Sparse supervised principal component analysis (SSPCA) for dimension reduction and variable selection

Principal component analysis (PCA) is one of the main unsupervised pre-processing methods for dimension reduction. When the training labels are available, it is worth using a supervised PCA strategy. In cases that both dimension reduction and variable selection are required, sparse PCA (SPCA) methods are preferred. In this paper, a sparse supervised PCA (SSPCA) method is proposed for pre-processing. This method is appropriate especially in problems where, a high dimensional input necessitates the use of a sparse method and a target label is also available to guide the variable selection strategy. Such a method is valuable in many Engineering and scientific problems, when the number of training samples is also limited. The Hilbert Schmidt Independence Criteria (HSIC) is used to form an objective based on minimization of a loss function and an L1 norm is used for regularization of the Eigen vectors. While the proposed objective function allows a sparse low rank solution for both linear and non-linear relationships between the input and response matrices, other similar methods in this case are only based on a linear model. The objective is solved based on penalized matrix decomposition (PMD) algorithm. We compare the proposed method with PCA, PMD-based SPCA and supervised PCA. In addition, SSPCA is also compared with sparse partial least squares (SPLS), due to the similarity between the two objective functions. Experimental results from the simulated as well as real data sets show that, SSPCA provides an appropriate trade-off between accuracy and sparsity. Comparisons show that, in terms of sparsity, SSPCA performs the highest level of variable reduction and also, in terms of accuracy it is one of the most successful methods. Therefore, the Eigen vectors found by SSPCA can be used for feature selection in various high dimensional problems. © 2017 Elsevier Ltd. All rights reserved.
Spatial models for probabilistic prediction of wind power with application to annual-average and high temporal resolution data

Producing accurate spatial predictions for wind power generation together with a quantification of uncertainties is required to plan and design optimal networks of wind farms. Toward this aim, 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 demonstrate the proposed methods on wind farm data from Western Denmark, and compare the results to those obtained with standard geostatistical methods. The results show that our method makes it possible to obtain fast and accurate predictions from posterior marginals for wind power generation. The proposed method is applicable in scientific areas as diverse as climatology, environmental sciences, earth sciences and epidemiology.

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), Pinson, P. (Intern), Clemmensen, L. K. H. (Intern), Guillot, G. (Intern)
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.
rates is explored. A response surface design is performed to optimize the algorithm using a data set from Fall 2010. Finally, the results are tested against a data set from Fall 2011. It is shown that all elements of the original algorithm (the base rule and the three additional conditions) play a role in the algorithm's performance and should be included in the algorithm. Because there is significant interaction between the base rule and the additional conditions, many acceptable combinations that balance the FPR and FNR can be found, but no true optimum seems to exist. The performance of the best optimizations and the original algorithm are similar. Therefore, it should be possible to choose new coefficient values for jury populations in other cultures and contexts logically and empirically without a full optimization as long as the algorithm assumptions are valid and the limitations for its use are well understood.
Does correlated color temperature affect the ability of humans to identify veins?

In the present study we provide empirical evidence and demonstrate statistically that white illumination settings can affect the human ability to identify veins in the inner hand vasculature. A special light-emitting diode lamp with high color rendering index (CRI 84–95) was developed and the effect of correlated color temperature was evaluated, in the range between 2600 and 5700 K at an illuminance of 40 9 lx on the ability of adult humans to identify veins. It is shown that the ability to identify veins can, on average, be increased up to 24% when white illumination settings that do not resemble incandescent light are applied. The illuminance reported together with the effect of white illumination settings on direct visual perception of biosamples are relevant for clinical investigations during the night. © 2015 Optical Society of America
From university research to innovation Detecting knowledge transfer via text mining

Knowledge transfer by universities is a top priority in innovation policy and a primary purpose for public research funding, due to being an important driver of technical change and innovation. Current empirical research on the impact of university research relies mainly on formal databases and indicators such as patents, collaborative publications and license agreements, to assess the contribution to the socioeconomic surrounding of universities. In this study, we present an extension of the current empirical framework by applying new computational methods, namely text mining and pattern recognition. Text samples for this purpose can include files containing social media contents, company websites and annual reports. The empirical focus in the present study is on the technical sciences and in particular on the case of the Technical University of Denmark (DTU). We generated two independent text collections (corpora) to identify correlations of university publications and company webpages. One corpus representing the company sites, serving as sample of the private economy and a second corpus, providing the reference to the university research, containing relevant publications. We associated the former with the latter to obtain insights into possible text and semantic relatedness. The text mining methods are extrapolating the correlations, semantic patterns and content comparison of the two corpora to define the document relatedness. We expect the development of a novel tool using contemporary techniques for the measurement of public research impact. The approach aims to be applicable across universities and thus enable a more holistic comparable assessment. This rely less on formal databases, which is certainly beneficial in terms of the data reliability. We seek to provide a supplementary perspective for the detection of the dissemination of university research and hereby enable policy makers to gain additional insights of (informal) contributions of knowledge dissemination by universities.

Identifying Drug–Drug Interactions by Data Mining: A Pilot Study of Warfarin-Associated Drug Interactions

Background—Knowledge about drug–drug interactions commonly arises from preclinical trials, from adverse drug reports, or based on knowledge of mechanisms of action. Our aim was to investigate whether drug–drug interactions were discoverable without prior hypotheses using data mining. We focused on warfarin–drug interactions as the prototype.

Methods and Results—We analyzed altered prothrombin time (measured as international normalized ratio [INR]) after initiation of a novel prescription in previously INR-stable warfarin-treated patients with nonvalvular atrial fibrillation. Data
sets were retrieved from clinical work. Random forest (a machine-learning method) was set up to predict altered INR levels after novel prescriptions. The most important drug groups from the analysis were further investigated using logistic regression in a new data set. Two hundred and twenty drug groups were analyzed in 61,190 novel prescriptions. We rediscovered 2 drug groups having known interactions (β-lactamase-resistant penicillins [dicloxacillin] and carboxamide derivatives) and 3 antithrombotic/anticoagulant agents (platelet aggregation inhibitors excluding heparin, direct thrombin inhibitors [dabigatran etexilate], and heparins) causing decreasing INR. Six drug groups with known interactions were rediscovered causing increasing INR (antiarrhythmics class III [amiodarone], other opioids [tramadol], glucocorticoids, triazole derivatives, and combinations of penicillins, including β-lactamase inhibitors) and two had a known interaction in a closely related drug group (opipavine derivatives [buprenorphine] and natural opium alkaloids). Antipropulsives had an unknown signal of increasing INR.

Conclusions—We were able to identify known warfarin–drug interactions without a prior hypothesis using clinical registries. Additionally, we discovered a few potentially novel interactions. This opens up for the use of data mining to discover unknown drug–drug interactions in cardiovascular medicine.
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 environments 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, Statistics and Data Analysis
Authors: Gronskyte, R. (Intern), Kulahci, M. (Intern), Clemmensen, L. K. H. (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: 370
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd370_Gronskyte_R.pdf

Relations
Projects:
Monitoring Animal Well-being
Publication: Research › Ph.D. thesis – Annual report year: 2016

Monitoring pig movement at the slaughterhouse using optical flow and modified angular histograms
We analyse the movement of pig herds through video recordings at a slaughterhouse by using statistical analysis of optical flow (OF) patterns. Unlike the previous attempts to analyse pig movement, no markers, trackers nor identification of individual pigs are needed. Our method handles the analysis of unconstrained areas where pigs are constantly entering and leaving. The goal is to improve animal welfare by real-time prediction of abnormal behaviour through proper interventions. The aim of this study is to identify any stationary pig, which can be an indicator of an injury or an obstacle. In this study, we use the OF vectors to describe points of movement on all pigs and thereby analyse the herd movement. Subsequently, the OF vectors are used to identify abnormal movements of individual pigs. The OF vectors, obtained from the pigs, point in multiple directions rather than in one movement direction. To accommodate the multiple directions of the OF vectors, we propose to quantify OF using a summation of the vectors into bins according to their angles, which we call modified angular histograms. Sequential feature selection is used to select angle ranges, which identify pigs that are moving abnormally in the herd. The vector lengths from the selected angle ranges are compared to the corresponding median, 25th and 75th percentiles from a training set, which contains only normally moving pigs. We show that the method is capable of locating stationary pigs in the recordings regardless of the number of pigs in the frame.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Danish Meat Research Institute
Authors: Gronskyte, R. (Intern), Clemmensen, L. K. H. (Intern), Hviid, M. S. (Ekstern), Kulahci, M. (Intern)
Pages: 19-30
Publication date: 2016
Main Research Area: Technical/natural sciences

Publication information
Novelty detection of foreign objects in food using multi-modal X-ray imaging

In this paper we demonstrate a method for novelty detection of foreign objects in food products using grating-based multimodal X-ray imaging. With this imaging technique three modalities are available with pixel correspondence, enhancing organic materials such as wood chips, insects and soft plastics not detectable by conventional X-ray absorption radiography. We conduct experiments, where several food products are imaged with common foreign objects typically found in the food processing industry. To evaluate the benefit from using this multi-contrast X-ray technique over conventional X-ray absorption imaging, a novelty detection scheme based on well known image- and statistical analysis
techniques is proposed. The results show that the presented method gives superior recognition results and highlights the advantage of grating-based imaging.

**General information**

State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, Technische Universität München
Authors: Einarsdottir, H. (Intern), Emerson, M. J. (Intern), Clemmensen, L. K. H. (Intern), Scherer, K. (Ekstern), Willer, K. (Ekstern), Bech, M. (Ekstern), Larsen, R. (Intern), Ersbøll, B. K. (Intern), Pfeiffer, F. (Ekstern)
Pages: 39-47
Publication date: 2016
Main Research Area: Technical/natural sciences

**Publication information**

Journal: Food Control
Volume: 67
Issue number: September
ISSN (Print): 0956-7135
Ratings:

BFI (2018): BFI-level 1
BFI (2017): BFI-level 1
Web of Science (2017): Indexed yes

BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 3.86 SJR 1.462 SNIP 1.719
Web of Science (2016): Indexed yes

BFI (2015): BFI-level 1
Scopus rating (2015): SJR 1.509 SNIP 1.72 CiteScore 3.65
Web of Science (2015): Indexed yes

BFI (2014): BFI-level 1
Scopus rating (2014): SJR 1.389 SNIP 1.718 CiteScore 3.27
Web of Science (2014): Indexed yes

BFI (2013): BFI-level 1
Scopus rating (2013): SJR 1.273 SNIP 1.745 CiteScore 3.14

ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes

BFI (2012): BFI-level 1
Scopus rating (2012): SJR 1.264 SNIP 1.916 CiteScore 3.1

ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes

BFI (2011): BFI-level 1
Scopus rating (2011): SJR 1.2 SNIP 1.726 CiteScore 2.9

ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes

BFI (2010): BFI-level 1
Scopus rating (2010): SJR 1.214 SNIP 1.683
Web of Science (2010): Indexed yes

BFI (2009): BFI-level 1
Scopus rating (2009): SJR 1.193 SNIP 1.685
Web of Science (2009): Indexed yes

BFI (2008): BFI-level 1
Scopus rating (2008): SJR 1.136 SNIP 1.394
Web of Science (2008): Indexed yes

Scopus rating (2007): SJR 0.89 SNIP 1.549
Web of Science (2007): Indexed yes

Scopus rating (2006): SJR 0.776 SNIP 1.323
Scopus rating (2005): SJR 0.584 SNIP 1.244
Scopus rating (2004): SJR 0.566 SNIP 1.074
Characterizing Digital Light Processing (DLP) 3D Printed Primitives

The resolution and repeatability of 3D printing processes depends on a number of factors including the software, hardware, and material used. When printing parts with features that are near or below the nominal printing resolution, it is important to understand how the printer works. For example, what is the smallest unit shape that can be produced? And what is the reproducibility of that process? This paper presents a method for automatically detecting and characterizing the height, width, and length of micro scale geometric primitives produced via a digital light processing (DLP) 3D printing process. An upper limit, lower limit, and best estimate for each dimension is reported for each primitive. Additionally, the roughness, rectangularity, and tilt of the top of each primitive is estimated. The uncertainty of the best estimate is indicated using standard deviations for a series of primitives. The method generalizes to unseen primitives, and the results illustrate that the dimension estimates converge as the size of the primitives increases. The primitives’ rectangularity also increases as the size increases. Finally, the primitives specified with 5 to 68μm varying heights have been estimated to group into five different heights with fairly low variance of the best estimates of the heights. This reflects how the requested geometry is parsed and produced by the printer.

Development of LED Light Sources for Improved Visualization of Veins: a statistical approach

The present statistical study investigates the difference of diffuse reflectances between skin and vein (defined as contrast indicator) under different visible wavelengths of a population of 39 adult participants. The purpose of the study is to examine if there is a group of wavelengths-color combinations that could explain most of the variance (of the contrast indicator) in the data set. Moreover the effect of gender and age on the contrast indicator is explored.
Generalized requirements and decompositions for the design of test parts for micro additive manufacturing research

The design of experimental test parts to characterize micro additive manufacturing (AM) processes is challenging due to the influence of the manufacturing and metrology processes. This work builds on the lessons learned from a case study in the literature to derive generalized requirements and high level decompositions for the design of test parts and the design of experiments to characterize micro additive manufacturing processes. While the test parts and the experiments described are still work in progress, the generic requirements derived from them can serve as a starting point for the design of other micro additive manufacturing related studies and their decompositions can help structure future work.

General information
State: Published
Organizations: Department of Applied Mathematics and Computer Science, Department of Mechanical Engineering, Manufacturing Engineering, Statistics and Data Analysis
Authors: Thompson, M. K. (Intern), Clemmensen, L. K. H. (Intern)
Pages: 229-235
Publication date: 2015
Conference: 9th International Conference on Axiomatic Design (ICAD 2015), Florence, Italy, 16/09/2015 - 16/09/2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Procedia CIRP
Volume: 34
ISSN (Print): 2212-8271
Ratings:
Scopus rating (2016): CiteScore 1.6 SNIP 1.297
Scopus rating (2015): SJR 0.572 SNIP 1.012
Scopus rating (2014): SJR 0.736 SNIP 1.419
Scopus rating (2013): SJR 0.515 SNIP 1.163
ISI indexed (2013): ISI indexed no
Original language: English
Additive manufacturing, micro manufacturing, metrology, design of experiments, requirements, decomposition

Electronic versions:
Thompson_and_Clemmensen_2015.pdf
DOIs:
10.1016/j.procir.2015.07.075

Bibliographical note
This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Source: PublicationPreSubmission
Source-ID: 115098587
Publication: Research - peer-review › Journal article – Annual report year: 2015

In silico modelling of permeation enhancement potency in Caco-2 monolayers based on molecular descriptors and random forest

General information
State: Published
Organizations: Department of Photonics Engineering, Diode Lasers and LED Systems, Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Authors: Argyraki, A. (Intern), Clemmensen, L. K. H. (Intern), Petersen, P. M. (Intern)
Number of pages: 1
Publication date: 2015
Main Research Area: Technical/natural sciences
Principal component analysis (PCA), LED, Vein visualization
Electronic versions:
Abstract_for_3rd_Asian_workshop.pdf
Structural traits of permeation enhancers are important determinants of their capacity to promote enhanced drug absorption. Therefore, in order to obtain a better understanding of structure–activity relationships for permeation enhancers, a Quantitative Structural Activity Relationship (QSAR) model has been developed. The random forest-QSAR model was based upon Caco-2 data for 41 surfactant-like permeation enhancers from Whitehead et al. (2008) and molecular descriptors calculated from their structure. The QSAR model was validated by two test-sets: (i) an eleven compound experimental set with Caco-2 data and (ii) nine compounds with Caco-2 data from literature. Feature contributions, a recent developed diagnostic tool, was applied to elucidate the contribution of individual molecular descriptors to the predicted potency. Feature contributions provided easy interpretable suggestions of important structural properties for potent permeation enhancers such as segregation of hydrophilic and lipophilic domains. Focusing on surfactant-like properties, it is possible to model the potency of the complex pharmaceutical excipients, permeation enhancers. For the first time, a QSAR model has been developed for permeation enhancement. The model is a valuable in silico approach for both screening of new permeation enhancers and physicochemical optimisation of surfactant enhancer systems.
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.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Authors: Sharifzadeh, S. (Intern), Clemmensen, L. K. H. (Intern), Ersbøll, B. K. (Intern)
Number of pages: 244
Publication date: 2015

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

Series: DTU Compute PHD-2015
Volume: 371
ISSN: 0909-3192
Main Research Area: Technical/natural sciences
Electronic versions:
phd371_Shariifzadeh_S.pdf
Publication: Research › Ph.D. thesis – Annual report year: 2015
Non-negative Matrix Factorization for Binary Data

We propose the Logistic Non-negative Matrix Factorization for decomposition of binary data. Binary data are frequently generated in e.g. text analysis, sensory data, market basket data etc. A common method for analysing non-negative data is the Non-negative Matrix Factorization, though this is in theory not appropriate for binary data, and thus we propose a novel Non-negative Matrix Factorization based on the logistic link function. Furthermore we generalize the method to handle missing data. The formulation of the method is compared to a previously proposed method (Tome et al., 2015). We compare the performance of the Logistic Non-negative Matrix Factorization to Least Squares Non-negative Matrix Factorization and Kullback-Leibler (KL) Non-negative Matrix Factorization on sets of binary data: a synthetic dataset, a set of student comments on their professors collected in a binary term-document matrix and a sensory dataset. We find that choosing the number of components is an essential part in the modelling and interpretation, that is still unresolved.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Authors: Larsen, J. S. (Intern), Clemmensen, L. K. H. (Intern)
Number of pages: 9
Publication date: 2015
Event: Paper presented at 7th International Conference on Knowledge Discovery and Information Retrieval (KDIR 2015), Lisbon, Portugal.
Main Research Area: Technical/natural sciences
Non-negative Matrix Factorization, Binary Data, Binary Matrix Factorization, Text Modelling
Source: PublicationPreSubmission
Source-ID: 118545190
Publication: Research - peer-review → Paper – Annual report year: 2015

Pig herd monitoring and undesirable tripping and stepping prevention

Humane handling and slaughter of livestock are of major concern in modern societies. Monitoring animal wellbeing in slaughterhouses is critical in preventing unnecessary stress and physical damage to livestock, which can also affect the meat quality. The goal of this study is to monitor pig herds at the slaughterhouse and identify undesirable events such as pigs tripping or stepping on each other. In this paper, we monitor pig behavior in color videos recorded during unloading from transportation trucks. We monitor the movement of a pig herd where the pigs enter and leave a surveyed area. The method is based on optical flow, which is not well explored for monitoring all types of animals, but is the method of choice for human crowd monitoring. We recommend using modified angular histograms to summarize the optical flow vectors. We show that the classification rate based on support vector machines is 93% of all frames. The sensitivity of the model is 93.5% with 90% specificity and 6.5% false alarm rate. The radial lens distortion and camera position required for convenient surveillance make the recordings highly distorted. Therefore, we also propose a new approach to correct lens and foreshortening distortions by using moving reference points. The method can be applied real-time during the actual unloading operations of pigs. In addition, we present a method for identification of the causes leading to undesirable events, which currently only runs off-line. The comparative analysis of three drivers, which performed the unloading of the pigs from the trucks in the available datasets, indicates that the drivers perform significantly differently. Driver 1 has 2.95 times higher odds to have pigs tripping and stepping on each other than the two others, and Driver 2 has 1.11 times higher odds than Driver 3. (C) 2015 Elsevier B.V. All rights reserved.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Danish Meat Research Institute
Authors: Gronskyte, R. (Intern), Clemmensen, L. K. H. (Intern), Hviid, M. S. (Ekstern), Kulahci, M. (Intern)
Pages: 51-60
Publication date: 2015
Main Research Area: Technical/natural sciences

Publication information
Journal: Computers and Electronics in Agriculture
Volume: 119
Issue number: November
ISSN (Print): 0168-1699
Ratings:
BFI (2018): BFI-level 1
BFI (2017): BFI-level 1
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 1
Regularized generalized eigen-decomposition with applications to sparse supervised feature extraction and sparse discriminant analysis

We propose a general technique for obtaining sparse solutions to generalized eigenvalue problems, and call it Regularized Generalized Eigen-Decomposition (RGED). For decades, Fisher's discriminant criterion has been applied in supervised feature extraction and discriminant analysis, and it is formulated as a generalized eigenvalue problem. Thus RGED can be applied to effectively extract sparse features and calculate sparse discriminant directions for all variants of Fisher discriminant criterion based models. Particularly, RGED can be applied to matrix-based and even tensor-based discriminant techniques, for instance, 2D-Linear Discriminant Analysis (2D-LDA). Furthermore, an iterative algorithm based on the alternating direction method of multipliers is developed. The algorithm approximately solves RGED with monotonically decreasing convergence and at an acceptable speed for results of modest accuracy. Numerical experiments based on four data sets of different types of images show that RGED has competitive classification performance with existing multidimensional and sparse techniques of discriminant analysis.

General information
Student employment and study effort for engineering students

The aim of this paper is to examine which factors effect student employment and study effort in a setting where engineering students are financially supported, such that their education is free of cost and that they receive financial support for living costs while studying. In addition, we wish to answer if the full-time student is under demise in these settings as opposed to settings without financial support [1, 2]. The research consisted of a web-based survey amongst all students at the Technical University of Denmark (DTU). The students in this survey had fewer employment hours and studied more than those in studies from e.g. UK and US [3, 4, 5]. A similar trend was seen in a study from Norway [6]. Government financial support seems to limit the amount of hours spent on paid work but not the percentage of students who take on paid work. Thus, full-time studies with benefits of increased capabilities and experience gained through employment could be aided by proper policies. Additionally, one of the highest impacts on study activity was the perceived study environment. As the engineering students have four hours per week of interaction with an instructor for each five ECTS, it is to be expected that the students generally spend a majority of their studying hours at the university. This study is to our knowledge the first to study student employment and study effort for Danish engineering students.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis, Copenhagen Business School
Authors: Clemmensen, L. K. H. (Intern), Harder, D. E. (Intern)
Number of pages: 8
Publication date: 2015

Host publication information
Title of host publication: Proceedings of the 43rd Annual SEFI Conference
Article number: 56598
BFI conference series: European Society for Engineering Education. Annual Conference proceedings (5010018)
Main Research Area: Technical/natural sciences
Conference: 43rd Annual Conference of the European Society for Engineering Education (SEFI 2015), Orléans, France, 29/06/2015 - 29/06/2015
Student employment, Study effort, Perceived study environment
Electronic versions:
SEFI2015_Paper_ClemmensenHarder_v4_2_1.pdf
Links:
http://www.sefi.be/?page_id=6043

Bibliographical note
CHAPTER 1: Eco Design sustainability in Engineering Education
Source: PublicationPreSubmission
Source-ID: 118544734
Publication: Research - peer-review › Article in proceedings – Annual report year: 2015

Can We Find Organic Materials in Food Using X-rays?

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis
Authors: Emerson, M. J. (Intern), Einarsdottir, H. (Intern), Clemmensen, L. K. H. (Intern), Erbsell, B. K. (Intern)
Number of pages: 1
Publication date: 2014
Main Research Area: Technical/natural sciences
Electronic versions:
PosterFAIM.pdf
Source: PublicationPreSubmission
Source-ID: 127748272
Publication: Research - peer-review › Poster – Annual report year: 2014

How do student evaluations of courses and of instructors relate?
Course evaluations are widely used by educational institutions to assess the quality of teaching. At the course evaluations, students are usually asked to rate different aspects of the course and of the teaching. We propose to apply canonical
correlation analysis (CCA) in order to investigate the degree of association between how students evaluate the course and how students evaluate the teacher. Additionally it is possible to reveal the structure of this association. Student evaluations data is characterized by high correlations between the variables within each set of variables, therefore two modifications of the CCA method; regularized CCA and sparse CCA, together with classical CCA were applied to find the most interpretable model. Both methods give results with increased interpretability over traditional CCA on the present student evaluation data. The method shows robustness when evaluations over several years are examined.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Authors: Sliusarenko, T. (Intern), Clemmensen, L. K. H. (Intern), Erbsøll, B. K. (Intern)
Pages: 280-287
Publication date: 2014

Host publication information
Title of host publication: Proceedings of the 6th International Conference on Computer Supported Education (CSEDU 2014)
Volume: 2
Publisher: SciTePress
Editors: Zvacek, S., Restivo, M. T., Uhomoibhi, J., Helfert, M.
BFI conference series: International Conference on Computer-Supported Education (5020033)
Main Research Area: Technical/natural sciences
Conference: 6th International Conference on Computer Supported Education (CSEDU 2014), Barcelona, Spain, 01/03/2014 - 01/03/2014
Course Evaluation, Teacher Evaluation, Canonical Correlation Analysis
Publication information: Research - peer-review > Article in proceedings – Annual report year: 2014

Hyperspectral imaging based on diffused laser light for prediction of astaxanthin coating concentration
We present a study on predicting the concentration level of synthetic astaxanthin in fish feed pellet coating using multi- and hyperspectral image analysis. This was done in parallel using two different vision systems. A new instrument for hyperspectral imaging, the SuperK setup, using a super-continuum laser as the light source was introduced. Furthermore, a parallel study with the commercially available multispectral VideometerLab imaging system was performed. The SuperK setup used 113 spectral bands (455–1,015 nm), and the VideometerLab used 20 spectral bands (385–1,050 nm). To predict the astaxanthin concentration from the spectral image data, the synthetic astaxanthin content in the pellets was measured with the established standard technique; high-pressure liquid chromatography (HPLC). Regression analysis was done using partial least squares regression (PLSR) and the sparse regression method elastic net (EN). The ratio of standard error of prediction (RPD) is the ratio between the standard deviation of the reference values and the prediction error, and for both PLSR and EN both devices gave RPD values between 4 and 24, and with mean prediction error of 1.4–8.0 parts per million of astaxanthin concentration. The results show that it is possible to predict the synthetic astaxanthin concentration in the coating well enough for quality control using both multi- and hyperspectral image analysis, while the SuperK setup performs with higher accuracy than the VideometerLab device for this particular problem. The spectral resolution made it possible to identify the most significant spectral regions for detection of astaxanthin. The results also imply that the presented methods can be used in general for quality inspection of various coating substances using similar coating methods.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, National Food Institute, Division of Industrial Food Research, Division of Toxicology and Risk Assessment, Statistics and Data Analysis
Authors: Ljungqvist, M. G. (Intern), Nielsen, O. H. A. (Intern), Frosch, S. (Intern), Nielsen, M. E. (Intern), Clemmensen, L. K. H. (Intern), Erbsøll, B. K. (Intern)
Number of pages: 17
Pages: 327-343
Publication date: 2014
Main Research Area: Technical/natural sciences

Publication information
Journal: Machine Vision & Applications
Volume: 25
Issue number: 2
ISSN (Print): 0932-8092
Ratings:
BFI (2018): BFI-level 1
On Weighted Support Vector Regression

We propose a new type of weighted support vector regression (SVR), motivated by modeling local dependencies in time and space in prediction of house prices. The classic weights of the weighted SVR are added to the slack variables in the objective function (OF-weights). This procedure directly shrinks the coefficient of each observation in the estimated functions; thus, it is widely used for minimizing influence of outliers. We propose to additionally add weights to the slack variables in the constraints (CF-weights) and call the combination of weights the doubly weighted SVR. We illustrate the differences and similarities of the two types of weights by demonstrating the connection between the Least Absolute Shrinkage and Selection Operator (LASSO) and the SVR. We show that an SVR problem can be transformed to a LASSO problem plus a linear constraint and a box constraint. We demonstrate the capabilities of the doubly weighted approach through an example of prediction of house prices. The weight functions in the house pricing model depend on the geographical distance to the house of interest and the difference in time of sale (CF-weights) as well as the differences lying in variables (OF-weights), such as house size and number of floors. The results illustrate that the combination of the two types of weights describes the relative importance of observations very well and lowers the influence of possible outliers. Therefore, it enables the SVR models to have good performance. Copyright © 2014 John Wiley & Sons, Ltd.
Selecting groups of covariates in the elastic net

This paper introduces a novel method to select groups of variables in sparse regression and classification settings. The groups are formed based on the correlations between covariates and ensure that for example spatial or spectral relations are preserved without explicitly coding for these. The preservation of relations gives increased interpretability. The method is based on the elastic net and adaptively selects highly correlated groups of variables and does therefore not waste time in grouping irrelevant variables for the problem at hand. The method is illustrated on a simulated data set and on regression of moisture content in multispectral images of sand. In both cases, the predictions were better or similar to existing regression and classification algorithms and the interpretation was enhanced using the grouping method. On top of that, the grouping method more consistently selects the important variables.

Sparse and shrunken estimates of MRI networks in the brain and their influence on network properties

Estimation of morphometric relationships between cortical regions is a widely used approach to identify and characterize structural connectivity. The elevated number of regions that can be considered in a whole-brain correlation analysis might lead to overfitted models. However, the overfitting can be avoided by using regularization methods. We found that, as expected, non-regularized correlations had low variability when a scarce number of variables were considered. However, a slight increase of variables led to an increase of variance of several magnitude orders. On the other hand, the regularized approaches showed more stable results with a relative low variance at the expense of a little bias. Interestingly, topological properties as local and global efficiency estimated in networks constructed from traditional non-regularized correlations also showed higher variability when compared to those from regularized networks. Our findings suggest that a population-based connectivity study can achieve a more robust description of cortical topology through regularization of the correlation estimates. Four regularization methods were examined: Two with shrinkage (Ridge and Schäfer’s shrinkage), one with sparsity (Lasso) and one with both shrinkage and sparsity (Elastic net). Furthermore, the different regularizations resulted in different correlation estimates as well as network properties. The shrunken estimates resulted in lower variance of the estimates than the sparse estimates.
Cortical network, Network properties, MRI, Partial correlation coefficients, Regularization, Shrinkage estimators, Sparse estimators, Structural connectivity

DOIs:
10.1117/12.2043009

Publication: Research - peer-review » Conference article – Annual report year: 2014

Study Delay

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Authors: Sliusarenko, T. (Intern), Gronskyte, R. (Intern), Clemmensen, L. K. H. (Intern), Kulahci, M. (Intern), Ersbøll, B. K. (Intern)
Number of pages: 65
Publication date: 2014
Supervised feature selection for linear and non-linear regression of Lab color from multispectral images of meat

In food quality monitoring, color is an important indicator factor of quality. The CIELab (Lab) color space as a device independent color space is an appropriate means in this case. The commonly used colorimeter instruments can neither measure the Lab color in a wide area over the target surface nor in a contact-less mode. However, developing algorithms for conversion of food items images into Lab color space can solve both of these issues. This paper addresses the problem of Lab color prediction from multispectral images of different types of raw meat. The efficiency of using multispectral images instead of the standard RGB is investigated. In addition, it is demonstrated that due to the fiber structure and transparency of raw meat, the prediction models built on the standard color patches do not work for raw meat test samples. As a result, multispectral images of different types of meat samples (430–970 nm) were used for training and testing of the Lab prediction models. Finding a sparse solution or the use of a minimum number of bands is of particular interest to make an industrial vision set-up simpler and cost effective. In this paper, a wide range of linear, non-linear, kernel-based regression and sparse regression methods are compared. In order to improve the prediction results of these models, we propose a supervised feature selection strategy which is compared with the Principal component analysis (PCA) as a pre-processing step. The results showed that the proposed feature selection method outperforms the PCA for both linear and non-linear methods. The highest performance was obtained by linear ridge regression applied on the selected features from the proposed Elastic net (EN) -based feature selection strategy. All the best models use a reduced number of wavelengths for each of the Lab components.
A sampling approach for predicting the eating quality of apples using visible–near infrared spectroscopy

BACKGROUND
Visible–near infrared spectroscopy remains a method of increasing interest as a fast alternative for the evaluation of fruit quality. The success of the method is assumed to be achieved by using large sets of samples to produce robust calibration models. In this study we used representative samples of an early and a late season apple cultivar to evaluate model robustness (in terms of prediction ability and error) on the soluble solids content (SSC) and acidity prediction, in the wavelength range 400–1100 nm.

RESULTS
A total of 196 middle–early season and 219 late season apples (Malus domestica Borkh.) cvs ‘Aroma’ and ‘Holsteiner Cox’ samples were used to construct spectral models for SSC and acidity. Partial least squares (PLS), ridge regression (RR) and elastic net (EN) models were used to build prediction models. Furthermore, we compared three sub-sample arrangements for forming training and test sets (‘smooth fractionator’, by date of measurement after harvest and random). Using the ‘smooth fractionator’ sampling method, fewer spectral bands (26) and elastic net resulted in improved performance for SSC models of ‘Aroma’ apples, with a coefficient of variation CVSSC = 13%. The model showed consistently low errors and bias (PLS/EN: R2cal=0.60/0.60; SEC = 0.88/0.88°Brix; Biascal=0.00/0.00; R2val=0.33/0.44; SEP = 1.14/1.03; Biasval=0.04/0.03). However, the prediction acidity and for SSC (CV = 5%) of the late cultivar ‘Holsteiner Cox’ produced inferior results as compared with ‘Aroma’.

CONCLUSION
It was possible to construct local SSC and acidity calibration models for early season apple cultivars with CVs of SSC and acidity around 10%. The overall model performance of these data sets also depend on the proper selection of training and test sets. The ‘smooth fractionator’ protocol provided an objective method for obtaining training and test sets that capture the existing variability of the fruit samples for construction of visible–NIR prediction models. The implication is that by using such ‘efficient’ sampling methods for obtaining an initial sample of fruit that represents the variability of the population and for sub-sampling to form training and test sets it should be possible to use relatively small sample sizes to develop spectral predictions of fruit quality. Using feature selection and elastic net appears to improve the SSC model performance in terms of R2, RMSECV and RMSEP for ‘Aroma’ apples. © 2013 Society of Chemical Industry

General information
State: Published
Correlation of iris biometrics and DNA

The presented work concerns prediction of complex human phenotypes from genotypes. We were interested in correlating iris color and texture with DNA. Our data consist of 212 eye images along with DNA: 32 single-nucleotide polymorphisms (SNPs). We used two types of biometrics to describe the eye images: one for iris color and one for iris texture. Both biometrics were high dimensional and a sparse principle component analysis (SPCA) reduced the dimensions and resulted in a representation of data with good interpretability. The correlations between the sparse principal components (SPCs) and the 32 SNPs were found using a canonical correlation analysis (CCA). The result was a single significant canonical correlation (CC) for both biometrics. Each CC comprised two correlated canonical variables, consisting of a linear combination of SPCs and a linear combination of SNPs, respectively. The significant canonical variables for color and texture were primarily explained by the first SPC (SPC1). Therefore, we made a visual inspection of the first SPCs. The color based SPC1 explained a blue to brown variation in iris color and the texture based SPC1 gave a general explanation of iris texture. The SNPs (rs12896399, rs3733542, rs6475555, rs12913832) and (rs12896399, rs3733542, rs12913832) had the highest correlation to the canonical variable for color and texture, respectively. Three of the most contributing SNPs were the same for both biometrics, revealing a covariance between iris color and texture.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Image Analysis & Computer Graphics, Statistics and Data Analysis, Technical University of Denmark, University of Copenhagen
Authors: Harder, S. (Intern), Clemmensen, L. K. H. (Intern), Dahl, A. B. (Intern), Andersen, J. D. (Forskerdatabase), Johansen, P. (Ekstern), Christoffersen, S. R. (Ekstern), Morling, N. (Forskerdatabase), Borsting, C. (Ekstern), Paulsen, R. R. (Intern)
Number of pages: 4
Publication date: 2013
Effects of mid-term student evaluations of teaching as measured by end-of-term evaluations: An empirical study of course evaluations

Universities have varying policies on how and when to perform student evaluations of courses and teachers. More empirical evidence of the consequences of such policies on quality enhancement of teaching and learning is needed. A study (35 courses at the Technical University of Denmark) was performed to illustrate the effects caused by different handling of mid-term course evaluations on student's satisfaction as measured by end-of-term evaluations. Midterm and end-of-term course evaluations were carried out in all courses. Half of the courses were allowed access to the midterm results. The evaluations generally showed positive improvements over the semester for courses with access, and negative improvements for those without access. Improvements related to: Student learning, student satisfaction, teaching activities, and communication showed statistically significant average differences of 0.1-0.2 points between the two groups. These differences are relatively large compared to the standard deviation of the scores when student effect is removed (approximately 0.7). We conclude that university policies on course evaluations seem to have an impact on the development of the teaching and learning quality as perceived by the students and discuss the findings.

Monitoring Motion of Pigs in Thermal Videos

We propose a new approach for monitoring animal movement in thermal videos. The method distinguishes movements as walking in the expected direction from walking in the opposite direction, stopping or lying down. The method utilizes blob detection combined with optical ow to segment the pigs and extract features which characterize a pig's movement (direction and speed). Subsequently a multiway principal component analysis is used to analyze the movement features and monitor their development over time. Results are presented in the form of quality control charts of the principal components. The method works on-line with pre-training.
On discriminant analysis techniques and correlation structures in high dimensions

This paper compares several recently proposed techniques for performing discriminant analysis in high dimensions, and illustrates that the various sparse methods differ in prediction abilities depending on their underlying assumptions about the correlation structures in the data. The techniques generally focus on two things: Obtaining sparsity (variable selection) and regularizing the estimate of the within-class covariance matrix. For high-dimensional data, this gives rise to increased interpretability and generalization ability over standard linear discriminant analysis. Here, we group the methods in two: Those who assume independence between the variables and thus use a diagonal estimate of the within-class covariance matrix, and those who assume dependence between the variables and thus use an estimate of the within-class covariance matrix, which also estimates the correlations between variables. The two groups of methods are compared and the pros and cons are exemplified using different cases of simulated data. The results illustrate that the estimate of the covariance matrix is an important factor with respect to choice of method, and the choice of method should thus be driven by the nature of the problem at hand.

Optimal vision system design for characterization of apples using US/Vis/NIR spectroscopy data

Quality monitoring of the food items by spectroscopy provides information in a large number of wavelengths including highly correlated and redundant information. Although increasing the information, the increase in the number of wavelengths causes the vision set-up to be more complex and expensive. In this paper, three sparse regression methods; lasso, elastic-net and fused lasso are employed for estimation of the chemical and physical characteristics of one apple cultivar using their high dimensional spectroscopic measurements. The use of sparse regression reduces the number of required wavelengths for prediction and thus, simplifies the required vision set-up. It is shown that, considering a tradeoff between the number of selected bands and the corresponding validation performance during the training step can result in a significant reduction in the number of bands at a small price in the test performance. Furthermore, appropriate regression methods for different number of bands and spectrophotometer design are determined.
PorkCAD: Case study of the design of a pork product prototyper

With the help of industry experts we developed porkCAD, an application intended to aid in the communication process between producer and retailer when developing new meat products for a constantly evolving market. The application interface allows the user to make planar cuts to a virtual pig formed from CT-scans of a real-world pig carcass. We present a case study of the design process from conceptualization to intended introduction into the work flow of a meat production company. We discuss critical design decisions during development and present perspectives for future development.

To determine the usability of porkCAD, we tested it with personnel from the pork industry, using two different controller interfaces, one being a traditional mouse and keyboard input, and the other a six degrees of freedom haptic feedback device. The accurate depiction of pig anatomy guided trained professionals to re-create standardized pig products using porkCAD. The quantitative results of the usability test with sales personnel did not lean significantly in favor of either interface.

Since one interface was extremely well known and the other highly unfamiliar, the fact that users did not express a clear preference for the known input modality is deemed important. We report on the observed user experience regarding the two interfaces.
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

Statistical Outlier Detection for Jury Based Grading Systems

This paper presents an algorithm that was developed to identify statistical outliers from the scores of grading jury members in a large project-based first year design course. The background and requirements for the outlier detection system are presented. The outlier detection algorithm and the follow-up procedures for score validation and appeals are described in detail. Finally, the impact of various elements of the outlier detection algorithm, their interactions, and the sensitivity of their numerical values are investigated. It is shown that the difference in the mean score produced by a grading jury before and after a suspected outlier is removed from the mean is the single most effective criterion for identifying potential outliers but that all of the criteria included in the algorithm have an effect on the outlier detection process.
Statistical Quality Assessment of Pre-fried Carrots Using Multispectral Imaging

Multispectral imaging is increasingly being used for quality assessment of food items due to its non-invasive benefits. In this paper, we investigate the use of multispectral images of pre-fried carrots, to detect changes over a period of 14 days. The idea is to distinguish changes in quality from spectral images of visible and NIR bands. High dimensional feature vectors were formed from all possible ratios of spectral bands in 9 different percentiles per piece of carrot. We propose to use a multiple hypothesis testing technique based on the Benjamini-Hachberg (BH) method to distinguish possible significant changes in features during the inspection days. Discrimination by the SVM classifier supported these results. Additionally, 2-sided t-tests on the predictions of the elastic-net regressions were carried out to compare our results with previous studies on fried carrots. The experimental results showed that the most significant changes occured in day 2 and day 14.
analysis. Application of factor analysis helped to reveal the important issues and the structure of the data hidden in the
students' written comments, while regression analysis showed that some of the revealed factors have a significant impact
on how students rate a course.

General information
State: Published
Organisations: Department of Applied Mathematics and Computer Science, Statistics and Data Analysis
Authors: Sliusarenko, T. (Intern), Clemmensen, L. K. H. (Intern), Erbsøll, B. K. (Intern)
Pages: 564-573
Publication date: 2013

Host publication information
Title of host publication: CSEDU 2013 - Proceedings of the 5th International Conference on Computer Supported
Education
Publisher: SciTePress
ISBN (Print): 9789898565532
BFI conference series: International Conference on Computer-Supported Education (5020033)
Main Research Area: Technical/natural sciences
Conference: 5th International Conference on Computer Supported Education (CSEDU 2013), Aachen, Germany,
06/05/2013 - 06/05/2013
Data mining, Extraction, Factor analysis, Multivariant analysis, Regression analysis, Surveys, Teaching, Students
Source: dtu
Source-ID: n:oai:DTIC-ART:compendex/425721941::34106
Publication: Research - peer-review › Article in proceedings – Annual report year: 2014

The Effect of Rubric Rating Scale on the Evaluation of Engineering Design Projects
This paper explores the impact of the rubric rating scale on the evaluation of projects from a first year engineering design
course. A small experiment was conducted in which twenty-one experienced graders scored five technical posters using
one of four rating scales. All rating scales tested produced excellent results in terms of inter-rater reliability and validity.
However, there were significant differences in the performance of each of the scales. Based on the experiment's results
and past experience, we conclude that increasing the opportunities for raters to deduct points results in greater point
deductions and lower overall scores. Increasing the granularity of the scale can reduce this effect. Rating scales that use
letter grades are less reliable than other types of scale. Assigning weights to individual criteria can lead to problems with
validity if the weights are improperly balanced. Thus, heavily weighted rubrics should be avoided if viable alternatives
exist. Placing more responsibility for the final score on the grader instead of the rubric seems to increase the validity at
the cost of rater satisfaction. Finally, rater discomfort can lead to intentional misuse of a rating scale. This, in turn, increases
the need to perform outlier detection on the final scores. Based on these findings, we recommend rating scale rubrics that
use simple 3 or 4-point ordinal rating scales (augmented checks) for individual criteria and that assign numerical scores to
groups of criteria.

General information
State: Published
Organisations: Department of Mechanical Engineering, Manufacturing Engineering, Department of Applied Mathematics
and Computer Science, Statistics and Data Analysis
Authors: Thompson, M. K. (Intern), Clemmensen, L. K. H. (Intern), Ahn, B. (Ekstern)
Pages: 1490-1502
Publication date: 2013
Main Research Area: Technical/natural sciences

Publication information
Journal: International Journal of Engineering Education
Volume: 29
Issue number: 6
ISSN (Print): 0949-149X
Ratings:
BFI (2018): BFI-level 2
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): SJR 0.386 SNIP 0.764 CiteScore 0.96
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 0.778 SNIP 1.166 CiteScore 0.95
Using multilevel systematic sampling to study apple fruit (Malus domestica Borkh.) quality and its variability at the orchard scale

We report on the performance of a novel sampling method for determining fruit quality variability and yield from an orchard, which focus on its applicability for the fruit industry. We used the ‘fractionator’ tree sampling method to investigate the quality variability of a small, representative sample of ‘Granny Smith’ (Malus x domestica cv. ‘Granny Smith’) apples obtained from a 17 ha orchard based on a final sample of 74 fruit. Estimates of fruit marketable yield and fruit size distribution agreed well with packing house records. The estimated marketable yield was 356.6 ± 89.2 t compared to 374.9 t of fruit packed for export. Distributions of starch (S), soluble solids content (SSC) and flesh firmness (F) were also estimated from the sample. The distribution of starch (S) and fruit mass (M) showed high variability (CVS = SD/mean = 0.32 and CVM = 0.23), whereas SSC and flesh firmness showed moderate variability (CVSSC = 0.11 and CVF = 0.10). The average within-tree variabilities were estimated as CVM = 0.04, CVSSC = 0.10, CVS = 0.15 and CVF = 0.07. Between-tree variabilities were similar to the within-tree variabilities, except for starch (CVM = 0.04, CVISSC = 0.13, CVIS = 0.29 and CVIF = 0.09). From the quality characteristics studied only fruit mass could be significantly related to position of the fruit in the canopy, represented by height of the fruit above ground, the fruit position along the branch and position relative to the tree row orientation in the orchard. Variations in starch, SSC and flesh firmness could not be explained by position of the fruit in the canopy. The methods used in this paper are proposed as tools for studies aimed at understanding sources of quality variability as well as for management purposes. Further research is needed to determine recommended sample sizes to accurately describe the distribution of various quality variables of apples at the orchard scale.

General information
State: Published
Adverse effects on sexual development in rat offspring after low dose exposure to a mixture of endocrine disrupting pesticides

The present study investigated whether a mixture of low doses of five environmentally relevant endocrine disrupting pesticides, epoxiconazole, mancozeb, prochloraz, tebuconazole and procymidone, would cause adverse developmental toxicity effects in rats. In rat dams, a significant increase in gestation length was seen, while in male offspring increased nipple retention and increased incidence and severity of genital malformations were observed. Severe mixture effects on gestation length, nipple retention and genital malformations were seen at dose levels where the individual pesticides caused no or smaller effects when given alone. Generally, the mixture effect predictions based on dose-additivity were in good agreement with the observed effects. The results indicate that there is a need for modification of risk assessment procedures for pesticides, in order to take account of the mixture effects and cumulative intake, because of the potentially serious impact of mixed exposure on development and reproduction in humans.

General information
State: Published
Organisations: National Food Institute, Division of Toxicology and Risk Assessment, Division of Food Chemistry, Department of Informatics and Mathematical Modeling, DTU Data Analysis
Pages: 261-274
Publication date: 2012
Conference: 40th Annual Conference of the European Teratology Society, Linz, Austria, 02/09/2012 - 02/09/2012
Main Research Area: Technical/natural sciences
Publication information
Journal: Reproductive Toxicology
Volume: 34
Issue number: 2
ISSN (Print): 0890-6238
Ratings:
BFI (2018): BFI-level 2
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.92 SJR 1.061 SNIP 1.015
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.219 SNIP 1.1 CiteScore 3.36
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.26 SNIP 1.109 CiteScore 3.28
Texture synthesis algorithms have been researched extensively in the past decade. However, most synthesis algorithms are governed by a set of parameters and produce different results depending on which parameter settings are chosen in conjunction with an exemplar used as a basis for synthesis. So far, automatically selecting parameters suitable for synthesis has been a relatively unexplored topic. In effect, this makes texture synthesis supervised rather than fully automatic.

In this technical paper, we propose automatic parameter optimization methods for example based texture synthesis. We cover research to directly estimate specific texture synthesis parameters, such as patch size and iteration convergence, based on input textures. We also examine various similarity measures and evaluate their effectiveness. The goal for each measure is to properly evaluate how well the resulting synthesis compares to the original input.

A good similarity measure will enable the search for the optimal texture synthesis parameters by maximizing the quality of the synthesis as a function of parameters.

We apply presented methods to a state of the art texture synthesis algorithm, namely the one proposed by Kopf et al [14]. It is easy to find a set of exemplars for which there is no single optimal set of settings. The results show a promising foundation for further research in establishing an automated optimal synthesis for a multitude of textures.
Classification of Astaxanthin Colouration of Salmonid Fish using Spectral Imaging and Tricolour Measurement
The goal of this study was to investigate if it is possible to differentiate between rainbow trout (Oncorhynchus mykiss) having been fed with natural or synthetic astaxanthin. Three different techniques were used for visual inspection of the surface colour of the fish meat: multi-spectral image capturing, tricolour CIELAB measurement, and manual SalmoFan inspection. Furthermore, it was tested whether the best predictions come from measurements of the steak or the fillet of the fish. Methods used for classification were linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), and sparse linear discriminant analysis (SLDA).

Data Driven Constraints for the SVM
We propose a generalized data driven constraint for support vector machines exemplified by classification of paired observations in general and specifically on the human ear canal. This is particularly interesting in dynamic cases such as tissue movement or pathologies developing over time. Assuming that two observations of the same subject in different states span a vector, we hypothesise that such structure of the data contains implicit information which can aid the classification, thus the name data driven constraints. We derive a constraint based on the data which allow for the use of the ℓ1-norm on the constraint while still allowing for the application of kernels. We specialize the proposed constraint to orthogonality of the vectors between paired observations and the estimated hyperplane. We show that imposing the constraint of orthogonality on the paired data yields a more robust classifier solution, compared to the SVM i.e. reduces variance and improves classification rates. We present a quantitative measure of the information level contained in the pairing and test the method on simulated as well as a high-dimensional paired data set of ear-canal surfaces.
Diagnosis and prognosis of Ostheoarthritis by texture analysis using sparse linear models

We present a texture analysis methodology that combines uncommitted machine-learning techniques and sparse feature transformation methods in a fully automatic framework. We compare the performances of a partial least squares (PLS) forward feature selection strategy to a hard threshold sparse PLS algorithm and a sparse linear discriminant model. The texture analysis framework was applied to diagnosis of knee osteoarthritis (OA) and prognosis of cartilage loss. For this investigation, a generic texture feature bank was extracted from magnetic resonance images of tibial knee bone. The features were used as input to the sparse algorithms, which denoted the best features to retain in the model. To cope with the limited number of samples, the data was evaluated using 10 fold cross validation (CV). The diagnosis evaluation using sparse PLS reached a generalization area-under-the-ROC curve (AUC) of 0.93 and the prognosis had AUC of 0.70, both superior to established cartilage based markers known to relate to OA diagnosis and prognosis.

General information
State: Published
Organisations: Department of Informatics and Mathematical Modeling, DTU Data Analysis, Biomediq, University of Copenhagen
Effects of network resolution on topological properties of human neocortex

Graph theoretical analyses applied to neuroimaging datasets have provided valuable insights into the large-scale anatomical organization of the human neocortex. Most of these studies were performed with different cortical scales leading to cortical networks with different levels of small-world organization. The present study investigates how resolution of thickness-based cortical scales impacts on topological properties of human anatomical cortical networks. To this end, we designed a novel approach aimed at determining the best trade-off between small-world attributes of anatomical cortical networks and the number of cortical regions included in the scale. Results revealed that schemes comprising 540–599 regions (surface areas spanning between 250 and 275mm²) at sparsities below 10% showed a superior balance between small-world organization and the size of the cortical scale employed. Furthermore, we found that the cortical scale representing the best trade-off (599 regions) was more resilient to targeted attacks than atlas-based schemes (Desikan–Killiany atlas, 66 regions) and, most importantly, it did not differ that much from the finest cortical scale tested in the present study (1494 regions). In summary, our study confirms that topological organization of anatomical cortical networks varies with both sparsity and resolution of cortical scale, and it further provides a novel methodological framework aimed at identifying cortical schemes that maximize small-worldness with the lowest scale resolution possible.

General information
State: Published
Organisations: DTU Data Analysis, Department of Informatics and Mathematical Modeling, University Pablo de Olavide
Authors: Romero-Garcia, R. (Ekstern), Atienza, M. (Ekstern), Clemmensen, L. K. H. (Intern), Cantero, J. L. (Ekstern)
Pages: 3522-3532
Publication date: 2012
Main Research Area: Technical/natural sciences

Publication information
Journal: NeuroImage
Volume: 59
Issue number: 4
ISSN (Print): 1053-8119
Ratings:
BFI (2018): BFI-level 2
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 6.31 SJR 3.823 SNIP 1.752
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 4.48 SNIP 1.84 CiteScore 6.71
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 4.201 SNIP 2.029 CiteScore 6.9
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 4.376 SNIP 2.026 CiteScore 7.06
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 3.922 SNIP 1.937 CiteScore 6.86
Multispectral Imaging of Wok-Fried Vegetables

Quality control in the food industry is often performed by measuring various chemical compounds in the food involved. The authors propose an imaging concept for acquiring high-quality multispectral images to evaluate optical reflection changes in carrots and celeriac over a period of 14 days. For comparison, sensory analysis was performed on the same samples. Prior to multispectral image recording, the vegetables were prefried and frozen at -30 °C for 4 months. During the 14 days
of image recording, the vegetables were kept at +5 °C. In this period, surface changes and thereby reflectance properties were very subtle. However, they noted statistically significant differences for some wavelengths and combinations of wavelengths. The corresponding sensory tests showed weak differences over the 14 days (significant at a 10% level of significance), which makes it the more important that the authors were able to detect minor changes using multispectral imaging. From our findings, it seems probable that oxidation caused the changes over time.

**General information**

- **State**: Published
- **Organisations**: Department of Informatics and Mathematical Modeling, DTU Data Analysis, National Food Institute, Division of Industrial Food Research
- **Authors**: Clemmensen, L. K. H. (Intern), Dissing, B. S. (Intern), Hyldig, G. (Intern), Løje, H. (Intern)
- **Publication date**: 2012
- **Main Research Area**: Technical/natural sciences

**Publication information**

- **Journal**: Journal of Imaging Science and Technology
- **Volume**: 56
- **Issue number**: 2
- **Article number**: 020404
- **ISSN (Print)**: 1062-3701
- **Ratings**:
  - BFI (2018): BFI-level 1
  - BFI (2017): BFI-level 1
  - Web of Science (2017): Indexed Yes
  - BFI (2016): BFI-level 1
  - Scopus rating (2016): SJR 0.199 SNIP 0.438 CiteScore 0.38
  - BFI (2015): BFI-level 1
  - Scopus rating (2015): SJR 0.191 SNIP 0.435 CiteScore 0.34
  - BFI (2014): BFI-level 1
  - Scopus rating (2014): SJR 0.223 SNIP 0.634 CiteScore 0.38
  - BFI (2013): BFI-level 1
  - Scopus rating (2013): SJR 0.309 SNIP 0.813 CiteScore 0.48
  - ISI indexed (2013): ISI indexed yes
  - BFI (2012): BFI-level 1
  - Scopus rating (2012): SJR 0.494 SNIP 1.309 CiteScore 0.8
  - ISI indexed (2012): ISI indexed yes
  - Web of Science (2012): Indexed yes
  - BFI (2011): BFI-level 1
  - Scopus rating (2011): SJR 0.495 SNIP 1.555 CiteScore 0.85
  - ISI indexed (2011): ISI indexed yes
  - BFI (2010): BFI-level 1
  - Scopus rating (2010): SJR 0.349 SNIP 0.973
  - Web of Science (2010): Indexed yes
  - BFI (2009): BFI-level 1
  - Scopus rating (2009): SJR 0.405 SNIP 0.942
  - BFI (2008): BFI-level 1
  - Scopus rating (2008): SJR 0.603 SNIP 1.246
  - Scopus rating (2007): SJR 0.835 SNIP 1.473
  - Scopus rating (2006): SJR 0.611 SNIP 1.317
  - Scopus rating (2005): SJR 0.44 SNIP 1.131
  - Scopus rating (2004): SJR 0.48 SNIP 1.355
  - Scopus rating (2003): SJR 0.373 SNIP 1.422
  - Scopus rating (2002): SJR 0.58 SNIP 1.039
  - Scopus rating (2001): SJR 0.597 SNIP 1.543
  - Scopus rating (2000): SJR 0.666 SNIP 1.118
  - Scopus rating (1999): SJR 0.455 SNIP 0.847
- **Original language**: English

**DOIs:**
Regression and Sparse Regression Methods for Viscosity Estimation of Acid Milk From It's Sls Features

Statistical solutions find widespread use in food and medicine quality control. We investigate the effect of different regression and sparse regression methods for a viscosity estimation problem using the spectro-temporal features from new Sub-Surface Laser Scattering (SLS) vision system. From this investigation, we propose the optimal solution for regression estimation in case of noisy and inconsistent optical measurements, which is the case in many practical measurement systems. The principal component regression (PLS), partial least squares (PCR) and least angle regression (LAR) methods are compared with sparse LAR, lasso and Elastic Net (EN) sparse regression methods. Due to the inconsistent measurement condition, Locally Weighted Scatter plot Smoothing (Loess) has been employed to alleviate the undesired variation in the estimated viscosity. The experimental results of applying different methods show that, the sparse regression lasso outperforms other methods. In addition, the use of local smoothing has improved the results considerably for all regression methods. Due to the sparsity of lasso, this result would assist to design a simpler vision system with less spectral bands.

Does an association between student evaluations of related CDIO courses exist?

This paper analyses routine course evaluation performed by students in the computer science related professional bachelor degree educations at DTU. Specifically, a set of two related courses are considered. The courses are: “Introductory Programming” and “Development Methods for IT-Systems”. Both courses include lectures and lab work. It is seen that both similarities and differences in the evaluations can be found. The similarities and differences can in part be used to assess if the CDIO concept has been implemented as it was intended and possible adjustments can be suggested.

Improving texture optimization with application to visualizing meat products

When inspecting food quality, CT Scanning is among the primary tools used to gain insight. It provides valuable volumetric data using a process, which leaves the product unspoiled and untouched. However, volumetric data is merely a measure of density and therefore contains no appearance information (such as color, translucency, reflective properties). One way of reintroducing this lost information back to the volume data is to synthesize an appropriate texture and apply this to the
A recent method within the field of texture synthesis is called Texture Optimization presented by Kopf et al. in 2007. This method accepts a number of 2D input exemplars, from which it generates a solid texture volume. The volume is iteratively improved via an expectation maximization algorithm. The bottleneck of Texture Optimization occurs during a nearest neighbor search, between texture patches from the 2D input exemplars and the generated texture volume. We examine the current procedures for minimizing the bottleneck and present a novel approach which increases the speed of the synthesis algorithm while minimizing loss of quality. The nearest neighbor search is performed in a high dimensional space. Applying a principal component analysis on the texture patches originating from the synthesized solid accelerates the process. These patches are then reduced in dimensionality until “only” 95% of their original variance remains. This usually results in a dimension reduction from 192 to about 60-80. The reduction in dimensionality speeds up the convergence of the Texture Optimization method considerably. We examine the impacts of reducing the dimensionality further by tweaking the parameters as well as introducing an alternative method to reducing the dimensionality.

Additionally, we study the possibility of selecting only a subsample of the neighborhoods available from the input exemplar without significantly impacting the overall synthesis quality.

General information
State: Published
Organisations: DTU Data Analysis, Department of Informatics and Mathematical Modeling, Image Analysis and Computer Graphics
Authors: Clemmensen, L. K. H. (Intern), Laursen, L. F. (Intern)
Number of pages: 98
Pages: 81-86
Publication date: 2011

Host publication information
Title of host publication: Scandinavian Workshop on Imaging Food Quality 2011: Ystad, May 27, 2011 - Proceedings
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Number: 15
Main Research Area: Technical/natural sciences
Workshop: Scandinavian Workshop on Imaging Food Quality 2011, Ystad, Sweden, 27/05/2011 - 27/05/2011
Subset Selection, Non-Negative Matrix Factorization, Dimension Reduction, Princible Component Analysis, Volumetric Rendering, Texture Synthesis
Electronic versions:
Pages from tr11_15-9.pdf
Links:
http://www2.imm.dtu.dk/projects/SWIFQ/
Source: orbit
Source-ID: 279487
Publication: Research - peer-review › Article in proceedings – Annual report year: 2011

Multispectral imaging of wok fried vegetables
This paper shows how multispectral images can be used to assess color change over time in wok fried vegetables. We present results where feature selection was performed with sparse methods from the multispectral images to detect the color changes of wok fried carrots and celeriac stored at +5°C over 14 days. A pairwise t-test was used to detect if the differences over days were significant. For both the original as well as a follow experiment significant differences were seen in particular for celeriac, but also to some extend for carrots.

General information
State: Published
Organisations: Mathematical Statistics, National Food Institute, Department of Informatics and Mathematical Modeling, DTU Data Analysis
Authors: Løje, H. (Intern), Dissing, B. S. (Intern), Clemmensen, L. K. H. (Intern), Ersbøll, B. K. (Intern), Adler-Nissen, J. (Intern)
Pages: 59-62
Publication date: 2011

Host publication information
Title of host publication: Scandinavian Workshop on Imaging Food Quality 2011: Ystad, May 27, 2011 - Proceedings
Place of publication: Kgs. Lyngby, Denmark
Publisher: Technical University of Denmark (DTU)
Number: 15
**Prediction of gestation length and nipple retention in rats after mixed exposure to low doses of dissimilarly acting endocrine disrupting pesticides**

**General information**
State: Published
Organisations: Division of Toxicology and Risk Assessment, National Food Institute, DTU Data Analysis, Department of Informatics and Mathematical Modeling, Technical University of Denmark
Authors: Jacobsen, P. R. (Intern), Christiansen, S. (Intern), Clemmensen, L. K. H. (Intern), Axelstad, M. (Ekstern), Ersbøll, B. K. (Ekstern), Hass, U. (Intern)
Pages: 168
Publication date: 2011
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Reproductive Toxicology
Volume: 32
Issue number: 2
ISSN (Print): 0890-6238
Ratings:
BFI (2018): BFI-level 2
BFI (2017): BFI-level 2
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 2.92 SJR 1.061 SNIP 1.015
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): SJR 1.219 SNIP 1.1 CiteScore 3.36
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): SJR 1.26 SNIP 1.109 CiteScore 3.28
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): SJR 1.02 SNIP 1.062 CiteScore 2.91
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): SJR 1.189 SNIP 1.07 CiteScore 3.28
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): SJR 1.113 SNIP 1.197 CiteScore 3.15
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Scopus rating (2010): SJR 1.298 SNIP 1.347
Web of Science (2010): Indexed yes
Sparse discriminant analysis

We consider the problem of performing interpretable classification in the high-dimensional setting, in which the number of features is very large and the number of observations is limited. This setting has been studied extensively in the chemometrics literature, and more recently has become commonplace in biological and medical applications. In this setting, a traditional approach involves performing feature selection before classification. We propose sparse discriminant analysis, a method for performing linear discriminant analysis with a sparseness criterion imposed such that classification
and feature selection are performed simultaneously. Sparse discriminant analysis is based on the optimal scoring interpretation of linear discriminant analysis, and can be extended to perform sparse discrimination via mixtures of Gaussians if boundaries between classes are nonlinear or if subgroups are present within each class. Our proposal also provides low-dimensional views of the discriminative directions. © 2011 American Statistical Association and the American Society for Qualitys.
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

A comparison of dimension reduction methods with application to multi-spectral images of sand used in concrete

This paper presents a comparison of dimension reduction methods based on a novel machine vision application for estimating moisture content in sand used to make concrete. For the application in question it is very important to know the moisture content of the sand so as to ensure good-quality concrete. In order to achieve a continuous in-line approach for the concrete mixing, digital image analysis is used. Multi-spectral images, consisting of nine spectral bands in the visible and near infrared (NIR) range, were acquired. Each image consists of approximately 9 million pixels. Five different sand types were examined with 20-60 images for each type. To reduce the amount of data, features were extracted from the multi-spectral images; the features were summary statistics on single bands and pairs of bands as well as morphological summaries. The number of features (2,016) is high in relation to the number of observations and, therefore, dimension reductive methods are needed. Furthermore, speed, which is an important consideration, is aided by the use of a small number of variables. On top of that, fewer dimensions tend to give more robust results. Two traditional statistical methods for dimension reduction (forward selection and principal components) combined with ordinary least squares and one sophisticated chemometrics algorithm (genetic algorithm-partial least squares) are compared to the recently proposed least angle regression-elastic net (LARS-EN) model selection method.

General information
State: Published
Organisations: DTU Data Analysis, Department of Informatics and Mathematical Modeling
Authors: Clemmensen, L. K. H. (Intern), Hansen, M. E. (Ekstern), Ersbøll, B. K. (Intern)
Canonical correlation analysis in education: associations between student evaluations of courses and instructors

At the Technical University of Denmark (DTU) course evaluations are performed by the students on a questionnaire. On one form the students are asked specific questions regarding the course. On a second form they are asked specific questions about the teacher. We propose to apply canonical correlation analysis (CCA) to investigate the association...
between how students evaluate the course and how students evaluate the teacher and to reveal the structure of this association. Student’s evaluation data is characterized by high correlation between the variables (questions) and insufficient sample size, which can lead to inaccurate estimates of parameters, non-generalizable and hardly interpretable results of CCA. Newly developed regularized CCA and sparse CCA methods are used to address weaknesses of CCA. Regularized CCA is uses an $L_2$ penalization to shrink the weights by imposing penalty on their size; highly correlated variables get similar weights, resulting in a grouping effect. Sparse CCA incorporates variable selection in both sets of variables by introducing Lasso penalization.

**General information**
State: Published
Organisations: DTU Data Analysis, Department of Informatics and Mathematical Modeling
Authors: Sliusarenko, T. (Intern), Clemmensen, L. K. H. (Intern)
Publication date: 2010
Event: Abstract from XXVth International Biometric Conference, Florianopolis, Brazil, .
Main Research Area: Technical/natural sciences
Source: orbit
Source-ID: 274400
Publication: Research - peer-review › Conference abstract for conference – Annual report year: 2010

**Investigating salt frost scaling by using statistical methods**
A large data set comprising data for 118 concrete mixes on mix design, air void structure, and the outcome of freeze/thaw testing according to SS 13 72 44 has been analysed by use of statistical methods. The results show that with regard to mix composition, the most important parameter is the equivalent water to cement ratio (w/ceq). Regarding the importance of the air void structure, the most important parameter turned out to be the total surface area of air voids. The impact of this parameter is statistically more significant than the impact of the spacing factor, and it raises the question if total surface area of air voids is a more appropriate evaluation criterion than the spacing factor, which is normally used.

**General information**
State: Published
Organisations: Section for Construction Materials, Department of Civil Engineering, DTU Data Analysis, Department of Informatics and Mathematical Modeling
Authors: Hasholt, M. T. (Intern), Clemmensen, L. K. H. (Intern)
Pages: 1-16
Publication date: 2010
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Nordic Concrete Research
Issue number: 2
ISSN (Print): 0800-6377
Ratings:
BFI (2018): BFI-level 1
BFI (2017): BFI-level 1
BFI (2016): BFI-level 1
BFI (2015): BFI-level 1
BFI (2014): BFI-level 1
BFI (2013): BFI-level 1
ISI indexed (2013): ISI indexed no
BFI (2012): BFI-level 1
ISI indexed (2012): ISI indexed no
BFI (2011): BFI-level 1
ISI indexed (2011): ISI indexed no
BFI (2010): BFI-level 1
BFI (2009): BFI-level 1
BFI (2008): BFI-level 1
Original language: English
Source: orbit
Source-ID: 271288
Publication: Research - peer-review › Journal article – Annual report year: 2010
NetRaVE is a small suite of R functions for generating dependency networks using sparse regression methods. Such networks provide an alternative to interpreting 'top n lists' of genes arising out of an analysis of microarray data, and they provide a means of organizing and visualizing the resulting information in a manner that may suggest relationships between genes.

**General information**
State: Published
Organisations: DTU Data Analysis, Department of Informatics and Mathematical Modeling
Authors: Phatak, A. (Ekstern), Kiiveri, H. (Ekstern), Clemmensen, L. K. H. (Intern), Wilson, W. (Ekstern)
Pages: 1576-1577
Publication date: 2010
Main Research Area: Technical/natural sciences

**Publication information**
Journal: Bioinformatics
Volume: 26
Issue number: 12
ISSN (Print): 1367-4803
Ratings:
BFI (2018): BFI-level 2
BFI (2017): BFI-level 2
Web of Science (2017): Indexed yes
BFI (2016): BFI-level 2
Scopus rating (2016): CiteScore 6.42
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 2
Scopus rating (2015): CiteScore 6.06
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 2
Scopus rating (2014): CiteScore 5.5
Web of Science (2014): Indexed yes
BFI (2013): BFI-level 2
Scopus rating (2013): CiteScore 5.78
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 2
Scopus rating (2012): CiteScore 6.73
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 2
Scopus rating (2011): CiteScore 5.61
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 2
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 2
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 2
Web of Science (2008): Indexed yes
Web of Science (2007): Indexed yes
Web of Science (2006): Indexed yes
Web of Science (2005): Indexed yes
Web of Science (2004): Indexed yes
Web of Science (2003): Indexed yes
Web of Science (2002): Indexed yes
Web of Science (2001): Indexed yes
Temporal reflectance changes in vegetables

Quality control in the food industry is often performed by measuring various chemical compounds of the food involved. We propose an imaging concept for acquiring high quality multispectral images to evaluate changes of carrots and celeriac over a period of 14 days. Properties originating in the surface chemistry of vegetables may be captured in an integrating sphere illumination which enables the creation of detailed surface chemistry maps with a good combination of spectral and spatial resolutions. Prior to multispectral image recording, the vegetables were prefried and frozen at -30°C for four months. During the 14 days of image recording, the vegetables were kept at +5°C in refrigeration. In this period, surface changes and thereby reflectance properties were very subtle. To describe this small variation we employed advanced statistical techniques to search a large featurespace of variables extracted from the chemistry maps. The resulting components showed a change in both the carrot and celeriac samples. We were able to deduct from the resulting components that oxidation caused the changes over time.

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, Division of Food Production Engineering, National Food Institute
Authors: Dissing, B. S. (Intern), Clemmensen, L. K. H. (Intern), Ersbøll, B. K. (Intern), Leje, H. (Intern), Adler-Nissen, J. (Intern)
Pages: 1917-1922
Publication date: 2009

Host publication information
Title of host publication: 2009 IEEE 12th International Conference on Computer Vision Workshops (ICCV Workshops)
Publisher: IEEE
ISBN (Print): 978-1-4244-4442-7
Main Research Area: Technical/natural sciences
Electronic versions:
Dissing.pdf
DOIs: 10.1109/ICCVW.2009.5457516

Bibliographical note
Copyright 2009 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Source: orbit
Source-ID: 250334
Publication: Research › Article in proceedings – Annual report year: 2009

Sparse Discriminant Analysis

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling, Stanford University
Authors: Clemmensen, L. K. H. (Intern), Hastie, T. (Ekstern), Ersbøll, B. K. (Intern)
Publication date: 2008

Publication information
Place of publication: Lyngby
Publisher: Technical University of Denmark, DTU Informatics, Building 321
Original language: English
A method for comparison of growth media in objective identification of Penicillium based on multi-spectral imaging

We consider the problems of using excessive growth media for identification and performing objective identification of fungi at the species level. We propose a method for choosing the subset of growth media, which provides the best discrimination between several fungal species. Furthermore, we propose the use of multi-spectral imaging as a means of objective identification. Three species of the fungal genus Penicillium are subject to classification. To obtain an objective classification we use multi-spectral images. Previously, RGB images have proven useful for the purpose. We use multi-spectral bands as they provide additional information about the chemistry of the fungal colonies. In this study three media [Czapek yeast extract agar (CYA), oatmeal agar (OAT), and yeast extract sucrose agar (YES)] have been compared on their ability to discriminate between the three species. We propose a statistical method to test which medium or combination of media gives the best discrimination. Statistical tests indicate that YES combined with CYA is the best choice of media in this case. However, for the objective identification one medium is sufficient to discriminate between the species. Statistical tests show that there are significant differences between the species on all individual media, and that these differences are largest on YES. The objective identification has been performed solely by means of digital image analysis. The features obtained from the image analysis merely correspond to macro-morphological features. The species have been classified using only 3–4 of the spectral bands with a 100% correct classification rate using both leave-one-out cross-validation and test set validation.
Filamentous fungi, Penicillium, Objective identification, Macro-morphology, Multi-spectral images, Choice of growth media

Original language: English

Ratings:
BFI (2018): BFI-level 1
BFI (2017): BFI-level 1
Web of Science (2017): Indexed Yes
BFI (2016): BFI-level 1
Scopus rating (2016): CiteScore 2.05 SJR 0.723 SNIP 0.8
Web of Science (2016): Indexed yes
BFI (2015): BFI-level 1
Scopus rating (2015): SJR 0.816 SNIP 0.873 CiteScore 2.04
Web of Science (2015): Indexed yes
BFI (2014): BFI-level 1
Scopus rating (2014): SJR 0.903 SNIP 1.037 CiteScore 2.28
BFI (2013): BFI-level 1
Scopus rating (2013): SJR 0.917 SNIP 1.019 CiteScore 2.5
ISI indexed (2013): ISI indexed yes
Web of Science (2013): Indexed yes
BFI (2012): BFI-level 1
Scopus rating (2012): SJR 0.87 SNIP 1.004 CiteScore 2.32
ISI indexed (2012): ISI indexed yes
Web of Science (2012): Indexed yes
BFI (2011): BFI-level 1
Scopus rating (2011): SJR 0.9 SNIP 0.972 CiteScore 2.29
ISI indexed (2011): ISI indexed yes
Web of Science (2011): Indexed yes
BFI (2010): BFI-level 1
Scopus rating (2010): SJR 0.945 SNIP 1.05
Web of Science (2010): Indexed yes
BFI (2009): BFI-level 1
Scopus rating (2009): SJR 0.993 SNIP 1.156
Web of Science (2009): Indexed yes
BFI (2008): BFI-level 1
Scopus rating (2008): SJR 0.926 SNIP 1.031
Web of Science (2008): Indexed yes
Scopus rating (2007): SJR 0.942 SNIP 1.111
Web of Science (2007): Indexed yes
Scopus rating (2006): SJR 1.138 SNIP 1.251
Web of Science (2006): Indexed yes
Scopus rating (2005): SJR 0.964 SNIP 1.141
Web of Science (2005): Indexed yes
Scopus rating (2004): SJR 0.921 SNIP 1.037
Web of Science (2004): Indexed yes
Scopus rating (2003): SJR 0.933 SNIP 1.217
Web of Science (2003): Indexed yes
Scopus rating (2002): SJR 0.821 SNIP 1.002
Scopus rating (2001): SJR 0.867 SNIP 0.894
Web of Science (2001): Indexed yes
Scopus rating (2000): SJR 0.721 SNIP 0.864
Web of Science (2000): Indexed yes
Scopus rating (1999): SJR 0.724 SNIP 0.729

DOIs:
Multiplicative updates for the LASSO
Multiplicative updates have proven useful for non-negativity constrained optimization. Presently, we demonstrate how multiplicative updates also can be used for unconstrained optimization. This is for instance useful when estimating the least absolute shrinkage and selection operator (LASSO), i.e. least squares minimization with $\ell_1$-norm regularization, since the multiplicative updates (MU) can efficiently exploit the structure of the problem traditionally solved using quadratic programming (QP). We derive an algorithm based on MU for the LASSO and compare the performance to Matlab's standard QP solver as well as the basis pursuit denoising algorithm (BP) which can be obtained from www.sparselab.stanford.edu. The algorithms were tested on three benchmark bio-informatic datasets: A small scale data set where the number of observations is larger than the number of variables estimated ($M$.

Precise acquisition and unsupervised segmentation of multi-spectral images
In this work, an integrated imaging system to obtain accurate and reproducible multi-spectral images and a novel multi-spectral image segmentation algorithm are proposed. The system collects up to 20 different spectral bands within a range that vary from 395 nm to 970 nm. The system is designed to acquire geometrically and chromatically corrected images in homogeneous and diffuse illumination, so images can be compared over time. The proposed segmentation algorithm combines the information provided by all the spectral bands to segment the different regions of interest. Three experiments are conducted to show the ability of the system to acquire highly precise, reproducible and standardized multi-spectral images and to show its applicabilities in different situations.
Multispectral recordings and analysis of psoriasis lesions
An objective method to evaluate the severeness of psoriasis lesions is proposed. In order to obtain objectivity multi-spectral imaging is used. The multi-spectral images give rise to a large p, small n problem which is solved by use of elastic net model selection. The method is promising for further studies of larger data sets including more patients than the four regarded here.

General information
State: Published
Organisations: Image Analysis and Computer Graphics, Department of Informatics and Mathematical Modeling
Authors: Clemmensen, L. K. H. (Intern), Ersbøll, B. K. (Intern)
Pages: 15-18
Publication date: 2006

Host publication information
Title of host publication: MICCAI ’06 - Workshop on Biophotonics Imaging for Diagnostics and Treatment, October 6, 2006 proceedings, 9th MICCAI Conference

Series: MM-Technical Report
Number: 17
Main Research Area: Technical/natural sciences
Workshop: 9th Medical Image Computing and Computer-Assisted Intervention Conference, Copenhagen, Denmark, 06/10/2006
Links:
Source: orbit
Source-ID: 207396
Publication: Research › Article in proceedings – Annual report year: 2006

Projects:

Designing a Real-time Tracking and Feedback System to use During Endoscopic Procedures
Department of Applied Mathematics and Computer Science
Period: 01/03/2016 → 26/04/2019
Number of participants: 4
Phd Student:
Norsk, David (Intern)
Supervisor:
Clemmensen, Line Katrine Harder (Intern)
Svendsen, Lars Bo (Ekstern)
Main Supervisor:
Paulsen, Rasmus Reinhold (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Optimisation of biodevice production
Master project
Department of Applied Mathematics and Computer Science
Statistics and Data Analysis
Scandinavian Micro Biodevice ApS
Period: 01/01/2016 → 04/06/2016
Number of participants: 3
Project participant:
Rabel, Mads Peter (Ekstern)
Supervisor:
Thyregod, Camilla (Intern)
Main Supervisor:
Clemmensen, Line Katrine Harder (Intern)

Project

Big Data Modelling with Applications to Airports
Department of Applied Mathematics and Computer Science
Period: 01/08/2015 → 31/12/2018
Number of participants: 3
Phd Student:
Nielsen, Agnes Martine (Intern)
Supervisor:
Dahl, Anders Bjorholm (Intern)
Main Supervisor:
Clemmensen, Line Katrine Harder (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Methods and tools for the statistical data analysis for large datasets collected from bio-based manufacturing processes
Department of Applied Mathematics and Computer Science
Period: 01/06/2015 → 31/05/2018
Number of participants: 4
Phd Student:
Spooner, Max Peter (Intern)
Supervisor:
Clemmensen, Line Katrine Harder (Intern)
Ersbøll, Bjarne Kjær (Intern)
Main Supervisor:
Kulahci, Murat (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Forskningsrådsfinansiering
Project: PhD

Computerised Quantification of Motions Associated with Psychiatric Disorders
Department of Applied Mathematics and Computer Science
Period: 15/08/2014 → 28/02/2018
Number of participants: 5
Phd Student:
Einarsson, Gudmundur (Intern)
Supervisor:
Clemmensen, Line Katrine Harder (Intern)
Fink-Jensen, Anders (Ekstern)
Pagsberg, Anne Katrine (Ekstern)
Main Supervisor:
Paulsen, Rasmus Reinhold (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Samfinansieret - Andet
Project: PhD

Statistical modelling of space-time processes with
Department of Applied Mathematics and Computer Science
Period: 01/11/2013 → 16/08/2017
Number of participants: 7
Phd Student: 
Lenzi, Amanda (Intern)
Supervisor:
Clemmensen, Line Katrine Harder (Intern)
Pinson, Pierre (Intern)
Main Supervisor: 
Ersbøll, Bjarne Kjær (Intern)
Examiner: 
Stockmarr, Anders (Intern)
Girard, Robin (Ekstern)
Thorarinsdottir, Thordis L. (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: Science Without Borders, Brasi

Relations
Publications:
Statistical modelling of space-time processes with application to wind power.
Project: PhD

Characterization of absorption enhancers for orally administered therapeutic peptides in tablet formulations - applying statistical learning
Department of Applied Mathematics and Computer Science
Period: 01/05/2013 → 30/09/2016
Number of participants: 9
Phd Student: 
Welling, Søren Havelund (Intern)
Supervisor:
Buckley, Stephen T. (Ekstern)
Clemmensen, Line Katrine Harder (Intern)
Hovgaard, Lars (Ekstern)
Refsgaard, Hanne (Intern)
Main Supervisor: 
Brockhoff, Per B. (Intern)
Examiner: 
Kulahci, Murat (Intern)
Arvastson, Lars Johan (Intern)
Genuer, Robin (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: ErhvervsPhD-ordningen VTU

Relations
Publications:
Characterization of absorption enhancers for orally administered therapeutic peptides in tablet formulations - Applying statistical learning
Project: PhD

Monitoring Animal Wellbeing
Department of Applied Mathematics and Computer Science
Period: 15/12/2011 → 31/03/2016
Number of participants: 7
Phd Student: 
Gronskyte, Ruta (Intern)
Supervisor:
Clemmensen, Line Katrine Harder (Intern)
Hviid, Marchen Sonja (Ekstern)
Main Supervisor:
Kulahci, Murat (Intern)
Examiner:
Ersbøll, Bjarne Kjær (Intern)
Bergquist, Bjarne (Ekstern)
Christensen, Lars Bager (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Institut, samfinansiering

Relations
Publications:
Monitoring Animal Well-being
Project: PhD

Spectral imaging of meat quality - color, texture and structure
Department of Applied Mathematics and Computer Science
Period: 01/12/2011 → 19/03/2015
Number of participants: 6
Phd Student:
Trinderup, Camilla Himmelstrup (Intern)
Supervisor:
Dahl, Anders Bjorholm (Intern)
Main Supervisor:
Conradsen, Knut (Intern)
Examiner:
Clemmensen, Line Katrine Harder (Intern)
Christensen, Lars Bager (Intern)
Parker, Alan (Ekstern)

Financing sources
Source: Internal funding (public)
Name of research programme: Forskningsrådsfinansiering
Project: PhD

Multivariate Analysis Techniques for Optimal Vision System Design
Department of Applied Mathematics and Computer Science
Period: 01/06/2011 → 21/09/2015
Number of participants: 6
Phd Student:
Sharifzadeh, Sara (Intern)
Supervisor:
Ersbøll, Bjarne Kjær (Intern)
Main Supervisor:
Clemmensen, Line Katrine Harder (Intern)
Examiner:
Conradsen, Knut (Intern)
Ames, Brendan P. W. (Ekstern)
vanden Berg, Frans W.J. (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Forskningsrådsfinansiering
Project: PhD
Control & Surveillance of Automated Production Steps (a part of the inSPIRe Food)

Summary of project: Automation of many manual operations in the food industry is difficult, because the criteria for process control are often based on tacit knowledge of the operator. Our hypothesis is that a route to optimal automation of such operations is to register how the trained process operator makes decisions from observations of the process and combining this knowledge with predictive modelling of input/output of the process units.

Department of Applied Mathematics and Computer Science
Statistics and Data Analysis
National Food Institute
Research Group for Food Production Engineering
Image Analysis & Computer Graphics
Period: 01/01/2011 → 31/12/2016
Number of participants: 5
Project participant:
Larsen, Rasmus (Intern)
Ersbøll, Bjarne Kjær (Intern)
Frosch, Stina (Intern)
Clemmensen, Line Katrine Harder (Intern)
Larsen, Anders Boesen Lindbo (Intern)

Financing sources
Source: Public research council
Name of research programme: Danish Council for Strategic Research and the Danish Council for Technology (now The Danish Innovation Foundation)
Amount: 5,218,000.00 Danish Kroner

Multivariate Analysis Techniques for Optimal Vision Design
Department of Informatics and Mathematical Modeling
Period: 01/10/2010 → 10/01/2011
Number of participants: 3
Phd Student:
Mazzaretto, Andrea (Intern)
Supervisor:
Clemmensen, Line Katrine Harder (Intern)
Main Supervisor:
Ersbøll, Bjarne Kjær (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: Forskningsrådsfinansiering
Project: PhD

Facial recognition
Department of Informatics and Mathematical Modeling
Period: 01/09/2010 → 26/02/2014
Number of participants: 5
Phd Student:
Fagertun, Jens (Intern)
Main Supervisor:
Pauksen, Rasmus Reinhold (Intern)
Examiner:
Clemmensen, Line Katrine Harder (Intern)
Cootes, Timothy F. (Ekstern)
Hansen, Dan Witzner (Intern)
Financing sources
Source: Internal funding (public)
Name of research programme: Institut stipendie (DTU)
Project: PhD

Data-analyse i sparse, høj-dimensionale rum
Department of Informatics and Mathematical Modeling
Period: 01/04/2006 → 31/03/2010
Number of participants: 5
Phd Student:
Clemmensen, Line Katrine Harder (Intern)
Main Supervisor:
Ersbøll, Bjarne Kjær (Intern)
Examiner:
Larsen, Rasmus Werner (Intern)
Bigun, Josef (Ekstern)
Bro, Rasmus (Intern)

Financing sources
Source: Internal funding (public)
Name of research programme: DTU-lønnet stipendie
Project: PhD