

# A Noise Robust Statistical Texture Model

Klaus B. Hilger, Mikkel B. Stegmann, and Rasmus Larsen

Informatics and Mathematical Modelling, Technical University of Denmark,  
DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby  
{kbh,mbs,rl}@imm.dtu.dk, <http://www.imm.dtu.dk>

**Abstract.** This paper presents a novel approach to the problem of obtaining a low dimensional representation of texture (pixel intensity) variation present in a training set after alignment using a Generalised Procrustes analysis. We extend the conventional analysis of training textures in the Active Appearance Models segmentation framework. This is accomplished by augmenting the model with an estimate of the covariance of the noise present in the training data. This results in a more compact model maximising the signal-to-noise ratio, thus favouring subspaces rich on signal, but low on noise. Differences in the methods are illustrated on a set of left cardiac ventricles obtained using magnetic resonance imaging.

## 1 Introduction

Over the past few years, models capable of synthesising complete images of objects have proven very useful when interpreting images. One example is the Active Appearance Models (AAMs) [1, 2]. Applications of AAMs include recovery and variation analysis of anatomical structures in medical images, such as magnetic resonance images (MRIs) [3], radiographs [4, 5] and ultrasound images [6].

Images can be synthesised in many ways, e.g. [7] uses a linear combination of shape-compensated training images. To reduce dimensionality, AAMs uses a Principal Component (PC) analysis of the training set to synthesise new images. By maximising the variance only, the PC is modelling any noise present in the training set along with the uncontaminated hidden image data. In this paper, we propose to extend the AAM framework by augmenting the image representation with noise characteristics. This is accomplished by applying the Minimum Noise Fraction (MNF) transformation [8].

The ancestor of AAMs, the Active Shape Models [9] have previously been extended by means of a variant of MNF in the analysis of shapes, see [5]. Here, we extend this work to pixel intensities, henceforth denoted *texture*.

The MNF extracts important otherwise occluded information in the correlation structures of the data, and aims at obtaining a low dimensional model representation. As opposed to the PC transform, the MNF transform takes the spatial nature of the image into account. Whereas the PC transform only requires knowledge of the dispersion (covariance) matrix, the MNF transform requires an estimate of the dispersion matrix of the noise structure as additional information. The MNF transform was originally proposed as a transformation for ordering multispectral data in terms of image quality with applications for noise removal.

This paper is organised as follows. Section 2 summarises AAMs and describes the applied statistical models. Section 3 describes the data analysed, and Section 4 presents a comparative study of the PC and MNF. In Section 5 we summarise and give some concluding remarks.

## 2 Methods

In the following AAMs are summarised along with a description of the traditional AAM texture model; the PC transform, and the proposed alternative; the MNF transform.

### 2.1 Active Appearance Models

Active Appearance Models [1, 2] establish a compact parameterisation of object variability, as learned from a training set by estimating a set of latent variables. From these quantities new images similar to the training set can be generated. Objects are defined by marking up each example with points of correspondence over the set either by hand, or by semi- to completely automated methods. Exploiting prior knowledge about the local nature of the optimisation space, these models can be rapidly fitted to unseen images, given a reasonable initialisation.

Shape and texture variability is conventionally modelled by means of PC transforms. Let there be given  $P$  training examples for an object class, and let each example be represented by a set of  $N$  landmark points and  $M$  texture samples. The  $P$  shape examples are aligned to a common mean using a Generalised Procrustes analysis. The Procrustes shape coordinates are subsequently projected into the tangent plane of the shape manifold, at the pole denoted by the mean shape. The  $P$  textures are warped into correspondence using a suitable warp function and subsequently sampled from this *shape-free* reference. Typically, this geometrical reference is the Procrustes mean shape. Let  $\mathbf{s}$  and  $\mathbf{t}$  denote a synthesized shape and texture and let  $\bar{\mathbf{s}}$  and  $\bar{\mathbf{t}}$  denote the corresponding means. New instances are now generated by the adjusting PC scores,  $\mathbf{b}_s$  and  $\mathbf{b}_t$  in

$$\mathbf{s} = \bar{\mathbf{s}} + \Phi_s \mathbf{b}_s \quad , \quad \mathbf{t} = \bar{\mathbf{t}} + \Phi_t \mathbf{b}_t \quad (1)$$

where  $\Phi_s$  and  $\Phi_t$  are eigenvectors of the shape and texture dispersions estimated from the training set. To regularise the model and improve speed and compactness,  $\Phi_s$  and  $\Phi_t$  are truncated, usually such that a certain amount of variance in the training set is explained. To obtain a combined shape and texture parameterisation,  $\mathbf{c}$ , the values of  $\mathbf{b}_s$  and  $\mathbf{b}_t$  over the training set are combined

$$\mathbf{b} = \begin{pmatrix} \mathbf{W}_s \mathbf{b}_s \\ \mathbf{b}_t \end{pmatrix} = \begin{pmatrix} \mathbf{W}_s \Phi_s^T (\mathbf{s} - \bar{\mathbf{s}}) \\ \Phi_t^T (\mathbf{t} - \bar{\mathbf{t}}) \end{pmatrix}. \quad (2)$$

Notice that a suitable weighting between pixel distances and pixel intensities is done through the diagonal matrix  $\mathbf{W}_s$ . To recover any correlation between shape and texture a third PC transform is applied

$$\mathbf{b} = \Phi_c \mathbf{c} \quad (3)$$

obtaining the combined appearance model parameters,  $\mathbf{c}$ , that generate new object instances by

$$\mathbf{s} = \bar{\mathbf{s}} + \Phi_s \mathbf{W}_s^{-1} \Phi_{c,s} \mathbf{c} \quad , \quad \mathbf{t} = \bar{\mathbf{t}} + \Phi_t \Phi_{c,t} \mathbf{c} \quad , \quad \Phi_c = \begin{pmatrix} \Phi_{c,s} \\ \Phi_{c,t} \end{pmatrix}. \quad (4)$$

The object instance,  $(\mathbf{s}, \mathbf{t})$ , is synthesised into an image by warping the pixel intensities of  $\mathbf{t}$  into the geometry of the shape  $\mathbf{s}$ . Given a suitable measure of fit the model is matched to an unseen image using an iterative updating scheme based on a fixed Jacobian estimate [10, 11] or a reduced rank regression [2].

## 2.2 Principal Components Transformation

Consider a set of  $P$  texture vectors  $\{\mathbf{t}_i\}_{i=1}^P$  laid out as a set of  $P$  shape-free images with grey levels  $r_i(\mathbf{x})$ ,  $i = 1, \dots, P$ , where  $\mathbf{x}$  is the coordinate vector denoting the grid point of the sample. Let  $\mathbf{r}(\mathbf{x}) = [r_1(\mathbf{x}) \cdots r_P(\mathbf{x})]^T$  and assume first and second order stationarity, i.e.  $E\{\mathbf{r}(\mathbf{x})\} = \mathbf{0}$  and  $D\{\mathbf{r}(\mathbf{x})\} = \Sigma$ . The PC transformation thus chooses  $P$  linear transformations  $z_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{r}(\mathbf{x})$ ,  $i = 1, \dots, P$  such that the variance for  $z_i(\mathbf{x})$  is maximum among all linear transforms orthogonal to  $z_j(\mathbf{x})$ ,  $j = 1, \dots, i - 1$ . The variance in the  $i$ th PC is given by

$$\text{Var}\{\mathbf{a}_i^T \mathbf{r}\} = \lambda_i = \mathbf{a}_i^T \Sigma \mathbf{a}_i. \quad (5)$$

We see that the basis for the PCs is identified as the conjugate eigenvectors of the dispersion matrix. Let  $\lambda_1 \geq \dots \geq \lambda_P \geq 0$  be the eigenvalues with the corresponding conjugate eigenvectors  $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_P]$ .

Above, the PC problem is solved in Q-mode. Using the Eckart-Young's Theorem the R-mode solution becomes  $\Phi_t = \mathbf{R}^T \Lambda^{-1/2} \mathbf{A}$ , where  $\mathbf{R} = [\mathbf{r}_1 \cdots \mathbf{r}_M]$  with  $\mathbf{r}_j$  containing spatially corresponding intensities over the training set, and  $\Lambda$  a diagonal matrix of the eigenvalues.

## 2.3 Minimum Noise Fractions Transformation

Consider the random signal variable  $\mathbf{r}(\mathbf{x})$  from above. Assuming that an additive noise structure applies  $\mathbf{r}(\mathbf{x}) = \mathbf{s}(\mathbf{x}) + \mathbf{n}(\mathbf{x})$  with  $\text{Corr}\{\mathbf{s}(\mathbf{x}), \mathbf{n}(\mathbf{x})\} = 0$ , the dispersion structure can be separated into

$$D\{\mathbf{r}(\mathbf{x})\} = \Sigma = \Sigma_s + \Sigma_n. \quad (6)$$

The Minimum Noise Fractions transformation chooses  $P$  linear combinations  $z_i(\mathbf{x}) = \mathbf{a}_i^T \mathbf{r}(\mathbf{x})$ ,  $i = 1, \dots, P$  which maximise the signal-to-noise ratio (SNR) for the  $i$ th component

$$\text{SNR}_i = \frac{V\{\mathbf{a}_i^T \mathbf{s}(\mathbf{x})\}}{V\{\mathbf{a}_i^T \mathbf{n}(\mathbf{x})\}} = \frac{\mathbf{a}_i^T \Sigma_s \mathbf{a}_i}{\mathbf{a}_i^T \Sigma_n \mathbf{a}_i} - 1. \quad (7)$$

and the problem reduces to solving a generalized eigenproblem,  $\Sigma_s \mathbf{a}_i = \lambda_i \Sigma_n \mathbf{a}_i$ . Let  $\lambda_1 \geq \dots \geq \lambda_P$  be the eigenvalues of  $\Sigma$  with respect to  $\Sigma_n$  with the corresponding conjugate eigenvectors  $\mathbf{a}_1, \dots, \mathbf{a}_P$ . Then  $z_i(\mathbf{x})$  is the  $i$ th MNF. A high

order component has a high noise fraction and thus little signal. A low order component has a high SNR, hence the name Minimum Noise Fraction transform. The central issue in obtaining good MNF components is the estimation of the dispersion matrix of the noise. Using the difference between a pixel and its neighbours as a noise estimate, MNF maximises the spatial autocorrelation. Let  $\mathbf{\Delta}^T = [\Delta_1 \ \Delta_2]$  represent a spatial shift. Introducing  $\mathbf{\Sigma}_{\Delta} = \text{D}\{\mathbf{r}(\mathbf{x}) - \mathbf{r}(\mathbf{x} + \mathbf{\Delta})\}$  which, when considered as a function of  $\mathbf{\Delta}$ , is a multivariate variogram and assuming a proportional covariance model [13] the covariance of the noise can be estimated by  $\mathbf{\Sigma}_n = \mathbf{\Sigma}_{\Delta}/2$ . When the covariance structure for the noise is proportional to the identity matrix, the MNF transform reduces to the PC transform. In [12] several other models are presented for estimating image noise.

When maximising autocorrelation the MNF analysis qualifies as an Independent Components Analysis (ICA) similar to the Molgedy-Schuster algorithm [14], see [5]. A comparative study of the PC and MNF can be found in [15, 16].

### 3 Data

Short-axis, end-diastolic cardiac MRIs were selected from 28 subjects. MRIs were acquired using a whole-body MR unit (Siemens Impact) operating at 1.0 Tesla. The chosen MRI slice position represented low and high morphologic complexity, judged by the presence of papillary muscles. Images were acquired using an ECG-triggered breath-hold fast low angle shot (FLASH) cinematographic pulse sequence. Slice thickness=10 mm; field of view=263x350 mm; matrix 256x256. The endocardial and epicardial contours of the left ventricle were annotated manually by placing 33 landmarks along each contour, see Figure 3.

## 4 Results and Discussion

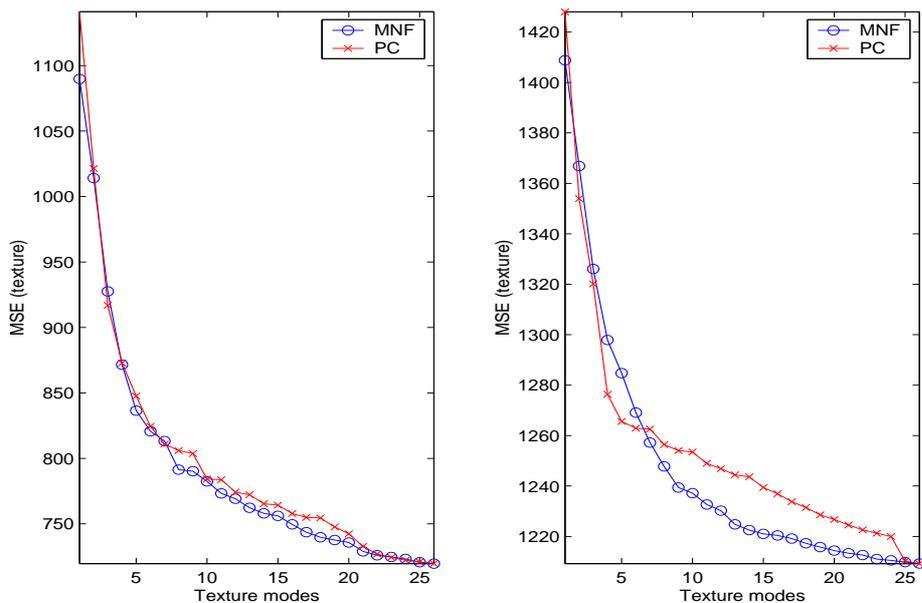
Noise is added to the training data simulating different SNRs, i.e. different quality of the MRIs due to inter-patient, inter-operator variation etc. This is done in order to examine the robustness of the texture representation in the MNF basis compared to the PC basis. Gaussian noise is applied with a standard deviation randomly chosen to produce training images with an SNR down to 6 dB. This knowledge of the noise structure is not used in the subsequent analyses.

### 4.1 Learning Based Image Representation

To examine the robustness of the MNF transform, 101 leave-one-out studies were carried out. One on the uncorrupted and 100 on the noise degraded shape-free sets of 28 MRIs. Results of the cross-validation analyses are presented in Figure 1. The left plot corresponds to uncorrupted MRIs and the right to a randomly chosen analysis on a degraded training set. The curves with  $\circ/x$  symbols marks the performance of the MNF/PC models and provides the mean squared texture error (MSE) as a function of the model rank. For the scenario without the

performance of the MNF and the PC transform is very similar. Notice, however that the MNF is better for almost all number of modes. The general trend for the noise degraded data is reflected in the MSE curves in Figure 1 (right). The MNF and PC are competing for low rank models, but for an intermediate number of modes the MNF outperforms the PC transform. The MNF thus does a better job of separating important signal from noise in the training data.

Figure 2 shows the PC and the MNF eigenvectors (the  $\Phi_i$ 's) of the mean shape aligned 28 noise degraded cardiac data for which the leave-one-out texture representation curve in Figure 1 (right) was generated. All images in Figure 2 are stretched between mean  $\pm 3$  std. The top four rows correspond to the PC eigenvectors and the four bottom rows to the MNFs. The components are ordered row-wise according to the amount of variance/SNR they explain. The last component in both shows the mean texture sample,  $\bar{t}$ . Notice that the MNF gives a better ordering of components in terms of texture quality. A higher degree of speckle noise is present in all PC components compared to the MNF components. Moreover, the last components of the PC analysis appear to include a relatively higher amount of auto-correlated signal. This explains the better performance of the MNF representation in the cross-validation study.



**Fig. 1.** Leave-one-out study on cardiac MRIs. Without noise (left). With noise (right).

## 4.2 Cardiac Segmentation

Hitherto, the PC and the MNF transform have been evaluated w.r.t. representation. To assess the transforms capabilities in a de facto segmentation setting, a cross-validation study was carried out on the cardiac data set.

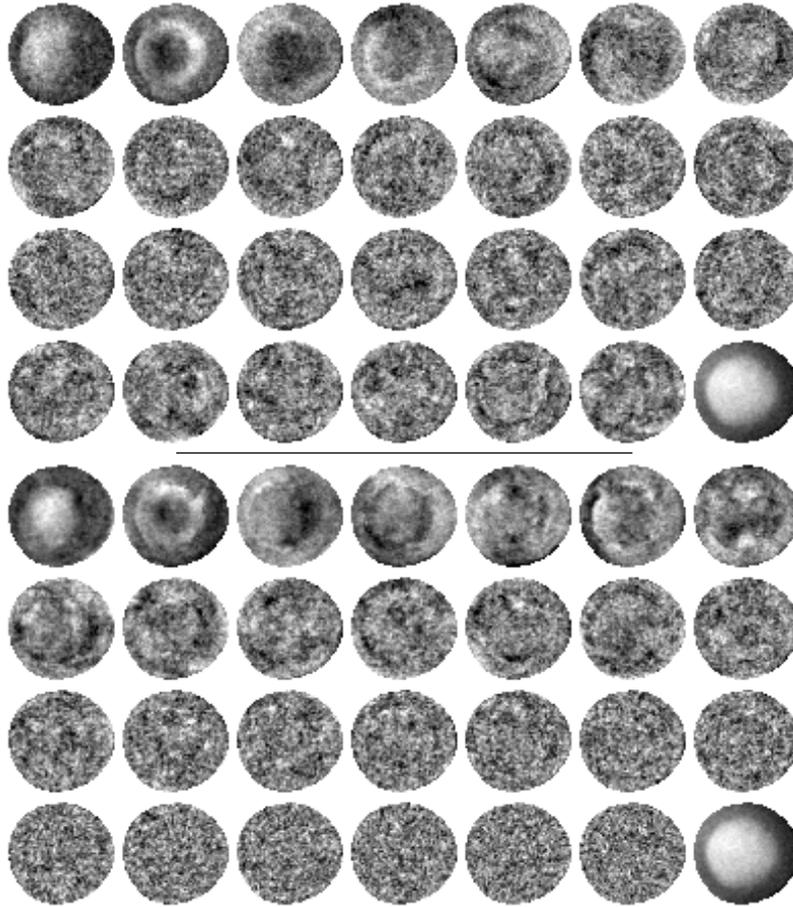


Fig. 2. PC (top) and MNF (bottom) decomposition of noise degraded cardiac MRIs.

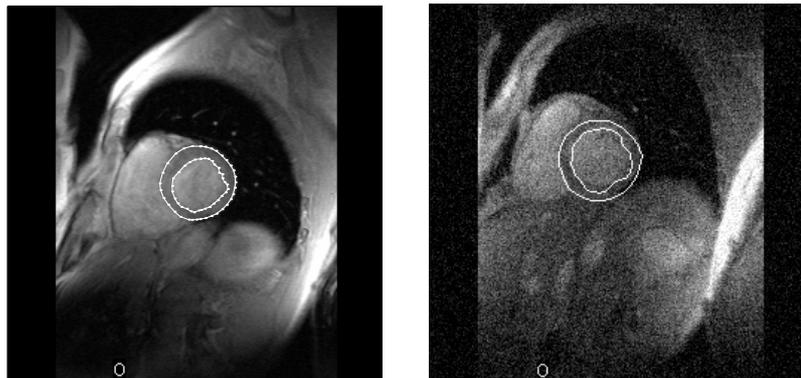


Fig. 3. Example annotation of the left ventricle using 66 landmarks (left). Segmentation result on noise contaminated cardiac MRI (right).

To maximise the effective size of the training set, validation was performed using a leave-one-out evaluation on the set of 28 short-axis cardiac MRIs. A total of 56 AAMs were built on noise-contaminated versions of the 28 cardiac MRIs; i.e. 28 PC AAMs and 28 MNF AAMs. In both transforms the largest 14 texture modes were included into the models. This model rank was chosen as half the maximum basis size producing a cut-off point where an average of 85% of the total amount of variation is explained. Each model was initialised on the image that was left out, in its mean configuration (i.e. mean shape and mean texture) and displaced  $\pm 8$  pixels from the ground truth centre in image coordinates. From this position the AAM search was started. Refer to Figure 3 (right) for an example segmentation.

Two performance measures were evaluated: normalised texture MSE (MSE) and mean point-to-point distance between corresponding landmarks of the model and the ground truth. Segmentations with a pt.-pt. distance larger than ten pixels were deemed outliers and removed. The PC/MNF AAMs yielded a mean normalised MSE of  $3.55 \pm 3.35$  /  $3.43 \pm 2.67$  and a pt.-pt. landmark error of  $5.03 \pm 1.60$  /  $4.79 \pm 1.51$  pixels, respectively. In the two PC/MNF runs 2 / 1 outliers were removed. Thus, a modest improvement in both performance measures and corresponding uncertainties is observed for the MNF AAMs. Notice the rather high MSE standard deviations, due to the large inhomogeneity in the noise characteristics.

## 5 Conclusion

We have shown that a more compact representation of texture can be obtained by extending the PC to the MNF transformation in the AAM framework. The novel approach shows better performance in leave-one-out representation studies both on original and on noise degraded cardiac MRIs. Thus, by separating important signal from noise the MNF transform generalises better than the PC transform.

The MNF texture representation is applied in a leave-one-out AAM segmentation study in comparison to applying a conventional PC basis of equal rank. Even though the MNF extension only affects the texture- and not the shape representation, and the texture model rank is chosen relatively high compared to the amount of noise present in the data; improvements in both landmark and texture error and corresponding uncertainties are observed for the MNF AAMs.

In contrast to the PC analysis, the new approach by maximizing SNR is invariant to linear transformations such as scaling of the individual components in the training set. As a consequence, the MNF decomposition is expected to be useful in future AAM studies involving data fusion of multiple features of different nature measured at different scales. This includes derived physiological measures, textual quantities, and multiple imaging modalities.

Moreover, the MNF analysis in itself can be applied as a data driven method probing for uncorrelated modes of biological variation in non-Euclidean space, and thus constitute a useful tool in exploratory analysis of medical data.

### Acknowledgments

MRIs were provided by M.D., Jens Chr. Nilsson and M.D., Bjørn A. Grønning, Danish Research Centre of Magnetic Resonance, H:S Hvidovre Hospital.

### References

1. Edwards, G., Taylor, C.J., Cootes, T.F.: Interpreting face images using active appearance models. In: Proc. 3rd IEEE Int. Conf. on Automatic Face and Gesture Recognition, IEEE Comput. Soc (1998) 300–5
2. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Proc. European Conf. on Computer Vision. Volume 2., Springer (1998) 484–498
3. Mitchell, S., Lelieveldt, B., van der Geest, R., Bosch, H., Reiver, J., Sonka, M.: Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac mr images. Medical Imaging, IEEE Transactions on **20** (2001) 415–423
4. Stegmann, M.B., Fisker, R., Ersbøll, B.K.: Extending and applying active appearance models for automated, high precision segmentation. In: Proc. 12th Scandinavian Conf. on Image Analysis. Volume 1. (2001) 90–97
5. Larsen, R., Eiriksson, H., Stegmann, M.B.: Q-MAF shape decomposition. In: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2001, 4th International Conference, Utrecht, The Netherlands. Volume 2208 of Lecture Notes in Computer Science., Springer (2001) 837–844
6. Bosch, H., Mitchell, S., Lelieveldt, B., Nijland, F., Kamp, O., Sonka, M., Reiber, J.: Active appearance-motion models for endocardial contour detection in time sequences of echocardiograms. Proceedings of SPIE **4322** (2001) 257–268
7. Jones, M., Poggio, T.: Multidimensional morphable models: a framework for representing and matching object classes. International Journal of Computer Vision **29** (1998) 107–31
8. Green, A.A., Berman, M., Switzer, P., Craig, M.D.: Transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing **26** (1988) 65–74
9. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models – their training and application. Comp. Vision and Image Understanding **61** (1995) 38–59
10. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. on Pattern Recognition and Machine Intelligence **23** (2001) 681–685
11. Cootes, T.F., Taylor, C.J.: Statistical Models of Appearance for Computer Vision. Tech. Report. Oct 2001, University of Manchester (2001)
12. Olsen, S.I.: Estimation of noise in images: An evaluation. Graphical Models and Image Processing **55** (1993) 319–323
13. Switzer, P., Green, A.A.: Min/max autocorrelation factors for multivariate spatial imagery. Technical Report 6, Dept. of statistics, Stanford University (1984)
14. Molgedey, L., Schuster, H.G.: Separation of a mixture of independent signals using time delayed correlations (1994)
15. Nielsen, A.A.: Analysis of Regularly and Irregularly Sampled Spatial, Multivariate, and Multi-temporal Data. PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby (1994)
16. Hilger, K.B.: Exploratory Analysis of Multivariate Data, Unsupervised Image Segmentation and Data Driven Linear and Nonlinear Decomposition. PhD thesis, Informatics and Mathematical Modelling, Technical University of Denmark, Kgs. Lyngby (2001) 186 pp.