

# Mutual Interdependence Analysis (MIA)

Heiko Claussen<sup>1,2</sup>, Justinian Rosca<sup>1</sup>, and Robert Dampier<sup>2</sup>

<sup>1</sup> Siemens Corporate Research Inc.,  
755 College Road East, Princeton, New Jersey 08540  
{Heiko.Claussen,Justinian.Rosca}@siemens.com

<sup>2</sup> University of Southampton,  
School of Electronics and Computer Science,  
Southampton SO17 1BJ, UK  
{hc05r,rid}@ecs.soton.ac.uk

**Abstract.** Functional Data Analysis (FDA) is used for datasets that are more meaningfully represented in the functional form. Functional principal component analysis, for instance, is used to extract a set of functions of maximum variance that can represent the data. In this paper, a method of Mutual Interdependence Analysis (MIA) is proposed that can extract an equally correlated function with a set of inputs. Formally, the MIA criterion defines the function whose mean variance of correlations with all inputs is minimized. The meaningfulness of the MIA extraction is proven on real data. In a simple text independent speaker verification example, MIA is used to extract a signature function per each speaker, and results in an equal error rate of 2.9 % in the set of 168 speakers.

## 1 Introduction

Principal Component Analysis (PCA) , discussed in [1] for nonfunctional and [2] for functional data, is an essential feature extraction method. Principal component functions are such that they can approximate new data with minimum mean square error even if only a subset of all components is used. The principal component functions are given by the directions of maximum variance in the projections of the data. The directions of minimum variance, or minor components, have received much less attention in the literature. Minor components correspond to directions that are either orthogonal to the inputs or minimize the variance of their projections. Hence, they represent an invariant or "common" direction of the inputs. Further information about PCA and MCA, including their probabilistic models, can be found in [3], [4], [5] and [6].

[7] acknowledged that minor components are important in some signal processing applications. Examples are spectral estimation, curve and hyper-surface fitting, cognitive perception and computer vision. This work coined the name Minor Component Analysis (MCA) to refer to neural network fitting methods that compute minor components. The authors also suggested that MCA solves the Total Least Square (TLS) problem [8].

Recently, a method called Extreme Component Analysis (XCA) was introduced that combines minor and principal components in order to represent a

dataset "optimally". As discussed in [9], the "optimal" representation, if by minor or principal components, is dependent on the dataset. By freely choosing between minor, principal components and their combinations, datasets can be represented with possibly a lower number of components.

As previously mentioned, MCA can be used to extract minor components that represent invariants of the data. However, when it is desired to extract the mutual interdependencies of a set of input functions, MCA suffers because of the necessary preprocessing step of input data centering. In the usual case, where the input functions are linearly independent, centering reduces the span of the data. Therefore, these centered functions can no longer fully represent the inputs. Furthermore, as discussed in [10] the TLS and therefore the MCA solution is non trivial and can usually not be found in closed form.

In this paper, we propose Mutual Interdependence Analysis (MIA) to solve a TLS-like optimization problem in the span of the original inputs in order to extract a mutually interdependent function from the input function set. We prove that in the case of linearly independent input functions, a closed form solution can be found that minimizes the MIA criterion.

In section 2 we define the MIA problem, derive its solution and illustrate its properties. In the experimental section 3, MIA is used for simple text independent speaker recognition. We end the paper with conclusions and directions for further work.

## 2 Mutual Interdependence Analysis (MIA)

Throughout the paper, we use  $\mathbf{x}^T$  to denote the transpose of column vector  $\mathbf{x}$ . Also, we denote  $\mathbf{X}$  to be the matrix whose columns are  $\mathbf{x}_i$  with  $i = 1, \dots, D$ . We use  $\underline{\mathbf{1}}$  to represent a vector of ones,  $\underline{\underline{\mathbf{1}}}$  to represent a matrix of ones and  $I$  to be the identity matrix. The dimension will be clear from the context.

Consider  $D$  real inputs,  $x_i(t_j)$  with  $i = 1, \dots, D$  and  $j = 1, \dots, N$ , where each input  $\mathbf{x}_i = [x_i(t_1), \dots, x_i(t_N)]^T$  is viewed as a single entity (i.e. has the intrinsic structure of a function) rather than a series of individual observations. Functional Data Analysis (FDA) normally treats data this way, therefore we refer to each  $\mathbf{x}_i$  as an input. In our case,  $N$  is typically larger or much larger than  $D$ . TLS solves a linear equation  $\mathbf{X}^T \cdot \mathbf{s} = \mathbf{b}$  by finding a direction  $\mathbf{s}$  that minimizes the squared, orthogonal distances to the data points  $\mathbf{x}_i$ :  $\min_{\mathbf{s}} \sum_{i=1}^D \frac{|\mathbf{x}_i^T \cdot \mathbf{s} - b_i|^2}{\mathbf{s}^T \cdot \mathbf{s} + 1}$ . On the other hand, the ordinary Least Square (LS) problem finds a direction  $\mathbf{s}$  that minimizes  $\min_{\mathbf{s}} \sum_{i=1}^D |\mathbf{x}_i^T \cdot \mathbf{s} - b_i|^2$ . While LS assumes that only vector  $\mathbf{b}$  to contain errors, the TLS approach models uncertainties in both  $\mathbf{b}$  and in the data  $\mathbf{X}$ . Our goal is different: Extract a new vector (function)  $\mathbf{s}$  that "optimally" represents the  $D$  input functions. An optimal solution is a function that is maximally correlated with all inputs with the constraint that it is a linear combination of the inputs. We call this problem Mutual Interdependence Analysis (MIA). Formally, the optimality criterion  $J(\mathbf{X}|\mathbf{s})$ , given a functional data series  $\mathbf{s}$ , is as follows:

$$J(\mathbf{X}|\mathbf{s}) = \sum_{i=1}^D \left( \mathbf{s}^T \cdot \mathbf{x}_i - \frac{1}{D} \sum_{k=1}^D \mathbf{s}^T \cdot \mathbf{x}_k \right)^2 \quad (1)$$

Our problem is to find the maximum likelihood vector  $\hat{\mathbf{s}}$  of norm one, in the span of the inputs, that minimizes  $J(\mathbf{X}|\mathbf{s})$ :

$$\hat{\mathbf{s}} = \underset{\mathbf{s}, \|\mathbf{s}\|=1, \mathbf{s}=\sum_{k=1}^D c_k \cdot \mathbf{x}_k}{\operatorname{arg\,min}} J(\mathbf{X}|\mathbf{s}) \quad (2)$$

## 2.1 Solution to MIA

Let us find an equivalent formulation for the MIA problem in (2). Consider the mean function  $\mathbf{x}^{(m)} = \frac{1}{D} \sum_{i=1}^D \mathbf{x}_i$  and the centered functions  $\mathbf{x}_i^{(c)} = \mathbf{x}_i - \mathbf{x}^{(m)}$ . It can be easily shown that  $\mathbf{s}^T \cdot \mathbf{x}_i - \frac{1}{D} \sum_{k=1}^D \mathbf{s}^T \cdot \mathbf{x}_k = \mathbf{s}^T \cdot \mathbf{x}_i^{(c)}$ . Hence, (2) becomes:

$$\hat{\mathbf{s}} = \underset{\mathbf{s}, \|\mathbf{s}\|=1, \mathbf{s}=\sum_{k=1}^D c_k \cdot \mathbf{x}_k}{\operatorname{arg\,min}} \left\| \mathbf{s}^T \cdot \mathbf{X}^{(c)} \right\|^2, \quad (3)$$

where  $\mathbf{X}^{(c)}$  has as columns  $\mathbf{x}_i^{(c)}$  with  $i = 1, \dots, D$ . Then,  $\mathbf{x}^{(m)} = \frac{1}{D} \mathbf{X} \cdot \mathbf{1}$  and  $\mathbf{X}^{(c)} = [\mathbf{x}_1 - \mathbf{x}^{(m)} | \dots | \mathbf{x}_D - \mathbf{x}^{(m)}] = \mathbf{X} - \mathbf{x}^{(m)} \cdot \mathbf{1}^T$ . It follows that

$$\mathbf{X}^{(c)} = \mathbf{X} \cdot \mathbf{P} \quad \text{with} \quad \mathbf{P} = \mathbf{I} - \frac{1}{D} \mathbf{1} \mathbf{1}^T. \quad (4)$$

Obviously,  $\sum_{i=1}^D \mathbf{x}_i^{(c)} = \mathbf{0}$ . Hence, the nullspace  $\mathcal{NUL}\mathcal{L}(\mathbf{x}_1^{(c)}, \mathbf{x}_2^{(c)}, \dots, \mathbf{x}_D^{(c)})$  is non trivial. All vectors  $\mathbf{s} \in \mathcal{NUL}\mathcal{L}(\mathbf{x}_1^{(c)}, \mathbf{x}_2^{(c)}, \dots, \mathbf{x}_D^{(c)})$  will minimize  $J(\mathbf{X}|\mathbf{s})$ . In the next theorem, we show that the problem given by (3) has at least one solution.

**Theorem 1.** *Assume  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_D$  are linearly independent. Then, there exists  $\mathbf{s} \neq \mathbf{0}$  in  $\mathcal{NUL}\mathcal{L}(\mathbf{x}_1^{(c)}, \mathbf{x}_2^{(c)}, \dots, \mathbf{x}_D^{(c)})$  such that  $\mathbf{s}$  is in the span of the inputs  $\mathbf{x}_i$ ,  $i = 1, \dots, D$ .*

*Proof.* A solution  $\mathbf{s} \neq \mathbf{0}$  and  $\mathbf{c} = [c_1, c_2, \dots, c_D]^T \in \mathbb{R}^D$  of the system of equations:

$$\mathbf{s}^T \cdot \mathbf{X}^{(c)} = \mathbf{0} \quad (5)$$

$$\mathbf{s} = \mathbf{X} \cdot \mathbf{c} \quad (6)$$

will also satisfy the theorem and solve the optimization criterion of problem (3). Indeed, (5) is equivalent to the existence of  $\mathbf{s}$  such that  $\mathbf{s} \in \mathcal{NUL}\mathcal{L}(\mathbf{x}_1^{(c)}, \mathbf{x}_2^{(c)}, \dots, \mathbf{x}_D^{(c)})$  and (6) specifies that  $\mathbf{s}$  is in the span of the inputs  $\mathbf{x}_i$  and  $\mathbf{s} \neq \mathbf{0}$ . Let us substitute  $\mathbf{s}$  from (6) and  $\mathbf{X}^{(c)}$  from (4) in (5):

$$\Rightarrow \mathbf{c}^T \cdot \mathbf{X}^T \cdot \mathbf{X} \cdot \mathbf{P} = \mathbf{0} \quad (7)$$

Given that  $\mathbf{G} = (\mathbf{X}^T \cdot \mathbf{X})$  is a Gram matrix formed by linearly independent vectors  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_D$ ,  $\mathbf{G}$  is invertible (see theorem 7.2.10 in [11]). Let

$$\mathbf{c}^T = \mathbf{d}^T \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1}. \quad (8)$$

Therefore, (7) becomes:  $\mathbf{d}^T \cdot \mathbf{P} = \mathbf{0}$  with

$$\mathbf{d} = \zeta \mathbf{1} \quad (9)$$

and  $\zeta \in \mathbb{R}$ . When substituting (9) into (8):  $\mathbf{c} = \zeta (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{1}$ . Hence:

$$\hat{\mathbf{s}} = \zeta \mathbf{X} \cdot (\mathbf{X}^T \cdot \mathbf{X})^{-1} \cdot \mathbf{1}. \quad (10)$$

Then,  $\frac{\hat{\mathbf{s}}}{\|\hat{\mathbf{s}}\|}$  is a solution to (6) and (5) for all  $\zeta \in \mathbb{R}$ . Therefore, any MIA problem with linearly independent inputs has a solution given by (10). An alternative interpretation of the MIA solution  $\frac{\hat{\mathbf{s}}}{\|\hat{\mathbf{s}}\|}$  is that of a direction in the  $N$ -dimensional space which minimizes the variance of the projections of all points  $\mathbf{x}_i$ ,  $i = 1, \dots, D$ .  $\square$

Should the inputs be translated by a constant  $\lambda \in \mathbb{R}$ , i.e.

$$\mathbf{x}'_i \leftarrow \mathbf{x}_i - \lambda \mathbf{1} \quad \forall i \quad (11)$$

then the solution of  $J(\mathbf{X}'|s)$  changes. Indeed, it can be easily proven that the criterion (1) itself is invariant to a translation (11), however the constraint of (2) will require that the solution  $\hat{\mathbf{s}}$  is in the span of  $\mathbf{x}'_i$ .

## 2.2 Example with Synthetic Data

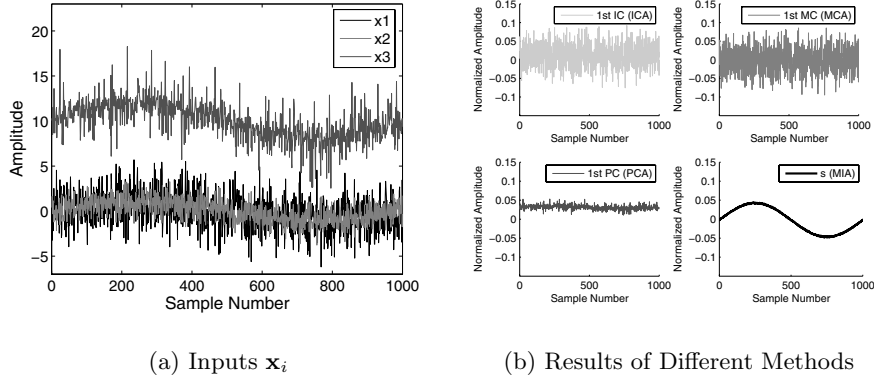
A synthetic example is given below to compare MIA, MCA, PCA and Independent Component Analysis (ICA) (see [12]). Assume three inputs given by:

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{f}_1 + \mathbf{f}_2 \\ \mathbf{x}_2 &= \mathbf{f}_1 - 0.5 \mathbf{f}_2 \\ \mathbf{x}_3 &= 2 \mathbf{f}_1 + 2 \mathbf{f}_3 + 10 \mathbf{1} \end{aligned}$$

where  $\mathbf{f}_1 = \sin(\frac{2\pi i}{N})$   $i = 1, \dots, N$ ,  $\mathbf{f}_2$  is Gaussian noise  $\mathcal{N}(0, 1)$  and  $\mathbf{f}_3$  is Laplacian noise. The inputs and results of the methods are illustrated in Fig. 1. The MIA solution closely approximates  $\mathbf{f}_1$ , in contrast to the other methods.

## 3 Application: Text Independent Speaker Verification

In this section, we apply MIA to the problem of extracting signatures from speech data for the purpose of text independent speaker recognition. This problem is challenging when we need to verify the identity of a person but can not control the way data is acquired (i.e. recording equipment, environment, etc.). For this study, we have used the TIMIT database [13]. Data from 168 speakers was partitioned 50-50 for training and testing. The data was preprocessed by silence removal and normalization of each recording. Data for a given speaker was used as input to MIA in order to generate a speaker signature as described below. We compare the equal error rate (EER) results for speaker verification obtained with MIA versus the PCA and ICA-based methods described in [14].



**Fig. 1.** (a) Inputs  $\mathbf{x}_i$  are linear combinations of three basis functions. (b) Signals extracted using ICA, MCA, PCA and MIA. The MIA result, with  $\lambda = 0$ , is meaningful.

### 3.1 Data Model

A speech signal can be modeled as an excitation that is convolved with a linear dynamic filter which represents the vocal tract. The excitation signal can be modeled for voiced speech as a periodic signal and for unvoiced speech as random noise. It is common to analyze the voiced and unvoiced speech separately [15]. In this example, only the voiced speech is used for speaker recognition. Let  $\mathbf{E}^{(p)}$ ,  $\mathbf{H}^{(p)}$  and  $\mathbf{V}^{(p)}$  be the spectral representations of the excitation, vocal tract filter and the voiced signal parts of a person  $p$  respectively. Moreover, let  $\mathbf{M}$  represent speaker independent signal parts in the spectral domain (i.e. recording equipment, environment, etc.). Therefore, the data can be modeled as:

$$\mathbf{V}^{(p)} = \mathbf{E}^{(p)} \cdot \mathbf{H}^{(p)} \cdot \mathbf{M}. \quad (12)$$

By cepstral deconvolution, the model can be represented as a linear combination of its basis functions:

$$\mathbf{x}_i^{(p)} = \log \mathbf{V}^{(p)} = \log \mathbf{E}^{(p)} + \log \mathbf{H}^{(p)} + \log \mathbf{M}. \quad (13)$$

This model suggests that we could use MIA to extract a function that represents the speaker's signature. In practice, we take speech segments of about one second as MIA inputs  $\mathbf{x}_i^{(p)}$  in order to achieve spectral accuracy. An example of inputs  $\mathbf{x}_i^{(p)}$  is shown in Fig. 2(a). Therefore, MIA will extract signatures that capture typical speaker dependent correlations in the logarithmic spectral domain. Speaker independent signal parts  $\mathbf{M}$  will be minimized if they are not equally present in all MIA inputs.

### 3.2 MIA-Based Text Independent Speaker Verification

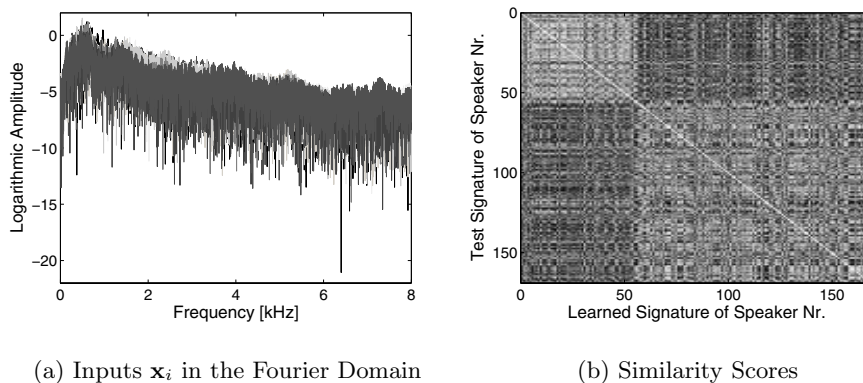
We partition the training and testing data for each person  $p$  into  $D = 8$  segments  $\{x_i^{(p)}\}_{i=1, \dots, D}$  of one second. For each person, we extract a voice signature  $s^{(p)}$

using MIA. The cosine distances between the testing data signatures and training data signatures is used as a measure of similarity. A matrix that represents the cosine distances between all signatures in the database is illustrated in Fig. 2(b).

Let  $N_{FA}$  and  $N_{CA}$  represent the number of false and correct acceptances respectively.  $N_U$  is the number of registered users. The speaker recognition results are evaluated by a comparison of the false acceptance rate FA with the false rejection rate FR, calculated as follows:

$$FA = \frac{N_{FA}}{N_U^2 - N_U} \quad \text{and} \quad FR = \frac{N_U - N_{CA}}{N_U}. \quad (14)$$

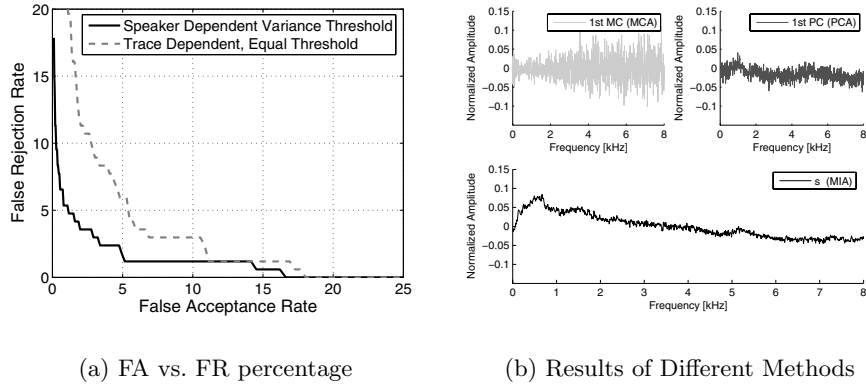
$N_{FA}$  and  $N_{CA}$  result from the entire database by testing each learned speaker signature against all test speaker signatures. This means that a number of  $N_U^2$  tests is performed, including  $N_U$  correct combinations and  $N_U^2 - N_U$  impostors. By changing a threshold value, more people can be accepted which results in a trade off between FA and FR. The FA versus FR plot of this example is illustrated in Fig. 3(a). The equal error rate (EER), where FA equals FR, is used to compare to previous results in [14].



**Fig. 2.** MIA applied to text independent speaker recognition. (a) Representation of input data  $\{\mathbf{x}_i^{(p)}\}_{i=1,\dots,8}$ , given by speech segments of a single speaker  $p$  in the logarithmic Fourier domain. (b) Matrix of similarity scores between different signatures. Bright gray stands for high and dark gray for low similarity between signatures.

### 3.3 Results

We used a set of 168 speakers from the TIMIT database [13]. For each signature extraction from the training and testing data, we used 8 seconds of voiced speech. The data was partitioned into 8 windows with non overlapping, nearly rectangular windowing functions of one second lengths and Gaussian tails of  $\frac{1}{20}$  second. The input functions had their mean subtracted. Thereafter, each extracted signature was down-sampled to 256 points. The mean signature was subtracted



**Fig. 3.** Results of MIA-based text independent speaker verification. (a) False rejection (FR) versus false acceptance (FA) rate. (b) Speaker signature extracted by MIA from  $\{\mathbf{x}_i^{(p)}\}_{i=1,\dots,8}$ . Also plotted are the first minor component (MCA) and the first principal component (PCA) of the data.

from all signatures to focus on the evaluation of differences. The comparison of the MIA-based signature with the 1st minor component and the 1st principal component is illustrated in Fig. 3(b). Note visually that only MIA extracts signal amplitudes in accordance with the well known result that low frequencies contain most information about a speaker. In order to alter the proportion between FA and FR, two different thresholds were used. First, a trace dependent threshold treats every user the same. Secondly, the variance dependent threshold uses information about the similarity between test cases to learn a speaker dependent weighting. The EER of this MIA-based text independent speaker recognition system was 2.9 % for the variance dependent threshold and 5.4 % for a trace dependent threshold. For a similar experiment, [14] reports EER's between 4.3 % and 6.1 % using ICA and PCA features. Here it has to be noted that this test was done with 462 speakers. However, one person was represented by 16 or 32 features of 128 or 256 samples length. On the other hand, MIA only uses a single signature of 256 samples length per speaker.

## 4 Conclusion

We proposed a novel feature extraction method, Mutual Interdependence Analysis (MIA), which finds an invariant function of the input function set, representing the direction of minimum variance of input projections. Intuitively, this function is mutually interdependent with all inputs. The proof of the minimization problem exploits the unconstrained span of the original input data to infer a closed form solution. Furthermore, we showed the effect of input data translation by a constant value  $\lambda$ . In this way, one can control the degree of correlatedness with outlier functions in the input set. Indeed, one can choose a value of  $\lambda$  to

discriminate between inputs and bias towards a result which correlates only with a subset of them. Further work will analyze the robustness of MIA to noise. Moreover, the effect of changes in the span constraint, or the choice of a basis function set, will be explored.

## Acknowledgments

We would like to thank Radu Balan for his invaluable input to this paper.

## References

1. Jolliffe, I.T.: *Principal Component Analysis*, 2nd edn. Springer, Heidelberg (2002)
2. Ramsay, J.O., Silverman, B.W.: *Functional Data Analysis*, 2nd edn. Springer, Heidelberg (2006)
3. Oja, E.: Principal components, minor components, and linear neural networks. *Neural Networks* 5, 927–935 (1992)
4. Tipping, M.E., Bishop, C.M.: Probabilistic principal component analysis. *Journal of the Royal Statistical Society* 61(3), 611–622 (1999) (Series B)
5. Williams, C., Agakov, F.: Products of gaussians and probabilistic minor components analysis. *Neural Computation* 14(5), 1169–1182 (2002)
6. Thompson, P.A.: An adaptive spectral analysis technique for unbiased frequency estimation in the presence of white noise. In: *Systems and Computers*, pp. 529–533 (1980)
7. Xu, L., Oja, E., Suen, C.Y.: Modified hebbian learning for curve and surface fitting. *Neural Networks* 5, 441–457 (1992)
8. Moon, T.K., Stirling, W.C.: *Mathematical methods and algorithms for signal processing*. Prentice-Hall, Upper Saddle River, NJ (2000)
9. Welling, M., Agakov, F., Williams, C.K.I.: Extreme components analysis. In: *Thrun, S., Saul, L., Schölkopf, B. (eds.) Advances in Neural Information Processing Systems 16*, MIT Press, Cambridge, MA (2004)
10. Rao, Y.N., Principe, J.: Efficient total least squares method for system modeling using minor component analysis. In: *Proceedings of international workshop on neural networks for signal processing*, pp. 259–268 (2002)
11. Horn, R.A., Johnson, C.R.: *Matrix Analysis*. Cambridge University Press, Cambridge (1999)
12. Hyvärinen, A., Karhunen, J., Oja, E.: *Independent component analysis*. John Wiley and Sons, Chichester (2001)
13. Garofolo, J.S., Lamel, L.F., Fisher, W.M., Fiscus, J.G., Pallett, D.S., Dahlgren, N.L., Zue, V.: *Timit acoustic-phonetic continuous speech corpus*. CDROM (1993)
14. Rosca, J., Kofmehl, A.: Cepstrum-like ica representations for text independent speaker recognition. In: *ICA*, pp. 999–1004 (2003)
15. Deng, L., O’Shaughnessy, D.: *Speech Processing: A Dynamic and Optimization-Oriented Approach*. Signal Processing and Communications. Marcel Dekker, Inc. (2003)