronic versions:
phd243_mwp_rev_net.pdf
Source: orbit
Source-ID: 267387
Publication: Research › Ph.D. thesis – Annual report year: 2010

Modeling media as latent semantics based on cognitive components

Though one might think of media as an audiovisual stream of consciousness, we frequently encode frames of video sequences and waves of sound into strings of text. Language allows us to both share the internal representations of what we perceive as mental concepts, as well as categorizing them as distinct states in the continuous ebb and flow of emotions underlying consciousness. Whether it being a soundscape of structured peaks or tiny black characters lined up across a page, we rely on syntax for parsing sequences of symbols, which based on hierarchically nested structures allow us to express and share the meaning contained within a sentence or a melodic phrase. As both low-level semantic structure of texts and our affective responses can be encoded in words, a simplified cognitive model can be constructed which uses LSA latent semantic analysis to emulate how we perceive the emotional context of media based on lyrics, synopses, subtitles, blogs or web pages associated with the content. In the proposed model the bottom-up generated sensory input is a matrix of tens of thousands of words co-occurring within multiple contexts, that are in turn represented as vectors in a semantic space of reduced dimensionality. While top-down, patterns of emotional categorization emerge by defining term vector distances to affective adjectives, that constrain the latent semantic structures according to the neurophysiological dimensions of valence and arousal. The thesis thus combines elements of machine learning with aspects of cognitive linguistics that potentially could be utilized in applications ranging from information retrieval and media personalization, to emotional brand building or neuroscientific modeling of syntax and semantics.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Petersen, M. K. (Intern), Hansen, L. K. (Intern)
Publication date: Sep 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
On the time required for identification of visual objects

The starting point for this thesis is a review of Bundesen's theory of visual attention. This theory has been widely accepted as an appropriate model for describing data from an important class of psychological experiments known as whole and partial report. Analysing data from this class of experiments with the help of the theory of visual attention – have proven to be an effective approach to examine cognitive parameters that are essential for a broad range of different patient groups. The theory of visual attention relies on a psychometric function that describes the ability to identify a stimulus as a function of exposure duration. An important contribution of the thesis is that it investigates whether other psychometric functions than the one originally used with the theory of visual attention could be more appropriate at describing data. The thesis points to two psychometric functions that seem more appropriate. Further the thesis shows that it is possible to incorporate any desired psychometric into the theory of visual attention. Common to the two psychometric functions suggested is that they both have a hazard function that is non-monotonic; a neural argument for this is also presented in the thesis. For the psychometric function it is further investigated how this depends on stimulus contrast. In this respect, we find that the type of psychometric function is independent of contrast, but that the parameters for the psychometric function vary systematically as a function of contrast. An analysis of the psychometric function for the individual letters of the alphabet shows that there are significant differences in the parameters of the psychometric function depending on letter identity. Here we should note that in many cases (also for Bundesen’s theory of visual attention) it has been customary to average performance over the entire set of stimuli, consisting for instance of the 26 alphabetic letters. The fact that each letter is perceived in a different way possibly reflects that each letter is represented differently in our brain. This might have to do with a difference in the set of features representing the individual letters. It is possible that some features are processed faster than others and that overlapping features representing more than one letter in the alphabet play a certain role for the tendency to confuse letters. Hopefully it should be possible, with the dataset that we collected, to directly analyse how confusability develops as a certain letter is exposed for increasingly longer time. An important scientific question is what shapes the psychometric function. It is conceivable that the function reflects both limitations and structure of the physical mechanism underlying perception. For this reason we argue that the alternative psychometric functions that we have suggested are also relevant for models trying to simulate the mechanism leading to perception. The thesis reviews a selection of stochastic models that are well-known candidates when it comes to modelling mechanisms of perception. These candidates include the Ornstein-Uhlenbeck model and the leaky competing accumulator model. A further contribution of the thesis is a demonstration that the leaky competing accumulator model (see Usher & Cohen, 1999) is able to explain a perceptual limit that characterises how many objects can in parallel be perceived. Finally, the thesis suggests five concrete topics for future work. These include as diverse themes as: determination of the visual features representing the individual letters in our brain, neurodynamical modelling of visual perception, investigation of the duration of visual short-term memory as well as psychometric functions and assumptions along with application areas in cognitive diagnostics.

General information

State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Petersen, A. (Intern), Andersen, T. (Intern), Hansen, L. K. (Intern), Kyllingsbæk, S. (Intern)
Publication date: Apr 2010

Publication information

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark, DTU Informatics, Building 321
Original language: English

Series: IMM-PHD-2010-224
Main Research Area: Technical/natural sciences
Electronic versions:
phd224_ap.pdf
Source: orbit
Source-ID: 253290
Publication: Research › Ph.D. thesis – Annual report year: 2010
Nonlinear Stochastic Modelling of Antimicrobial resistance in Bacterial Populations

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Philipsen, K. R. (Intern), Christiansen, L. E. (Intern), Madsen, H. (Intern)
Publication date: Jun 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2009-226
Main Research Area: Technical/natural sciences
Electronic versions:
phd226_krp-new.pdf
Source: orbit
Source-ID: 253754
Publication: Research › Ph.D. thesis – Annual report year: 2010

Support for Programming Models in Network-on-Chip-based Many-core Systems
This thesis addresses aspects of support for programming models in Network-on-Chip-based many-core architectures. The main focus is to consider architectural support for a plethora of programming models in a single system. The thesis has three main parts. The first part considers parallelization and scalability in an image processing application with the aim of providing insight into parallel programming issues. The second part proposes and presents the tile-based Clupea many-core architecture, which has the objective of providing configurable support for programming models to allow different programming models to be supported by a single architecture. The architecture features a specialized network interface processor which allows extensive configurability of the memory system. Based on this architecture, a detailed implementation of the cache coherent shared memory programming model is presented. The third part considers modeling and evaluation of the Clupea architecture configured for support for cache coherent shared memory. An analytical model and the MC sim simulator, which provides detailed cycle-accurate simulation of many-core architectures, have been developed for the evaluation of the Clupea architecture. The evaluation shows that configurability causes a moderate increase of the application execution time. Considering the improved flexibility, this impact is considered acceptable as the architecture can potentially exploit application-specific optimizations and offers a valuable platform for comparing programming models.

General information
State: Published
Organisations: Embedded Systems Engineering, Department of Informatics and Mathematical Modeling
Authors: Rasmussen, M. S. (Intern), Sparsø, J. (Intern), Karlsson, S. (Intern), Madsen, J. (Intern)
Publication date: Sep 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2010-235
Main Research Area: Technical/natural sciences
Electronic versions:
phd235_msr-a4.pdf

Relations
Projects:
Support for Programming Models in Network-on-Chip-based Many-core Systems
Source: orbit
Source-ID: 264623
Publication: Research › Ph.D. thesis – Annual report year: 2010

Homotopy Based Reconstruction from Acoustic Images
This thesis presents work in the direction of generating smooth surfaces from linear cross sections embedded in R2 and R3 using homotopy continuation. The methods developed in this research are generic and can be applied to higher dimensions as well. Two types of problems addressed in this research are reconstruction from an organised set of linear cross sections and reconstruction from an arbitrary set of linear cross sections. The first problem is looked upon in the
context of acoustic signals wherein the cross sections show a definite geometric arrangement. A reconstruction in this case can take advantage of the inherent arrangement. The problem of reconstruction from arbitrary cross sections is a generic problem and is also shown to be solved here using the mathematical tool of continuous deformations. As part of a complete processing, segmentation using level set methods is explored for acoustic images and fast GPU (Graphics Processing Unit) based methods are suggested for a streaming computation on large volumes of data. Validation of results for acoustic images is not straightforward due to unavailability of ground truth. Accuracy figures for the suggested methods are provided using phantom object with known geometry. The results of the methods shown here can be used to gain objective knowledge about the reconstructed features. It is envisioned that due to the generic nature of the algorithms developed in this research, domains other than fisheries research can benefit from the reconstruction algorithms.

**General information**
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Sharma, O. (Intern), Anton, F. (Intern), Christensen, N. J. (Intern)
Publication date: Sep 2010

**Longitudinal MRI studies of brain morphometry**
High resolution MR images acquired at multiple time points of the brain allow quantification of localized changes induced by external factors such as maturation, ageing or disease progression/recovery. High-dimensional warping of such MR images incorporates changes induced by external factors into the accompanying deformation field. Deformation fields from high dimensional warping founds tensor based morphometry (TBM), and provides unique opportunities to study human brain morphology and plasticity. In this thesis, specially adapted image processing streams utilizing several image registration techniques to characterize differences between brains, demonstrate the versatility and specificity of the employed voxel-wise morphometric methods. More specifically TBM is used to study neurodegenerative changes following severe traumatic brain injuries. Such injuries progress for months, perhaps even years postinjury. Little information is known about the spatial distribution and the clinical significance of this late atrophy. TBM revealed a large coherent cluster of significant atrophy consisting of the brain stem and cerebellar peduncles extending bilaterally through the thalamus, internal and external capsules, putamen, inferior and superior longitudinal fasciculus, corpus callosum and corona radiata. This indicates that the long-term atrophy is attributable to consequences of traumatic axonal injury. Despite progressive atrophy, remarkable clinical improvement occurred in most patients. The other study utilized TBM and voxel based morphometry (VBM) in two separate papers concerning antipsychotic-naïve first episode schizophrenia. Volume reductions of hippocampal and caudate regions were found in patients compared to controls using VBM. Six months later, TBM revealed continued volume loss in striatum and hippocampus, despite treatment with quetiapine. The mechanisms underlying these progressive brain dynamics, specific antipsychotic compounds and clinical symptoms warrant further clarification.

**General information**
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, Copenhagen University Hospital
Authors: Skimminge, A. J. M. (Intern), Larsen, R. (Intern), Baare, W. (Ekstern)
Publication date: Dec 2010

**Publication information**
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2010-238
Main Research Area: Technical/natural sciences
Electronic versions: phd238_as-net.pdf
Programming Models and Tools for Intelligent Embedded Systems

Design automation and analysis tools targeting embedded platforms, developed using a component-based design approach, must be able to reason about the capabilities of the platforms. In the general case where nothing is assumed about the components comprising a platform or the platform topology, analysis must be employed to determine its capabilities. This kind of analysis is the subject of this dissertation. The main contribution of this work is the Service Relation Model used to describe and analyze the flow of service in models of platforms and systems composed of reusable components. Fundamental to the service relation model is the novel concept of service aggregation that simply states that one service is accessible through another. The usefulness and versatility of the Service Relation Model is demonstrated by means of three different applications. In the first application, the model is used for checking the consistency of a design with respect to the availability of services and resources. In the second application, a tool for automatically implementing the communication infrastructure of a process network application, the Service Relation Model is used for analyzing the capabilities of a platform and as a basis for efficient code generation. In the third application, the Service Relation Model and the concept of consistency are used to guide an automated procedure for designing systems composed of components.

Modelling, Synthesis, and Configuration of Networks-on-Chips

This thesis presents three contributions in two different areas of network-on-chip and system-on-chip research: Application modelling and identifying and solving different optimization problems related to two specific network-on-chip architectures. The contribution related to application modelling is an analytical method for deriving the worst-case traffic pattern caused by an application and the cache-coherence protocol in a cache-coherent shared-memory system. The contributions related to network-on-chip optimization problems consist of two parts: The development and evaluation of six heuristics for solving the network synthesis problem in the MANGO network-on-chip, and the identification and formalization of the ReNoC configuration problem together with three heuristics for solving it.
Computing an Ontological Semantics for a Natural Language Fragment

The key objective of the research that has been carried out has been to establish theoretically sound connections between the following two areas: • Computational processing of texts in natural language by means of logical methods • Theories and methods for engineering of formal ontologies We have tried to establish a domain independent "ontological semantics" for relevant fragments of natural language. The purpose of this research is to develop methods and systems for taking advantage of formal ontologies for the purpose of extracting the meaning contents of texts. This functionality is desirable e.g. for future content–based search systems in contrast to today's keyword based search systems (viz., Google) which rely chiefly on recognition of stated keywords in the targeted text. Logical methods were introduced into semantic theories for natural language already during the 60's in what is today known as Montague semantics. However, this well–established tradition addresses mainly the domain independent logical structures of language such as quantifiers/determiners by means of logic [18], such as type theory [2]. By contrast this project focuses on the domain–specific parts of language (nouns, verbs, adjectives) introducing formal so–called generative ontologies as semantic target domains for noun– and verb phrases. Such a logico–semantic theory links the meaning of a sentence phrases to nodes in the chosen ontology for the domain.

General information
State: Published
Organisations: Algorithms and Logic, Department of Informatics and Mathematical Modeling, Copenhagen Business School
Authors: Szymczak, B. A. (Intern), Nilsson, J. F. (Intern), Jensen, P. A. (Ekstern)
Number of pages: 295
Publication date: 2010

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2010
Number: 242
Main Research Area: Technical/natural sciences
Electronic versions:

Wind turbine control and model predictive control for uncertain systems

This thesis presents both an applied study and a theoretical study within the field of control theory. Control theory is an interdisciplinary branch between mathematics and engineering dealing with the manipulation of systems to produce a desired output. The applied study deals with wind turbine control. Wind turbines are controlled to optimize energy extraction from the wind. This must be done while respecting physical restrictions and ensuring that loads on the wind turbine structure do not seriously reduce the lifetime of components. This poses a trade-off in the design and the wind turbine problem is hence a complex multivariable problem. In this thesis the main focus is on design of controllers which optimally attenuates the impact of the variability in the wind. The angles of the wind turbine blades have been used as the primary control variable to achieve this goal. Strategies have been studied in which the blades are controlled collectively and individually. The wind has both temporal and spatial variations with a stochastic nature. Furthermore, the wind has deterministic (or slowly varying) trends. Large parts of the thesis hence deals with developing wind models which can be used as disturbance models for controller design. The theoretical study deals with Model Predictive Control (MPC). MPC is an optimal control method which is characterized by the use of a receding prediction horizon. MPC has risen in popularity due to its inherent ability to systematically account for time-domain constraints on signals. During the last decades several theoretical advances have been made, so that it can handle a wide variety of system structures. In this thesis, the focus is on handling uncertain linear system description. To this end the so-called Youla parameterizations have been used. Two methods are proposed: The first method exploits the modularity of the parameterizations so that the uncertainty can be identified and the MPC controller can be reconfigured in a modular setting. The second method is a robust MPC method in which the Youla parameters are used as an integral part of the online optimization. In this way stability can be guaranteed given an assumed bound on the uncertainty. The contributions of the thesis have been documented in a series of scientific papers. The papers form the main part of this thesis.
Statistical Process Control in a Modern Production Environment

Paper 1 is aimed at practitioners to help them test the assumption that the observations in a sample are independent and identically distributed. An assumption that is essential when using classical Shewhart charts. The test can easily be performed in the control chart setup using the samples gathered here and standard statistical software.

In Paper 2 a new method for process monitoring is introduced. The method uses a statistical model of the quality characteristic and a sliding window of observations to estimate the probability that the next item will not respect the specifications. If the estimated probability exceeds a pre-determined threshold the process will be stopped. The method is flexible, allowing a complexity in modeling that remains invisible to the end user. Furthermore, the method allows to build diagnostic plots based on the parameters estimates that can provide valuable insight into the process. The method is explored numerically and a case study is provided. In Paper 3 the method is explored in a bivariate setting.

Paper 4 is a case study on a problem regarding missing values in an industrial process. The impact of the missing values on the quality measures of the process is assessed. Furthermore, guidelines along with software is provided to handle similar problems.

Aspects with Program Analysis for Security Policies

Enforcing security policies in IT systems, especially for a mobile distributed system, is challenging. As society becomes more IT-savvy, our expectations about security and privacy evolve. This is usually followed by changes in regulation in the form of standards and legislation. In many cases, small modification of the security requirement might lead to substantial changes in a number of modules within a large mobile distributed system. Indeed, security is a crosscutting concern which can spread to many business modules within a system, and is difficult to be integrated in a modular way. This dissertation explores the principles of adding challenging security policies to existing systems with great flexibility and modularity. The policies concerned cover both classical access control and explicit information flow policies. We built our solution by combining aspect-oriented programming techniques with static program analysis techniques. The former technique can
separate security concerns out of the main logic, and thus improves system modularity. The latter can analyze the system behavior, and thus helps detect software bugs or potential malicious code. We present AspectKE, an aspect-oriented extension of the process calculus KLAfM that excels at modeling mobile, distributed systems. A novel feature of our approach is that advices are able to analyze the future use of data, which is achieved by using program analysis techniques. We also present AspectKE to propose other possible aspect-oriented extensions based on KLAfM, followed by a discussion of open joinpoints that commonly exist in coordination languages such as KLAfM. Based on the idea of AspectKE, we design and implement a proof-of-concept programming language AspectKE*, which enables programmers to easily specify analysis-based security policies with the help of high-level program analysis predicates and functions. The prototype is efficiently realized by a two-stage implementation strategy and a static-dynamic dual value evaluation mechanism. We have performed two case studies to evaluate our programming model and language design. One application is based on a electronic health care workflow system. The other is a distributed chat system. We considered a number of security policies for both primary and secondary use of data, classical access control and predictive access control - control access based on the future behavior of a program. Some of the above mentioned policies can only be enforced by analysis of process continuations.

General information
State: Published
Organisations: Language-Based Technology, Department of Informatics and Mathematical Modeling
Authors: Yang, F. (Intern), Nielson, F. (Intern), Nielson, H. R. (Intern)
Publication date: 2010

Developing Virtual Power Plant for Optimized Distributed Energy Resources Operation and Integration
Distributed Energy Resources (DER) are small-scale power generation and storage technologies, (typically in the range of a few kW to tens of kW) located close to the customer side. They are right now under heavy development and have a great market potential in the near future. However, these sources are usually deployed in the way of “fit and forget” which to a great extent confines their value and presents challenges in relation to:
- Optimized DER operation related to time varying onsite demand requirements,
- Ambient conditions and electricity prices, etc.
- Coordinated control of many small units in the electric power system
- Efficient electricity market participation to benefit both power system operation and DER owners
To address these issues, an innovative concept Virtual Power Plant (VPP) is investigated in this PhD study. Based on a comprehensive overview of the state of the art of VPP, the Market-Based VPP (MBVPP) concept is proposed. The function-based MBVPP provides a generic and flexible solution for the DER integration by connecting the DER to the bulk power system operation via market participation.
Two schemes for managing the DER generation and trading portfolios, direct control and price signal control have been discussed and simulated. Due to their prevalence and controllability, the μCHP systems are modeled to represent the general DER technology in the corresponding studies. For the direct controlled VPP, all the μCHP units are optimally controlled by the VPP operator based on forecasted market and demand information. For the proposed price signal scheme, an Artificial Neural Network (ANN) is developed to characterize and estimate the price responsiveness of a μCHP group. It is found that although the prognosis result is relatively good, the price signal controlled scheme is still challenged by many uncertainties which reside in the nature of price signal control such as jumpy response. To demonstrate the feasibility of the VPP, a prototyped VPP with two Dachs μCHP systems is set up in the laboratory as a proof of concept. It has shown that, on the premise of an advanced Information and Communication Technology (ICT) infrastructure, the VPP represents a feasible solution to be implemented.

General information
State: Published
Organisations: Electric Components, Department of Electrical Engineering, Software Engineering, Department of Informatics and Mathematical Modeling
Authors: You, S. (Intern), Traeholt, C. (Intern), Poulsen, B. (Intern)
Number of pages: 148
Publication date: Sep 2010
Enhancing Media Personalization by Extracting Similarity Knowledge from Metadata

The world of media today can be characterized by us being exposed to vast amounts of content, both produced professionally and user generated. Ever since the digital technologies in the form of computers and video cameras have diminished the production costs and the Internet has significantly lowered the costs of distribution, we became more and more overwhelmed with the choice of media. In such conditions the focus falls on the available mechanisms to filter and recommend media to users, thus resulting in the growing need for personalization. Media personalization is a complex process with many interrelated parts – recommendation engines, content metadata, contextual information and user profiles. In the center of any type of recommendation lies the notion of similarity. The most popular way to approach similarity is to look for the feature overlaps. This results often in recommending only “more of the same” type of content which does not necessarily lead to the meaningful personalization. Another way to approach similarity is to find a similar underlying meaning in the content. Aspects of meaning in media can be represented using Gardenfors Conceptual Spaces theory, which can be seen as a cognitive foundation for modeling concepts. Conceptual Spaces is applied in this thesis to analyze media in terms of its dimensions and knowledge domains, which in return defines properties and concepts. One of the most important domains in terms of describing media is the emotional one, especially when we talk about the contents of music. Therefore the main focus in the thesis is how to extract such emotional information from media, and how to use it to enhance media personalization. This dissertation proposes a novel method to extract emotional information from text (unstructured metadata) using Latent Semantic Analysis (one of the unsupervised machine learning techniques). It presents three separate cases to illustrate the similarity knowledge extraction from the metadata, where the emotional components in each case represents different abstraction levels – genres, synopsis and lyrics. The emotional value is extracted by first creating a conceptual space for emotions based on a semantic differential which divides the underlying plane along two psychological dimensions – arousal and valence. Then the space is divided into regions serving as emotional markers – a selection of affective terms. After that LSA is used to calculate the cosine similarity between the text (synopsis or lyrics) and each of the chosen affective terms. As a result we can plot emotional correlation in the content as patterns, which we can then use to find emotional similarity among media items. By being able to compare media items on the basis of their emotional patterns, we add a new level to how we can evaluate the similarity between two media items. Which in return might improve media recommendation since it provides a novel approach to recommendation that goes beyond traditional genre boundaries, and thereby improves media personalization.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling, CICT
Authors: Butkus, A. (Intern), Olesen, H. (Intern)
Publication date: Mar 2009

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2008-198
Main Research Area: Technical/natural sciences
Electronic versions:
phd198_ab-final.pdf
Source: orbit
Source-ID: 220256
Publication: Research › Ph.D. thesis – Annual report year: 2009

Computer Vision for Timber Harvesting
The goal of this thesis is to investigate computer vision methods for timber harvesting operations. The background for developing computer vision for timber harvesting is to document origin of timber and to collect qualitative and quantitative parameters concerning the timber for efficient harvest planning. The investigations in this thesis is done as initial work on a planning and logistic system for timber harvesting called logTracker. In this thesis we have focused on three methods for the logTracker project, which includes image segmentation, image classification, and image retrieval. Segmentation is to partition an image based on image characteristics and in our study we have focused on image texture. Our segmentation
method is inspired by iterative function systems and contractive maps, which makes the basis for both our texture
characterization and our method for obtaining the image segments. The purpose of image segmentation is to make the
basis for more advanced computer vision methods like object recognition and classification. Our second method concerns
image classification and we present a method where we classify small timber samples to tree species based on Active
Appearance Models and texture characteristics. The last method is image retrieval based on the so called "bag of visual
words" procedure. An image is characterized as a distribution of local image descriptors, which is the basis for effective
image search. These methods are described and discussed in relation to the logTracker project and ideas for further
development of the system is provided. Building a complete logTracker system is a very demanding task and the
conclusion is that it is important to focus on the elements that can bring most value to timber harvest planning. Besides
crating to the development of the logTracker system the described methods have a general applicability making them
useful for many other computer vision problems.

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, DTU Data
Analysis
Authors: Dahl, A. L. (Intern), Ersbøll, B. K. (Intern), Aanæs, H. (Intern)
Publication date: May 2009

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2009-214
Main Research Area: Technical/natural sciences
Electronic versions:
phd214_abd.pdf
Source: orbit
Source-ID: 238816
Publication: Research › Ph.D. thesis – Annual report year: 2009

Shape and Deformation Analysis of the Human Ear Canal

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Darkner, S. (Intern), Olsen, O. F. (Ekstern), Paulsen, R. R. (Intern), Larsen, R. (Intern)
Publication date: Jan 2009

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark, DTU Informatics, Building 321
Original language: English
Series: IMM-PHD-2008-204
Main Research Area: Technical/natural sciences
Electronic versions:
phd204_pdf
Source: orbit
Source-ID: 222732
Publication: Research › Ph.D. thesis – Annual report year: 2009

Modeling the Biological Diversity of Pig Carcasses
This thesis applies methods from medical image analysis for modeling the biological diversity of pig carcasses. The
Danish meat industry is very focused on improving product quality and productivity by optimizing the use of the carcasses
and increasing productivity in the abattoirs. In order to achieve these goals there is a need for more detailed information
about pig carcasses in relation to measures of quality. Non-invasive imaging such as X-ray Computed Tomography (CT)
can provide this very detailed information discerning the major tissue types. Medical image analysis provides the tools for
extracting and modeling meaningful information from the vast amount of information available from non-invasive imaging
data. The lean meat percentage (LMP) is a common standard for measuring the quality of pig carcasses. Measuring the
LMP using CT and using this as a reference for calibration of online equipment is investigated, without the need for a
calibration against a less accurate manual dissection. The rest of the contributions regard the construction and use of
point distribution models (PDM). PDM’s are able to capture the shape variation of a population of shapes, in this case a
3D surface of a specific bone structure in the ham. These models can assist developers of robotic tools by enabling
population based testing before actual construction of the tools. Sparse models are compared to the standard PCA based model and a method for fitting PDM’s to sparse data is proposed. The former provides more spatially localized modes of variation that are easier interpretable and the latter enables the use of PDM’s without the need for full point correspondence of new data. There is great potential in applying CT as non-invasive modality in the meat industry, e.g. in population based studies, for shape modeling and for analyzing carcass composition. In the future online CT applications can be used to make decisions on the use of each specific carcass by obtaining improved quality measures.

**General information**

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, DTU Data Analysis, Department of Photonics Engineering
Authors: Erbou, S. G. H. (Intern), Ersbøll, B. K. (Intern), Larsen, R. (Intern), Christensen, L. B. (Intern)
Publication date: Feb 2009

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2008-207
Main Research Area: Technical/natural sciences
Electronic versions:

phd207_sge.pdf
Source: orbit
Source-ID: 228351
Publication: Research › Ph.D. thesis – Annual report year: 2009

---

**Data Assimilation in Marine Models**

This thesis consists of six research papers published or submitted for publication in the period 2006-2009 together with a summary report. The main topics of this thesis are nonlinear data assimilation techniques and estimation in dynamical models. The focus has been on the nonlinear filtering techniques for large scale geophysical numerical models and making them feasible to work with in the data assimilation framework. The filtering techniques investigated are all Monte Carlo simulation based. Some very nice features that can be exploited in the Monte Carlo based data assimilation framework from a computational point of view, e.g. low storage cost, no linearizations of the numerical models, etc. However, this also gives rise to many unforeseen difficulties, e.g. the curse of dimensionality, huge computational costs, etc. The challenge faced in this thesis was finding filters that could handle the nonlinearities encountered in data assimilation and at the same time are robust and reliable enough given the constraints and difficulties that can arise. These problems were addressed in the papers A, E and D. The other topic of this thesis is estimation in dynamical geophysical numerical models. The challenge of estimating model parameters for well establish geophysical dynamical systems is that these models are not formulated in a way that incorporates the necessary stochastic assumptions that make estimation possible in a maximum likelihood sense. The maximum likelihood approach is selected due to its unique performance in data rich situations. The estimations are often based on output from the model and the raw observations which lead to suboptimal estimates. The challenge is to give a meaningful description of the model errors through diffusion processes that can be identified and incorporated into the existing maximum likelihood framework. These issues are discussed in paper B. The third part of the thesis falls a bit out of the above context is work published in papers C, F. In the first paper, a simple data assimilation scheme was investigated to examine the potential benefits of incorporating a data assimilation concept into an atmospheric chemical transport model. This paper deals with the results and conclusions obtained through some of the first experiments with the Optimal Interpolation filter in a geophysical model. The second paper F, deals with the construction of a finite element solver for the Fokker-Planck equation on a 2 dimensional flexible mesh system. The report details the construction of the finite element solver and investigates the potential benefits of a parallel FORTRAN implementation through a series of experiments.

**General information**

State: Published
Organisations: Scientific Computing, Department of Informatics and Mathematical Modeling, Mathematical Statistics
Authors: Frydendall, J. (Intern), Madsen, H. (Intern), Sørensen, J. V. T. (Intern)
Publication date: Nov 2009

**Publication information**

Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English

Series: IMM-PHD-2009-217
Main Research Area: Technical/natural sciences
Electronic versions:
On parameterized deformations and unsupervised learning

The work presented here consists of contributions in three areas. An efficient algorithm for calculating the entire regularization path of the support vector domain description (SVDD) is presented. The ability to calculate the entire path with a complexity in the same order as solving the original quadratic problems gives inspiration to utilize the extra information available from the entire path. A method for hierarchical support vector clustering, based on information from the entire regularization path, and multiple Gaussian kernels is described. Bayesian methods are applied in the attempt to draw direct statistical conclusions from the SVDD analysis.

In the context of image registration, different assumptions on the warp fields, namely diffeomorphism and a linear elastic potential in the form of regularization are discussed. A new warp representation which allows statistical analysis on an unrestricted linear parameter space, where all derivatives are defined, is introduced. Furthermore, it is shown that L2-norm the parameter space introduces a reasonable metric in the actual space of modelled diffeomorphisms. A new parametrization of 3D deformation fields, using potentials and Helmholtz decomposition is also presented. The representation can be considered a natural parametrization for both elastic and fluid image registration due to the decoupling of the parameters. The determinant gradient field is shown to be the first-order small-deformation approximation to the determinant of the Jacobian matrix.

Spline approximations of functions and in particular image registration warp fields are discussed. It is shown how spline bases may be learned from the optimization process, i.e. image registration optimization, and how this may contribute with a reasonable prior, or regularization in the method. A new formula, based on the multivariate divided difference, for explicit calculation of the simplex splines is presented. The formula additionally admits easy calculation of derivatives, both spatial, and with respect to the position of the knots. It is demonstrated that conditions may be set on the knot movements, which ensures that the splines form a partition of unity, even if the knots are not Delaunay. A subdivision scheme is also presented, which requires no recalculation of the configurations of the splines. The use of the splines for image registration is demonstrated, and the inherent smoothing or averaging cost, of selecting warp parameterizations at a specific kernel resolution, has been analyzed. A refinement measure has been derived, which is shown to be efficient for guiding the local mesh layout. With the combination of the refinement measure and the local flexibility of the multivariate B-splines, the warp field is automatically refined in areas where it results in the minimization of the registration cost function.

General information

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, DTU Data Analysis
Authors: Hansen, M. S. (Intern), Larsen, R. (Intern), Ersbøll, B. K. (Intern)
Number of pages: 144
Publication date: Nov 2009

Publication information

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2009-219
Main Research Area: Technical/natural sciences
Electronic versions:
phd219_Hansen_MS.pdf
Source: orbit
Source-ID: 249932
Publication: Research › Ph.D. thesis – Annual report year: 2009

The virtual knife

Since post World War II and until 2008 the Danish pig producing industry (DPPi) has been in a continuing state of growth. In spite of an ever fiercer competition DPPi has managed to protect its position as export leader by maintaining a focus on research and development. Today, DPPi is in a state of recession and must increase the efficiency if not to reduce the production capacity further. The industry recognizes that a more efficient use of the raw materials is one of the largest and most important challenges. To meet this challenge it is a necessity to get a better understanding of the biological variation of pigs. The development of models for describing the biological variation of pigs is one of the key components needed to attain a better sorting of the pig carcasses and an improved cutting in the abattoirs. Such models can be related to possible products, which can be related to potential yield and order books. The Danish Meat Research Institute (DMRI) is
Predictive tools for designing new insulins and treatment regimens

The thesis deals with the development of "Predictive tools for designing new insulins and treatments regimens" and consists of two parts: A model based approach for bridging properties of new insulin analogues from glucose clamp experiments to meal tolerance tests (MTT) and a second part that describes an implemented software program able to handle stochastic differential equations (SDEs) with mixed effects. The thesis is supplemented with scientific papers published during the PhD. Developing an insulin analogue from candidate molecule to a clinical drug consists of a development programme including different phases targeting safety and efficacy. The focus of this thesis is the shift from Phase I, targeting safety, to Phase II, targeting efficacy. An insulin analogue is typically tested for safety in glucose clamp experiments in Phase I clinical trials and progresses into Phase II where dose and efficacy are investigated. Numerous methods are used to quantify dose and efficacy in Phase II - especially of interest is the 24-hour meal tolerance test as it tries to portray near normal living conditions. Part I describes an integrated model for insulin and glucose which is aimed at simulating 24-hour glucose profiles from a MTT with treatments based on the new insulin analogue that previously only has been tested in clamps. The bridge between insulin analogue properties determined in clamp experiments to meal tolerance test outcomes in Phase II trials is not simple and is complicated by shifts in experimental setup, time horizon and treatment regimen. A bridging strategy was introduced where an integrated model simulating MTTs was extended with models developed on clamp data that described PK and PD for the new insulin analogue. The bridging strategy was tested by building an integrated model based on human insulin trials which was then evaluated using insulin Aspart (IAsp). The integrated model was estimated in two separate sub models due to computational complexity. Insulin model challenges were faced at the estimation step regarding separability of insulin input pathways (exogenous/secretion) which resulted in several fixed parameters but also an insulin delivery model as opposed to a prehepatic insulin secretion model coupled with hepatic extraction. The glucose model was an extended version of the oral glucose minimal model [Man et al., 2002] which had a meal function incorporated. The two sub models were combined into an integrated model which was evaluated in different scenarios: An iso-glucemic glucose clamp, an insulin tolerance test and comparing derived measures of glucose effectiveness. The model evaluation pinpointed insulin sensitivity issues which were accommodated with a change in model building towards a more insulin sensitive model type. Conclusively, the integrated model fitted estimation data well both for insulin and glucose. Furthermore, the evaluation scenarios showed overall correspondence with literature with only minor discrepancies. The evaluation on insulin Aspart required a PK model for IAsp and a model describing IAsp action in MTTs. The IAsp PK model was available from a different Novo Nordisk project and the action transfer function was estimated on cross-over clamp data with human insulin and insulin Aspart. The two components were then embedded into the integrated model. The extended integrated model was used to simulate 24-hour profiles of insulin and glucose from meal tolerance tests including treatments with biphasic insulin Aspart. The evaluation showed that the extended integrated model was able to predict insulin levels reasonably both mean profile and variation whereas glucose proles were not predicted accurately. Post modelling analysis targeting both insulin and glucose components showed that preconditions for the bridging strategy which implied the use of a mean IAsp PK model, could be the cause for the mis-predictions. Future research should look into ways for individualising the insulin treatment when no information on individual level is present. The model building process could have benefitted from the use of SDEs. Unfortunately, availability of a software program able to handle mixed effects and SDEs resulted in a modelling approach based on ordinary differential equations. The absence of such a program motivated the development of new a tool with PK/PD features, SDEs and mixed effects. Part II presents a software package which was developed in order to be able to handle...
SDEs with mixed effects. The package was implemented in R which allowed for a single environment for data preparation, model building and results handling but also provided accessibility for users and ease of installation. The R-package implements the (Extended) Kalman Filter for handling SDEs and uses the FOCE approximation to calculate the marginal likelihood for parameters used in maximum likelihood estimation. A number of applications of PSM are presented in which deconvolution is the topic for most. Deconvolution based on SDEs was used to determine pre-hepatic insulin secretion rates; hepatic insulin extraction rates using both insulin and C-peptide measurements, and glucose appearance rates constrained to be in the positive range in a simulated minimal model setting. More applications included an insulin secretion model based on an intervention model type and an analysis of in fluence from input error propagation as estimated with ODEs and SDEs.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling, Department of Chemical and Biochemical Engineering
Authors: Klim, S. (Intern), Madsen, H. (Intern), Kristensen, N. R. (Intern)
Publication date: Dec 2009

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark, DTU Informatics, Building 321
Original language: English
Series: IMM-PHD-2009-216
Main Research Area: Technical/natural sciences
Electronic versions:
phd216_skl-net.pdf

Bibliographical note
The PhD project was an industrial PhD programme with a collaboration between Department for informatics and mathematical modelling (IMM), Technical University of Denmark and Biomodelling, Novo Nordisk A/S (NN)
Source: orbit
Source-ID: 248455
Publication: Research › Ph.D. thesis – Annual report year: 2009

Modelling Strategies for Functional Magnetic Resonance Imaging
This thesis collects research done on several models for the analysis of functional magnetic resonance neuroimaging (fMRI) data. Several extensions for unsupervised factor analysis type decompositions including explicit delay modelling as well as handling of spatial and temporal smoothness and generalisations to higher order arrays are considered. Additionally, an application of the natural conjugate prior for supervised learning in the general linear model to efficiently incorporate prior information for supervised analysis is presented. Further extensions include methods to model nuisance effects in fMRI data thereby suppressing noise for both supervised and unsupervised analysis techniques.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Madsen, K. H. (Intern), Sidaros, K. (Intern), Hansen, L. K. (Intern)
Publication date: Jul 2009

Publication information
Original language: English
Series: IMM-PHD-2008-203
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd203_khm-net-uden_artikler.pdf
Source: orbit
Source-ID: 220839
Publication: Research › Ph.D. thesis – Annual report year: 2009

Context based multimedia information retrieval
The large amounts of digital media becoming available require that new approaches are developed for retrieving, navigating and recommending the data to users in a way that reflects how we semantically perceive the content. The thesis investigates ways to retrieve and present content for users with the help of contextual knowledge. Our approach to model the context of multimedia is based on unsupervised methods to automatically extract meaning. We investigate two paths of context modelling. The first part extracts context from the primary media, in this case broadcast news speech, by
extracting topics from a large collection of the transcribed speech to improve retrieval of spoken documents. The context modelling is done using a variant of probabilistic latent semantic analysis (PLSA), to extract properties of the textual sources that reflect how humans perceive context. We perform PLSA through an approximation based on non-negative matrix factorisation NMF. The second part of the work tries to infer the contextual meaning of music based on extramusical knowledge, in our case gathered from Wikipedia. The semantic relations between artists are inferred using linking structure of Wikipedia, as well as text-based semantic similarity. The final aspect investigated is how to include some of the structured data available in Wikipedia to include temporal information. We show that a multiway extension of PLSA makes it possible to extract temporally meaningful topics, better than using a stepwise PLSA approach to topic extraction.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Mølgaard, L. L. (Intern), Hansen, L. K. (Intern), Larsen, J. (Intern)
Publication date: Dec 2009

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2009-218
Main Research Area: Technical/natural sciences
Electronic versions:
phd218_lim.pdf
Source: orbit
Source-ID: 248008
Publication: Research › Ph.D. thesis – Annual report year: 2009

A Verifiable Language for Cryptographic Protocols
We develop a formal language for specifying cryptographic protocols in a structured and clear manner, which allows verification of many interesting properties; in particular confidentiality and integrity. The study sheds new light on the problem of creating intuitive and human readable languages, that are analysable with respect to interesting properties. Furthermore it motivates and is an example of, a novel, more general methodology of language design by first verbosely describing the semantics in a mathematical language, e.g. a logic, then restricting the properties of interest to be computable, and finally systematically transforming it into a more intuitive specification language, maintaining this tractability.

General information
State: Published
Organisations: Language-Based Technology, Department of Informatics and Mathematical Modeling
Authors: Nielsen, C. R. (Intern), Nielson, F. (Intern), Nielson, H. R. (Intern)
Publication date: May 2009

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2008-210
Main Research Area: Technical/natural sciences
Electronic versions:
phd210.pdf
Source: orbit
Source-ID: 229299
Publication: Research › Ph.D. thesis – Annual report year: 2009

Structure Learning In Audio
By having information about the setting a user is in, a computer is able to make decisions proactively to facilitate tasks for the user. Two approaches are taken in this thesis to achieve more information about an audio environment. One approach is that of classifying audio, and a new approach using pitch dynamics is suggested. The other approach is finding structures between the mixings of multiple sources based on an assumption of statistical independence of the sources. Three different audio classification tasks have been investigated. Audio classification into three classes, music, noise and speech, using novel features based on pitch dynamics. Within instrument classification two different harmonic models have been compared. Finally voiced/unvoiced segmentation of popular music is done based on MFCC’s and AR coefficients. The structures in the mixings of multiple sources have been investigated. A fast and computationally simple
approach that compares recordings and classifies if they are from the same audio environment have been developed, and shows very high accuracy and the ability to synchronize recordings in the case of recording devices which are not connected. A more general model is proposed based on Independent Component Analysis. It is based on sequential pruning of the parameters in the mixing matrix and a version based on a fixed source distribution as well as a parameterized distribution is found. The parameterized version has the advantage of modeling both sub- and super-Gaussian source distributions allowing a much wider use of the method. All methods uses a variety of classification models and model selection algorithms which is a common theme of the thesis.
Adaptive Text Entry for Mobile Devices
The reduced size of many mobile devices makes it difficult to enter text with them. The text entry methods are often slow or complicated to use. This affects the performance and user experience of all applications and services on the device. This work introduces new easy-to-use text entry methods for mobile devices and a framework for adaptive context-aware language models. Based on analysis of current text entry methods, the requirements to the new text entry methods are established. Transparent User guided Prediction (TUP) is a text entry method for devices with one dimensional touch input. It can be touch sensitive wheels, sliders or similar input devices. The interaction design of TUP is done with a combination of high level task models and low level models of human motor behaviour. Three prototypes of TUP are designed and evaluated by more than 30 users. Observations from the evaluations are used to improve the models of human motor behaviour. TUP-Key is a variant of TUP, designed for 12 key phone keyboards. It is introduced in the thesis but has not been implemented or evaluated. Both text entry methods support adaptive context-aware language models. YourText is a framework for adaptive context-aware language models that is introduced in the thesis. YourText enables different language models to be combined to a new common language model. The framework is designed so it can be adapted to different text entry methods, thereby enabling the language model to be transferred between devices. YourText is evaluated with a corpus of mobile text messages. The corpus is created by collecting all sent and received messages from 12 persons in four weeks. The corpus contains 25,000 messages. A model of text entry speed for TUP is created from the observations in the evaluations. The model is used to predict the performance of TUP, used together with different YourText language models.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Proschowsky, M. S. (Intern), Schultz, N. (Intern), Hansen, L. K. (Intern)
Publication date: Jul 2009

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2008-209
Main Research Area: Technical/natural sciences
Electronic versions:
phd209.pdf
Source: orbit
Source-ID: 231661
Publication: Research › Ph.D. thesis – Annual report year: 2009

Single-channel source separation using non-negative matrix factorization
Single-channel source separation problems occur when a number of sources emit signals that are mixed and recorded by a single sensor, and we are interested in estimating the original source signals based on the recorded mixture. This problem, which occurs in many sciences, is inherently under-determined and its solution relies on making appropriate assumptions concerning the sources. This dissertation is concerned with model-based probabilistic single-channel source separation based on non-negative matrix factorization, and consists of two parts: i) three introductory chapters and ii) five published papers. The first part introduces the single-channel source separation problem as well as non-negative matrix factorization and provides a comprehensive review of existing approaches, applications, and practical algorithms. This serves to provide context for the second part, the published papers, in which a number of methods for single-channel source separation based on non-negative matrix factorization are presented. In the papers, the methods are applied to separating audio signals such as speech and musical instruments and separating different types of tissue in chemical shift imaging.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Schmidt, M. N. (Intern), Larsen, J. (Intern)
Publication date: Jan 2009

Publication information
Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark, DTU Informatics, Building 321
Original language: English
Cryptographic Hash Functions

Cryptographic hash functions are commonly used in many different areas of cryptography: in digital signatures and in public-key cryptography, for password protection and message authentication, in key derivation functions, in pseudo-random number generators, etc. Recently, cryptographic hash functions have received a huge amount of attention due to new attacks on widely used hash functions. This PhD thesis, having the title "Cryptographic Hash Functions", contains both a general description of cryptographic hash functions, including their applications and expected properties as well as some well-known designs, and also some design and cryptanalysis in which the author took part. The latter includes a construction method for hash functions and four designs, of which one was submitted to the SHA-3 hash function competition, initiated by the U.S. standardisation body NIST. It also includes cryptanalysis of the construction method MDC-2, and of the hash function MD2.

Image Registration and Optimization in the Virtual Slaughterhouse

This thesis presents the development and application of algorithms for the analysis of pig carcasses. Focus is on the simulation and quality estimation of meat products produced in a Danish slaughterhouse. Computed Tomography scans of pig carcasses provide the data used in the application. Image analysis is applied in order to imitate some of the cutting processes found in a slaughterhouse but also to give a quantitative measure of the composition of each carcass. The basis of the algorithms is non-linear image registration. This method finds the anatomical correspondence between a reference carcass and a template carcass. By iteratively comparing the transformed template with the reference a resulting dense deformation field is found. Propagating a set of landmarks from the reference coordinate system onto the template enables the simulation of slaughtering processes. Non-invasively estimating the quality of the slaughtering products provides a very valuable tool for use in the slaughterhouse in the future.
Agent Based Individual Traffic Guidance
This thesis investigates the possibilities in applying Operations Research (OR) to autonomous vehicular traffic. The explicit difference to most other research today is that we presume that an agent is present in every vehicle - hence Agent Based Individual Traffic guidance (ABIT). The next evolutionary step for the in-vehicle route planners is the introduction of two-way communication. We presume that the agent is capable of exactly this. Based on this presumption we discuss the possibilities and define a taxonomy and use this to discuss the ABIT system. Based on a set of scenarios we conclude that the system can be divided into two separate constituents. The immediate dispersion, which is used for small areas and quick response, and the individual alleviation, which considers the longer distance decision support. Both of these require intrinsinc models and cost functions which at the beginning of the project were not previously considered. We define a special inseparable cost function and develop a solution complex capable of using this cost function. In relation to calibration and estimation of statistical models used for dynamic route guidance we worked with generating random number sequences. During this work we made significant findings related to random numbers.

Constraint Solver Techniques for Implementing Precise and Scalable Static Program Analysis
As people rely on all kinds of software systems for almost every aspect of their lives, how to ensure the reliability of software is becoming more and more important. Program analysis, therefore, becomes more and more important for the software development process. Static program analysis helps developers to build reliable software systems more quickly and with fewer bugs or security defects. While designing and implementing a program analysis remains a hard work, making it both scalable and precise is even more challenging. In this dissertation, we show that with a general inclusion constraint solver using unification we could make a program analysis easier to design and implement, much more scalable, and still as precise as expected. We present an inclusion constraint language with the explicit equality constructs for specifying program analysis problems, and a parameterized framework for tuning a constraint system. Implementing an analysis is simplified as generating a set of constraints to be complied with. The equality constraints specifies equivalent analysis variables and thereby could be taken advantage of by a constraint solver to reduce a problem space and improve performance. We show a good balance between performance and precision could be achieved with the framework. We also introduce off-line optimizations for a general constraint solver. The optimizations automatically efficiently detect (potential) equivalent classes. With our case studies on a C pointer analysis, and two data flow analyses for C language, we demonstrate a large amount of equivalences could be detected by off-line analyses, and they could then be used by a constraint solver to significantly improve the scalability of an analysis without sacrificing any precision.
Minimum Mean-Square Error Single-Channel Signal Estimation

This topic of this thesis is MMSE signal estimation for hearing aids when only one microphone is available. The research is relevant for noise reduction systems in hearing aids. To fully benefit from the amplification provided by a hearing aid, noise reduction functionality is important as hearing-impaired persons in some noisy situations need a higher signal to noise ratio for speech to be intelligible when compared to normal-hearing persons. In this thesis two different methods to approach the MMSE signal estimation problem is examined. The methods differ in the way that models for the signal and noise are expressed and in the way the estimator is approximated. The starting point of the first method is prior probability density functions for both signal and noise and it is assumed that their Laplace transforms (moment generating functions) are available. The corresponding posterior mean integral that defines the MMSE estimator is rewritten into an inverse Laplace transform integral over an integrand involving the moment generating functions. This integral is approximated using saddlepoint approximation. It is found that the saddlepoint approximation becomes inaccurate when two saddlepoints coalesce and a saddlepoint approximation based on two coalescing saddlepoints is derived. For practical reasons the method is limited to low dimensional problems and the results are not easily extended to the multivariate case. In the second approach signal and noise are specified by generative models and approximate inference is performed by particle filtering. The speech model is a time-varying auto-regressive model reparameterized by formant frequencies and bandwidths. The noise is assumed non-stationary and white. Compared to the case of using the AR coefficients directly then it is found very beneficial to perform particle filtering using the reparameterized speech model because it is relative straightforward to exploit prior information about formant features. A modified MMSE estimator is introduced and performance of the particle filtering algorithm is compared to a state of the art hearing aid noise reduction algorithm. Although performance of the two algorithms is found comparable then the particle filter algorithm is doing a much better job tracking the noise.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Beierholm, T. (Intern), Winther, O. (Intern)
Publication date: Apr 2008

Publication information
Original language: English
Series: IMM-PHD-2007-185
Main Research Area: Technical/natural sciences

Detection of Weather Radar Clutter

Weather radars provide valuable information on precipitation in the atmosphere but due to the way radars work, not only precipitation is observed by the weather radar. Weather radar clutter, echoes from non-precipitating targets, occur frequently in the data, resulting in lowered data quality. Especially in the application of weather radar data in quantitative precipitation estimation and forecasting a high data quality is important. Clutter detection is one of the key components in achieving this goal. This thesis presents three methods for detection of clutter. The methods use supervised classification and use a range of different techniques and input data. The first method uses external information from multispectral satellite images to detect clutter. The information in the visual, near-infrared, and infrared parts of the spectrum can be used to distinguish between cloud and cloud-free areas and precipitating and non-precipitating clouds. Another method uses the difference in the motion field of clutter and precipitation measured between two radar images. Furthermore, the direction of the wind field extracted from a weather model is used. The third method uses information about the refractive index of the atmosphere as extracted from a numerical weather prediction model to predict the propagation path of the radar’s electromagnetic energy. This facilitates the prediction of areas of clutter caused by anomalous propagation of the radar’s rays. The methods are evaluated using a large independent test set, and to illustrate the performance on individual radar images three typical case examples are also evaluated. The results of the evaluation of the methods show that each method has good skill in detection of clutter with an average classification accuracy of 95%. The methods thus have the potential for increasing the quality of weather radar data in their operational use.

General information
User Involvement in the innovation process: In a mobile service and application development perspective

User involvement in the innovation process is not a new phenomenon. However, combined with the growing individualisation of demand and with highly competitive and dynamic environments, user involvement in the innovation process and thereby in the design, development, and manufacturing process, can nevertheless provide a competitive advantage. This is the case as an intensified user involvement in the innovation process potentially results in a more comprehensive understanding of the user needs and requirements and the context within which these are required, and thereby provides the possibility of developing better and more suitable products. The theoretical framework of this thesis is based on user involvement in the innovation process and how user involvement in the innovation process can be deployed in relation to deriving and collecting user needs and requirements, and thereby serves as a foundation for developing better and more supportive service and application concepts within the information and communication technology domain. Three interrelated research areas are deployed within the theoretical framework, which combined constitute and highlight the intertwined and complex interaction of people and their use of information and communication technologies: mobile system requirements, mobility, and the concept of Personal Networks. The mobile system requirement perspective is related to providing a more user oriented research approach, which historically has not been the case. The mobility perspective is related to categorising and conceptualising the term mobility in a combined user need and requirement perspective and service and application development perspective. The Personal Network concept perspective is seen in relation to the introduction of new information and communication technologies, and in particular in relation to derive and collect user needs and requirements. Within the theoretical framework of user involvement in the innovation process the concepts of lead users, sticky information, and toolkits, and the usages and perspectives of these root in a method development and testing perspective, have been deployed to derive user needs and requirements within two case studies: a diabetes case and a journalist case. The diabetes case has been conducted in collaboration with a diabetes treatment centre and diabetics and the journalist case in collaboration with the sports department at a large Danish broadcasting company, both with the main objective of deriving and collecting real user needs and requirements and based on these to develop service and application concepts which support diabetics and journalists in their activities and tasks. The lead user theory has been deployed as it indicates that users residing on the leading edge of any given market, technology, etc. are more likely to develop innovations compared to more ordinary users as the lead users will be experiencing needs and requirements presently, which the ordinary users will not experience until later. Sticky information denotes the transferability of a given unit of information, which in this context is related to the transferability of user needs and requirements. The deployment of different toolkits has been related to transferring sticky information (user needs and requirements) into less sticky information and thereby shifting the deriving and collecting of user needs and requirements into the user domain. This thesis shows how the deployment of the lead user, sticky information, and toolkit methods combined with some more traditional approaches and in relation to the two case studies have proven to provide a more detailed and context related understanding of the user needs and requirements within the two case segments. Furthermore the mobility and context related aspects of user needs and requirements have been deployed and incorporated into the gathering and collecting process, and have provided valuable insights in relation to the developed future service and application concepts, which are based on real user needs, requirements, mobility, and contexts. All with the purpose of deriving user needs and requirements and thereby develop and describe the concepts for future services and applications, which support these users in their everyday life, tasks, and contexts – value innovation.
Modelling Brain Tissue using Magnetic Resonance Imaging

Diffusion MRI, or diffusion weighted imaging (DWI), is a technique that measures the restricted diffusion of water molecules within brain tissue. Different reconstruction methods quantify water-diffusion anisotropy in the intra- and extracellular spaces of the neural environment. Fibre tracking models then use the directions of greatest diffusion as estimates of white matter fibre orientation. Several fibre tracking algorithms have emerged in the last few years that provide reproducible visualizations of three-dimensional fibre bundles. One class of these algorithms is probabilistic tractography. Although probabilistic tractography currently holds great promise as a powerful non-invasive connectivity-measurement tool, its accuracy and limitations remain to be evaluated. Probabilistic tractography was assessed post mortem in an in vitro environment. Postmortem DWI benefits from the possibility of using high-field experimental MR scanners and long scanning times, thereby significantly improving the signal-to-noise ratio (SNR) and anatomical resolution. Moreover, many of the degrading effects observed in vivo, such as physiological noise, are no longer present. However, the post mortem environment differs from that of in vivo both due to a lowered environmental temperature and due to the fixation process itself. We argue that the perfusion fixation procedure employed in this thesis ensures that the postmortem tissue is as close to that of in vivo as possible. Different fibre reconstruction models were tested on a range of different b-values (a b-value is a summary measurement of the strength of the applied diffusion gradients). We conclude that for robust reconstruction of fibre directions, and subsequently for tractography, b-values in the range of ~2000 s/mm² and ~8000 s/mm² should be used. Within a two year period, no statistical inter- or intra-brain difference in the diffusion coefficient was found in perfusion fixated minipig brains. However, a decreasing tendency in the diffusion coefficient was found at the last time points about 24 months post mortem and might be explained by an ongoing chemical reaction due to the fixative used. Short-term instabilities within the first 15 hours of DWI scanning were observed and found likely to be caused by the preparation of the postmortem tissue prior to MR scanning. This artefact can be avoided e.g. by simply excluding DWI-volumes obtained in the first time period of the scanning session. Probabilistic tractography was validated against two invasive in vivo neuronal tracers that were used to derive a gold standard. A high spatial agreement between tractography and the gold standard was found, and some of the widely known limitations of tractography methods could be confirmed e.g. uncertainty in regions containing crossing fibres, and definition of tract termination. In the thesis we delve behind the published results to describe all the practical issues that had to be considered in order to ensure a reliable outcome, and a successful experiment. This includes the selection of independent anatomical data to be used to derive a gold standard, the selection of a gyrated animal model in place of the human brain, objective selection of the seed region to initiate, and a waypoint region to constrain the tractography results.

Cognitive Component Analysis

This dissertation concerns the investigation of the consistency of statistical regularities in a signaling ecology and human cognition, while inferring appropriate actions for a speech-based perceptual task. It is based on unsupervised Independent Component Analysis providing a rich spectrum of audio contexts along with pattern recognition methods to map components to known contexts. It also involves looking for the right representations for auditory inputs, i.e. the data analytic processing pipelines invoked by human brains. The main ideas refer to Cognitive Component Analysis, defined as the process of unsupervised grouping of generic data such that the ensuing group structure is well-aligned with that resulting from human cognitive activity. Its hypothesis runs ecologically: features which are essentially independent in a context defined ensemble, can be efficiently coded as sparse independent component representations. The focus has been to construct a preprocessing pipeline for COCA to search for the ‘cognitive structure’, and to measure the alignment.
of the resulting from unsupervised learning and human cognition. Based on the nature of human auditory system and psychoacoustics, we have constructed the pipeline: feature extraction; feature integration; energy based sparsification; and principal component analysis. To test whether human uses information theoretically optimal ICA methods in higher cognitive functions, is the main concern in this thesis. It is well-documented that unsupervised learning discovers statistical regularities. However human cognition is too complicated and not yet fully understood. Nevertheless, in our approach we represent human cognitive processes as a classification rule in supervised learning. Thus we have devised a testable protocol to test the consistency of statistical properties and human cognitive activity, i.e. unsupervised learning of perceptual inputs and supervised learning of inputs together with manually obtained labels. The comparison has been carried out at different levels. This protocol has successfully revealed the consistency of two classifications via several speech-based cognitive tasks.

Light, Matter, and Geometry: The Cornerstones of Appearance Modelling
This thesis is about physically-based modelling of the appearance of materials. When a material is graphically rendered, its appearance is computed by considering the interaction of light and matter at a macroscopic level. In particular, the shape and the macroscopic optical properties of the material determine how it will interact with incident illumination. In this thesis the macroscopic optical properties are connected to the microscopic physical theories of light and matter. This enables prediction of the macroscopic optical properties of materials, and, consequently, also prediction of appearance based on the contents and the physical conditions of the materials. Physically-based appearance models have many potential input and output parameters. There are many choices that must be made: How many material components to include in the model, how many physical conditions to take into account, whether the shape of the material should be coupled to the appearance model or not, etc. A generalised concept of shape and geometry is presented to provide a framework for handling these many degrees of freedom. Constraints between input and output parameters are modelled as multidimensional shapes. This gives the opportunity to use the appearance models not only for prediction, but also for analysis of the contents and the physical conditions of a material given information about its macroscopic optical properties. Since it is possible to measure these properties using camera technology, the presented framework enables analysis of material contents and conditions using camera technology. Three detailed appearance models are presented as to exemplify the applicability of the theory: (1) A model which finds the appearance of water given temperature, salinity, and mineral and algal contents of the water; (2) a model which finds the appearance of ice given temperature, salinity, density, and mineral and algal contents of the ice; and (3) a model which finds the appearance of milk given fat and protein contents of the milk.
Analysis of Security Protocols by Annotations
The trend in Information Technology is that distributed systems and networks are becoming increasingly important, as most of the services and opportunities that characterise the modern society are based on these technologies. Communication among agents over networks has therefore acquired a great deal of research interest. In order to provide effective and reliable means of communication, more and more communication protocols are invented, and for most of them, security is a significant goal. It has long been a challenge to determine conclusively whether a given protocol is secure or not. The development of formal techniques, e.g. control flow analyses, that can check various security properties, is an important tool to meet this challenge. This dissertation contributes to the development of such techniques. In this dissertation, security protocols are modelled in the process calculus LYSA. A variety of interesting security properties that protocols are often expected to have are formalised: authentication, confidentiality, freshness, absence of simple and complex type flaws. Those security properties are explicitly specified as annotations embedded in the LYSA syntax. Finally, a number of automatic techniques for the analysis of system behaviour are developed. These techniques are specified as control flow analyses and are, therefore, guaranteed to terminate. The perspectives for the analysis techniques are discussed. Thus the dissertation marks a step forward both for scientists, who gain a general framework for the study of several interesting security properties, and developers, who get a collection of tools that can validate protocols with respect to various aspects of security.

General information
State: Published
Organisations: Computer Science and Engineering, Department of Informatics and Mathematical Modeling, Language-Based Technology
Authors: Gao, H. (Intern), Nielson, H. R. (Intern), Nielson, F. (Intern)
Publication date: May 2008

Publication information
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Original language: English
Series: IMM-PHD-2008-190
Main Research Area: Technical/natural sciences
Electronic versions:
phd190_hg.pdf
Source: orbit
Source-ID: 210837
Publication: Research › Ph.D. thesis – Annual report year: 2008

Monitoring angiogenesis using magnetic resonance methods
When a tumor reaches a certain size it can no longer rely on passive perfusion for nutrition. The tumor therefore emits signaling molecules which stimulating surrounding vessels to divide and grow towards the tumor, a process known as angiogenesis. Very little angiogenesis is present in healthy adults where it is primarily found in wound healing, pregnancy and during the menstrual cycle. This thesis focus on the negative consequences of angiogenesis in cancer. It consists of an initial overview followed by four manuscripts. The overview gives a short introduction to the process of angiogenesis and the involved signaling molecules. Subsequently, a short review of contrast agents and perfusion measurements is given. Finally, methods for monitoring angiogenesis using magnetic resonance imaging are reviewed. A method for monitoring early stages of angiogenesis as well as the effect of anti-angiogenic treatment is presented in the first manuscript. In the second and third manuscript, two separate methods of quantifying perfusion, blood volume and vessel permeability are presented. The methods are used to show that drug delivery to a xenografted tumor is plausible and to show possible vascular maturation in a transgenic mouse model. The last manuscript presents a new method for in vivo cell labeling. This method could find use in studying the metastatic spread of cancer cells throughout the body.

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Holm, D. A. (Intern), Sidaros, K. (Intern), Larsen, R. (Intern)
Number of pages: 130
Publication date: Nov 2008

Publication information
Original language: English
Series: IMM-PHD-2008-199
Main Research Area: Technical/natural sciences
Decomposition methods for unsupervised learning
This thesis presents the application and development of decomposition methods for Unsupervised Learning. It covers topics from classical factor analysis based decomposition and its variants such as Independent Component Analysis, Non-negative Matrix Factorization and Sparse Coding to their generalizations to multi-way array, i.e. tensor decomposition, through models such as the CanDecomp/ PARAFAC and the Tucker model. Extensions for these types of decomposition models to incorporate shift, reverberation and general transformations are also described. Finally, a connection between decomposition methods and clustering problems is derived both in terms of classical point clustering but also in terms of community detection in complex networks. A guiding principle throughout this thesis is the principle of parsimony. Hence, the goal of Unsupervised Learning is here posed as striving for simplicity in the decompositions. Thus, it is demonstrated how a wide range of decomposition methods explicitly or implicitly strive to attain this goal. Applications of the derived decompositions are given ranging from multi-media analysis of image and sound data, analysis of biomedical data such as electroencephalography to the analysis of social network data.

General information
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Mørup, M. (Intern), Hansen, L. K. (Intern), Winther, O. (Intern)
Publication date: Sep 2008

Analysis of Craniofacial Images using Computational Atlases and Deformation Fields
The topic of this thesis is automatic analysis of craniofacial images. The methods proposed and applied contribute to the scientific knowledge about different craniofacial anomalies, in addition to providing tools for detailed and robust analysis of craniofacial images for clinical and research purposes. The basis for most of the applications is non-rigid image registration. This approach brings one image into the coordinate system of another resulting in a deformation field describing the anatomical correspondence between the two images. A computational atlas representing the average anatomy of a group may be constructed and brought into correspondence with a set of images of interest. Having established such a correspondence, various analyses may be carried out. This thesis discusses two types of such analyses, i.e. statistical deformation models and novel approaches for the quantification of asymmetry. The analyses are applied to the study of three different craniofacial anomalies. The craniofacial applications include studies of Crouzon syndrome (in mice), unicoronal synostosis plagiocephaly and deformational plagiocephaly. Using the proposed methods, the thesis reveals novel findings about the craniofacial morphology and asymmetry of Crouzon mice. Moreover, a method to plan and evaluate treatment of children with deformational plagiocephaly, based on asymmetry assessment, is established. Finally, asymmetry in children with unicoronal synostosis is automatically assessed, confirming previous results based on manual reference points and providing a higher level of detail.

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, DTU Data Analysis
Authors: Ólafsdóttir, H. (Intern), Ersbøll, B. K. (Intern), Larsen, R. (Intern)
Number of pages: 159
Publication date: Apr 2008
Mathematical models and methods for analysis of distributed power generation on market conditions

The liberalisation of electricity markets around the world which has taken place in recent years – and is still ongoing – has had several consequences for the various players in the markets affected. Typically, the tasks of production, transmission, and distribution of electricity which were often handled by so-called vertically integrated monopolies have been separated to varying degrees and are in liberalised systems handled by different players. In the Nordic system, electricity is traded as a commodity on a day-ahead spot market where suppliers and consumers submit their bids for the following day and a common hourly electricity spot price is found. Intra-day markets for balancing power also exist. The raison d’etre for this type of market is that although supply and demand are balanced on a day-ahead basis, actual demand is impossible to forecast with complete accuracy. Thus on the day of operation actual demand and planned supply never match precisely. The system operator must then procure so-called balancing power in the intra-day market to maintain the physical balance of the system at all times. The present thesis considers the effects of large amounts of distributed electricity generation in a power system subject to a liberalised market. In particular, the Danish electricity system is analysed in terms of four different focus topics which are considered in the six research papers presented and commented on in the thesis. The analyses range from planning the operation and/or bidding of single-technology units such as wind power turbines and local combined heat and power plants to analyses from a system point of view such as the interaction between the natural gas, district heating, and electricity systems, and the system operator dilemma of procuring reserve power well in advance as opposed to purchasing the needed volumes in the intra-day balancing market. The thesis itself provides an introduction to the Nordic power system and market with emphasis on the Danish situation. After presenting a few classic topics in power system operation, the situation post-liberalisation of the electricity markets is analysed and a literature review is given of the major topics of the thesis, setting the contributions of the thesis into perspective of previous work on related topics. Subsequently, the papers included in the thesis are summarised and commented upon and the main contributions are listed, before the thesis is concluded upon.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Operations Research, Department of Management Engineering
Authors: Schaumburg-Müller, C. (Intern), Clausen, J. (Intern), Ravn, H. V. (Intern)
Number of pages: 167
Publication date: Jul 2008

Publication information
Original language: English

Series: IMM-PHD-2008-177
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd177_csm-web.pdf
Source: orbit
Source-ID: 211368
Publication: Research › Ph.D. thesis – Annual report year: 2008

Mobile location services for the next generation wireless network

Mobile location services exploit mobile location technologies for determining where a mobile user is geographically located. This information can then be used for providing location-specific content to the mobile user. The mobile location services can be used, for example, for finding points of interest, getting weather information, and tracking the whereabouts of a child. Mobile location services gained a great deal of interest in 2000, and they were envisioned by the business players in the mobile service market as one of the few service categories where the mobile users would be willing to pay for the usage. Since 2000, we have seen countless mobile location services commercially deployed in different parts of the world, and the services have been adopted more enthusiastically by the mobile users in Asia, especially in Japan and South Korea, compared to other parts of the world. However, the overall usage of the mobile location services is still not very high compared to other entertainment and messaging services. The mobile location services are currently not the important part of the mobile data services, and the services have obviously not yet met the hyped expectation of the mass-market adoption that was expressed in 2000. This thesis examines and analyzes the existing mobile location technologies and services to identify the factors that inhibit the take-off of the existing mobile location services. These factors provide
indications and ideas of, e.g., what to emphasize, what to avoid, and what to improve when developing a mobile location technology and a mobile location service in the future. Based on the qualitative studies of the existing location methods and services made in this thesis, the lack of location methods that can provide accurate location information in closed environments and dense urban areas and the lack of adaptability and offerings tailored to different users’ requirements in different contexts of use are the main inhibitors to the take-off of the existing mobile location services. Based on these findings, a new conceptual location method has been proposed in this thesis to resolve the lack of indoor location capability, and a conceptual service architecture for adaptive mobile location services has been developed to facilitate the provision of compelling mobile location services for the future network. The developed service architecture allows the mobile location service to be adapted to best fit with the user requirements/preferences in the current contexts of use, which is one of the missing parts that limit the adoption of the mobile location services available today.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, CICT
Authors: Schou, S. (Intern), Olesen, H. (Intern), Sørensen, L. T. (Intern)
Publication date: Jul 2008

Publication information
Original language: English
Series: IMM-PHD-2008-202
Main Research Area: Technical/natural sciences
Electronic versions:
phd202_saowane.pdf
Source: orbit
Source-ID: 220529
Publication: Research › Ph.D. thesis – Annual report year: 2008

Computational aspects of electronic transport in nanoscale devices
This thesis is concerned with the modeling of electronic properties of nano-scale devices. In particular the computational aspects of calculating the transmission and current-voltage characteristics of Landauer-Büttiker two-probe systems are in focus. To begin with, the main existing methods are described in detail and benchmarked. These are the Green’s function method and the wave function matching method. The methods are subsequently combined in a hybrid scheme in order to benefit from a common formalism. The most time demanding stages of common electronic transport calculations are identified. For systems of more than about a hundred atoms, two specific tasks stand out; the evaluation of self-energy matrices to describe the coupling between the electrodes and the device, and the solution of the central region Schrödinger equation either by matrix inverse of by solving a system of linear equations. In this work the objective is to develop new efficient algorithms for these tasks in order to model nano-scale systems of larger size in the future. The starting point of the new methods is the combined formalism of the Green’s function and wave function matching methods. The first new algorithm described is for the calculation of the block tridiagonal matrix inverse of a block tridiagonal matrix in O(N) operations. This algorithm also leads to an optimal evaluation of the frequently used Caroli transmission formula. A modified wave function matching scheme is then developed which allows for a significant reduction in the cost of the self-energy matrix calculations when combined with an iterative eigensolver. Finally, such an iterative eigensolver is developed and implemented based of a shift-and-invert Krylov subspace approach. The method is applied to a selection of nano-scale systems and speed-ups of up to an order of magnitude are achieved.

General information
State: Published
Organisations: Scientific Computing, Department of Informatics and Mathematical Modeling, Department of Micro- and Nanotechnology
Authors: Sørensen, H. H. B. (Intern), Hansen, P. C. (Intern), Stokbro, K. (Intern)
Publication date: Jul 2008

Publication information
Original language: English
Series: IMM-PHD-2008-195
Main Research Area: Technical/natural sciences
Electronic versions:
phd195_hhs.pdf
Source: orbit
Source-ID: 219861
Publication: Research › Ph.D. thesis – Annual report year: 2008
System-Level Design Methodologies for Networked Multiprocessor Systems-on-Chip

The first part of the thesis presents an overview of the existing theories and practices of modeling and simulation of multiprocessor systems-on-chip. The systematic categorization of the plethora of existing programming models at various levels of abstraction is the main contribution here which is the first such attempt in the published literature. The second part of the thesis deals with the issues related to the development of system-level design methodologies for networked multiprocessor systems-on-chip at various levels of design abstraction with special focus on the modeling and design of wireless integrated sensor networks which are an emerging class of networked embedded computer systems. The work described here demonstrates how to model multiprocessor systems-on-chip at the system level by abstracting away most of the lower-level details albeit retaining the parameters most relevant at the system-level. The multiprocessor modeling framework is then extended to include models of networked multiprocessor systems-on-chip which is then employed to model wireless sensor networks both at the sensor node level as well as the wireless network level. In the third and the final part, the thesis covers the issues related to the design, implementation and testing of a system-on-chip-based wireless sensor node development platform, specifically, for the Hogthrob project. This part also deals with the cycle-accurate model of the multiprocessor system-on-chip and its possible extensions to the transaction-level model. The thesis, as a whole makes contributions by describing a design methodology for networked multiprocessor embedded systems at three layers of abstraction from system-level through transaction-level to the cycle accurate level as well as demonstrating it practically by implementing a wireless sensor node design.
Signal Processing for Improved Wireless Receiver Performance
This thesis is concerned with signal processing for improving the performance of wireless communication receivers for well-established cellular networks such as the GSM/EDGE and WCDMA/HSPA systems. The goal of doing so, is to improve the end-user experience and/or provide a higher system capacity by allowing an increased reuse of network resources. To achieve this goal, one must first understand the nature of the problem and an introduction is therefore provided. In addition, the concept of graph-based models and approximations for wireless communications is introduced along with various Belief Propagation (BP) methods for detecting the transmitted information, including the Turbo principle. Having established a framework for the research, various approximate detection schemes are discussed. First, the general form of linear detection is presented and it is argued that this may be preferable in connection with parameter estimation. Next, a realistic framework for interference whitening is presented, allowing flexibility in the selection of whether interference is accounted for via a discrete or a Gaussian distribution. The approximate method of sphere detection and decoding is outlined and various suggestions for improvements are presented. In addition, methods for using generalized BP to perform approximate joint detection and decoding in systems with convolutional codes are outlined. One such method is a natural generalization of the traditional Turbo principle and a generalized Turbo principle can therefore be established. For realistic wireless communication scenarios, a multitude of parameters are not known and must instead be estimated. A general variational Bayesian EM-algorithm is therefore presented to provide such estimates. It generalizes previously known methods for communication systems by estimating parameter densities instead of point-estimates and can therefore account for uncertainty in the parameter estimates. Finally, an EM-algorithm for band-Toeplitz covariance estimation is presented as such an estimate is desirable for noise and interference whitening. Using simulations, the method is shown to be near-optimal in the sense that it achieves the unbiased Cramer-Rao lower-bound for medium and large sample-sizes.

Specifying Geographic Information - Ontology, Knowledge Representation, and Formal Constraints
This thesis deals with the specification of geographic information. On the basis of the role of geographic information as an infrastructure element, a method is developed for the making of specifications which are well-structured and ensure the connection between the data collections being part of a joint infrastructure. The motivation for the presented work is to meet the need for topical geographic information at any time, so that the requirements for data content and quality are fulfilled, and the information can thus form actively part of the task performance in public administration as well as in the private sector. The theoretical background is the establishment of a representational system, which ontologically comprises a representation of notions in the “real world” and notions which include the representation of these. Thus, the thesis leans towards a traditional division between modeling of domains and conceptualization of these. The thesis contributes a formalization of what is understood by domain models and conceptual models, when the focus is on geographic information. Moreover, it is shown how specifications for geographic information are related to this
representational system. The starting point of the thesis is an analysis mapping the elements in a specification for geographic information. The basis of this empirical investigation is TOP10DK's data content specification, version 3.2 of the National Survey and Cadastre. The basic idea is to view a specification as a collection of requirements and rules, building on terms from the domain and concept ontologies. In combination with the theoretical basis the analysis is used for developing an underlying model of notions, which defines the individual elements in a specification and the relations between them. In the chapters of the thesis this underlying model is extended to include a number of components, which each contribute to the model being able to form the basis of a strong and productive specification tool for the making and maintenance of specifications for geographic information. These components among others include description of quality requirements and formalization of rules, so that they can be used for verification of produced information. An essential contribution is a formal specification language dedicated to the formulation of formal rules to be observed by the information. The language is based on a formal semantic model which makes translation into other languages possible. In the thesis it is shown how statements can be translated into SQL and thus form the basis of direct implementation in the production environments where the geographic information is procured. To be able to describe requirements for the quality of geographic information is an essential part of a specification. The thesis contributes a structure of quality descriptions by introducing two notions: "Acceptable Quality Levels" (AQL) and "Quality Element Requirements" (QER), which designate respectively the minimum quality requirements for information produced according to a given specification and the requirements for the quality parameters used to describe this information. The two notions are incorporated and related to the developed system of notions for specification for geographic information. It is an important part of an infrastructure for geographic information that there is a connection between the individual data collections. This thesis argues for ensuring the connection by first and foremost describing these as an integrated part of the specification work. The thesis contributes a model which describes relations and dependencies by writing specifications in the context of one or more other specifications. As an illustration of the applications of specifications written in the developed specification language, a concept is developed in the thesis to make possible a decentralized collection and distribution of information about changes to be used for updating geographic information.
Hemodynamic modelling of BOLD fMRI - A machine learning approach
This Ph.D. thesis concerns the application of machine learning methods to hemodynamic models for BOLD fMRI data. Several such models have been proposed by different researchers, and they have in common a basis in physiological knowledge of the hemodynamic processes involved in the generation of the BOLD signal. The BOLD signal is modelled as a non-linear function of underlying, hidden (non-measurable) hemodynamic state variables. The focus of this thesis work has been to develop methods for learning the parameters of such models, both in their traditional formulation, and in a state space formulation. In the latter, noise enters at the level of the hidden states, as well as in the BOLD measurements themselves. A framework has been developed to allow approximate posterior distributions of model parameters to be learned from real fMRI data. This is accomplished with Markov chain Monte Carlo (MCMC) sampling techniques, including ‘parallel tempering’, an improvement of basic MCMC sampling. On top of this, a method has been developed that allows comparisons to be made of the quality of these models. This is based on prediction of test data, and comparisons of learnt parameters for different training data. This gives estimates of the generalization ability of the models, as well as of their reproducibility. The latter is a measure of the robustness of the learnt parameters to variations in training data. Together, these measures allow informed model comparison, or model choice. Using resampling techniques, a measure of the uncertainty about the generalization ability and reproducibility of the models is also obtained. The results show that for some of the data, the standard so-called ‘balloon’ model is sufficient. More complex data have also been designed, however, and for these, the stochastic state space version of the standard balloon model is shown to be superior, although an augmented version of the standard balloon model is not found to be an improvement for either data set.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Jacobsen, D. J. (Intern), Hansen, L. K. (Intern)
Publication date: May 2007

Publication information
Original language: English
Series: IMM-PHD
Number: 173
Main Research Area: Technical/natural sciences
Electronic versions:
imm4956.pdf
Source: orbit
Source-ID: 200800
Publication: Research › Ph.D. thesis – Annual report year: 2007

Three Dimensional Measurements in Medical Data

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics, Computer Science and Engineering
Authors: Jakobsen, B. (Intern), Christensen, N. J. (Intern), Pedersen, S. (Intern), Madsen, J. (Intern)
Publication date: Sep 2007

Publication information
Original language: English
Series: IMM-PHD-2008-183
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 210320
Publication: Research › Ph.D. thesis – Annual report year: 2007

Stochastic Modelling of Hydrologic Systems
In this PhD project several stochastic modelling methods are studied and applied on various subjects in hydrology. The research was prepared at Informatics and Mathematical Modelling at the Technical University of Denmark. The thesis is divided into two parts. The first part contains an introduction and an overview of the papers published. Then an introduction to basic concepts in hydrology along with a description of hydrological data is given. Finally an introduction to stochastic modelling is given. The second part contains the research papers. In the research papers the stochastic methods are described, as at the time of publication these methods represent new contribution to hydrology. The second part also contains additional description of software used and a brief introduction to stiff systems. The system in one of the papers is stiff.
The Structure of Complex Networks

The present thesis consists of a summary report, four research articles, one technical report and one manuscript. The subject of the thesis is individual-based stochastic models. The summary report is composed of three parts and a brief history of some basic models in population biology. This history is included in order to provide a reader that has no previous exposure to models in population biology with a sufficient background to understand some of the biological models that are mentioned in the thesis. The first part of the rest of the summary is a description of the dramatic changes in the degree of aggregation of sprat or herring in the Baltic during the day, with special focus on the dispersion of the fish from schools at dusk. The next part is a brief introduction to Markovian arrival processes, a type of stochastic processes with potential applications as sub-models in population dynamical models. The last part introduces Markovian building blocks for individual-based modelling.

Markovian Building Blocks for Individual-Based Modelling

The present thesis consists of a summary report, four research articles, one technical report and one manuscript. The subject of the thesis is individual-based stochastic models. The summary report is composed of three parts and a brief history of some basic models in population biology. This history is included in order to provide a reader that has no previous exposure to models in population biology with a sufficient background to understand some of the biological models that are mentioned in the thesis. The first part of the rest of the summary is a description of the dramatic changes in the degree of aggregation of sprat or herring in the Baltic during the day, with special focus on the dispersion of the fish from schools at dusk. The next part is a brief introduction to Markovian arrival processes, a type of stochastic processes with potential applications as sub-models in population dynamical models. The last part introduces Markovian building blocks for individual-based modelling.

The present thesis consists of a summary report, four research articles, one technical report and one manuscript. The subject of the thesis is individual-based stochastic models. The summary report is composed of three parts and a brief history of some basic models in population biology. This history is included in order to provide a reader that has no previous exposure to models in population biology with a sufficient background to understand some of the biological models that are mentioned in the thesis. The first part of the rest of the summary is a description of the dramatic changes in the degree of aggregation of sprat or herring in the Baltic during the day, with special focus on the dispersion of the fish from schools at dusk. The next part is a brief introduction to Markovian arrival processes, a type of stochastic processes with potential applications as sub-models in population dynamical models. The last part introduces Markovian building blocks for individual-based modelling.
simple case we provide a time-scale for when the heterogeneous environment can be assumed to be well-mixed, or close to a Poisson process, for the predator. It is also shown that in some cases the variability may be more important than the mean, thus the mean rate does not necessarily provide sufficient information for the population dynamics. Article B provides the mathematical apparatus for evaluating any moment of a MAP, and also the means for evaluating the conditional moments of a transient or terminating MAP. Transient MAPs are suitable as modelling tools when an important property of the system is that it can stop. This is the case for the young of many animals, where most of a large clutch die rather quickly, and yet it is the survivors that are interesting. The conditional moments can for instance be constructed such that one can evaluate the mean or the variance of the ingestion rate given that the animal did not die. Several different methods are used to obtain the formulas, which is an interesting aspect since some of these methods may be more suitable in situations where it is problematic to proceed using the standard formalism. I provide material on how to model periodic MAPs in paper E. These are, or could be, important since most animals live in a periodic environment and a periodic system generally have dynamics that are different from the corresponding system with mean rates. The technical report F concerns how to model Markovian stomachs. Both aspects can be used in more advanced functional or numerical responses. The second part of this thesis deals with parameter estimation for diffusions. The first idea is in an optimal way to incorporate prior information in the estimation equation \( G(X_{t1} ; \ldots ; X_{tn}) = 0 \), used to nd an estimator of the unknown parameter \( \theta \). The general idea is to introduce an new optimality criterion which optimizes the correlation with the posterior score function. From an application point of view this methodology is easy to apply, since the optimal estimating function \( G(X_{t1} ; \ldots ; X_{tn}) \) is equal to the classical optimal estimating function, plus a correction term which takes into account the prior information. The methodology is particularly useful in situations where prior information is available and only few observations are present. The resulting estimators in some sense have better properties than the classical estimators. The second idea is to formulate Michael Sørensens method "prediction based estimating function" for measurement error models. This is obtained by constructing an estimating function through projections of some chosen function of \( Y_t + 1 \) onto functions of previous observations \( Y_t ; \ldots ; Y_0 \). The process of interest \( X_t + 1 \) is partially observed through a measurement equation \( Y_t + 1 = h(X_t + 1) + \text{noise} \), where \( h(\cdot) \) is restricted to be a polynomial. Through a simulation study we compare for the CIR process the obtained estimator with an estimator derived from utilizing the extended Kalman filter. The simulation study shows that the two estimation methods perform equally well.
Algorithms for Source Separation - with Cocktail Party Applications

In this thesis, a number of possible solutions to source separation are suggested. Although they differ significantly in shape and intent, they share a heavy reliance on prior domain knowledge. Most of the developed algorithms are intended for speech applications, and hence, structural features of speech have been incorporated. Single-channel separation of speech is a particularly challenging signal processing task, where the purpose is to extract a number of speech signals from a single observed mixture. I present a few methods to obtain separation, which rely on the sparsity and structure of speech in a time-frequency representation. My own contributions are based on learning dictionaries for each speaker separately and subsequently applying a concatenation of these dictionaries to separate a mixture. Sparse decompositions required for the decomposition are computed using nonnegative matrix factorization as well as basis pursuit. In my work on the multi-channel problem, I have focused on convolutive mixtures, which is the appropriate model in acoustic setups. We have been successful in incorporating a harmonic speech model into a greater probabilistic formulation. Furthermore, we have presented several learning schemes for the parameters of such models, more specifically, the expectation-maximization (EM) algorithm and stochastic and Newton-type gradient optimization.

Language Based Techniques for Systems Biology

Process calculus is the common denominator for a class of compact, idealised, domain-specific formalisms normally associated with the study of reactive concurrent systems within Computer Science. With the rise of the interactioncentred science of Systems Biology a number of bio-inspired process calculi have similarly been used for the study of bio-chemical reactive systems. In this dissertation it is argued that techniques rooted in the theory and practice of programming languages, language based techniques if you will, constitute a strong basis for the investigation of models of biological systems as formalised in a process calculus. In particular it is argued that Static Program Analysis provides a useful approach to the study of qualitative properties of such models. In support of this claim a number of static program analyses are developed for Regev’s BioAmbients – a bio-inspired variant of Cardelli’s Ambient Calculus that incorporates all features of Milner’s π-calculus: The property of spatial reachability, which is related to the function of cellular transport mechanisms, is addressed by two traditional Control Flow Analyses (CFAs). The simpler of the two, a mono-variant analysis (0CFA), is context insensitive, while the other, a poly-variant analysis (2CFA), is context-sensitive. These analyses compute safe approximations to the set of spatial configurations that are reachable according to a given model. This is useful in the qualitative study of cellular self-organisation and, e.g., the effects of receptor defects or drug delivery mechanisms. The property of sequential realisability, which is closely related to the function of biochemical pathways, is addressed by a variant of traditional Data Flow Analysis (DFA). This so-called ‘Pathway Analysis’ computes safe approximations to the set of reaction sequences that is realisable according to given model. This is useful in the qualitative...
study of the metabolic pathways that emerge from a group of connected biochemical agents. Technically, these approaches are complementary, but the analyses all overapproximate the set of run-time enabled reactions. This is used in an iterative narrowing scheme that achieves considerable synergy between CFA and DFA, and dramatically improves the results of both. The specified analyses are proved correct with respect to the semantics of Bio-Ambients, and their strength is illustrated by application to abstract models of biological phenomena: One is a model of the LDL degradation pathway, where it is shown that the analyses are able to pinpoint the effects of certain genetic defects that are known to be associated to cardiovascular disease. The other is a model of genetic transcription that relies only on the τ calculus fragment of BioAmbients. In both cases the analyses compute very precise estimates of the temporal structure of the underlying pathways; hence they are applicable across a family of widely used bio-ware languages that descend from Milner’s Calculus of Communicating Systems. The presented set of analyses constitutes a nice toolbox for the analysis of biological models. The individual tools range in complexity from low polynomial to exponential, while the precision scales similarly. Thus, the toolbox may provide useful information at all stages of a models lifetime, including development, where one is interested in frequent quick estimates, verification, and prediction, where one is willing to wait longer for more precise estimates.

**General information**

State: Published
Organisations: Computer Science and Engineering, Department of Informatics and Mathematical Modeling, Language-Based Technology
Authors: Pilegaard, H. (Intern), Nielson, F. (Intern), Nielson, H. R. (Intern)
Publication date: Dec 2007

**Publication Information**

Place of publication: Kgs. Lyngby
Publisher: Technical University of Denmark, DTU Informatics, Building 321
Original language: English

Series: IMM-PHD-2008-184
Main Research Area: Technical/natural sciences
Electronic versions:
phd184_hopi-update.pdf
Source: orbit
Source-ID: 210319
Publication: Research › Ph.D. thesis – Annual report year: 2007

**Decision Support System for Fighter Pilots**

During a mission over enemy territory a fighter aircraft may be engaged by ground based threats. The pilot can use different measures to avoid the aircraft from being detected by e.g. enemy radar systems. If the enemy detects the aircraft a missile may be fired to seek and destroy the aircraft. Such a missile will almost always be either radar guided or heat seeking. It will be launched from a permanent launch pad, or it will be man portable and small enough to fit in the boot of a car. The probability of a missile being detected by onboard sensors depends on the type of missile. If a missile is detected the pilot may choose to deploy electronic countermeasures to avoid the impact of the missile. The countermeasures to choose depends on e.g. the type of missile and guidance system, distance and direction between the missile and the aircraft, an assessment of the environment hostility, aircraft altitude and airspeed, and the availability of countermeasures. Radar systems, guidance of missiles, and electronic countermeasures are all parts of the electronic warfare domain. A brief description of this domain is given. It contains an introduction to both systems working on-board the aircraft and countermeasures that can be applied to mitigate threats. This work is concerned with finding proper evasive actions when a fighter aircraft is engaged by ground based threats. To help the pilot in deciding on these actions a decision support system may be implemented. The environment in which such a system must work is described, as are some general requirements to the design of the system. Decisions suggested by the system are based on information acquired from different sources. The process of providing information from sources such as intelligence, on-board sensor systems, and tactical data from other platforms (aircraft, ships, etc.) is described. Different approaches to finding the combination of countermeasures and manoeuvres improving the pilots survivability is investigated. During training a fighter pilot will learn a set of rules to follow when threat occurs. For the pilot these rules will be formulated in natural language. An expert system can be build by translating these rules into a language understandable by a computer program. This is done in the development of a Prolog based decision support system. A decision support system will base its decisions on input from non-perfect sources. Warnings from on-board sensor can be false and intelligence reports deficient. A Bayesian net is modelled to address this. Building the dependency tables of a Bayesian net requires a large number of cells to be filled with relevant probabilities. Not having sufficient knowledge about these probabilities makes the work with developing a Bayesian net cumbersome. Therefore a method for structural learning is investigated. Here a Bayesian net is build using a set of sample data from a number of missile flight simulations. Knowledge about threats in the current combat scenario may influence the choice of evasive manoeuvres and proper countermeasures. If at any given time more expendables are dispensed than necessary, and none is left for a later necessity, the pilots survivability may decrease. A mathematical model is developed to describe this problem. It is solved to optimality using solver software. When new threats occur the decision support system must be able to provide suggestions within a fraction of a second. Since the time it takes to find an optimal solution to the mathematical model can not comply with this requirement solutions are sought using a
metaheuristic.

**General information**
State: Published
Organisations: Operations Research, Department of Informatics and Mathematical Modeling, Operations Research, Department of Management Engineering
Authors: Randleff, L. R. (Intern), Clausen, J. (Intern)
Publication date: Sep 2007

**Publication information**
Original language: English
Series: IMM-PHD-2007-172
Main Research Area: Technical/natural sciences
Electronic versions:
 phd172_lrr.pdf
Source: orbit
Source-ID: 200847
Publication: Research › Ph.D. thesis – Annual report year: 2007

---

**Techno-Economics of Residential Broadband Deployment: Multimedia Services in Residential Broadband Networks**

**General information**
State: Published
Organisations: Department of Informatics and Mathematical Modeling, CICT
Authors: Sigurdsson, H. M. (Intern), Tadayoni, R. (Intern), Olesen, H. (Intern)
Number of pages: 278
Publication date: Sep 2007

**Publication information**
ISBN (Print): 87-643-0238-5
Original language: English
Series: IMM-PhD
Number: 186
Main Research Area: Technical/natural sciences
Electronic versions:
070902_Halldor Sigurdsson_Phd_thesis_imm_format_FINAL.pdf
Source: orbit
Source-ID: 202889
Publication: Research › Ph.D. thesis – Annual report year: 2007

---

**Regularized Statistical Analysis of Anatomy**

**General information**
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Sjöstrand, K. (Intern), Larsen, R. (Intern), Hansen, L. K. (Intern)
Publication date: Sep 2007

**Publication information**
Original language: English
Series: IMM-PHD
Number: 182
Main Research Area: Technical/natural sciences
Electronic versions:
phd182_kas.pdf
Links:
http://www2.imm.dtu.dk/pubdb/views/publication_details.php?id=5303
Source: orbit
Source-ID: 202531
Publication: Research › Ph.D. thesis – Annual report year: 2007
Music Genre Classification Systems - A Computational Approach

Automatic music genre classification is the classification of a piece of music into its corresponding genre (such as jazz or rock) by a computer. It is considered to be a cornerstone of the research area Music Information Retrieval (MIR) and closely linked to the other areas in MIR. It is thought that MIR will be a key element in the processing, searching and retrieval of digital music in the near future. This dissertation is concerned with music genre classification systems and in particular systems which use the raw audio signal as input to estimate the corresponding genre. This is in contrast to systems which use e.g. a symbolic representation or textual information about the music. The approach to music genre classification systems has here been system-oriented. In other words, all the different aspects of the systems have been considered and it is emphasized that the systems should be applicable to ordinary real-world music collections. The considered music genre classification systems can basically be seen as a feature representation of the song followed by a classification system which predicts the genre. The feature representation is here split into a Short-time feature extraction part followed by Temporal feature integration which combines the (multivariate) time-series of short-time feature vectors into feature vectors on a larger time scale. Several different short-time features with 10-40 ms frame sizes have been examined and ranked according to their significance in music genre classification. A Consensus sensitivity analysis method was proposed for feature ranking. This method has the advantage of being able to combine the sensitivities over several resamplings into a single ranking. The main efforts have been in temporal feature integration. Two general frameworks have been proposed; the Dynamic Principal Component Analysis model as well as the Multivariate Autoregressive Model for temporal feature integration. Especially the Multivariate Autoregressive Model was found to be successful and outperformed a selection of state-of-the-art temporal feature integration methods. For instance, an accuracy of 48% was achieved in comparison to 57% for the human performance on an 11-genre problem. A selection of classifiers were examined and compared. We introduced Cooccurrence models for music genre classification. These models include the whole song within a probabilistic framework which is often an advantage compared to many traditional classifiers which only model the individual feature vectors in a song.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Ahrendt, P. (Intern), Hansen, L. K. (Intern)
Publication date: Jun 2006

Publication information
Original language: English
Series: IMM-PHD-2006-164
Main Research Area: Technical/natural sciences
Electronic versions:
imm4438.pdf
Source: orbit
Source-ID: 191702
Publication: Research › Ph.D. thesis – Annual report year: 2006
The MANGO clockless network-on-chip: Concepts and implementation

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, System-on-Chip Hardware
Authors: Bjerregaard, T. (Intern), Sparsø, J. (Intern)
Publication date: Feb 2006

Publication information
Original language: English
Series: IMM-PHD-2005-153
Main Research Area: Technical/natural sciences
Electronic versions:
imm4025.pdf
Source: orbit
Source-ID: 185919
Publication: Research › Ph.D. thesis – Annual report year: 2006

Data assimilation on atmospheric dispersion of radioactive materials

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Mathematical Statistics
Authors: Drews, M. (Intern), Madsen, H. (Intern)
Publication date: 2006

Publication information
Original language: English
Series: IMM-PHD-2008-136
Main Research Area: Technical/natural sciences
Links:
http://www2.imm.dtu.dk/pubdb/p.php?3236
Source: orbit
Source-ID: 185920
Publication: Research › Ph.D. thesis – Annual report year: 2005

Independent Component Analysis in a convoluted world
This thesis is about convolutive ICA with application to EEG. Two methods for convolutive ICA are proposed. One method, the CICAP algorithm, uses a linear predictor in order to formulate the convolutive ICA problem in two steps: linear deconvolution followed by instantaneous ICA. The other method, the CICAAR algorithm, generalizes Infomax ICA to include the case of convolutive mixing. One advantage to the CICAAR algorithm is that Bayesian model selection is made possible, and in particular, it is possible to select the optimal order of the filters in a convolutive mixing model. A protocol for detecting the optimal dimensions is proposed, and verified in a simulated data set. The role of instantaneous ICA in context of EEG is described in physiological terms, and in particular the nature of dipolar ICA components is described. It is showed that instantaneous ICA components of EEG lacks independence when time lags are taken into consideration. The CICAAR algorithm is shown to be able to remove the delayed temporal dependencies in a subset of ICA components, thus making the components "more independent". A general recipe for ICA analysis of EEG is proposed: first decompose the data using instantaneousICA, then select a physiologically interesting subspace, then remove the delayed temporal dependencies among the instantaneous ICA components by using convolutive ICA. By Bayesian model selection, in a real world EEG data set, it is shown that convolutive ICA is a better model for EEG than instantaneous ICA.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Dyrholm, M. (Intern), Hansen, L. K. (Intern)
Publication date: Sep 2006

Publication information
Original language: English
Series: IMM-PHD-2006-158
Main Research Area: Technical/natural sciences
Electronic versions:
**Tomographic reconstruction using anatomy and regularization**

**General information**
State: Published
Organisations: Cognitive Systems, Department of Informatics and Mathematical Modeling
Authors: Høgh-Rasmussen, E. (Intern), Hansen, L. K. (Intern)
Publication date: Mar 2006

**Publication information**
Original language: English
Series: IMM-PHD-2005-156
Main Research Area: Technical/natural sciences
Links:
http://www2.imm.dtu.dk/pubdb/p.php?4049

**In Silico ADME Classification of Chemical Drug Candidates**

**General information**
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Jensen, B. F. (Intern), Brockhoff, P. B. (Intern)
Publication date: Sep 2006

**Publication information**
Original language: English
Series: IMM-PHD-2006-166
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 185926
Publication: Research › Ph.D. thesis – Annual report year: 2006

**Stabilization Algorithms for Large-Scale Problems**
The focus of the project is on stabilization of large-scale inverse problems where structured models and iterative algorithms are necessary for computing approximate solutions. For this purpose, we study various iterative Krylov methods and their abilities to produce regularized solutions. Some of the Krylov methods have previously been studied and identified as iterative regularization methods, whereas others have been proposed in the literature, but only sparsely studied in practice. This thesis considerably improves the understanding of these methods. Image deblurring problems constitute a nice class of large-scale problems for which the various methods can be tested. Therefore, this present work includes a separate study of the matrix structures that appear in this connection – not least to create a common basis for discussions. Another important part of the thesis is regularization matrices for the formulation of inverse problems on general form. Special classes of regularization matrices for large-scale problems (among these also two-dimensional problems) have been analyzed. Moreover, the above mentioned Krylov methods have also been analyzed in connection with the solution of problems on general form, and a new extension to the methods has been developed for this purpose. The L-curve method is one among several parameter choice methods that can be used in connection with the solution of inverse problems. A part of the work has resulted in a new heuristic for the localization of the corner of a discrete L-curve. This heuristic is implemented as a part of a larger algorithm which is developed in collaboration with G. Rodriguez and P. C. Hansen. Last, but not least, a large part of the project has, in different ways, revolved around the object-oriented Matlab toolbox MOORe Tools developed by PhD Michael Jacobsen. New implementations have been added, and several bugs and shortcomings have been fixed. The work has resulted in three papers that are all included in an appendix for convenience.

**General information**
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Scientific Computing
Simulation-based Modeling Frameworks for Networked Multi-processor System-on-Chip

This thesis deals with modeling aspects of multi-processor system-on-chip (MpSoC) design affected by the on-chip interconnect, also called the Network-on-Chip (NoC), at various levels of abstraction. To begin with, we undertook a comprehensive survey of research and design practices of networked MpSoC. The survey presents the challenges of modeling and performance analysis of the hardware and the software components used in such devices. These challenges are further exasperated in a mixed abstraction workspace, which is typical of complex MpSoC design environment. We provide two simulation-based frameworks: namely ARTS and RIPE, that allows to model hardware
Thereby, we can realistically model the application executing on the architecture. This includes e.g. accurate modeling of synchronization, cache refills, context switching effects, so on, which are critically dependent on the architecture and the performance of the NoC. The foundation of the ARTS model is abstract tasks, while the foundation of the RIPE model is cycle-count. For ARTS, using different case-studies with over one hundred tasks (five applications) from the mobile multimedia domain, we show the potential of the framework under real-time constraints. For RIPE, first using six applications we derive the requirements to model the application and the architecture properties independent of the NoC, and then use these applications to successfully validate the approach against a reference cycle-true system. The presence of a standard socket at the intellectual property (IP) and the NoC interface in both the ARTS and the RIPE frameworks allows easy incorporation of IP cores from either frameworks, into a new instance of the design. This could pave the way for seamless design evaluation from system-level to cycle-true abstraction in future component-based MpSoC design practice.

Temporal Feature Integration for Music Organisation
This Ph.D. thesis focuses on temporal feature integration for music organisation. Temporal feature integration is the process of combining all the feature vectors of a given time-frame into a single new feature vector in order to capture relevant information in the frame. Several existing methods for handling sequences of features are formulated in the temporal feature integration framework. Two datasets for music genre classification have been considered as valid test-beds for music organisation. Human evaluations of these, have been obtained to access the subjectivity on the datasets. Temporal feature integration has been used for ranking various short-time features at different time-scales. This include short-time features such as the Mel frequency cepstral coefficients (MFCC), linear predicting coding coefficients (LPC) and various MPEG-7 short-time features. The ‘consensus sensitivity ranking’ approach is proposed for ranking the short-time features at larger time-scales according to their discriminative power in a music genre classification task. The multivariate AR (MAR) model has been proposed for temporal feature integration. It effectively models local dynamical structure of the short-time features. Different kernel functions such as the convolutive kernel, the product probability kernel and the symmetric Kullback Leibler divergence kernel, which measures similarity between frames of music have been investigated for aiding temporal feature integration in music organisation. A special emphasis is put on the product probability kernel for which the MAR model is derived in closed form. A thorough investigation, using robust machine learning methods, of the MAR model on two different music genre classification datasets, shows a statistical significant improvement using this model in comparison to existing temporal feature integration models. This improvement was more pronounced for the larger and more difficult dataset. Similar findings where observed using the MAR model in a product probability kernel. The MAR model clearly outperformed the other investigated density models: the multivariate Gaussian model and the Gaussian mixture model.
Pharmacokinetic/Pharmacodynamic modelling with a stochastic perspective. Insulin secretion and Interleukin-21 development as case studies

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Mathematical Statistics
Authors: Overgaard, R. V. (Intern), Madsen, H. (Intern)
Publication date: Oct 2006

Publication information
Original language: English
Series: IMM-PHD-2006-169
Main Research Area: Technical/natural sciences
Electronic versions:

Source Separation for Hearing Aid Applications

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Pedersen, M. S. (Intern), Larsen, J. (Intern)
Publication date: Nov 2006

Publication information
Original language: English
Series: IMM-PHD-2006-167
Main Research Area: Technical/natural sciences
Electronic versions:

Mean Field ICA
This thesis describes investigations and improvements of a technique for Independent Component Analysis (ICA), called "Mean Field ICA". The main focus of the thesis is the optimization part of the algorithm, the so-called "EM algorithm". Using different approaches it is demonstrated that the EM algorithm is inefficient and therefore an improper choice for a certain class of models. As an alternative, the so-called "Easy Gradient Algorithm" is presented, which makes it possible to apply advanced optimization techniques using the computationally simple E-step and M-step of the EM algorithm. The Easy Gradient Recipe is applicable to a wide selection of models. Furthermore, the Mean Field ICA model is extended to incorporate filtering over time in a so-called "convolutive ICA" model. Finally, by using mixture of Gaussians as source priors, the generative and filtering approach to ICA is compared in the overcomplete setting, i.e. the situation in which there is more sources than sensors.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Petersen, K. B. (Ekstern), Winther, O. (Intern)
Publication date: Jan 2006

Publication information
Original language: English
Risk Assessments of Minefields in Humanitarian Mine Action - a Bayesian Approach

During the last 10-15 years, the international community has become aware of the devastating mine contamination problems experienced in many post-conflict countries. As a consequence, a considerable amount of money and time is spent on research and development in new ways of locating buried mines and unexploded ordnance in a fast and secure way. A major breakthrough is however still waiting, and a large fraction of the mine clearance, which still remains to be done, will therefore hinge on slow and dangerous procedures based on prodders and metal detectors. Realizing that landmine contamination is a phenomenon which cannot be eliminated overnight but is a problem which has to managed in several years to come, it is essential that the resources a national government in a mine affected country spends on mine clearance are used on the right projects. However, the identification of the mine clearance projects with the greatest impact is a delicate task. More systematic approaches to the ranking of minefields with respect to mine clearance can be found in the literature, but these methods are either founded on simple scoring rules or are of a more qualitative nature. Thus nobody seems yet to have examined the usefulness of the analytical tools which might be provided by operations research and statistics in order to support decision makers involved in national mine clearance programmes. In February 2002, the Danish Defence Research Establishment initiated in collaboration with the Technical University of Denmark a Ph.D.-project to investigate whether the application of operations research and statistics can support decision makers in Humanitarian Mine Action to make the prioritization of mine clearance operations more effective. The main part of that project, which is presented in the enclosed thesis, has concentrated on the development of a risk model quantifying to what extent a minefield poses a risk to a society. The risk model is derived in two steps: First, a general model, which requires detailed information about the mined area in question, is derived. Secondly, by the introduction of two additional assumptions, the general model is turned into a simple binomial model depending on two parameters m and q. In this context the integer m denotes the number of so-called functional mines in the minefield under consideration, and the parameter q denotes the probability of a randomly selected mine being encountered by a person, a vehicle, etc… during a predefined observation period. The true values of the binomial parameters, which jointly characterize the state of the mined area, will rarely be known in advance, but beliefs about these based on whatever information is available can conveniently be expressed in terms of probability distributions p(m) and p(q). This prepares the way for the introduction of Bayesian data analysis by which updates of the probability distributions can be generated from incoming accident statistics. The major obstacle to a real-life application of the derived risk model seems to be the lack of actual information about the binomial parameter q. A considerable part of the enclosed thesis focuses therefore on ways to provide information about q through statistical modelling. Depending on the level of historical information available to a hypothetical decision maker, two different proposed models are examined as ways of extracting information about q : 1) A simple hierarchical model which as input requires accident statistics and clearance reports from already cleared minefields; 2) A finite mixture model where only accident statistics and the specification of certain prior distributions are needed as input data. Common to both models is the generation of posterior distributions of the parameter q. To extract information about q from these distributions various simulation techniques are applied including importance sampling and Markov Chain simulation. The possibility of making updates of the entering probability distributions p(m) and p(q) through incoming accident statistics by the use of Bayes’ rule makes the suggested risk model dynamic. Moreover, the application of Bayesian data analysis gives the derived risk model a very flexible structure which allows an accommodation to the varied circumstances found in Humanitarian Mine Action with respect to the amount of accessible information. The present thesis closes with an overall prescription for the synthesis of different pieces of information based on the concept of reference priors.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Operations Research, Department of Management Engineering
Authors: Vistisen, J. B. (Intern), Clausen, J. (Intern)
Publication date: Mar 2006

Publication information
Original language: English
Development of an image based system to objectively score the severity of phoriasis

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics
Authors: Gomez, D. D. (Intern), Erbsøll, B. K. (Intern), Carstensen, J. M. (Intern)
Publication date: 2005

Publication information
Original language: English
Series: IMM-PHD-2005-152
Main Research Area: Technical/natural sciences
Links:
http://www2.imm.dtu.dk/pubdb/p.php?4013
Source: orbit
Source-ID: 185922
Publication: Research › Ph.D. thesis – Annual report year: 2005

A Level Set Discontinuous Galerkin Method for Free Surface Flows - and Water-Wave Modeling

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Scientific Computing
Authors: Grooss, J. (Intern), Thomsen, P. G. (Intern)
Publication date: 2005

Publication information
Original language: English
Series: IMM-PHD-2005-145
Main Research Area: Technical/natural sciences
Electronic versions:
imm3865.ps
imm3865.pdf
Links:
http://www2.imm.dtu.dk/pubdb/p.php?3865
Source: orbit
Source-ID: 185923
Publication: Research › Ph.D. thesis – Annual report year: 2005

Flow logic for language-based safety and security

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Language-Based Technology
Authors: Hansen, R. R. (Intern), Nielson, F. (Intern)
Publication date: 2005

Publication information
Original language: English
Series: IMM-PHD-2005-143
Main Research Area: Technical/natural sciences
Electronic versions:
phd143_rrh-ny.pdf
Links:
http://www2.imm.dtu.dk/pubdb/p.php?3585
Source: orbit
Source-ID: 185925
Indexing and Analysis of Fungal Phenotypes Using Morphology and Spectrometry

**General information**
State: Published
Organisations: Center for Microbial Biotechnology, Department of Systems Biology, Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Hansen, M. A. E. (Intern), Carstensen, J. M. (Intern), Frisvad, J. C. (Intern)
Publication date: 2005

**Publication information**
Original language: English
Series: IMM-PHD-2005-151
Main Research Area: Technical/natural sciences
Links:
http://www2.imm.dtu.dk/pubdb/p.php?3325
Source: orbit
Source-ID: 185924
Publication: Research › Ph.D. thesis – Annual report year: 2005

Making Faces - State-Space Models Applied to Multi-Modal Signal Processing

**General information**
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Lehn-Schiøler, T. (Intern), Hansen, L. K. (Intern)
Publication date: 2005

**Publication information**
Original language: English
Series: IMM-PHD-2005-146
Main Research Area: Technical/natural sciences
Electronic versions:
imm4047.pdf
Links:
http://www2.imm.dtu.dk/pubdb/p.php?4047
Source: orbit
Source-ID: 185927
Publication: Research › Ph.D. thesis – Annual report year: 2005

Design and analysis of environmental monitoring programs

**General information**
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Department of Environmental Science and Engineering
Authors: Lophaven, S. N. (Intern), Carstensen, N. J. (Intern)
Publication date: Feb 2005

**Publication information**
Original language: English
Main Research Area: Technical/natural sciences
Electronic versions:
imm3370.ps
imm3370.pdf
Links:
Source: orbit
Source-ID: 154832
Publication: Research › Ph.D. thesis – Annual report year: 2005
Behavioral synthesis of asynchronous circuits

General information
State: Published
Organisations: Computer Science and Engineering, Department of Informatics and Mathematical Modeling, System-on-Chip Hardware, Embedded Systems Engineering
Authors: Nielsen, S. F. (Intern), Sparsø, J. (Intern), Madsen, J. (Intern)
Publication date: 2005

Publication information
Original language: English
Series: IMM-PHD-2005-144
Main Research Area: Technical/natural sciences
Electronic versions:
imm3866.ps
imm3866.pdf
Links:
http://www2.imm.dtu.dk/pubdb/p.php?3866
Source: orbit
Source-ID: 185930
Publication: Research › Ph.D. thesis – Annual report year: 2005

Distributed nonlinear optical response

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Nikolov, N. I. (Intern), Christiansen, P. L. (Intern)
Publication date: 2005

Publication information
Original language: English
Series: IMM-PHD-2005-133
Main Research Area: Technical/natural sciences
Electronic versions:
imm3147.pdf
Links:
http://www2.imm.dtu.dk/pubdb/p.php?3147
Source: orbit
Source-ID: 185931
Publication: Research › Ph.D. thesis – Annual report year: 2005

Maintenance of Digital Topographical Map Databases

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Olsen, B. P. (Intern), Jacobi, O. I. (Intern)
Publication date: 2005

Publication information
Original language: English
Series: IMM-PHD-2005-142
Main Research Area: Technical/natural sciences
Links:
http://www2.imm.dtu.dk/pubdb/p.php?3867
Source: orbit
Source-ID: 185932
Publication: Research › Ph.D. thesis – Annual report year: 2005
Automated Analysis of Security in Networking Systems

It has for a long time been a challenge to built secure networking systems. One way to counter this problem is to provide developers of software applications for networking systems with easy-to-use tools that can check security properties before the applications ever reach the market. These tools will both help raise the general level of awareness of the problems and prevent the most basic flaws from occurring. This thesis contributes to the development of such tools. Networking systems typically try to attain secure communication by applying standard cryptographic techniques. In this thesis such networking systems are modelled in the process calculus LySa. On top of this programming language based formalism an analysis is developed, which relies on techniques from data and control flow analysis. These are techniques that can be fully automated, which make them an ideal basis for tools targeted at non-experts users. The feasibility of the techniques is illustrated by a proof-of-concept implementation of a control flow analysis developed for LySa. From a technical point of view, this implementation also interesting because it encodes in finite sets of algebraic terms, which denote encryption, as a nite number of tree grammar rules. The security of any software application relies crucially on the scenario in which the application is deployed. In contrast to many related analysis approaches, this thesis provides an explicit mechanism for specifying deployment scenarios. Even though these scenarios may be arbitrarily large the analysis techniques can be extended to cope with such scenarios. The analysis techniques are furthermore capable of tackling security issues that arise because of attacks from arbitrary attackers: the analysis can deal with confidentiality and authentication properties, parallel session attacks, and attacks launched by insiders. Finally, the perspectives for the application of the analysis techniques are discussed, thereby, coming a small step closer to providing developers with easy-to-use tools for validating the security of networking applications.

General Information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Language-Based Technology
Authors: Buchholtz, M. (Intern), Nielson, H. R. (Intern), Nielson, F. (Intern)
Publication date: 2004

Publication Information
Original language: English
Series: IMM-PHD-2004-141
Main Research Area: Technical/natural sciences
Electronic versions:
imm3419.pdf
imm3419.ps
Links:
Source: orbit
Source-ID: 154824
Publication: Research › Ph.D. thesis – Annual report year: 2004

Methods for Measurement and Analysis of Craniofacial Morphology and Growth in Children with Cleft Lip and Palate

General Information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics
Authors: Darvann, T. A. (Intern), Ersbøll, B. K. (Intern)
Publication date: 2004

Publication Information
Original language: English
Series: IMM-PHD-2004-129
Main Research Area: Technical/natural sciences
Links:
Source: orbit
Source-ID: 154826
Publication: Research › Ph.D. thesis – Annual report year: 2004

Real-time Global Illumination by Simulating Photon Mapping

This thesis introduces a new method for simulating photon mapping in realtime. The method uses a variety of both CPU and GPU based algorithms for speeding up the different elements in global illumination. The idea behind the method is to calculate each illumination element individually in a progressive and efficient manner. This has been done by analyzing
the photon mapping method and by selecting efficient methods, either CPU based or GPU based, which replaces the original photon mapping algorithms. We have chosen to focus on the indirect illumination and the caustics. In our method we first divide the photon map into several photon maps in order to make local updates possible. Then indirect illumination is added using light maps that are selectively updated by using selective photon tracing on the CPU. The final gathering step is calculated by using fragment programs and GPU based mipmapping. Caustics are calculated by using photon tracing on the CPU and the filtering which is performed on the GPU. Direct illumination is calculated by using shading on the GPU. We achieve real-time frame rates for simple scenes with up to 133,000 polygons. The scenes include standard methods for reflection and refraction and hard shadows. Furthermore, the scenes include our methods for progressively updated caustics and progressively updated indirect illumination. We have compared the image quality of our method to the standard photon mapping method and the results are very similar.

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Larsen, B. D. (Intern), Christensen, N. J. (Intern)
Publication date: Sep 2004

Publication information
Original language: English
Series: IMM-PHD-2004-130
Main Research Area: Technical/natural sciences
Electronic versions:
imm4115.pdf
Source: orbit
Source-ID: 200817
Publication: Research › Ph.D. thesis – Annual report year: 2004

Analysis and Synthesis of Communication-Intensive Heterogeneous Real-Time Systems
Embedded computer systems are now everywhere: from alarm clocks to PDAs, from mobile phones to cars, almost all the devices we use are controlled by embedded computer systems. An important class of embedded computer systems is that of real-time systems, which have to fulfill strict timing requirements. As realtime systems become more complex, they are often implemented using distributed heterogeneous architectures. The main objective of this thesis is to develop analysis and synthesis methods for communication-intensive heterogeneous hard real-time systems. The systems are heterogeneous not only in terms of platforms and communication protocols, but also in terms of scheduling policies. Regarding this last aspect, in this thesis we consider time-driven systems, event-driven systems, and a combination of both, called multi-cluster systems. The analysis takes into account the heterogeneous interconnected nature of the architecture, and is based on an application model that captures both the dataflow and the flow of control. The proposed synthesis techniques derive optimized implementations of the system that fulfill the design constraints. An important part of the system implementation is the synthesis of the communication infrastructure, which has a significant impact on the overall system performance and cost. To reduce the time-to-market of products, the design of real-time systems seldom starts from scratch. Typically, designers start from an already existing system, running certain applications, and the design problem is to implement new functionality on top of this system. Hence, in addition to the analysis and synthesis methods proposed, we have also considered mapping and scheduling within such an incremental design process. The analysis and synthesis techniques proposed have been thoroughly evaluated using a solid experimental platform. Besides the evaluations, performed using a large number of generated example applications, we have also validated our approaches using a realistic case study consisting of a vehicle cruise controller.

General information
State: Published
Organisations: Computer Science and Engineering, Department of Informatics and Mathematical Modeling
Authors: Pop, P. (Intern)
Publication date: 2003

Publication information
Original language: English
Series: Linköping Studies in Science and Technology
Number: 833
Main Research Area: Technical/natural sciences

Bibliographical note
An extended version of the thesis has been published by Kluwer Academic Publishers in 2004. If you want a copy of the PhD thesis, please send me an e-mail.
Publisher: Dept. of Computer and Information Science, Linköping University
Manipulation of volumetric solids with applications to sculpting

The topic of this thesis is volume graphics, and in particular techniques which are applicable to volume sculpting. A volume sculpting system is an interactive computer program for shape modelling where the shape is represented volumetrically in a 3D lattice of so-called voxels. It is argued that it is reasonable to classify the tools in a sculpting system according to whether the tools tend to deform the sculpted object or work according to the paradigm of Constructive Solid Geometry (CSG). Existing volume sculpting systems are surveyed, and it is found that almost all systems provide sculpting tools belonging exclusively to either or both categories. It is also found that existing systems have a number of important deficiencies. For instance, none of the systems provide a generic methodology for deformation. Rather they provide specific solutions for concrete deformation tasks, e.g. smoothing or the creation of small protrusions or dents. Moreover, most of the existing systems are based on a volume representation where the value of a voxel is construed as a pseudo-density with no precise meaning. More precisely, we can tell from a voxel whether it is on the inside or the outside of a represented solid, but nothing more. In this thesis it is argued, that it is useful to be able to give a voxel a more precise meaning. This leads to a cleaner volume representation, and if we choose (as the precise meaning of a voxel) the shortest distance from the voxel position to the closest surface point, we reap additional benefits: It becomes trivial to find surface points, and it becomes much easier to find offset surfaces and to compute various geometric properties such as curvature. Generic techniques for constructive (CSG based) and deformative tools have been implemented. Both sets of tools maintain a volume representation where the meaning of a voxel is shortest distance. The deformative tools are based on a specialization of the Level Set Method. The main advantage of using the Level Set Method is that it is a very generic technique as opposed to methods previously proposed. The main task here has been to restrict the effect of the Level Set Method to a local region of influence and to ensure a smooth transition between the affected region and the unaffected. The theoretical problem of what shapes that are suitable for volume representation has been considered. I reach the conclusion that a shape is suitable if we can roll a ball on either side of the surface in such a way that no point on (either side of) the surface is untouched. Here, the size of the ball depends on the scale of the voxel lattice. The intuitive quality that the ball can roll on either side of the surface of the solid has been formulated more precisely using concepts from mathematical morphology. Essentially, if the solid is unchanged by a morphological opening using the ball as structuring element, then the ball rolls on the interior side. Likewise, invariance with respect to closing implies that the ball can roll on the exterior. These results are, of course, of theoretical interest, but not exclusively: A technique for constructive manipulation which maintains the properties of openness and closedness has been developed. A technique for fast volume visualization is an essential part of a sculpting system. Two techniques for interactive visualization have been implemented: A novel technique based on point rendering and the well-known Marching Cubes Method. The point rendering technique is compared to marching cubes and to texture based visualization. A ray casting method has also been implemented for the generation of high quality images. The most important disadvantage of the volume representation is its lack of support for features at different scales. By choosing a volume representation, we implicitly choose a scale, and features that are very small with respect to that scale are essentially un-representable. As a solution, I propose an adaptive framework, where voxels are no longer stored in a regular grid but in adaptive grid. This allows for higher concentrations of voxels in some parts of the volume than others, and this, in turn, allows for features at vastly differing scales.

General information

State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Bærentzen, J. A. (Intern), Christensen, N. J. (Intern)
Publication date: 2002

Publication information

Original language: English

Series: IMM-PHD-2002-159
Main Research Area: Technical/natural sciences
Electronic versions:
imm1100.pdf
Source: orbit
Source-ID: 200773
Publication: Research › Ph.D. thesis – Annual report year: 2002

Stochastic Modelling of Energy Systems

In this thesis dynamic models of typical components in Danish heating systems are considered. Emphasis is made on describing and evaluating mathematical methods for identification of such models, and on presentation of component models for practical applications. The thesis consists of seven research papers (case studies) together with a summary report. Each case study takes it's starting point in typical heating system components and both, the applied mathematical modelling methods and the application aspects, are considered. The summary report gives an introduction to the scope of application and the applied modelling method and summarizes the research papers. The foundation of the identification
process is the grey box modelling method. The grey box modelling method is characterized by using information from measurements in conjunction with physical knowledge. The combination of statistical methods and physical interpretation is exploited in the modelling procedure, from the design of experiments to parameter estimation and model validation. The presented models are mainly formulated as state space models in continuous time with discrete time observation equations. The state equations are expressed in terms of stochastic differential equations. From a theoretical viewpoint the techniques for experimental design, parameter estimation and model validation are considered. From the practical viewpoint emphasis is put on how this methods can be used to construct models adequate for heating system simulations. Significant parts of the research work have been done in cooperation with leading companies from the Danish heating industry. The presented models have been developed for the purpose of analyzing typical heating system installations. The focal point of the developed models is that the model structure has to be adequate for practical applications, such as system simulation, fault detection and diagnosis, and design of control strategies. This also reflects on the methods used for identification of the component models. The main result from this research is the identification of component models, such as e.g. heat exchanger and valve models, adequate for system simulations. Furthermore, the thesis demonstrates and discusses the advantages and disadvantages of using statistical methods in conjunction with physical knowledge in establishing adequate component models of heating systems.

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Andersen, K. K. (Intern), Madsen, H. (Intern)
Number of pages: 198
Publication date: 2001

Publication information
Original language: English
Series: IMM-PHD-2001-79
Main Research Area: Technical/natural sciences
Electronic versions:
imm314.pdf
Links:
Source: orbit
Source-ID: 57994
Publication: Research › Ph.D. thesis – Annual report year: 2001

Scientific discovery using genetic programming

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Keijzer, M. (Intern), Hansen, L. K. (Intern)
Publication date: 2001

Publication Information
Original language: English
Series: IMM-PHD-2001-92
Main Research Area: Technical/natural sciences
Electronic versions:
phd92_MaartenKeijzer.ps
phd92_MaartenKeijzer.pdf
Links:
Source: orbit
Source-ID: 58000
Publication: Research › Ph.D. thesis – Annual report year: 2001

Numerical simulation of turbulent airflow in livestock buildings

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Bennetsen, J. C. (Intern)
Online correction of scanning probe microscopes with pixel accuracy

In this project "Online Control of Scanning Probe Microscopes with Pixel Accuracy", the development of an algorithm is described that enhances the measurement uncertainty of software controlled SPM by one order of magnitude from 2% to 0.2%. The SPM is globally used as a metrological instruments for dimensional and length measurements. The sample surfaces are scanned 3-dimensionally, typically within the ranges up to 150 µm x 150 µm x 6 µm. This is done by moving a sharp tip systematically across the sample while simultaneously recording the height of the tip. Typically, the tip has a radius of curvature of 10 nm and an opening angle of 30.0°. Even atomic resolution can be achieved. The scan movement of the tip is not linear however. This is caused by the propelling device of the SPM for the scan motion - a piezoelectric ceramic. The two major non-linear responses of the piezo to the applied control voltage are rate-independent hysteresis between the scanner's position and the voltage and time-dependent creep of the ceramic. Hysteresis leads to a non-linear mapping of distances while creep changes the actual sensitivity of the ceramic. The non-linearity of a hysteresis loop is in the order of 2-20% depending on the piezo material used and the scan range. The change in sensitivity is up to 20% as well, depending on the scan frequency. Current software controlled SPM are equipped with an algorithm that changes the shape of the control voltage online in a way to produce a linear piezo movement. The algorithm typically contains 5 - 7 parameters which have to be calibrated manually. Still, non-linear errors remain in the order of 1-2%. One pixel in a 512x512 image corresponds to 0.2% per direction. This goal of measurement accuracy is reached with the algorithm developed in this thesis. Three different SPM are analyzed for their non-linearity. Two commercial tube scanners are applied with a maximum scan range in x and y of 40.0 µm and 160.0 µm as well as one specially designed stack scanner with a maximum range of 5.0 µm. For the tube scanners, a 1-dimensional line pattern with a reference period of 3.0 µm and a 2-dimensional grating with a reference pitch distance of 200.0 nm are applied as length standards. The non-linearity of the scanner is then traceable to the distances on the samples. The stack scanner is equipped with capacitive sensors that measure the position of the scanner during the scan process. The signal of the sensors can be used as closed loop feedback signal. At first a model is set up to describe the measured hysteresis. An ordinary linear differential equation proves to yield the desired accuracy of 0.2% when simulating the measured hysteresis. This is done with 5 model parameters and verified for 99% of the scan range of the SPM. In addition to this, the model is not restricted to a periodic scan movement in the lateral plane as most online models are. It is flexible enough to even describe the random rate-independent movement in the z-direction. After simulating the hysteresis, the model is enhanced in order to describe time-dependent creep during the scan motion. The new model contains 7 parameters and yields the desired accuracy of 0.2% for a large choice of scan ranges and scan frequencies. The parameters are determined in a numerically optimum way by using a least-squares-fitting technique. After having successfully simulated the measured non-linearities, the model is inverted in order to form an algorithm for online correction during the scan process. Also the online algorithm is tested on two different scanners. The residual non-linearity of online corrected images is in the order of 0.2% for both scanners: The error in length changes between ± 1% from experiment to experiment. Within one experiment however, the variation of the errors is 0.3%. Therefore it is concluded that the online algorithm is stable within the set goal of 0.2% measuring uncertainty, but the piezo changes arbitrarily in the its sensitivity. Further results of this thesis include the simulation of transient hysteresis as occurs at a change of scan conditions. This is also applied to the z-direction. Here an overshoot at a large step is qualitatively simulated and explained by hysteretic behaviour.
Making Deformable Template Models Operational

Deformable template models are a very popular and powerful tool within the field of image processing and computer vision. This thesis treats this type of models extensively with special focus on handling their common difficulties, i.e. model parameter selection, initialization and optimization. A proper handling of the common difficulties is essential for making the models operational by a non-expert user, which is a requirement for intensifying and commercializing the use of deformable template models. The thesis is organized as a collection of the most important articles, which has been published during the Ph.D. project. To put these articles into the general context of deformable template models and to pass on an overview of the deformable template model literature, the thesis starts with a compact survey of the deformable template model literature with special focus on representation, model parameter estimation, initialization, optimization and performance measures. The original articles - aligned a bit in notation and corrected from discovered spelling errors and other typos - are enclosed in the appendices. Compared to the literature one contribution is a general scheme for estimation of the model parameters, which applies a combination of a maximum likelihood and minimum distance criterion. Another contribution is a very fast search based initialization algorithm using a filter interpretation of the likelihood model. These two methods can be applied to most deformable template models making a non-expert user able to use the model. A comparative study of a number of optimization algorithms is also reported. In addition a general polygon-based model, an ellipse model and a textile model are proposed and a number of applications have been solved. Finally the Grenander model and the Active Appearance Model have been explored and some extensions are presented.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Fisker, R. (Intern)
Publication date: 2000

Publication information
Original language: English
Series: IMM-PHD-2000-77
Main Research Area: Technical/natural sciences
Electronic versions:
imm1199.pdf

Bibliographical note
ISSN 0909-3192
Source: orbit
Source-ID: 200781
Publication: Research › Ph.D. thesis – Annual report year: 2000

Elastic wave propagation in anisotropic inhomogeneous materials

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Halkjær, S. (Intern)
Number of pages: 133
Publication date: 2000

Publication information
Original language: English
Series: IMM-PHD-2000-74
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 202997
Publication: Research › Ph.D. thesis – Annual report year: 2000

Topics in computational linear optimization

Linear optimization has been an active area of research ever since the pioneering work of G. Dantzig more than 50 years ago. This research has produced a long sequence of practical as well as theoretical improvements of the solution techniques available for solving linear optimization problems. Linear optimization problems cover both linear programming problems, which are polynomially solvable, and mixed integer linear programming problems, which belong to the class of NP-hard problems. The three main reasons for the practical success of linear optimization are: wide applicability, availability of high quality solvers and the use of algebraic modelling systems to handle the communication between the modeller and the solver. This dissertation features four topics in computational linear optimization: A) automatic reformulation of mixed
0/1 linear programs, B) direct solution of sparse unsymmetric systems of linear equations, C) reduction of linear programs and D) integration of algebraic modelling of linear optimization problems in C++. Each of these topics is treated in a separate paper included in this dissertation. The efficiency of solving mixed 0-1 linear programs by linear programming based branch-and-bound algorithms depends heavily on the formulation of the problem. In seeking a better formulation, automatic reformulation employs a range of different techniques for obtaining an equivalent formulation of a given pure or mixed integer programming problem, such that the new formulation has a tighter LP-relaxation than the original formulation. The first paper surveys the use of automatic reformulation techniques for general mixed 0-1 linear programs. An implementation, Spiky, of a direct method for solving large sparse unsymmetric systems of linear equations is presented in the second paper. The method is based on a matrix modification approach applied to a spiked triangular permutation of the system. The factorization consists of 3 steps. 1) a reordering of the matrix, such that only a small number, 's', of spike columns reach above the diagonal is determined. 2) a block LU factorization of an augmented system is computed. This involves the solution of a sparse triangular system with 's' right-hand sides. Solution sparsity is exploited in the sparse triangular solves of the block LU factorization. 3) a factorization of the Schur complement matrix, of order 's', is computed. The idea of this factorization method comes from Gondzio (Gondzio:94c), but has been improved in several ways. Most importantly: a) A new fast reordering algorithm is used. b) Sparsity of the Schur complement is exploited. Simple techniques for reducing the size of linear programs prior to the application of a solution algorithm are incorporated as important parts of most LP solvers. The usual LP reduction techniques require a very limited effort but nevertheless often result in substantial reductions. One of the reasons for this good performance is the way LP models are typically formulated using algebraic modelling languages. The third paper presents a common framework for LP reduction algorithms and shows how the usual LP reduction techniques fit into this framework. It also demonstrate how a stronger bound strengthening and the use of primal information to improve the dual bounds can be used to obtain further reductions. In the fourth and last paper, a prototype implementation of a C++ class library, FLOPC++, for formulating linear optimization problems is presented. Using FLOPC++, linear optimization models can be specified in a declarative style, similar to algebraic modelling languages such as GAMS and AMPL. While preserving the traditional strengths of algebraic modelling languages, FLOPC++ eases the integration of linear optimization models with other software components. The class library implements a full-fledged algebraic modelling language with indexed variables and constraints, repeated sums, index arithmetic and conditional exceptions. Besides the articles the thesis features six introductory chapters, including a description of two real-world applications of linear optimization, in which I have been involved during my Ph.D. study.
Phosphorus diffusion in float zone silicon crystal growth
This Ph.D thesis encompasses a global numerical simulation of the needle-eye oat zone process, used to grow silicon single crystals. The numerical models includes coupled electromagnetic and free surface models and a global heat transfer model, with moving boundaries. An axisymmetric uidow model, including centrifugal, buoyancy, thermocapillary and electromagnetic forces, is used to determine flow field, after the phase boundaries have been determined, by the heat transfer model. A finite element model for calculating dopant transport, using the calculated unsteady flow field, has been developed within this project. This model has furthermore been expanded to two equations coupled by a non-zero right hand side, for simulating transport of point defects in the crystal during growth. Free surface shapes and induced electric surface current are calculated for two different 4'' configurations and a 0.8'' configuration. The heat transfer calculations of the same three configurations, yields the global temperature field, from which temperature gradients are determined. The heat transfer model is furthermore expanded to study convective cooling of the crystal from natural convection in the pressurized surrounding gas, for one of the 4'' configurations. The depth of the lower phase boundaries of both 4'' configurations have with good agreement to the calculations been measured, to provide experimental verification of the heat transfer model. Calculations of melt convection within the floating zone is done for the two 4'' configurations, for four different setups. The flow is unsteady, but laminar and is seen to be repressed by increased rates of rotation. Gas doping is simulated by prescribing a flux of dopant through the free surface, resulting in unsteady dopant concentration fields, with distinct concentration boundary layers at the lower phase boundaries. The dopant concentrations, at the lower phase boundaries, are used to determine radial resistivity profiles, which with fair agreement are compared to measurements. Simulations of defect transport are conducted for both of the 4'', as well as the 0.8"" conformation, for two different values of the recombination factor. The calculation of the 0.8"" crystal is compared to DLTS measurements, revealing good agreement for one of the recombination factors, which however does not fit the Voronkov theory. Both factors are used in the simulation of the two 4"" configurations, which both have so high growth parameters, that the vacancy domination is relatively unaffected by the recombination.
Signal processing in the Dolphin sonar system

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Andersen, L. N. (Intern)
Number of pages: 214
Publication date: 1999

Publication information
Original language: English
Series: IMM-PHD-1999-68
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 274540
Publication: Research › Ph.D. thesis – Annual report year: 2000

Statistical models for standardized preclinical studies

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Andersen, H. (Intern)
Publication date: 1999

Publication information
Original language: English
Series: IMM-PHD-1999-67
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200760
Publication: Research › Ph.D. thesis – Annual report year: 1999

Surface-bounded growth modeling applied to human mandibles.
This thesis presents mathematical and computational techniques for three dimensional growth modeling applied to human mandibles. The longitudinal shape changes make the mandible a complex bone. The teeth erupt and the condylar processes change direction, from pointing predominantly backward to pointing more upward. The full dataset consists of 31 mandibles from six patients. Each patient is longitudinally CT scanned between three and seven times. Age range is 1 month to 12 years old for the scans. Growth modeling consists of three overall steps: 1.extraction of features. 2.registration of the common features. 3.model the process that moves the matched points (growth modeling). A local shape feature called crest line has shown itself to be structurally stable on mandibles. Registration of crest lines (from different mandibles) results in a sparse deformation field, which must be interpolated to yield a spatially dense field. Different methods for constructing the sparse field are compared. Adaptive Gaussian smoothing is the preferred method since it is parameter free and yields good results in practice. A new method, geometry-constrained diffusion, is used to simplify The most successful growth model is linear and based on results from shape analysis and principal component analysis. The growth model is tested in a cross validation study with good results. The worst case mean modeling error in the cross validation study is 3.7 mm. It occurs when modeling the shape and size of a 12 years old mandible based on the 3 month old scan. When using successively more recent scans as basis for the model the error drops to 2.0 mm for the 11 years old scan. Thus, it seems reasonable to assume that the mandibular growth is linear.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Andresen, P. R. (Intern)
Publication date: 1999

Publication information
Modelling of wastewater systems
In this thesis, models of pollution fluxes in the inlet to 2 Danish wastewater treatment plants (WWTPs) as well as of suspended solids (SS) concentrations in the aeration tanks of an alternating WWTP and in the effluent from the aeration tanks are developed. The latter model is furthermore used to analyze and quantify the effect of the Aeration Tank Settling (ATS) operating mode, which is used during rain events. Furthermore, the model is used to propose a control algorithm for the phase lengths during ATS operation. The models are mainly formulated as state space model in continuous time with discrete-time observation equations. The state equations are thus expressed in stochastic differential equations. Hereby it is possible to use the maximum likelihood estimation method to estimate the parameters of the models. A Kalman filter is used to estimate the one-step ahead predictions that are used in the evaluation of the likelihood function. The proposed models are of the grey-box type, where the most important physical relations are combined with stochastic terms to describe the deviations between model and reality as well as measurement errors. The pollution flux models are models of the COD (Chemical Oxygen Demand) flux and SS flux in the inlet to the WWTP. COD is measured by means of a UV absorption sensor while SS is measured by a turbidity sensor. These models include a description of the deposit of COD and SS amounts, respectively, in the sewer system, and the models can thus be used to quantify these amounts as well as to describe possible first flush effects. The buildup and flush out of the deposits are modelled by differential equations, thus the models are dynamic models. The dynamic models are furthermore compared to simpler static models and it is found that the dynamic models are better at modelling the fluxes in terms of the multiple correlation coefficient R2. The model of the SS concentrations in the aeration tanks of an alternating WWTP as well as in the effluent from the aeration tanks is a mass balance model based on measurements of SS in one aeration tank and in the common outlet of all the aeration tanks, respectively. This model is a state space model with the SS concentrations and the sludge blanket depths in the aeration tanks as state variables and with the SS concentrations in one aeration tank and in the common outlet as observations. The SS concentration model is used to quantify the benefits of ATS operation in terms of increased hydraulic capacity. The model is furthermore used to propose a control algorithm for the phase lengths during ATS operation. The quantification of the benefits of ATS operation as well as the proposal for a control algorithm is based on the assumption that if the SS concentration in the secondary clarifier increases beyond a plant and situation specific amount above the normal dry weather level, the SS concentration in the effluent increases to an unacceptable level. It was found that ATS increases the hydraulic capacity of the WWTP considered by more than 167%, while the proposed control algorithm is yet to be implemented in full scale.
Spatial solitons in quasi-phase matched structures

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Clausen, C. A. B. (Intern)
Publication date: 1999

Publication information
Original language: English
Series: IMM-PHD-1999-63
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200776
Publication: Research › Ph.D. thesis – Annual report year: 1999

On chromatic and geometrical calibration
The main subject of the present thesis is different methods for the geometrical and chromatic calibration of cameras in various environments. For the monochromatic issues of the calibration we present the acquisition of monochrome images, the classic monochrome aberrations and the various sources of non-uniformity of the illumination of the image plane. Only the image deforming aberrations and the non-uniformity of illumination are included in the calibration models. The topics of the pinhole camera model and the extension to the Direct Linear Transform (DLT) are described. It is shown how the DLT can be extended with non-linear models of the common lens aberrations/errors some of them caused by manufacturing defects like decentering and thin prism distortion. The relation between a warping and the non-linear defects are shown. The issue of making a good resampling of an image by using the correct interpolation method is described. For the chromatic issues of calibration we present the acquisition of colour and multi-spectral images, the chromatic aberrations and the various lens/camera based non-uniformities of the illumination of the image plane. It is described how the monochromatic calibrations are extended to multi channel images. Since accurate colour images require equal (uniform) intensity levels in all channels, the various sources, in both open and closed scenes, for the non-uniform intensities and some corresponding calibration methods are described. The various possibilities to design calibration targets for both geometrical and chromatic calibration are described. We present some possible systematical errors on the detection of the objects in the calibration targets, if viewed in a non orthogonal angle, if the intensities are uneven or if the image blurring is uneven. Finally we present the implementation of a complete calibration method for an accurate colour texture measurement device called VMX2000, the calibration for uneven laser sheet illumination in a flow measuring system and the use of automatic detection of calibration targets for a DLT/warping in a 3D PIV system.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Folm-Hansen, J. (Intern)
Publication date: 1999

Publication information
Original language: English
Series: IMM-PHD-1999-61
Main Research Area: Technical/natural sciences
Electronic versions:
imm1210.pdf

Bibliographical note
ISSN 0909-3192
Source: orbit
Source-ID: 200788
Publication: Research › Ph.D. thesis – Annual report year: 1999

A probabilistic framework for classification of dermatoscopic images

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Hintz-Madsen, M. (Intern), Hansen, L. K. (Intern), Larsen, J. (Intern)
Number of pages: 156
Publication date: Jul 1999
Signal processing for distribution network monitoring

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Jensen, K. J. (Intern), Sørensen, J. A. (Intern), Munk, S. M. (Intern)
Publication date: Nov 1999

Statistical modelling of fish stocks
In this thesis uncertainty associated with stock assessment has been considered, especially uncertainty associated with the input data to the model. The thesis provides new approaches to analyse the sources of variation in the input data and their magnitude, and an alternative approach for modelling the dynamics of a fish population is suggested. A new approach is introduced to analyse the sources of variation in age composition data, which is one of the most important sources of information in the cohort based models for estimation of stock abundancies and mortalities. The approach combines the continuation-ratio logits, which can take the ordinal and multinomial characteristics of the response into account, and the generalized linear mixed models, which allow for fixed as well as random effects to be analysed. Catch at age data and the associated uncertainties have been estimated, by separating the statistical analysis into separate analyses of the various data sources. The results were combined into estimates of the catch at age data and the associated uncertainties for the sandeel landings from the North Sea in 1989 and 1991. An overview of age-structured stock assessment models is given and it is argued that an approach utilising stochastic differential equations might be advantageous in fish stock assessments.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Kvist, T. (Intern)
Number of pages: 175
Publication date: 1999
Parallelization of the Vehicle Routing Problem with Time Windows

This dissertation presents a number of algorithms for solving the Vehicle Routing Problem with Time Windows (VRPTW). The VRPTW is a generalization of the well known capacity constrained Vehicle Routing Problem (VRP). In the VRP a fleet of vehicles based at a central depot must service a set of customers. In the VRPTW each customer has a time window. Service of a customer must begin within the interval given by the time window. The objective is to minimize some aspect of operating costs (e.g. total distance traveled, number of vehicles needed or a combination of parameters). Since the late 80’s and the beginning of the 90’s optimal methods for the VRPTW have appeared in the literature. Methods have basically been based on three approaches: dynamic programming, Lagrange relaxation and column generation (Dantzig-Wolfe). The most successful approaches rely on column generation. Good results have also been obtained using Lagrange relaxation. This dissertation is divided into three parts. First the theoretical framework is described. Thereafter a number of techniques to improve the performance of the column-generation framework are proposed and analyzed. Finally a parallel algorithm based on the sequential algorithm developed in the previous part of the dissertation is developed and analyzed.

General information
State: Published
Organisations: Operations Research, Department of Informatics and Mathematical Modeling
Authors: Larsen, J. (Intern)
Publication date: May 1999

Publication information
Original language: English
Series: IMM-PHD-1999-62
Main Research Area: Technical/natural sciences
Electronic versions:
imm140.pdf
Source: orbit
Source-ID: 200819
Publication: Research › Ph.D. thesis – Annual report year: 1999

Parametric and Non-Parametric System Modelling
The present thesis consists of ten research papers published in the period 1996-1999 together with a summary report. The thesis deals with different aspects of mathematical modelling of systems using data and, if possible, partial knowledge about the systems. In the first part of the thesis the focus is on combinations of parametric and non-parametric methods of regression. This combination can be in terms of additive models where e.g. one or more non-parametric term is added to a linear regression model. It can also be in terms of conditional parametric models where the coefficients of a linear model are estimated as functions of some explanatory variable(s). Also, software for handling the estimation is presented. The software runs under S-PLUS and R and contains also a number of tools useful when doing model diagnostics or interpreting the results. Adaptive estimation is also considered. It is shown that adaptive estimation in conditional parametric models can be performed by combining the well known methods of local polynomial regression and recursive least squares with exponential forgetting. The approach used for estimation in conditional parametric models also highlights how recursive least squares with exponential forgetting can be generalized and improved by approximating the time-varying parameters with polynomials locally in time. In one of the papers well known tools for structural identification of linear time series are generalized to the non-linear time series case. For this purpose non-parametric methods together with additive models are suggested. Also, a new approach specifically designed to detect non-linearities is introduced. Confidence intervals are constructed by use of bootstrapping. As a link between non-parametric and parametric methods a paper dealing with neural networks is included. In this paper, neural networks are used for predicting the electricity production of a wind farm. The results are compared with results obtained using an adaptively estimated ARX-model. Finally, two papers on stochastic differential equations are included. In the first paper, among other aspects, the properties of a method for parameter estimation in stochastic differential equations is considered within the field of heat dynamics of buildings. In the second paper a lack-of-fit test for stochastic differential equations is presented. The test can be applied to both linear and non-linear stochastic differential equations. Some applications are presented in the papers. In the summary report references are made to a number of other applications. Resumé på dansk: Nærværende afhandling består af ti artikler publiceret i perioden 1996-1999 samt et sammendrag og en perspektivering heraf. I afhandlingen behandles aspekter af matematisk modellering af systemer vha. data og, såfremt det er muligt, delvis viden om disse systemer. Den første del af artiklerne fokuserer på kombinationer af parametriske og ikke-parametriske regressionsmetoder. Sådanne kcombinationer kan være additive, hvor f.eks. et eller flere ikke-parametriske led adderes til en lineær regressionsmodel. En anden mulighed er at bruge parametriske modeller, hvor koefficienterne i en lineær model estimeres som funktioner af en eller flere forklarende variable. Endvidere præsenteres et EDB-program til håndtering af og estimation i sådanne modeller. Programmet er en udvidelse til S-PLUS og R. Programmet inkluderer også en række værktøjer, der er nyttige i forbindelse med diagnostic og fortolkning af resultater. Endvidere behandles adaptiv estimation. Det vises, at der ved at kombinere adaptiv estimation i lineære modeller med lokal polynomial regression, som begge er velkendte metoder, fås en metode, der kan håndtere adaptiv estimation i betinget parametriske modeller. Den anvendte metode til estimation i betinget parametriske modeller lydeligger også, hvorledes den rekursive mindste kvadraters metode med eksponentiel glemsel kan generaliseres og forbedres ved at approksimere de tidsvarierende parametre med polymorier, der er lokale i tid. Ikke-parametriske metoder bruges sammen med additive modeller til at generalisere

General information
State: Published
Organisations: Mathematical Statistics, Department of Informatics and Mathematical Modeling
Authors: Nielsen, H. A. (Intern)
Publication date: 1999

Publication information
Original language: English
Series: IMM-PHD-1999-70
Main Research Area: Technical/natural sciences
Electronic versions:
imm1703.pdf
imm1703.ps
Source: orbit
Source-ID: 200830
Publication: Research › Ph.D. thesis – Annual report year: 1999

Stochastic modelling of dynamic systems

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Nielsen, J. N. (Intern)
Publication date: 1999

Publication information
Original language: English
Series: IMM-PHD-1999-72
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200831
Publication: Research › Ph.D. thesis – Annual report year: 1999

Reconstruction and restoration of PET images

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Cognitive Systems
Authors: Philipsen, P. A. (Intern), Hansen, L. K. (Intern)
Number of pages: 134
Publication date: May 1999

Publication information
Original language: English
Series: IMM-PHD-1998-55
Main Research Area: Technical/natural sciences
Electronic versions:
Thesis.pdf
Source: orbit
Source-ID: 200844
Hidden Markov models and neural networks for speech recognition

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Riis, S. K. (Intern), Hansen, S. D. (Intern)
Number of pages: 223
Publication date: Feb 1999

Publication information
Original language: English
Series: IMM-PHD-1998-46
Main Research Area: Technical/natural sciences
Electronic versions:
thesis.200498.pdf
Source: orbit
Source-ID: 200849
Publication: Research › Ph.D. thesis – Annual report year: 1999

Acquisition and analysis of complex dynamic intra- and intercellular signaling events

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Skyggebjerg, O. (Intern)
Number of pages: 83
Publication date: 1999

Publication information
Original language: English
Series: IMM-PHD-1999-59
Main Research Area: Technical/natural sciences

Bibliographical note
ISSN 0909-3192
Source: orbit
Source-ID: 200859
Publication: Research › Ph.D. thesis – Annual report year: 1999

Spatial distribution maps for benthic communities
The application of hydroacoustic measurements for preparation of spatial distribution maps of benthic communities is reported. For the present study common mussels (Mytilus edulis), neptune grass (Posidonia oceanica) and Cymodocea nodosa, serving as canonical species of many European marine ecosystems, were selected. These species are supposed to be good indicators of marine ecosystem health. The hydroacoustic measurements comprise preprocessed echo sounder recordings and side-scan sonar data forming a large and unique collection of datasets based on 4 field campaigns in Øresund and the Mediterranean. combination of geostatistical methods for spatial interpolation of the echo sounder observations and a set of classification rules, based on discriminant analysis of the feature space of=20 the observations, is found to yield reliable distribution maps when compared to groundtruth data. The data-driven methodology developed is shown to be adaptive to instationarities in the echo sounder observations and is recommended as a substantial improvement of existing methods of sea floor mapping based on echo sounder data. Elaborations of the developed methodology are studied, comprising the use of geostatistical simulation, Markov random fields and Boolean models. Geostatistical simulation provides a means of assessing the variability of random field functionals such as the estimated distribution area of a benthic species. The Markov random field allows the spatial distribution of the benthic communities to be modelled as a less smooth or regular phenomena than assumed when using geostatistical models. The use of Markov random fields in a Markov chain Monte Carlo simulation framework enables an alternative means of assessing variability of image functionals that is based on a sound theoretical basis. The estimates of variability obtained for estimated distribution areas with the two approaches compare satisfactorily. The Boolean models are suggested as a point of departure for embedding models of spatial patterns on the minor scales of observations to be used in up-scaling approaches to enhance the quality of the distribution maps and to be combined with biogeochemical models describing spatiotemporal population dynamics. Finally, the use of side-scan sonar data is illustrated in a data fusion exercise combining side-scan sonar data with the results based on echo sounder measurements. The feasible use of side-scan
sonar for mapping of benthic communities remains an open task to be studied in the future. The data processing methodology developed is a contribution to the emerging field of hydroacoustic marine biology. The method of penalised maximum pseudo-likelihood for estimation of the Ising model under a huge amount of missing pixel data is a contribution to statistical image analysis. Furthermore, the estimation method developed for non-stationary Boolean models that combines scale-space kernel smoothing with the so-called method-of-moments applied to stationary Boolean models is a contribution to stochastic geometry.

Numerical Solution of Differential Algebraic Equations: A PhD course

Statistical analysis of biotests - Applied to complex polluted samples

Multicriteria planning and optimization - Heuristic approaches
Condition monitoring and fault diagnosis in marine diesel engines

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Fog, T. L. (Intern)
Number of pages: 178
Publication date: 1998

Publication information
Original language: English
Series: IMM-PHD-1998-52
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200782
Publication: Research › Ph. D. thesis – Annual report year: 1998

Quality inspection of sugar beets using vision.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Frydendal, I. (Intern)
Publication date: 1998

Publication information
Original language: English
Series: IMM-PHD-1998-48
Main Research Area: Technical/natural sciences

Bibliographical note
ISSN 0909-3192
Source: orbit
Source-ID: 200783
Publication: Research › Ph. D. thesis – Annual report year: 1998

Metaheuristics for multiple objective combinatorial optimization

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Hansen, M. P. (Intern)
Number of pages: 163
Publication date: 1998

Publication information
Original language: English
Signal subspace methods for speech enhancement

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Hansen, P. S. K. (Intern), Hansen, S. D. (Intern), Sørensen, J. A. (Intern)
Number of pages: 214
Publication date: Jun 1998

Bayesian signal processing and interpretation of brain scans

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Kjems, U. (Intern)
Number of pages: 125
Publication date: 1998

Industrial vision
This dissertation is concerned with the introduction of vision-based applications in the ship building industry. The industrial research project is divided into a natural sequence of developments, from basic theoretical projective image generation via CAD and subpixel analysis to a description of an implementation in real production environments. The theory for projection of world points into images is concentrated upon the direct linear transformation (DLT), also called the Extended Pinhole model, and the stability of this method. A complete list of formulas for calculating all parameters in the model is presented, and the variability of the parameters is examined and described. The concept of using CAD together with vision information is based on the fact that all items processed at OSS have an associated complete 3D CAD model that is accessible at all production states. This concept gives numerous possibilities for using vision in applications which otherwise would be very difficult to automate. The requirement for low tolerances in production is, despite the huge dimensions of the items involved, extreme. This fact makes great demands on the ability to do robust sub pixel estimation. A new method based on cross correlation is presented. Working with vision in harsh environments with few possibilities for controlling light, vibrations, electrical noise etc. requires knowledge about all factors and components in the vision system, which can possibly influence the image generated. A description of the experience achieved during the project is provided. The project is industrial oriented. An essential part of the project has been focused on the possibilities for immediate use of the results. A full implemented application doing vision based positioning is described. It is concluded that vision-based applications in ship building are not only possible, but also holds great potential in the area of quality control and automation. The effort involved is not necessarily very great, at least not, if the vision-based information can
Parallel computation of rotating flows

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Lundin, L. K. (Intern)
Number of pages: 106
Publication date: 1998

Publication information
Original language: English
Series: IMM-PHD-1998-49
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200823
Publication: Research › Ph.D. thesis – Annual report year: 1998

A Multivariate Approach to Functional Neuro Modeling

This Ph.D. thesis, A Multivariate Approach to Functional Neuro Modeling, deals with the analysis and modeling of data from functional neuro imaging experiments. A multivariate dataset description is provided which facilitates efficient representation of typical datasets and, more importantly, provides the basis for a generalization theoretical framework relating model performance to model complexity and dataset size. Briefly summarized the major topics discussed in the thesis include: - An introduction of the representation of functional datasets by pairs of neuronal activity patterns and overall conditions governing the functional experiment, via associated micro- and macroscopic variables. The description facilitates an efficient microscopic re-representation, as well as a handle on the link between brain and behavior; the latter is achieved by hypothesizing variations in the micro- and macroscopic variables to be manifestations of an underlying system. - A review of two microscopic basis selection procedures, namely principal component analysis and independent component analysis, with respect to their applicability to functional datasets. - Quantitative model performance assessment via a generalization theoretical framework centered around measures of model generalization error. - Only few, if any, examples of the application of generalization theory to functional neuro modeling currently exist in the literature. - Exemplification of the proposed generalization theoretical framework by the application of linear and more flexible, nonlinear microscopic regression models to a real-world dataset. The dependency of model performance, as quantified by generalization error, on model flexibility and training set size is demonstrated, leading to the important realization that no uniformly optimal model exists. - Model visualization and interpretation techniques. The simplicity of this task for linear models contrasts the difficulties involved when dealing with nonlinear models. Finally, a visualization technique for nonlinear models is proposed. A single observation emerges from the thesis as particularly important; optimal model flexibility is a function of both the complexity and the size of the dataset at hand. This is something that has not received appropriate attention by the functional neuro modeling community so far. The observation implies that optimal model performance rarely is achieved with black-box models; rather, model flexibility must be matched to the specific functional dataset. The potential advantage is a model that more precisely approximates the true nature of the relationship between brain and behavior, thus paving the way for increased insight into the function of the human brain.
New developments in the theory of wheel/rail contact mechanics
Today many simulation routines concerning railway dynamics employ rather primitive contact models which are not necessarily suited for the specific wheel/rail contact problem. The objective of the present thesis is to derive a more flexible contact model which can be applied on a variety of contact problems. When it comes to the modelling of the wheel/rail contact it is always a compromise between computational speed and accuracy. Many numerical methods provide a very good accuracy, but since most railway simulations necessitates the evaluation of many consecutive contact situations the relative slow computational speed is extremely critical. To avoid this problem the present model is based on an analytical approach. The model derived in the thesis is a two-dimensional contact model based on elastic half spaces. It is demonstrated that the solution to a three-dimensional contact problem with no spin has many similarities with the two-dimensional solution. Thus, the results obtained with the present model can qualitatively be extended to the three-dimensional contact problem. The thesis is divided into two parts: one containing the derivation of the contact model and one containing examples of application. The model is applied on four different types of contact problems which cannot be treated with the most common contact models: - contact between corrugated surfaces - contact with velocity dependent friction coefficient - contact between rough surfaces - non-steady contact The calculations demonstrate with much clearness that the solution to the contact problem is very sensitive to the choice of contact model. This illustrates how crucial it is to employ an adequate contact model in a given simulation routine in order to obtain a realistic result. If the assumptions of the contact model do not fulfill the actual contact situation the result can be most erroneous and thus misleading.

Wavelets in scientific computing
Wavelet analysis is a relatively new mathematical discipline which has generated much interest in both theoretical and applied mathematics over the past decade. Crucial to wavelets are their ability to analyze different parts of a function at different scales and the fact that they can represent polynomials up to a certain order exactly. As a consequence, functions with fast oscillations, or even discontinuities, in localized regions may be approximated well by a linear combination of relatively few wavelets. In comparison, a Fourier expansion must use many basis functions to approximate such a function well. These properties of wavelets have lead to some very successful applications within the field of signal processing. This dissertation revolves around the role of wavelets in scientific computing and it falls into three parts: Part I
gives an exposition of the theory of orthogonal, compactly supported wavelets in the context of multiresolution analysis. These wavelets are particularly attractive because they lead to a stable and very efficient algorithm, namely the fast wavelet transform (FWT). We give estimates for the approximation characteristics of wavelets and demonstrate how and why the FWT can be used as a front-end for efficient image compression schemes. Part II deals with vector-parallel implementations of several variants of the Fast Wavelet Transform. We develop an efficient and scalable parallel algorithm for the FWT and derive a model for its performance. Part III is an investigation of the potential for using the special properties of wavelets for solving partial differential equations numerically. Several approaches are identified and two of them are described in detail. The algorithms developed are applied to the nonlinear Schrödinger equation and Burgers’ equation. Numerical results reveal that good performance can be achieved provided that problems are large, solutions are highly localized, and numerical parameters are chosen appropriately, depending on the problem in question.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, Scientific Computing
Authors: Nielsen, O. M. (Intern), Sørensen, M. P. (Intern), Barker, V. A. (Intern)
Publication date: Aug 1998

Publication information
Original language: English
Series: IMM-PHD-1998-43
Main Research Area: Technical/natural sciences
Electronic versions:
thesis.pdf
Source: orbit
Source-ID: 200832
Publication: Research › Ph.D. thesis – Annual report year: 1998

Environmental reference materials methods and case studies
This thesis introduces the reader to the concept of chemical environmental reference materials and their role in traceability and chemical analyses of the environment. A number of models and principles from the literature are described. Some suggestions are made as to how stability studies can be modelled when the length of the study is unknown. Experimental data has been collected from two stability studies of aqueous matrices. The first study regards the stability of TN NO²⁻, N and n in autoclaved wastewater samples over a period of 22 months. Data was collected specifically for this study with two purposes: 1) to investigate the stability of selected analytes in the chosen matrices and 2) to explore the applicability of various statistical models for the description of stability studies. Three univariate and three multivariate stability models have been applied to these data sets. The methods have been evaluated with regard to their robustness towards variations in the chemical analytical method and with regard to the number of times a significant out of control situation is indicated. The second study regards the stability of NH₄-N and total phosphorous in autoclaved seawater samples. This study lasted 22 months as well. The samples were produced and stored according to a 2³ factorial design. The influences of storage temperature, UV radiation and ultra-filtration on the stability of NH₄-N and total phosphorous have been investigated. A Youden plot method is suggested for the graphical evaluation of certification data. The plots illustrate consistency between replicate measurements on samples from a batch of reference material, carried out in a number of laboratories according to a staggered nested design. The development of a reference material is illustrated by a series of experiments with wastewater. The purpose was to improve ortho-phosphate (and total phosphorous) homogeneity. A procedure is suggested which includes freeze-drying and redissolving. All calculations have been performed in SAS® primarily by means of elementary procedures, analyses of variance procedures, SAS Insight and SAS IML (Interactive Matrix Language).

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Schramm-Nielsen, K. E. (Intern)
Number of pages: 261
Publication date: 1998

Publication information
Original language: English
Series: IMM-PHD-1998-58
Main Research Area: Technical/natural sciences
Electronic versions:
imm2492.ps
imm2492.pdf
Source: orbit
Robust Performance And Dissipation of Stochastic Control Systems

The topic of the present dissertation is robustness and performance issues in nonlinear control systems. The control systems in our study are described by nominal models consisting of nonlinear deterministic or stochastic differential equations in Euclidean state space. The nominal models are subject to perturbations which are completely unknown dynamic systems, except that they are known to possess certain properties of dissipation. A dissipation property restricts the dynamic behaviour of the perturbation to conform with a bounded resource; for instance energy. The main contribution of the dissertation is a number of sufficient conditions for robust performance of such systems. Since the perturbations in these uncertain models possess several dissipation properties simultaneously, we study fundamental properties of such multi-dissipative systems. These properties are related to convexity and topology on the space of supply rates. For instance, we give conditions under which the available storage is a continuous convex function of the supply rate.

Dissipation theory in the existing literature applies only to deterministic systems. This is unfortunate since robust control applications typically also contain uncertainty which is better modelled in a probabilistic framework, such as measurement noise. This motivates an extension of the theory of dissipative dynamic systems to stochastic systems. This dissertation presents such an extension: We propose a definition and generalize fundamental results from deterministic dissipation theory to stochastic systems. Furthermore, we argue that stochastic dissipation is a natural fundament for a theory of robust performance of stochastic systems. To this end, we present a number of performance requirements to stochastic systems which can be formulated in terms of dissipation, after which we give sufficient conditions for these requirements to be robust towards multi-dissipative perturbations. A final contribution of the dissertation concerns the problem of simultaneous H-infinity control of a finite number of linear time invariant plants. This problem is a prototype of robust adaptive control problems. We show that the optimal (minimax) controller for this problem is finite dimensional but not based on certainty equivalence, and we discuss the heuristic certainty equivalence controller.

Modelling colliding-pulse mode-locked semiconductor lasers

It is the goal of applied mathematics to study mathematical models of physical systems to elucidate their operation principles. The development of simple theoretical models requires a strong interaction between experimentalists and theoreticians. especially in the case of phenomenological models, where parameter values in a proposed model are fitted,
so that the model output is in good agreement with experiments. In general many parameter values entering a derived or proposed model are determined by experiments while the theoretical models are typically used to improve the device design or to determine the optimum operation conditions. The purpose of this thesis is to elucidate some of the physics of interest in the field of semiconductor laser modelling, semiconductor optics and fiber optics. To be more specific we will investigate: The Colliding-Pulse Mode-Locked (CPM) Quantum Well (QW) laser diode; the excitonic semiconductor response for varying material thickness in the case of linear optics; and modulational instability of electromagnetic waves in media with spatially varying non-linearity.

**Stochastic modelling of central heating systems**

This thesis has been prepared at the Department for Mathematical Modelling (IMM, Institut for Matematisk Modellering) at the Technical University of Denmark (DTU, Danmarks Tekniske Universitet), and at Grundfos A/S in partial fulfillment of the requirements for the degree of Ph.D. in engineering and the degree Erhvervsforsker (a special Danish degree, equivalent to "Industrial Ph.D."). The thesis is mainly concerned with experimental design and system identification for individual components in water based central heating systems. The main contribution to this field is on the nonlinear dynamic modelling of the following components: pipe, radiator, thermostatic valve and house. Each model is based on prior knowledge, on physical properties and experimental conditions. The estimated parameters are directly the physical model parameters. The model performance of each model has been assessed based on a number of validation methods.

**Dynamic modelling of processes in rivers affected by precipitation runoff**

In this thesis, models for the dynamics of oxygen and organic matter in receiving waters (such as rivers and creeks), which are affected by rain, are developed. A time series analysis framework is used, but presented with special emphasis on continuous time state space models. Also, the concept of model identifiability is attended. For estimation of the parameters in the models the maximum likelihood method is used and the Kalman filter employed to evaluate the likelihood function. In the case of non-linear models, the extended Kalman filter is used. To evaluate the models, various residual analysis methods and model validation tools are employed. To develop the water quality model, including hydraulic relations and the states of oxygen and organic matter, the qualitative concepts of the physical, biological and chemical models are introduced. The model types used in this thesis are one-dimensional stochastic models. Most of the models are based on
measurements from one measuring station, though models based on the linear reservoir description applied to
measurements from two measuring stations are also considered. Both time varying and time invariant models are
employed. In the models, the oxygen dynamic complex includes reaeration, photosynthesis, respiration and degradation of
organic matter. The effect of pre-filtering data is investigated, as is various functions for photosynthesis as a function of
solar radiation. For the degradation of organic matter delayed reactions have been studied. In most models, precipitation
in the form of rain have been included to study the impact from this. Finally, the future and industrial perspectives are
presented, along with a list of suggestions for future research related to the subjects considered in this thesis.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Jacobsen, J. L. (Intern)
Publication date: 1997

Publication information
Original language: English
Series: IMM-PHD-1997-30
Main Research Area: Technical/natural sciences
Electronic versions:
imm2439.ps
Source: orbit
Source-ID: 200801
Publication: Research › Ph.D. thesis – Annual report year: 1997

Delta-Domain Predictive Control and Identification for Control
The present thesis is concerned with different aspects of modelling, control and identification of linear systems.
Traditionally, discrete-time sampled-data systems are represented using shift-operator parametrizations. Such
parametrizations are not suitable at fast sampling rates. An alternative parametrization using the so-called delta-operator
is examined. It is shown how to maintain a close correspondence to continuous-time when sampling a system described in
continuous-time by stochastic differential equations. Using delta-operator parametrizations makes it possible to unify
discrete-time and continuous-time theory. In addition these parametrizations possess certain numerical advantages
compared to shift-operator representations. A new prediction method is developed. It is based on ideas from continuous-
time but derived from discrete-time delta-operator models. It is shown to include the optimal minimum-variance predictor
as a special case and to have a well-defined continuous-time limit. By means of this new prediction method a unified
framework for discrete-time and continuous-time predictive control algorithms is developed. This contains a continuous-
time like discrete-time predictive controller which is insensitive to the choice of sampling period and has a well-defined limit
in the continuous-time case. Also more conventional discrete-time predictive control methods may be described within the
unified approach. The predictive control algorithms are extended to frequency weighted criterion functions. Also a state-
space approach is described which extends straightforwardly to the multi-variable case. Finally, aspects on the connection
between system identification and control design are discussed. Several approaches to improve this interconnection have
been proposed. The frequency-distribution of the estimation error with low-complexity models is treated and proves to be
important for the development of control-relevant prefilters in estimation. Iterative approaches are presented, both using
standard estimation methods with prefiltering and non-standard control-relevant estimation methods. New combined
adaptive/iterative techniques are proposed.

General information
State: Published
Adaptive Extremum Control and Wind Turbine Control

This thesis is divided into two parts, i.e., adaptive extremum control and modelling and control of a wind turbine. The first part of the thesis deals with the design of adaptive extremum controllers for some processes which have the behaviour that process should have as high efficiency as possible. Firstly, it is assumed that the nonlinear processes can be divided into a dynamic linear part and static nonlinear part. Consequently the processes with input nonlinearity and output nonlinearity are treated separately. With the nonlinearity at the input it is easy to set up a model which is linear in parameters, and thus directly lends itself to parameter estimation and adaptive control. The extremum control law is derived based on static optimization of a performance function. For a process with nonlinearity at output the intermediate signal between the linear part and nonlinear part plays an important role. If it can be emphasis on control design. The models have been validated by experimental data obtained from an existing wind turbine. The effective wind speed experienced by the rotor of a wind turbine, which is often required by some control methods, is estimated by using a wind turbine as a wind measuring device. The investigation of control design is divided into below rated operation and above rated operation. Below rated power, the aim of control is to extract maximum energy from the wind. The pitch angle of the rotor blades is fixed at its optimal value and turbine speed is adjusted to follow the changes in wind speed. Above rated power, the control design problem is to limit and smooth the output electrical power. The pitch control is investigated for both constant speed and variable speed wind turbines. The minimization of the turbine transient loads is focussed in both cases.

Optimization of recurrent neural networks for time series modeling

The present thesis is about optimization of recurrent neural networks applied to time series modeling. In particular is considered fully recurrent networks working from only a single external input, one layer of nonlinear hidden units and a linear output unit applied to prediction of discrete time series. The overall objectives are to improve training by application of second-order methods and to improve generalization ability by architecture optimization accomplished by pruning. The major topics covered in the thesis are: 1. The problem of training recurrent networks is analyzed from a numerical point of view. Especially it is analyzed how numerical ill-conditioning of the Hessian matrix might arise. 2. Training is significantly improved by application of the damped Gauss-Newton method, involving the Hessian. This method is found to outperform gradient descent in terms of both quality of solution obtained as well as computation time required. 3. A theoretical definition of the generalization error for recurrent networks is provided. This definition justifies a commonly adopted approach for estimating generalization ability. 4. The viability of pruning recurrent networks by the Optimal Brain Damage (OBD) and Optimal Brain Surgeon (OBS) pruning schemes is investigated. OBD is found to be very effective whereas
OBS is severely influenced by numerical problems which leads to pruning of important weights. 5. A novel operational tool for examination of the internal memory of recurrent networks is proposed. The tool allows for assessment of the length of the effective memory of previous inputs built up in the recurrent network during application. Time series modeling is also treated from a more general point of view, namely modeling of the joint probability distribution function of the observed series. Two recurrent models rooted in statistical physics are considered in this respect, namely the "Boltzmann chain" and the "Boltzmann zipper" and a comprehensive tutorial on these models is provided. Boltzmann chains and zippers are found to benefit as well from second-order training and architecture optimization by pruning which is illustrated on artificial problems and a small speech recognition problem.
localized states in the vicinity of an impurity. Studying the discrete nonlinear Schrödinger model, we first analyze the intrinsically localized excitations supported by this model in one dimension. This analysis is accomplished using analytical methods developed for nonlinear maps. It is demonstrated how the nonanalyticity of the map through homoclinic and heteroclinic connections permits the existence of localized states on the lattice. The pinning effect of the discrete lattice is also investigated, constructing a Melnikov function describing qualitatively the difference between on-site and inter-site states. Since the intrinsically localized excitations are rather robust we further study the implications of fluctuations and nonlinear damping in this discrete model. The fluctuations are found always to destroy the localized states. Existence and dynamics of the intrinsically localized excitations in the two-dimensional discrete model are also studied. It is found that in two dimensions a bistability phenomenon of the localized states appears. The bistability expresses itself by allowing localized states of various width to have equal norms. We find in the two-dimensional discrete model that the interplay of the collapse effect and the discrete pinning allows dynamical creation of a spatially distributed set of localized states from a broad initial excitation. The last kind of models studied in the Thesis is nonlinear Schrödinger models with nonlocal dispersive interaction. First a continuum model with an exponential dependence of dispersive interaction is studied. This model shows in contrast to the ordinary continuum nonlinear Schrödinger models that the nonlocality imposes an upper bound on the norm of a possible localized excitation. The model is also shown to support a cusp soliton. A similar discrete nonlocal model is discussed. This model has an algebraic dependence of the dispersive interaction. There exists no upper limit of the norm in the discrete model, but the possibility of a bistability phenomenon similar to that of the two-dimensional model is shown to occur. Finally, we show that a two-dimensional Kronig-Penney model describing for example propagation of electromagnetic waves in photonic bandgap materials can be reduced to a one-dimensional nonlocal nonline Schrödinger model, which is similar to the nonlocal models considered previously.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Rasmussen, K. (Intern), Christiansen, P. L. (Intern)
Publication date: 1997

Publication information
Original language: English
Series: IMM-PHD-1997-33
Main Research Area: Technical/natural sciences
Electronic versions: imm2442.ps
Source: orbit
Source-ID: 200848
Publication: Research › Ph.D. thesis – Annual report year: 1997

Statistical methods for segmentation and classification of images
The central matter of the present thesis is Bayesian statistical inference applied to classification of images. An initial review of Markov Random Fields relates to the modeling aspect of the indicated main subject. In that connection, emphasis is put on the relatively unknown sub-class of Pickard Random Fields (PRF’s). The properties of this type of random fields are given a specially thorough treatment, including an investigation of a previously unresolved general parameterization issue. Novel insight into the parameterization of discrete versions of these fields is given. A visual evaluation of the properties of Pickard Random Fields has been enabled by simulations of both continuous and discrete fields. Simulations of a compared Potts model from the traditional MRF theory are also given. Variations of the surveyed Markov Random Fields are used as prior and observation models, respectively, in some presented Bayesian classification techniques. The applied techniques implement either the traditional MAP criterion or the MPM criterion, both from Bayesian decision theory. The latter criterion have previously been shown to have a fast approximate solution when used in combination with a Pickard Random Field modeling of a considered (categorical) image phenomenon. An extension of the fast PRF based classification technique is presented. The modification introduces auto-correlation into the model of an involved noise process, which previously has been assumed independent. The suitability of the extended model is documented by tests on controlled image data containing auto-correlated noise.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Rosholm, A. (Intern)
Publication date: 1997

Publication information
Original language: English
Series: IMM-PHD-1997-39
Main Research Area: Technical/natural sciences
Variability in microbiological degradation experiments, analysis and case study

The variability of parameter estimates in microbiological degradation models has not received much attention in the literature. This in spite of the fact that the parameters are used in models for predicting and controlling microbiological processes of commercial interest. Furthermore, the accuracy of the parameter estimates are depending of the choice of estimation method, this fact has not either received much attention, all though an unsuitable estimation method can lead to estimates which are quite different from the "true" values. The present thesis describes various nonlinear estimation techniques and describes analysis techniques for testing the reproducibility of a given experiment. The parameter estimation method employed for the experiments in this study is based on an iterative maximum likelihood method and the test statistic is an approximated likelihood ratio test. The estimations were carried out by the nonlinear estimation program Dekimo (developed at IMM by Bilbo and Sommer), available on request. The program successfully fitted all experiments. A few estimations were also carried out by the Lineweaver-Burk linearization, but the estimated parameters fitted the data poorly due to the inappropriate estimation method. The examination of reproducibility/variability were carried out for two kinds of experiments: A single substrate experiment with toluene and a dual substrate experiment with toluene and benzene. A pure culture, isolated from soil, grew with benzene and/or toluene as the only carbon and energy source. The substrates were degraded in batches under aerobic conditions. The Monod model was employed to describe the biological processes in the single substrate system, and 'Bailey & Ollis' model was employed to describe the processes in the dual substrate system. In the single substrate system 9 identical experiments were performed on three different days, and in the dual substrate system 12 identical experiments were performed on four different days. The data are available on the address: http://www.imm.dtu.dk/documents/ftp/phdliste/phd31.abstract.html Experimental observations indicate that these microbiological degradation experiments have a limited reproducibility, i.e. that a common set of parameter estimates could not be employed to describe all experiments. However, experiments carried out on the same days (within runs) were more uniform than experiment carried out on different days (between runs). In the single substrate system a common sets of parameter estimates for experiments within runs fitted the data very well, whereas common sets of parameter estimates for experiments between runs fitted the data poorly and were moreover strongly rejected to be identical by the likelihood ratio test. In the dual substrate system a common set of parameter estimates could not be accepted nor within the runs neither between the runs. Never the less, experiments within the runs were more uniform compared to experiments carried out on different days (between the runs). The lag phases within runs were thus exactly the same, but were quite different from experiments from different runs. The limited reproducibility is probably caused by variability in the precultures, more precisely, variations in the activity level of the precultures just before used as inoculum. Facing the fact that these microbiological degradation experiments have a limited reproducibility one must in general expect large variability on the parameter estimates.
FADBAD/TADIFF. Some examples showing how to use the three methods are presented. A feature of FADBAD/TADIFF not present in other automatic differentiation packages is the possibility to combine the three methods in an extremely flexible way. We examine some applications where this flexibility is very useful. A method for Taylor expanding solutions of ordinary differential equations is presented, and a method for obtaining interval enclosures of the truncation errors incurred, when truncating these Taylor series expansions is described. By combining the forward method and the Taylor expansion method, it is possible to implement the (extended) mean value enclosure of a truncated Taylor series expansion with enclosures of the truncation errors. A C++ program package ADIODES, using this method has been developed. (ADIODES is an abbreviation of ”Automatic Differentiation Interval Ordinary Differential Equation Solver“). ADIODES is used to prove existence and uniqueness of periodic solutions to specific ordinary differential equations occurring in dynamical systems theory. These proofs of existence and uniqueness are difficult or impossible to obtain using other known methods. Also, a method for solving boundary value problems is described. Finally a method for enclosing solutions to a class of integral equations is described. This method is based on the mean value enclosure of an integral operator and uses interval Bernstein polynomials for enclosing the solution. Two numerical examples are given, using two orders of approximation and using different numbers of discretization points.
**Holomorphic Dynamical Systems in the Complex Plane**

General information  
State: Published  
Organisations: Department of Mathematics  
Authors: Branner, B. (Intern)  
Number of pages: 20  
Publication date: 1996

Medical Image Registration and Surgery Simulation  
This thesis explores the application of physical models in medical image registration and surgery simulation. The continuum models of elasticity and viscous fluids are described in detail, and this knowledge is used as a basis for most of the methods described here. Real-time deformable models for surgery simulation Real-time deformable models, using finite element models of linear elasticity, have been developed for surgery simulation. The time consumption of the finite element method is reduced dramatically, by the use of condensation techniques, explicit inversion of the stiffness matrix, and the use of selective matrix vector multiplication. Fluid medical image registration A new and faster algorithm for non-rigid registration using viscous fluid models is presented. This algorithm replaces the core part of the original algorithm with multi-resolution convolution using a new filter, which implements the linear elasticity operator. Using the filter results in a speedup of at least an order of magnitude. Use of convolution hardware is expected to improve the performance even more. Mandibular growth for time registration of mandibles Non-rigid registration using a physically valid model of bone growth is also presented. Using medical knowledge about the growth processes of the mandibular bone, a registration algorithm for time sequence images of the mandible is developed. Since this registration algorithm models the actual development of the mandible, it is possible to simulate the development. Rigid medical image registration Rigid image registration using voxel similarity measures are reviewed, and new measures based on Grey Level Cooccurrence Matrices (GLCM) are introduced. These measures are evaluated extensively using CT, MR, and cryosection images from the Visible Human data set. The results show that mutual information remains the best generally applicable measure. But for specific modality combinations the new GLCM measures show considerable promise.

Predictive control and identification: Applications to steering dynamics  
The main objective of the present thesis is to enhance insight into predictive controller design and identification in connection with steering dynamics. In Chapter 2, the dynamics of ship steering are reviewed. Models of different complexity, suitable for control systems design are presented. The influence of wind, waves and currents on the ship motions are also discussed. Chapter 3 deals with the model reduction problem. Some basic concepts are explained, due to their role in the reduction of the dynamic models. Two model reductions techniques, based on singular values, are described. The theoretical properties of these methods are studied and their performance is examined via simulation on a stochastic linear Mariner Class Vessel model. In Chapter 4, the attention is focused on the derivation of an extended GPC.
This extended strategy implies a generalization of the model structure and of the loss function, which defines the optimality of the control. Some guidelines on how to choose the design parameters, depending on the type of process to be controlled and on the required control performance, are presented. A predictive track keeping system for a Mariner Class Vessel is formulated based on the minimization of the mean squares prediction errors of the ship's deviation from the desired track. Chapter 5 is concerned with constrained predictive control. The presented algorithm, which is based on Rosen's gradient projection method, minimizes a multi-step quadratic loss function, taking physical constrains systematically into account. The constraints may consist of amplitude constraints (signal level constraints) as well as rate constraints. The influences of the different parameters on the solution are illustrated via simulation experiments on a Mariner Class Vessel model. The results show that the proposed strategy leads to a significant better control than the ad-hoc control strategy. Chapter 6 gives a survey on the so-called forgetting factor methods designed for tracking slowly drifting system parameters. The goal of this chapter is to formulate the identification framework in order to support the understanding of the connection between identification and control, analysed in Chapter 7. Chapter 7 focuses on how to make the on-line identification for predictive control more robust towards unmodelled dynamics. The theory is verified via simulation studies on a Mariner Class Vessel. The effects and the need of a prefilter in the estimation are analysed and illustrated. Based on the idea that the control criterion must be dual to the estimation criterion, an iterative optimal prefilter is designed. This seems to be an appealing way to tune the model towards the objective for which the model is to be used.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Hansen, A. D. (Intern)
Publication date: 1996

Publication information
Original language: English
Series: IMM-PHD-1996-22
Main Research Area: Technical/natural sciences
Electronic versions:
imm2104.ps
Source: orbit
Source-ID: 200792
Publication: Research › Ph.D. thesis – Annual report year: 1996

Analysis of irregularly distributed points
The present thesis is on the analysis of irregularly distributed points. The main part of the thesis is concerned with interpolating and restoration of irregularly distributed points. The least squares methods of kriging and Kalman filtering and the Bayesian restoration method of iterated conditional modes are applied to this problem. The Kalman filter is described as a powerful tool for modelling two-dimensional data. Motivated by the development of the reduced update Kalman filter we propose a reduced update Kalman smoother which offers considerable computational savings. Kriging is described as a robust estimator which may be applied straightforwardly to a wide range of point patterns and processes when the correlation structure is known. We give a qualitative and quantitative comparison of kriging, Kalman filter and iterated conditional modes. The Kalman filter have in a case study on fusion of maps with different spatial resolution, shown to provide a powerful modelling of autocorrelated noise structures. The Kalman filter have shown to be superior to ordinary kriging in precision and computational speed. Simple kriging has same statistical properties as the Kalman filter, but the usage of simple kriging may lead to ill-conditioned matrices when applied to highly irregularly distributed points. Adaptive Kalman filter schemes are investigated. A new parallel Kalman filter algorithm based on windowing technique gives good results in a case study on the Igallic satellite scene and represents an interesting contextul classifier. Extended Kalman filtering on the other hand seems to be well suited for interpolation in gradually changing environments. Bayesian restoration is applied to a point matching problem, which consists of matching a grid to an image of (irregularly) distributed point observations. We present an extension to an existing grid model, which is based on a combined line- and point-process. A pseudolikelihood estimator for the parameters is introduced, which is defined in terms of the semivariance structure. The developed models have been applied to a case study on hybridisation analysis, which comprise matching a grid to an arrayed set of DNA- clones spotted onto a hybridisation filter. The line process has proven to perform a satisfactorily modelling of shifted fields (subgrids) in the hybridisation grid, and a two-staged hierarchical grid matching scheme which was designed to firstly locate the overall positions of "node-blocks" in the grid and secondly locate the individual positions of grid nodes has proven to work.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Hartelius, K. (Intern)
Publication date: 1996

Publication information
Original language: English
Experiment design and optimization in complex systems

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Sadegh, P. (Intern)
Number of pages: 162
Publication date: 1996

Publication information
Original language: English

Series: IMM-PHD-1996-23
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200855
Publication: Research › Ph.D. thesis – Annual report year: 1996

Quantitative methods for the analysis of electron microscope images
The topic of this thesis is an general introduction to quantitative methods for the analysis of digital microscope images. The images presented are primarily been acquired from Scanning Electron Microscopes (SEM) and interfermeter microscopes (IFM). The topic is approached though several examples in a number work cases. These mainly falls in the three categories: (i) Description of coarse scale measures to quantify surface structure or texture (topography); (ii) Characterization of fracture surfaces in steels (fractography); (iii) Grain boundary segmentation in sintered ceramics. The theoretical foundation of the thesis fall in the areas of: 1) Mathematical Morphology; 2) Distance transforms and applications; and 3) Fractal geometry. Image analysis opens in general the possibility of a quantitative and statistical well founded measurement of digital microscope images. Herein lies also the conditions to correlate the micro-structure of materials to macroscopic properties, such as tensile strength, durability, chemical resistance.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Skands, P. U. V. (Intern)
Number of pages: 198
Publication date: 1996

Publication information
Original language: English

Series: IMM-PHD-1996-24
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200858
Publication: Research › Ph.D. thesis – Annual report year: 1996

Modelling of packet traffic with matrix analytic methods
The dissertation is concerned with modelling various performance aspects pertaining to packet switched telecommunication networks. The emphasis has been put on versatile modelling of the packet arrival process which is a very relevant issue in cotext with the future Broadband Integrated Service Data Network (B-ISDN). As a case study of packet switched networks Common Channel Signalling System No. 7(SS7) has been used. SS7 is the packet switched network currently employed to command/control the public switched telephone network i.e. the setup and teardown of trunks and the basis for intelligent network services i.e. 800 and 900 calls and advanced mobile communication services.
The Markovian Arrival Process (MAP) has been used as a versatile tool to model the packet arrival process. Applying the MAP facilitates the use of Matrix Analytic methods to obtain performance measures associated with for example the single server queue with a MAP arrival process and a general service time distribution. Measured SS7 traffic data has been analyzed as a part of this study. Recently there has been expressed concern regarding adverse behaviour of measured SS7 traffic i.e. long range dependence. Our studies did not reveal any adverse behaviour. In fact the observed traffic seemed very close to what would be expected from Poisson traffic. The Changeover/Changeback procedure in SS7, which is used to redirect traffic in case of link failure, has been analyzed. The transient behaviour during a Changeover scenario was modeled using Markovian models. The Ordinary Differential Equations arising from these models were solved numerically. The results obtained seemed very similar to those obtained using a different method in previous work by Akinpelu & Skoog 1985. Recent measurement studies of packet traffic from Local Area Networks (LAN) carried out at Bell Communications Research (Bellcore), New Jersey, have revealed behaviour very different from what conventional teletraffic models assume. In fact the observed traffic displays behaviour with (second order) self-similar characteristics over several time-scales. In this study we show that 8-16 state MAPs are able to capture this very variable behaviour over several timescales. The queueing behaviour of these MAPs has been analyzed with Matrix Analytic methods. The results correspond to those obtained by trace driven simulations of measured LAN traffic. It is shown that two arrival processes with the same first and second order properties of their counting processes can yield substantially different queueing behaviour. This illustrates that one should by very careful when trying to assess queueing behaviour from the first and second order properties of the arrival process. A heuristic formula for the tail behaviour of a single server queue fed by a superposition of renewal processes has been evaluated. The evaluation was performed by applying Matrix Analytic methods. The heuristic formula has applications in the Call Admission Control (CAC) procedure of the future BISDN network. The heuristic formula did not seem to yield substantially better results than already available approximations. Finally, some results for the finite capacity BMAP/G/1 queue have been obtained. The steady state probability vector of the embedded chain is found by a direct method where a modified LU factorisation scheme is employed. The standard LU factorisation scheme is modified so that all computations are done cancellation free and in a manner that exploits the special structure of the embedded chain yielding substantial computational savings. The queue-length at an arbitrary time is found by noting the close relationship with the expressions for the corresponding infinite queue. For the special case of a batch Poisson arrival process this observation makes it possible to express the queue length at an arbitrary time in terms of the corresponding queue lengths for the infinite case.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Andersen, A. T. (Intern)
Number of pages: 242
Publication date: 1995

Publication information
Original language: English
Series: IMM-PHD-1995-18
Main Research Area: Technical/natural sciences
Electronic versions: imm2101.ps
Source: orbit
Source-ID: 200762
Publication: Research › Ph.D. thesis – Annual report year: 1995

Distributed-memory matrix computations
The main goal of this project is to investigate, develop, and implement algorithms for numerical linear algebra on parallel computers in order to acquire expertise in methods for parallel computations. An important motivation for analyzing and investigating the potential for parallelism in these algorithms is that many scientific applications rely heavily on the performance of the involved dense linear algebra building blocks. Even though we consider the distributed-memory as well as the shared-memory programming paradigm, the major part of the thesis is dedicated to distributed-memory architectures. We emphasize distributed-memory massively parallel computers - such as the Connection Machines model CM-200 and model CM-5/CM-5E - available to us at UNI-C and at Thinking Machines Corporation. The CM-200 was at the time this project started one of the few existing massively parallel computers. Several areas in the numerical linear algebra field are investigated and they illustrate the problems that arise as well as the techniques that are related to the use of massively parallel computers: 1. Study of Strassen's matrix-matrix multiplication on the Connection Machine model CM-200. What performance can we expect to achieve? Why? 2. Solving systems of linear equations using a Strassen-type matrix-inversion algorithm. A good way to solve systems of linear equations on massively parallel computers? 3. Aspects of computing the singular value decomposition on the Connection Machine CM-5/CM-5E. What are the guidelines to follow in order to achieve an efficient, highly parallel and scalable implementation of the considered algorithms? What are the numerical properties of our implementation? 4. A relatively new algorithm to compute the singular vectors of bidiagonal matrix via Rutishauser's qd algorithm is investigated. This algorithm is built on top of several scan-operations. What difficulties occur when implementing this algorithm to massively parallel computers?
Numerical studies of unsteady coherent structures and transport in two-dimensional flows

Modelling, estimation and control of fast sampled dynamical systems

Teoretiske og eksperimentelle dynamiske undersøgelser af jernbanekøretøjer

The present thesis is concerned with the dynamics of railway vehicles. In the thesis the dynamics of a mathematical model of the Danish IC3 train is treated. The mathematical model consists of half a carbody, one bogie and two wheelsets. The creepage-creepforce relation is approximated by the theory of Shen, Hedrick and Elkins. The derivation of the dynamical equations is described, and special focus is placed on the contact between wheel and rail. This contact is treated locally elastic which means that the wheel is allowed to penetrate the rail. This method is capable of taking into account multiple coexisting contact areas between the wheel and rail. Various examples of the geometrical contact between wheel and rail are given and the influence of track gauge and rail inclination on the contact is outlined. The contact-geometry for a worn wheel-profile on new rails is determined and used in the simulations of the dynamics of the IC3 train. In the thesis the dynamics are numerically investigated for the IC3 running on three different track configurations: 1) A perfectly straight...
track 2) A measured track with irregularities 3) A curve When running on a straight track, we follow the linear and nonlinear critical velocities in dependence of the coefficient of adhesion, the longitudinal stiffness of the primary suspension and the damping coefficient of the secondary yaw-damper. By the term "nonlinear critical velocity" is meant the lowest velocity where a stable oscillating solution vanishes in a saddle-node bifurcation. The critical velocities are computed for the track-gauges 1432 mm and 1435 mm and the rail-inclinations 1/20 and 1/40. The difference between linear and nonlinear analysis is discussed in detail. It is shown that the difference is quite substantial. The numerical calculations indicate that the track-gauge and rail-inclination has a great influence on the dynamics of the IC3 train. The computer simulations of the IC3 running on a measured track with irregularities are compared with the measured signals obtained during a test-run. The physical conditions when negotiating a curve are discussed. Finally a sum up of the experimental experiences obtained during test-runs are presented. A comparison and discussion of the test and computer results are given.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Jensen, J. C. (Intern)
Number of pages: 174
Publication date: 1995

Publication information
Original language: English
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200804
Publication: Research › Ph.D. thesis – Annual report year: 1995

Nonlinear Hamiltonian systems
It is generally very difficult to solve nonlinear systems, and such systems often possess chaotic solutions. In the rare event that a system is completely solvable, it is said to integrable. Such systems never have chaotic solutions. Using the Inverse Scattering Transform Method (ISTM) two particular configurations of the Discrete Self-Trapping (DST) system are shown to be completely solvable. One of these systems includes the Toda lattice in a certain limit. An explicit integration is carried through for this Near-Toda lattice. The Near-Toda lattice is then generalized to include singular boundary terms, while at the same time retaining the integrability. When quantizing products of momentum p and position q an ambiguity arises. This is discussed in detail and the need for choosing a particular ordering is shown. The Symmetric Ordering rule, which is equivalent to Weyl's rule, is considered in detail. Explicit formulae for quantizing arbitrary functions of p and q are derived. When the basis functions are chosen as eigenfunctions of the harmonic oscillator, explicit formulae are obtained for the matrix elements of the Hamiltonian. Properties of the solutions to the radially symmetric two-dimensional defocusing Nonlinear Schroedinger (NLS) equation are studied analytically and numerically. It is found that no bound states exist. When the initial condition is a dark ring on a background of finite amplitude, the ring initially shrinks until the curvature effects become dominant, forcing the ring to expand to infinity with constant velocity.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Jørgensen, M. F. (Intern)
Number of pages: 120
Publication date: 1995

Publication information
Original language: English
Series: IMM-PHD-1995-14
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200809
Publication: Research › Ph.D. thesis – Annual report year: 1995

Exact methods for time constrained routing and related scheduling problems
This dissertation presents a number of optimization methods for the Vehicle Routing Problem with Time Windows (VRPTW). The VRPTW is a generalization of the well known capacity constrained Vehicle Routing Problem (VRP), where a fleet of vehicles based at a central depot must service a set of customers. In the VRPTW customers must be serviced within a given time period - a so called time window. The objective can be to minimize operating costs (e.g. distance travelled), fixed costs (e.g. the number of vehicles needed) or a combination of these component costs. During the last decade optimization methods, i.e. methods guaranteeing to find a proven optimal solution, have been studied by a number
of researchers. The most successful approaches until now have been the Dantzig-Wolfe decomposition (column
generation) method of Desrochers, Desrosiers and Solomon (1992) and the Variable Splitting approach of Jørnsten,
Madsen and Sørensen (1986), which has been tested computationally by Halse (1992). Both methods decompose the
problem into a series of time and capacity constrained shortest path problems. This yields a tight lower bound on the
optimal objective, and the dual gap can often be closed by branch-and-bound methods. This work is closely related to
these methods. The aim is to synthesize the research as well as develop the methods further with some original
contributions. The dissertation is divided in to three parts. First, the theoretic framework is developed, and a number
of methods are proposed and analyzed. Second, some of the methods are tested computationally. In the third part,
generalizations of the VRPTW are considered. In the theoretic part of the dissertation, we outline the relationship between
the methods based on constrained shortest path decompositions and show that the only real difference is how the
coordinating master problem - a concave non-differentiable maximization problem - is solved. We show how the
constrained shortest path problem can be solved efficiently, and present a number of different strategies for solving the
master problem. The lower bound obtainable can be improved further by incorporation of valid inequalities. We show how
this can be done computationally and we introduce a number of valid inequalities for the VRPTW. Finally we present a
number of strategies, primarily branch-and-bound, to obtain integer solutions. In the computational part of the dissertation,
we test some of the proposed methods on the well known set of benchmark problems by Solomon (1987). On basis of
these experiments an optimization algorithm is constructed and executed on the benchmark problems. Of the 87 problems
70 were solved to optimality. This is to be compared to Desrochers, Desrosiers and Solomon (1992) who were able to
solve 50 of the problems, and Halse (1992) who solved 33. The main reason for the success of the algorithm is the
exploitation of valid inequalities. The increase in speed of computers since 1992 play only a minor role. In the last part of
the dissertation a number of generalizations of the VRPTW are considered. We discuss complex routing problems with
different constraints as well as important real world scheduling problems arising in transportation companies. We show
how these problems can be modelled in the same framework as the VRPTW. This means, that the methods developed
can be applied to a large number of important routing and scheduling problems. The dissertation gives a state-of-the-art
review of optimization methods for the VRPTW based on constrained shortest path decompositions. It also contains a
number of new theoretic results, and is the first application of valid inequalities on the VRPTW. The algorithm developed
represents a major step forward in terms of computational ability to solve the VRPTW. Solutions to a large number of
previously unsolved problems are reported.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Kohl, N. (Intern)
Number of pages: 234
Publication date: 1995

Publication information
Original language: English
Series: IMM-PHD-1995-16
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200813
Publication: Research › Ph.D. thesis – Annual report year: 1995

On automatic visual inspection of reflective surfaces
This thesis describes different methods to perform automatic visual inspection of reflective manufactured products, with the
aim of increasing productivity, reduce cost and improve the quality level of the production. We investigate two different
systems performing automatic visual inspection. The first is the inspection of highly reflective aluminum sheets, used by
the Danish company Bang & Olufsen, as a part of the exterior design and general appearance of their audio and video
products. The second is the inspection of IBM hard disk read/write heads for defects during manufacturing. We have
surveyed visual inspection system design methods and presented available image processing hardware to perform high
resolution image capture. We present general usable practical visual inspection system solutions, when performing high
resolution visual inspection of surfaces. We have presented known and new lighting methods in a framework, general
usable for inspecting reflective surfaces. Special attention has been given to the design of illumination techniques to
enhance defects of highly reflective aluminum sheets. The chosen optical system setup has been used to enhance
surface defects of other reflective surfaces, providing new and exciting applications subject to automated visual inspection.
Several contextual features have been surveyed along with introduction of novel methods to perform data-dependent
enhancement of local surface appearance. Morphological methods have been described and utilized in algorithms for
detecting 3-dimensional surface damages based on images from a novel structured lighting setup enhancing the
appearance of these defects in specular surfaces. A hardware implementable polynomial classifier structure has been
described and compared to better known techniques based on multidimensional Gaussian models and tree classifiers. We
have introduced a reject class definition for this classifier, and compared it against the classical Mahalanobis distance
approach. Finally, an evaluation of the total system performance in the case of inspecting reflective surface for scratches,
and inspection of hard disk sliders is presented.
Nonlinearity in superconductivity and Josephson Junctions

Spatial structure of selected porous media has been analysed in terms of the two first spatial moments (i.e. porosity and autocorrelation). Having established directional isotropy in the three spatial planes, multiple geometrical features measured in 2-d are attempted generalized to 3-d using stereological methods. The measured sample autocorrelations are modeled by analytical correlation functions. A method for simulating porous networks from their porosity and spatial correlation originally developed by Joshi (14) is presented. This method is based on a conversion between spatial autocorrelation functions of Gaussian fields and spatial autocorrelation functions of binary fields. An enhanced approach which embodies semi-analytical solutions for the conversions has been made. The scope and limitations of the method have been analysed in terms of realizability of different model correlation functions in binary fields. Percolation threshold of reconstructed porous media has been determined for different discretizations of a selected model correlation function. Also critical exponents such as the correlation length exponent \( \nu \), the strength of the infinite network and the mean size of finite clusters have been determined. We have obtained results which indicate that the effect of spatial correlation does affect not only the percolation threshold but also the exponents with respect to the values known for random media. We have attempted to predict key percolation values for a continuous medium (i.e. beyond discretization effects).

Porous media: Analysis, reconstruction and percolation

Spatial structure of selected porous media has been analysed in terms of the two first spatial moments (i.e. porosity and autocorrelation). Having established directional isotropy in the three spatial planes, multiple geometrical features measured in 2-d are attempted generalized to 3-d using stereological methods. The measured sample autocorrelations are modeled by analytical correlation functions. A method for simulating porous networks from their porosity and spatial correlation originally developed by Joshi (14) is presented. This method is based on a conversion between spatial autocorrelation functions of Gaussian fields and spatial autocorrelation functions of binary fields. An enhanced approach which embodies semi-analytical solutions for the conversions has been made. The scope and limitations of the method have been analysed in terms of realizability of different model correlation functions in binary fields. Percolation threshold of reconstructed porous media has been determined for different discretizations of a selected model correlation function. Also critical exponents such as the correlation length exponent \( \nu \), the strength of the infinite network and the mean size of finite clusters have been determined. We have obtained results which indicate that the effect of spatial correlation does affect not only the percolation threshold but also the exponents with respect to the values known for random media. We have attempted to predict key percolation values for a continuous medium (i.e. beyond discretization effects).
Segmentation and classification of biological objects
The present thesis is on segmentation and classification of biological objects using statistical methods. It is based on case studies dealing with different kinds of pork meat images, and we introduce appropriate statistical methods to solve the tasks in the case studies. The case studies concern classification of back bacon slices from images of back bacon, prediction of ham weight from images of the carcass, and estimation of meat percent from cross-sectional images of the carcass. The first case study investigates three different classifiers ability to classify the quality of back bacon slices. The back bacon slices are classified into four ordered classes representing the quality, and we use the sizes of different meat and fat areas of the slices as variables. The classifiers are Bayesian discriminant functions, Classification and Regression Trees, and feed-forward neural networks with back-propagation. We compare the classifiers in respect to the way they classify the back bacon slices. We predict the ham weight on pork carcasses, before the carcasses are divided into front, middle, and ham part, in the second case study. Principal component analysis applied to the shapes of the carcass is used to predict the ham weight. Given two images of a pork carcass the shapes are defined by landmarks based on anatomical knowledge. We explore how the principal components describe the variation on the carcass shapes. In the third case study we measure the sizes of meat and fat areas on cross-sectional cuts from pork carcasses and use them to estimate the meat percent. We describe a bimodal histogram transformation which is used to equalize the cross-sectional images. The meat and fat areas are segmented using deformable templates. The deformable templates are studied thoroughly, and the segmentation procedures are designed to the specific task.

General information
State: Published
Organisations: CICT, Department of Informatics and Mathematical Modeling
Authors: Schultz, N. (Intern)
Publication date: 1995

Publication information
Original language: English
Series: IMM-PHD-1995-13
Main Research Area: Technical/natural sciences

Bibliographical note
ISSN 0909-3192
Source: orbit
Source-ID: 200857
Publication: Research › Ph.D. thesis – Annual report year: 1995

On the interaction between wheels and rails in railway dynamics
The present thesis is concerned with the mathematical modelling of a railway vehicle. The modelling does not only deal with the vehicle but also the track it runs on. Different models are described and investigated as to how they affect the dynamics of the vehicle. The bulk of the investigations is focussed on the stability of the vehicle. For this we introduce two stability criteria: the linear critical speed and the nonlinear critical speed. The linear critical speed is the vehicle speed at which the vehicle becomes unstable in a linear analyses, while the nonlinear critical speed is such that no oscillating solutions occur below this vehicle speed. The difference between a linear and a nonlinear analysis is hereby pointed out. The oscillating solutions found are analysed by applying methods from the nonlinear dynamics. By this periodic and chaotic solutions are described, for instance a scenario of a period adding sequence. For the vehicle we use a two-car test vehicle with a prototype of a single-axle bogie (a so-called KERF bogie). The vehicle is from the Danish State Railways and runs on the Copenhagen network. What is special about this vehicle is that the single axle bogie is steered by a mechanical steering system. Interest is focussed on the single-axle bogie. For simplification a model of the single-axle bogie alone is analysed under different modelling conditions. The dynamics of a model of the whole vehicle are investigated on: A) a straight track, B) a curved track, C) a track flexible in the vertical direction, D) a track with irregularities. Among other things, the investigations lead to the understanding of the influence of the stiffness in the steering system. On the irregular track the simulations are compared with corresponding measurements. Furthermore two different models are developed for the track: A simple model of the whole track as one rigid body following each wheelset. An elastic model where the rails are modelled as Euler-Bernoulli beams discretely supported by rigid sleepers. The simple model is used to find the influence of a flexible track on the dynamics of the single-axle bogie, while the elastic model is more useful to for example study the effect of changes in flexibility along the track. Finally measurements of the flexibility of tracks are described.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Slivsgaard, E. C. (Intern), True, H. (Intern)
On the spatial and temporal correlations in experimentation with agricultural applications
The present thesis describes design and analysis of agricultural experiments utilizing the spatial and temporal correlation between the measurements. The thesis is organized in three parts, spatial experimental design in Part 1, analysis of temporally correlated measurements in Part 2 and a brief introduction to spatio-temporal models in part 3. Classical statistical analysis normally assumes independent observations. Therefore, knowledge concerning the spatial and temporal relation between plots and between measurements are not included in this kind of analysis. However, agricultural experiments often contain spatial correlations due to a spatial layout and/or temporal correlation due to repeated sampling of measurements at the same experimental unit. A method for design of field experiments is proposed in Part 1. The residual variance between plots in different layouts is used to compare different layouts. The optimal design and layout from a statistical point of view is the one with the smallest residual variance. The residual variance between plots consists of an error term which depends on the plot size (the dispersion variance) and an error term independent of the plot size (assumed to be the nugget variance). The two error terms are estimated using a semivariogram describing the variation between plots as a function of the distance between them. The method for calculation of the residual variance is based on a uniformity trial. Unfortunately, uniformity trials are seldomly performed. Therefore, an approach for estimating and removing treatment effects in ordinary field experiments is described. The treatment eliminated residuals obtained in this way can then be used as the base for calculating the residual variance. An example based on a uniformity trial showed a remarkable reduction of the residual variance by choosing among different possible layouts. In Part 2 different methods are described for the analysis of temporally correlated measurements in field trials. When the assumption of sphericity is satisfied the univariate analysis of variance is a valid, easy and comprehensive method to use. However, the assumption is seldomly satisfied for repeated measurements due to the temporal correlation between the measurements at the same experimental unit. Alternative methods have to be used in this case to obtain a valid analysis. A modified univariate analysis of variance with adjusted F-tests is a simple alternative to the usual univariate analysis of variance. Different multivariate analyses are given both with and without a structured variance-covariance matrix. Ante-dependence and autoregressive variance-covariance structures have been tried. The analysis with a structured variance-covariance matrix is in some sense a compromise between the univariate and multivariate analyses. The latter methods with a structured variance-covariance matrix often bring forth a very informative analysis with few restrictions and reasonable results. An analysis with a structured variance-covariance matrix using an ante-dependence structure is to be preferred to an autoregressive structure because the ante-dependence structure gives a model which can take the different variations and correlations into account at the cost of only a few extra parameters. The conclusion of the thesis is that design and analysis of agricultural experiments can be improved by utilization of the spatial and/or temporal correlation between measurements.

Automated determination of crystal orientations from electron backscattering patterns
The electron backscattering pattern (EBSP) technique is widely accepted as being an extremely powerful tool for measuring the crystallographic orientation of individual crystallites in polycrystalline materials. Procedures which allow
crystal orientations to be calculated on the bases of the position of the bands or the zone axes of EBSPs have existed for several years now. Until recently, however, the localization of either the bands or the zone axes of EBSPs has required the valuable time and attention of a human operator, thus obviously limiting the amounts of orientation data that can be collected by this method. This thesis describes the development and implementation of a system which enables crystallographic orientations to be obtained fully automatically through the use of computerized analysis and interpretation of EBSPs. More specifically, this thesis will describe the design of a pattern recognition procedure which enables 8 to 12 bands to be localized in typical EBSPs from a modern system. It will be described, how these automatically localized bands can be indexed and used for optimal estimation of the unknown crystal orientations. A necessary prerequisite for precise determination of crystallographic orientations from EBSPs is accurate knowledge of three calibration parameters which describe the position of the point from which the patterns are emitted relative to the phosphor screen on which they are recorded. This thesis will describe a novel method by which these calibration parameters can be estimated with high precision. The quality of EBSPs provides important information about the reliability of the measured crystal orientations and about the perfection of the lattice in which the pattern is generated. A measure which allows the quality of EBSPs to be evaluated quantitatively is therefore described. Presently, little is known about the uncertainty of the lattice orientations which can be measured from EBSPs. This subject will be discussed in detail in this thesis. With the application of newly developed statistical methods for analyzing orientation data it will be shown how the relative precision of lattice orientations measured from EBSPs can be described. By applying this methodology to a large number of EBSPs of varying quality it is demonstrated that the precision of automatically measured crystal orientations is comparable to the precision obtained, when the positions of four to five bands are supplied by an experienced and careful operator.

Identification of physical models
The problem of identification of physical models is considered within the frame of stochastic differential equations. Methods for estimation of parameters of these continuous time models based on discrete time measurements are discussed. The important algorithms of a computer program for ML or MAP estimation of the parameters of nonlinear stochastic differential equations are described and the implemented tool is validated with respect to bias and uncertainty of the estimated parameters. The different phases involved in identification of this type of models are considered in the thesis. This includes design of experiments, which is for instance the design of an input signal that are optimal according to a criterion based on the information provided by the experiment. Also model validation is discussed. An important verification of a physical model is to compare the physical characteristics of the model with the available prior knowledge. The methods for identification of physical models have been applied in two different case studies. One case is the identification of thermal dynamics of building components. The work is related to a CEC research project called PASSYS (Passive Solar Components and Systems Testing), on testing of building components related to passive solar energy conservation, tested under outdoor climate conditions. The second case study is related to the performance of a spark ignition car engine. A phenomenological model of the fuel flow is identified under various operating conditions of the engine. This engine submodel is important for controlling the air/fuel ratio, e.g. in a feed-forward controller.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Lassen, N. C. K. (Intern), Conradsen, K. (Intern), Juul Jensen, D. (Intern)
Publication date: Sep 1994

Publication information
Original language: English
Series: IMM-PHD-1994-3
Main Research Area: Technical/natural sciences
Electronic versions:
imm1200.pdf
Source: orbit
Source-ID: 200814
Publication: Research › Ph.D. thesis – Annual report year: 1994

Identification of physical models
The problem of identification of physical models is considered within the frame of stochastic differential equations. Methods for estimation of parameters of these continuous time models based on discrete time measurements are discussed. The important algorithms of a computer program for ML or MAP estimation of the parameters of nonlinear stochastic differential equations are described and the implemented tool is validated with respect to bias and uncertainty of the estimated parameters. The different phases involved in identification of this type of models are considered in the thesis. This includes design of experiments, which is for instance the design of an input signal that are optimal according to a criterion based on the information provided by the experiment. Also model validation is discussed. An important verification of a physical model is to compare the physical characteristics of the model with the available prior knowledge. The methods for identification of physical models have been applied in two different case studies. One case is the identification of thermal dynamics of building components. The work is related to a CEC research project called PASSYS (Passive Solar Components and Systems Testing), on testing of building components related to passive solar energy conservation, tested under outdoor climate conditions. The second case study is related to the performance of a spark ignition car engine. A phenomenological model of the fuel flow is identified under various operating conditions of the engine. This engine submodel is important for controlling the air/fuel ratio, e.g. in a feed-forward controller.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Melgaard, H. (Intern)
Publication date: 1994

Publication information
Original language: English
Series: IMM-PHD-1994-4
Main Research Area: Technical/natural sciences
Electronic versions:
Design and optimization of flexible manufacturing systems

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Rygaard, J. M. (Intern)
Number of pages: 232
Publication date: 1994

Publication information
Original language: English
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200854
Publication: Research › Ph.D. thesis – Annual report year: 1994

Stochastic differential equations and a biological system
The purpose of this Ph.D. study is to explore the property of a growth process. The study includes solving and simulating of the growth process which is described in terms of stochastic differential equations. The identification of the growth and variability parameters of the process based on experimental data is considered. As an example, the growth of bacteria Pseudomonas fluorescens is taken. Due to the specific features of stochastic differential equations, namely that their solutions do not exist in the general sense, two new integrals - the Ito integral and the Stratonovich integral - have been developed. Their properties and the relationship between them are discussed. The evolution of a dynamic system or process is usually of great practical interest. In order to simulate the evolution of the process, alternative methods are used to get numerical solutions. In this study, Euler, Milstein and Runge-Kutta methods are used. Because of the specific feature of the model for the growth process, that its solution does not exist in the general sense, we combine these numerical integration methods with a transformation technique, and the solutions are derived in the Ito sense. The simulated results are compared with the experimental data, and it is found that the Euler method is the most simple and efficient method for the stochastic growth model considered. Estimation of the parameters of the growth model is based on the stochastic Kalman filter and a continuous Markov process description. In order to identify the parameters, a Maximum likelihood estimation method is used together with a simplified truncated second order filter. Because of the continuity feature of the predictor equation, two numerical integration methods, called the Odeint and the Discretization method, are developed. By the simulation and identification of the parameters in the theoretical model, a satisfactory match is found.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Wang, C. (Intern)
Number of pages: 153
Publication date: 1994

Publication information
Original language: English
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 200867
Publication: Research › Ph.D. thesis – Annual report year: 1994

On the analysis of image data using simultaneous interaction models

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling
Authors: Grunkin, M. (Intern)
Application of computer-intensive data analysis methods to the analysis of digital images and spatial data

Computer-intensive methods for data analysis in a traditional setting have developed rapidly in the last decade. The application of and adaption of some of these methods to the analysis of multivariate digital images and spatial data are explored, evaluated, and compared to well-established classical linear methods. Different strategies for selecting projections (linear combinations) of multivariate images are presented. An exploratory, iterative method for finding interesting projections originated in data analysis is compared to principal components. A method for introducing spatial context into the projection pursuit is presented. Examples from remote sensing are given. The ACE algorithm for computing non-linear transformations for maximizing correlation is extended and applied to obtain a non-linear transformation that maximizes autocorrelation or 'signal' in a multivariate image. This is a generalization of the minimum/maximum autocorrelation factors (MAF's) which is a linear method. The non-linear method is compared to the linear method when analyzing a multivariate TM image from Greenland. The ACE method is shown to give a more detailed decomposition of the image than the MAF-transformation and there is a good agreement between the ACE-MAF's and geological structures known in the area studied. Geological units are easily recognized even at macro scale, implying potential use in geological mapping. Also the ACE algorithm is modified to finding transformations that minimize correlation which is of interest in change detection studies from two different images of the same area recorded at different time points. An example is given using a TM summer scene and a TM winter scene of an area in Spain. The non-parametric CART classification method is integrated with traditional geostatistical methods in computing structural images for heavy minerals based on irregularly sampled geochemical data. This methodology has proven useful in producing images that reflect real geological structures with potential application in mineral exploration. A method for removing laboratory-produced map-sheet patterns in spatial data by means of local histogram matching is presented and its use is demonstrated in the analysis of geochemical samples on a regional scale.

Transformations and classifications of remotely sensed data. Theory and geological cases

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Ersbøll, B. K. (Intern)
Number of pages: 297
Publication date: 1989