

Extraction of Myocardial Contractility Patterns from Short-Axes MR Images Using Independent Component Analysis

A. Suinesiaputra¹, A.F. Frangi², M. Üzümcü¹, J.H.C. Reiber¹, and
B.P.F. Lelieveldt¹

¹ Division of Image Processing, Department of Radiology,
Leiden University Medical Center, Leiden, the Netherlands

² Computer Vision Group, Aragon Institute of Engineering,
University of Zaragoza, Zaragoza, Spain
`a.suinesiaputra@lumc.nl`

Abstract. Regional myocardial wall motion analysis has been used in clinical routine to assess myocardial disease, such as infarction and hypertrophy. These diseases can be distinguished from normals by looking at the local abnormality of cardiac motion. In this paper, we present a first result of a feature extraction experiment using the Independent Component Analysis (ICA), where abnormal patterns of myocardial contraction from patients are recognizable and distinguishable from normal subjects.

1 Introduction

Myocardial contractility is an important quantitative indicator for the diagnosis of myocardial diseases. This function can be visually examined and quantified by using a cine MRI sequence. Two most important phases for myocardial contraction are the *end-diastole* (ED), or the start of contraction, and the *end-systole* (ES), or the end of contraction.

Abnormal myocardial contraction is mainly caused by the occlusion of coronary arteries, particularly in the infarcted regions. Figure 1 shows two examples of MRI images from a healthy volunteer and an infarct patient, both at ES phase. Note that the inferior region (indicated by a white arrow) of the infarct patient does not contract. This region has a small wall thickness value.

Subspace analysis techniques have been used recently in many areas, such as appearance-based modeling and recognition. Principal Component Analysis (PCA) is the common subspace analysis for dimensionality reduction. Independent Component Analysis (ICA) is another subspace analysis, which seeks statistically independent components of the observed data. ICA is commonly used for blind source separation of an observed signal.

In machine learning, both PCA and ICA can be used for feature extraction [1, 2, 3]. There has been some literature showing a comparison between both methods with different results. Moghaddam [4] shows no statistical differences

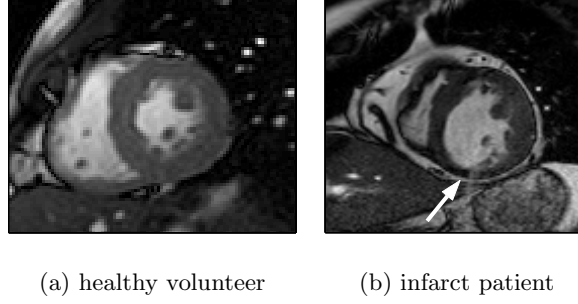


Fig. 1. MRI images of a healthy volunteer and an infarct patient at end-systole (the final contraction phase in the cardiac cycle). White arrow points to the infarcted tissue of the patient, where that myocardium region has a small contraction.

between PCA and ICA. Draper et al. [5] compared ICA and PCA for face recognition and reported that some ICA algorithms give better performance than PCA, but some do not.

Regardless of these comparisons, PCA and ICA are both linear generative models, because every training shape can be approximated by a linear combination of the components. An important difference between ICA and PCA lies in the shape variation. Independent components from ICA create local shape variation, while principal components from PCA give a global shape variation [6]. This indicates that ICA is more suitable for extracting local shape features, than PCA, which is desired in our study.

In this paper, we present an ICA-based local feature extraction method for the diagnosis of myocardial disease, especially for myocardial infarction. Section 2 describes our shape model, the ICA method and a new sorting method for ICA modes. Section 3 presents experimental results, followed by a discussion in Section 4 and some perspective on future work in Section 5.

2 Methodology

2.1 ICA Model

In this study, the observation data are left ventricular (LV) myocardial contours, manually drawn from short-axis cardiac MRI images at ED and ES phases. Samples for each observation are landmark points, defined by equal angular distance along each contour.

To model the contractility pattern between ED and ES, contours for each subject are combined serially into one *shape* vector. A shape $\mathbf{x} \in \mathbb{R}^{2m}$ is defined by m landmark points from 4 contours together in the following order: endocardium (inner) contour at ED, epicardium (outer) contour at ED, endocardium contour at ES and epicardium contour at ES. Thus the shape analysis

is performed on all concatenated contours together, preserving the aspect ratio between ED and ES because of the contraction.

The shape vector $\mathbf{x}$ consists of m pairs of (x, y) coordinates of landmark points:

$$\mathbf{x} = (x_1, y_1, x_2, y_2, \dots, x_m, y_m)^T \quad (1)$$

The mean shape $\bar{\mathbf{x}}$ from n shapes is defined by

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_{i=1}^n \mathbf{x}_i \quad (2)$$

Each observed data (shape) $\mathbf{x}$ can be generated by a linear combination of the independent components $\Phi \in \mathbb{R}^{2m \times p}$. This linear generative model is formulated as follows

$$\mathbf{x} \approx \bar{\mathbf{x}} + \Phi \mathbf{b} \quad (3)$$

where $\bar{\mathbf{x}}$ is the mean shape and $\mathbf{b} \in \mathbb{R}^p$ is the component weighting vector.

In ICA, the basis of the subspace is sought to be statistically independent, with the main assumption of the non-gaussian distribution of the observed data [7]. The resulting subspace is non orthogonal and unordered. There is no closed form solution for ICA. Several numerical algorithms to estimate ICA are available (see [8] for the survey of ICA algorithms).

When applied to shape modeling, there is an important property of ICA in its *modes*. As the number of computed independent components increases, the component gives more localized shape variations. On the contrary, if the number of independent components is too small, then the component gives global shape variation, much like PCA modes. A shape variation in ICA has a general shape of a local bump, whereas the remainder of the shape is unaffected (see Fig. 2(a)). This is the difference between ICA and PCA: PCA modes give global shape variations, distributed over the entire contour (see Fig. 2(b)). Üzümcü et. al. [6] have presented the comparison between PCA and ICA in the modelling of cardiac shapes.

2.2 Geometry-Based Sorting for ICA Modes

In subspace analysis such as ICA or PCA, the number of selected components is less than the number of dimensions of the observed data. This allows a lower dimensional representation that still covers enough information of the observed data, either for description, detection or recognition.

Principal components are ordered from higher variance to the lowest, making it easy to select which and how many components to retain for further analysis; this is however not the case in ICA. There is no natural sorting criteria for independent components. One needs to define a sorting method for independent components that is suitable for a specific application. Since ICA components are local, we can however sort them based on their local position along the contour and this sorting criterion gives a more intuitive interpretation of local shape variations.

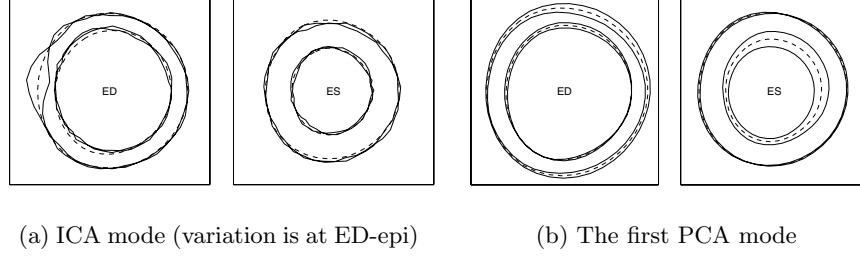


Fig. 2. Examples of local shape variation from an ICA mode. As a comparison, the first PCA mode is given. The mean shape is shown with dashed lines. The solid lines are shape variations, i.e. $\bar{\mathbf{x}} \pm \Phi \mathbf{a}$ (see Eq. 4).

Let i -th mode $\hat{\mathbf{x}}_i$ be defined as the shape variation at the i -th column of Φ :

$$\hat{\mathbf{x}}_i = \bar{\mathbf{x}} + \Phi \mathbf{e}_i \quad (4)$$

where $1 \leq i \leq p$ and $\mathbf{e}_i \in \mathbb{R}^p$ is a vector that has element 1 at the i -th position whereas the rest are 0. Thus, $\hat{\mathbf{x}}_i$ describes the shape variation of the i -th component.

To locate the position of each $\hat{\mathbf{x}}_i$ along a contour, we use a bank of Gaussian filters and perform normalized cross-correlation of each of the filters with a *distance vector* of each mode $\hat{\mathbf{x}}_i$. The distance vector of the i -th mode, $\mathbf{d}_i \in \mathbb{R}_+^m$, is defined as the distance of each landmark point in the shape variation $\mathbf{x}_i$ to the mean shape. Each element j of the i -th distance vector is defined by

$$\mathbf{d}_i^{(j)} = \sqrt{\sum_{k=2j-1}^{2j} \left(\hat{\mathbf{x}}_i^{(k)} - \bar{\mathbf{x}}^{(k)} \right)^2} \quad (5)$$

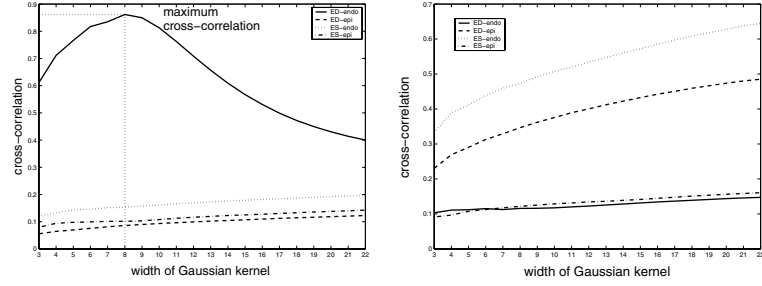
where $j = 1, 2, \dots, m$.

The cross-correlation is performed only on a particular contour, circularly. Thus there are four cross-correlation processes, because there are four contours for each shape.

The Gaussian filter giving the maximum cross-correlation for vector $\mathbf{d}_i$ is stored. The center of this filter defines the position of the i -th component; the width of the Gaussian filter represents the width of the component. Figure 3(a) shows an example of the cross-correlation response from a component.

There is an extra advantage of using the normalized cross-correlation for sorting ICA modes. Modes that consist of noise are automatically detected and thus can be eliminated. Noise modes have a global wrinkled shape variation along the whole contour, which correlates best with the widest Gaussian filter. Figure 3(b) shows an example of the cross-correlation response for a noise component.

After all modes have been cross-correlated, positions of all modes are determined. We sort ICA modes based on position along the contour.



(a) a component

(b) a noise component

Fig. 3. Example of maximum cross-correlation results of two components.

2.3 Cluster Measurement Metrics

To evaluate the cluster formation between normal and patient subjects, a number of q components ($q \leq p$) are selected from the weighting coefficient matrix $\mathbf{b}$.

Let $\mathcal{D} = \sum_i^c \mathcal{D}_i \subset \mathbb{R}^q$ be a subset of the weighting coefficient matrix $\mathbf{b}$, after q components are selected. Let c be the number of classes. In this case, $c = 2$, because there are only two classes, i.e. normals and patients.

The first measurement is called *within-cluster scatter matrix*, which measures the compactness of a cluster. The within-cluster scatter matrix, $\mathbf{S}_W$, is defined as the sum of scatter matrices for each group:

$$\mathbf{S}_W = \sum_{i=1}^c \sum_{\mathbf{x} \in \mathcal{D}_i} (\mathbf{x} - \mathbf{m}_i)(\mathbf{x} - \mathbf{m}_i)^T \quad (6)$$

where $\mathbf{m}_i$ is the mean ("center of gravity") of the cluster i .

A scalar value representing the measurement of the compactness from this metric, is simply its trace. The trace of a scatter matrix accounts for the square of the scattering radius, because it is actually the sum of the variances in each coordinate direction. This scalar value is equal to the sum-of-squared error. Thus one seeks the minimum of this value to get the best representation of a cluster. The compactness measurement, J_W , is defined as

$$J_W = \text{tr}[\mathbf{S}_W] \quad (7)$$

The second measurement is *between-cluster scatter matrix* measurement, $\mathbf{S}_B$, represents how far clusters are separated. It is defined as follows

$$\mathbf{S}_B = \sum_{i=1}^c n_i (\mathbf{m}_i - \mathbf{m})(\mathbf{m}_i - \mathbf{m})^T \quad (8)$$

where n_i is the number of subject in cluster i and $\mathbf{m}$ is the total mean:

$$\mathbf{m} = \frac{1}{n} \sum_{\mathbf{x} \in \mathcal{D}} \mathbf{x} \quad (9)$$

The scalar measurement value of the between-cluster scatter matrix is also its trace:

$$J_B = \text{tr}[\mathbf{S}_B] \quad (10)$$

The within-cluster and between-cluster scatter matrices are mostly used to design cluster validity indices for clustering methods [3]. In this study, we used these measurements to compare the quality of the cluster representation given by PCA and ICA components.

To visualize the cluster distribution, the Fisher discriminant line [3] is calculated and coefficient values from the selected components are projected to the Fisher line. Fisher linear discriminant accounts the ratio between the between-cluster and the within-cluster matrix measurements and it is given by:

$$\mathbf{w} = \mathbf{S}_W^{-1} (\mathbf{m}_1 + \mathbf{m}_2) \quad (11)$$

where $\mathbf{w}$ is a vector with the direction that maximizes the separation between two clusters.

3 Experimental Results

In this study, 42 normal subjects and 47 patients suffering from myocardial infarction were investigated. For each subject, endocardium and epicardium contours of the left ventricle myocardium at ED and ES phases from short-axis MRI were drawn manually by experts.

Contours were resampled to 40 landmarks defined by equi-angular sampling, starting from the intersection point between the lower right ventricular myocardium with the left ventricular myocardium. The total number of landmark points for each shape then becomes 160 points.

The calculation of ICA was performed using the JADE algorithm [9], implemented in Matlab. The number of ICA modes is selected carefully to 40 in this study, that gives enough local shape variations for each of the four contours. If the number is too small, then the shape variations become more global. If the number is too large, then too many local shape variations may occur, which look like noise components.

In the sorting of ICA modes, 20 Gaussian filters are used, ranging from width 3 to 22. Modes correlating with a Gaussian filter, which has width larger than 20 (half of a contour), were considered to be noise. From the original 40 ICA modes, the sorting method retains 35 modes, thus eliminating 5 noise modes.

3.1 Weighting Coefficient Matrix

Figure 4(a) shows the weighting matrix $\mathbf{b}$ of the ICA model that is constructed from shapes of normal subjects and infarct patients. The weighting coefficient matrix contains values that are needed to generate each training shape. These coefficient values are different for each subject. Thus the weighting matrix $\mathbf{b}$ is the most important value for classification purposes.

From Fig. 4(a), a clear difference can be seen distinguishing between the two groups in the endocardium at the ES phase. As a comparison, Figure 4(b) shows the PCA model from the same data. With PCA, the difference between the two groups is less pronounced, only clearly visible from the first component.

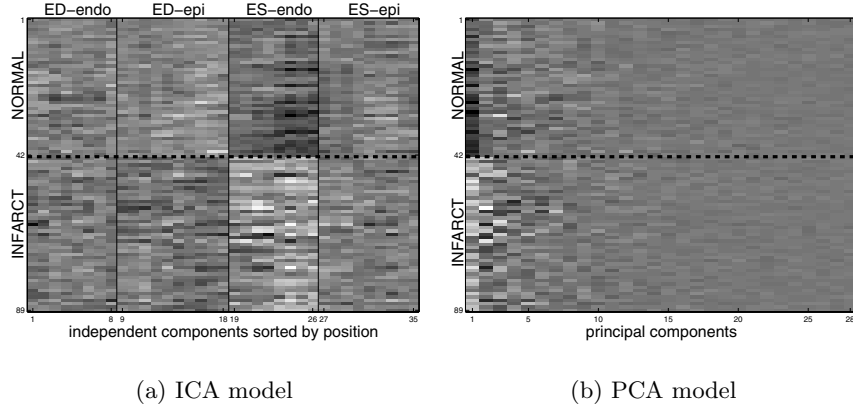


Fig. 4. The weighting coefficient between normal subjects and infarct patients.

3.2 Mean Cluster Distance

To enable the comparison between PCA and ICA, the weighting coefficient matrices for both models are normalized, such that $\|\mathbf{b}\| = 1$. Hence weighting coefficient matrices for PCA and ICA are both in the same unit.

The distance between means of normal and patient subjects for each component is calculated using the *mean cluster distance* (MCD), as given by:

$$d_i = |\mathbf{m}_{n,i} - \mathbf{m}_{p,i}| \quad (12)$$

where i is an index of a component, $m_{n,i}$ and $m_{p,i}$ are the mean of the weighting coefficient values at the i -th component for normal and patient subjects respectively. Figure 5 shows the bar plot of the MCD of PCA and ICA for each mode.

A t-test experiment was conducted on each of independent and principal component to see whether the two means from normal and patient coefficient values come from two different clusters. The result is illustrated in Figure 5. From 35 selected independent components, there are 27 components with each has statistically significant difference of two means, while PCA only gives 1 component (the first principal component). The t-tests were performed with 95% confidence interval.

It is evident that independent components at ES-endo are among the highest MCD value. Mean cluster distance of the first PCA mode is the highest among others, even compared with ICA.

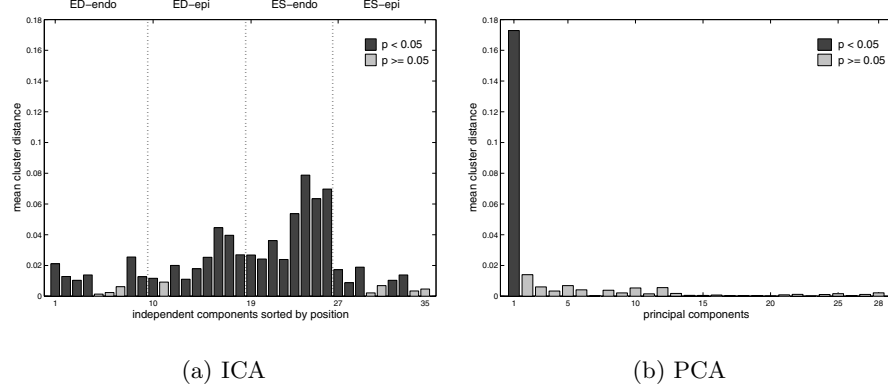


Fig. 5. Mean cluster distance of each component from ICA and PCA. The results of t-test experiment on each component are shown as dark gray for $p < 0.05$ and light gray bars for $p \geq 0.05$.

3.3 Cluster Analysis

In this study, we only present an analysis of cluster properties, but **not yet** a classification result. Clusters are defined by selecting all independent components from ICA and principal components that covers 95% of total variance from PCA. This gives 35 ICA components and 16 PCA components.

Table 1 shows the measurement results using Eq. 7 for the cluster compactness and Eq. 10 for the cluster separation. Figure 6 shows result of the projected coefficient values to their Fisher discriminant line.

Table 1. Cluster validity measurement results.

	compactness (J_W)	separation (J_B)
ICA	1.84	0.66
PCA	0.65	0.12

PCA gives better compactness than ICA, but less separable (see Tab. 1). However the projection to the Fisher discriminant line favors ICA (see Fig. 6). There is only one point of misclassification in ICA, if a threshold value is defined.



Fig. 6. Projection of independent components (above) and principal component (below) to their Fisher discriminant line.

However there are more overlaps in the projection of principal components to the Fisher discriminant line.

3.4 Separation Degree

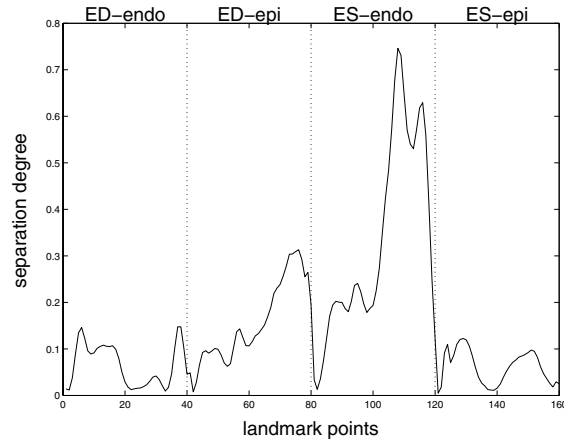
The MCD in Eq. 12 can be used to map the cluster separation for each component onto the same information for each landmark points. This enables a more intuitive regional interpretation of the differences between the two groups.

From the sorting of independent components, location and width of each component are retrieved. Thus we can generate the corresponding Gaussian function for each component and multiply it with its MCD, resulting a Gaussian mixture for each landmark point. The sum of the Gaussian mixture is called *separation degree*. Figure 7(a) shows the separation degree of the ICA model from normal and patient subjects. Figure 7(b) also shows the same visualization, but a more intuitive way using the bullseye plot, where the color denotes the separation degree.

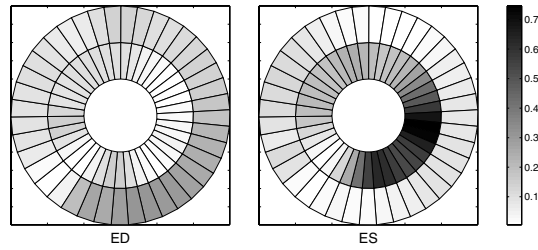
Figure 7(b) corresponds with Figure 4(a), where the most important feature to distinguish between normal and patient is the endocardium at ES phase. The least important features lie on the epicardium contour at ES phase, where there is a small separation degree.

4 Discussion

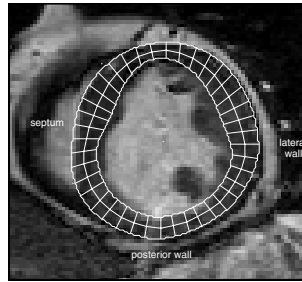
In this paper, we have investigated the potential of ICA in the computer-aided diagnosis of myocardial diseases, based on cardiac shapes. The first result indicates that the ICA method is a promising analysis tool to extract local shape deformations from observed data. The sorting method of independent components based on their position leads to an anatomically meaningful interpretation



(a) Separation degree for each ICA mode



(b) Bullseye plot of the separation degree. The inner wall is endo-contour and the outer is epi-contour.



(c) An illustration of the bullseye plot interpretation in the short-axes MRI.

Fig. 7. The separation degree of the ICA model between normal subjects and infarct patients.

for classification purposes. We have shown that the weighting coefficient matrix from the ICA model can clearly distinguish between the two different groups in the endo-contour at ES.

From the cluster analysis, projection of independent components to the Fisher discriminant line gives better cluster representation than principal components. Given the ability to classify globally and to extract local features, ICA is a powerful tool to detect and to localize shape abnormalities, comparing favorably to PCA.

In our study, we found that most of the infarction area affects the endocardium in the infero-lateral wall, because our data contains most patients who have infarction in the lateral and inferior regions. A few patients have infarction in the septum area. From this study, the endocardium at end-systole phase is the most distinguishable feature, because this is the part of myocardium having the most deformation process due to contraction.

The reason why we do not perform any classification experiment in this study is that the problem of classifying a patient versus normal is a toy problem. In clinical routine, it is not interesting to determine a subject as a patient. It is more important to detect if there is an anomaly, to localize it and then to quantify the disease.

5 Future Work

The number of computed independent components is a free parameter to choose. The smaller the number is, the more global the independent components are for a shape variation. On the other hand, the shape variation becomes more localized if this parameter is increased. Thus a method to find an optimal number of independent components is needed. An analysis of how sensitive this parameter is to the diagnostic performance in this case will be helpful to define the optimal value.

The next important clinical question for the diagnosis of myocardial infarction is at which particular region of myocardium a patient has an infarction. This basically to localize the local abnormality and to quantify the severity of the disease. These are the topics of ongoing research.

Acknowledgement This work is supported by the Dutch Science Foundation (NWO), under the 2nd batch of the innovational research incentive grant, vernieuwingsimpuls 2001. A.F. Frangi is supported by a Ramón y Cajal Research Fellowship and grants TIC2002-04495-CO2 and ISCIII G03/185 from the Spanish Ministries of Science & Technology, and Health, respectively.

References

- [1] Bartlett, M.S., Movellan, J.R., Sejnowski, T.J.: Face Recognition by Independent Component Analysis. *IEEE Trans. on Neural Networks* **13** (2002) 1450–1464

- [2] Lee, T.W., Lewicki, M.S., Sejnowski, T.J.: ICA Mixture Models for Unsupervised Classification of Non-Gaussian Classes and Automatic Context Switching in Blind Signal Separation. *IEEE Trans. on PAMI* **22** (2000)
- [3] Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*. 2 edn. John Wiley & Sons, Inc. (2001)
- [4] Moghaddam, B.: Principal Manifolds and Probabilistic Subspaces for Visual Recognition. *IEEE Trans. on PAMI* **24** (2002)
- [5] Draper, B.A., Baek, K., Bartlett, M.S., Beveridge, J.R.: Recognizing Faces with PCA and ICA. *Computer Vision and Image Understanding* **91** (2003) 115–137
- [6] Üzümcü, M., Frangi, A.F., Reiber, J.H., Lelieveldt, B.P.: Independent Component Analysis in Statistical Shape Models. In Sonka, M., Fitzpatrick, J.M., eds.: *Proc. of SPIE*. Volume 5032. (2003) 375–383
- [7] Hyvärinen, A., Oja, E.: Independent Component Analysis: Algorithms and Applications. *Neural Networks* **13** (2000) 411–430
- [8] Hyvärinen, A.: Survey on Independent Component Analysis. *Neural Computing Surveys* **2** (1999) 94–128
- [9] Cardoso, J., Souloumiac, A.: Blind Beamforming for Non Gaussian Signals. *IEEE Proceedings-F* **140** (1993) 362–370