OPTIMAL PERMUTATION CORRECTION BY MULTIOBJECTIVE GENETIC ALGORITHMS

Dorothea Kolossa, Bert-Uwe Köhler, Markus Conrath, Reinhold Orglmeister

Berlin University of Technology, Institute of Electronics, Einsteinufer 17, 10587 Berlin, Germany

ABSTRACT

The separation of convolved signals can be performed in the frequency domain by splitting the signal into frequency bands, applying ICA techniques to each band and reassembling all bands to obtain the separated output signals. As ICA cannot guarantee for any ordering of output signals, the reassembly of the frequency bands must use some additional information in order to assign the frequency bands consistently to the right output. A number of different criteria have been proposed to solve this permutation problem. We have applied genetic algorithms at this point, which have the advantage of coping well with the discrete, multimodal search space that characterizes this problem. Furthermore, applying a multiobjective genetic algorithm allows to take more than one criterion for optimal separation into account simultaneously. This can give an insight into the relative merit of various criteria of separation optimality.

1. INTRODUCTION

While the separation of instantaneous mixtures of signals can be performed satisfactorily by a number of different methods, e.g. information maximization or minimization of cross cumulants [1,2,3], the separation of convolved mixtures remains a challenging problem.

One often applied approach is to split the signals into a number of frequency bands, separate each frequency band by those methods applicable to additive mixtures, and to reassemble the thus separated frequency bands of the signals [4,5,6,7,8,9]. However, this reassembly still poses a problem, as source separation cannot guarantee for consistent assignment of the same sources to the same output channels, so that criteria are required to avoid permutations between frequency bands.

There are a number of different criteria which all capture some aspect of the correct permutation, for example amplitude modulation correlation, which is especially suitable for speech signal separation, or flatness criteria, which restrict the allowed variability of the demixing filter that is effectively composed by calculating different demixing matrices for different frequencies. However, each of the proposed criteria can only capture one aspect of the optimal permutation, therefore we propose an algorithm that pays regard to a number of different criteria simultaneously.

Since the relative merit of these criteria is not known beforehand, multiobjective genetic algorithms were used, thus leaving the relative weighting of the criteria open until the very end of optimization.

This also allows evaluating the relative merit of criteria by calculating the objective performance of all multiobjective optima in test cases with artificially introduced permutations.

1.1 Test Setup

The approach of using genetic algorithms for permutation correction was tested by artificially permuted signals for the case of two sources. In a first stage of processing, signals are windowed and an FFT is applied to split them into bands and an artificial random permutation is introduced. In the optimization loop, bands are reassembled according to a bit-sequence, in which there is one bit for each frequency that decides whether the bands are to be used in their given order or whether they are to be exchanged. Using this reassembled speech, the criteria for separation quality are calculated, some in the time and some in the frequency domain, which are then used in the genetic algorithm to decide on the next generation of bit-sequences. The performance of this algorithm was tested in a single-objective optimization for four criteria and all combinations of two of these criteria were used in the multiobjective algorithm.

1.2 Genetic Algorithms

Genetic algorithms are a means of stochastic optimization, which is especially useful for the permutation problem due to a couple of advantages:

- As is generally the advantage of stochastic optimization, genetic algorithms do not get trapped in local optima when their parameters are set appropriately. This is especially useful for permutation correction, as the criteria all have strong local minima.
- Genetic algorithms show quicker convergence than other stochastic optimization algorithms, especially for high-dimensional problems [10].
- Genetic algorithms are especially useful for discrete optimization, as they naturally use binary strings to code solutions.

- An efficient algorithm for multiobjective genetic optimization exists, which allows an entire set of multiobjectively optimal solutions to be evolved simultaneously [11].

For a detailed explanation of genetic algorithms, the reader is referred to [10] and [12]. The basic idea is taken from nature's means of "optimizing" its species – a population of individuals is left to fight for survival; the fittest individuals survive and get a chance to recombine their genetic information and pass them on to the next generation, subject to some additional mutations.

This natural optimization is modeled by genetic optimization – solutions are coded either in the form of a binary string (as it was done here, by coding frequency bins to be exchanged by a "1" and those to be left in the given order by a "0") or by vectors of floating point values. A number of solutions ("the population") is evaluated, and those solutions leading to the best objective values are selected to form the next population. For this, individual solutions are combined by randomly splitting the binary strings and recombining them, and some "mutations" are introduced by flipping a small number of bits.

2. MULTIOBJECTIVE OPTIMIZATION

For the problem of signal separation, a number of different optimality criteria have been proposed. These will be described in Section III. Each of these criteria have been used separately to perform source separation, but since each criterion only captures one aspect of optimal separation quality, an adequate combination of the criteria should prove more useful for permutation correction than the application of just one criterion.

2.1 Pareto Optimality

One widely accepted definition for optimality with respect to multiple criteria was introduced by V. Pareto at the beginning of the previous century, and is still widely accepted today.

One individual is considered superior to another in the multiobjective sense, if it is better regarding at least one criterion and at least equally good with regard to all other criteria. The inferior individual is then said to be "dominated" by the other.

An individual is considered "Pareto-optimal", if it is not dominated by any other individual - which means that no other individual exists, which improves upon one criterion without allowing another criterion to get worse.

To determine the standing of one individual of a population, it is necessary to count by how many individuals it is dominated - i.e. how many individuals are better in the Pareto-sense. For performing the selection in a multiobjective genetic algorithm, the standing or "Pareto-rank" of all individuals is determined in this way, and the resulting Pareto-rank is used to select the next generation in an appropriate random process. Figure 1 shows individuals of a 2-criterion multiobjective genetic algorithm with their corresponding Pareto-rank.

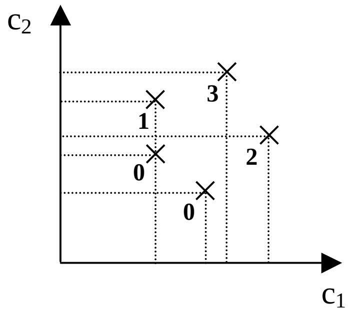

Fig1. Population of multiobjective genetic algorithm with corresponding Pareto-rank, where c_1 and c_2 are the two criteria considered.

However, in such multiobjective genetic algorithms, individuals tend to concentrate on one small part of the Pareto-surface (the set of all Pareto-optimal solutions to the problem) [13]. This problem, which is known as genetic drift, is also observed in single-objective genetic algorithms faced with a number of equivalent optima.

A number of approaches exist to counteract this problem, most notably fitness sharing, which penalizes individuals in "crowded" areas of the search space [13]. In our implementation, a variant described in [14] was employed, which calculates penalties in such a way that crowding is avoided but that any Pareto-optimal individual will remain higher in rank than any non-optimal individual after penalties are added.

3. CRITERIA

3.1 Overview

There are two common approaches to solve the permutation problem. One approach seeks for that permutation which maximizes the statistical dependencies between the time courses of the different frequency bands within the spectrogram of each recovered source signal [4,8,9,15]. The other approach tries to make the impulse response of the mixing filters finite, which corresponds to a smooth frequency response of the mixing filters [5,7].

Both approaches are based on the same assumption, i.e. the spectral smoothness of the recovered source signals. This assumption is well-justified for real-world audio signals.

With the first approach, the spectral smoothness assumption is translated into the evaluation of the dependencies between time courses of different frequency components of each source signal. The second approach considers the (multiplicative) impact of the mixing filters on the spectrum of the recovered signals and requires smooth frequency responses of the mixing filters.

For this investigation, three of the formerly proposed permutation correction criteria [7,9,15] were chosen, as well as a simple time domain dependency criterion for comparison. All four criteria are explained in detail below.

3.2 Cross-cumulants between adjacent frequency bands

The criterion proposed in [9] follows the first approach, i.e. the direct evaluation of the spectral smoothness of the recovered sources. The spectral smoothness is measured in terms of the dependencies between different frequency bands of the spectrogram by means of fourth order cross-cumulants:

$$\begin{aligned} Cum(Y_m(i),Y_l(j)) = \; & E(|Y_m(i)|^2 |Y_l(j)|^2) - \\ & E(|Y_m(i)|^2) E(|Y_l(j)|^2) - \\ & E(|Y_m(i)Y_l^*(j)|)^2 - \\ & E(|Y_m(i)^* Y_l(j)|)^2 \end{aligned} \quad (1)$$

where $m,l = 1..N$, N is the number of sensors and $i,j = 1..L$, with L being the number of frequency bins, and an asterisk denotes the complex conjugate of a signal.

For statistically independent signals, the cross-cumulants are zero. Dependent frequency components have non-zero cross-cumulants. Based on this property, a permutation correction is possible. Problems may occur in case of short signals where the estimation of cross-cumulants is difficult.

3.3 Amplitude Modulation Correlation

A related criterion, considering the amplitude modulation correlation (Amcor) structure of speech, is proposed in [15]. First, the covariances of the amplitudes are computed for all pairs of frequency bands:

$$c(X_i ,Y_j) \equiv E(|X_i (t')| - E\{|X_i (t')|\}, |Y_j (t')| - E\{|Y_j (t')|\}) \quad (2)$$

where $X_i (t')$ and $Y_j (t')$ are the short time spectra of the two signals $x(t)$ and $y(t)$ in the frequency bands i and j, respectively.

Then, all these covariance values, which show the extent to which amplitude modulation in the two channels at the two considered frequencies is correlated, are integrated into the criterion

$$H = \sum_{i,j \neq i} c(X_i, Y_j). \quad (3)$$

According to [15], the amplitude modulation is a distinct criterion for blind source separation, local permutations are penalized by the cost function, and, because a high number of constraints is imposed, the algorithm achieves a good separation quality on real world data.

3.4 Spectral smoothness of the mixing filters

This criterion was evaluated using an algorithm proposed in [7]. In this method, the permutations of the mixing filter spectral components are corrected according to the criterion

$$d = \sum_{i=1}^{nf} \left| a_{12}(i-1) - a_{12}(i) \right| + \left| a_{21}(i-1) - a_{21}(i) \right| \quad (4)$$

where $a_{12}(i)$ *and* $a_{21}(i)$ are the diagonal elements of the normalized mixing matrix $\boldsymbol{A}(i)=[1\ a_{12}(i);\ a_{21}(i)\ 1]$ in the i^{th} frequency band and nf is the number of frequency bands.

3.5 Cross-Correlation of Time Domain Signals

For comparison, the maximum value of the cross-correlation of the time domain separated signals was also included. This criterion is meant to serve as a simple test case of the optimization algorithm itself, though it cannot be considered very promising of a criterion for source separation.

3.6 Evaluation of Criteria

One basic assumption for the applied algorithm is that the different criteria will capture different aspects of optimality. To test this assumption, all criteria were evaluated for signals with artificially introduced permutations of frequency bands. Figures 2 to 5 show the values of the criteria, depending on the number of introduced permutations. For each number of permutations, 20 different random permutations were generated, and the plots show the mean, the standard deviation, the maximum and the minimum for each criterion.

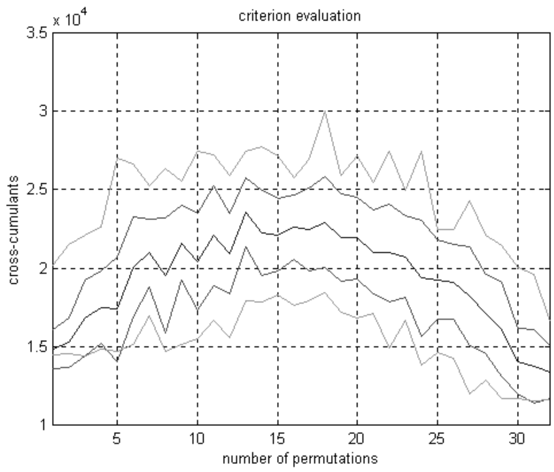

Fig 2. Mean, mean ± standard deviation, maximum and minimum of cross-cumulants, calculated according to (1), over number of permutations.

As can be seen from these figures, all criteria have their global minimum when no permutation is introduced, or when all bands are exchanged, and the maximum occurs, when there is a permutation of exactly half the frequency bands. This shows the feasibility of each criterion for optimization. However, all criteria have strong local minima, which can be expected to cause problems for gradient-based optimization.

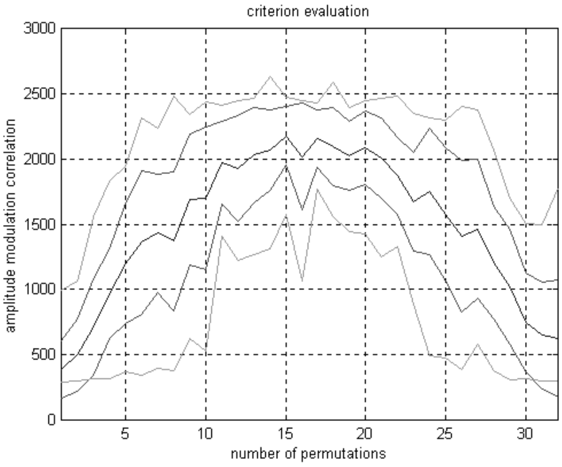

Fig 3. Mean, mean ± standard deviation, maximum and minimum of amplitude modulation correlation, calculated according to (3), over number of permutations.

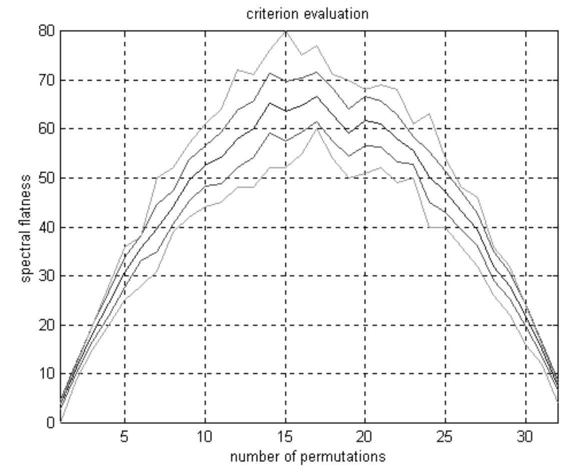

Fig 4. Mean, mean ± standard deviation, maximum and minimum of spectral flatness criterion, calculated according to (4), over number of permutations.

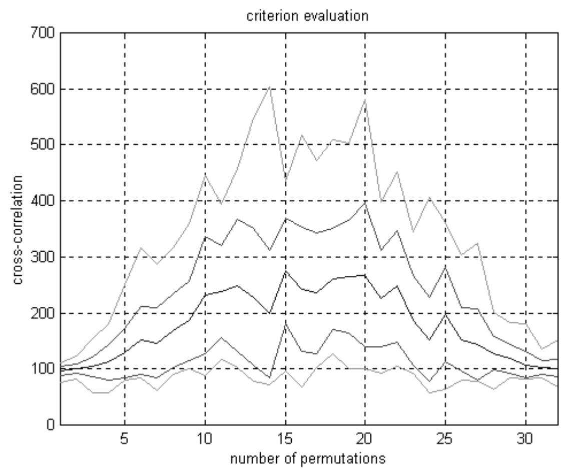

Fig 5. Mean, mean ± standard deviation, maximum and minimum of time domain cross-correlation over number of permutations.

Table 1 shows the sums of the correlation coefficients of all criteria value over all numbers of permutations, which were calculated as

$$cor_{xy} = \sum_{p=1}^{N} \frac{\sigma_{xy,p}}{\sigma_{x,p}\sigma_{y,p}} = \sum_{p=1}^{N} \frac{E[(x_p-\mu_{x,p})(y_p-\mu_{y,p})]}{\sqrt{E[(x_p-\mu_{x,p})^2]}\sqrt{E[(y_p-\mu_{y,p})^2]}} \quad (5)$$

Here, x and y are the criteria,and $\mu_{x,p}$ and $\mu_{y,p}$ are the mean values of x and y, respectively, for p permutations. N is the number of frequency bands, which was set to 64 for calculating the table, while the figures were generated at $N = 32$.

Table 1: Correlation coefficients cor_{xy} of criteria

	Cumulants	Amcor	Flatness	Correlation
Cumulants	1	0.5958	0.0181	0.1205
Amcor		1	0.5663	0.2205
Flatness			1	0.0828
Correlation				1

As can be seen from Table 1, the correlation between amplitude modulation correlation and the flatness as well as the cumulant criterion is especially high. However, at a correlation value of about 0.6, it can still be expected, that the other criteria will give significant new information above the sole evaluation of the Amcor criterion.

The flatness and the cross-cumulant criterion have a very low correlation, and it thus in an especially interesting case for joint optimization, as the information that is contained in both criteria will complement each other well.

4. RESULTS

To evaluate the performance of the criteria in multiobjective optimization, two sound signals of male German speakers were transformed to the frequency domain, using a hamming window and 256 frequency bands. An artificial random permutation of these signals was introduced. The parameters of the genetic algorithm were set to 256 Generations, 10% mutation and 90% of crossover.

An example for how optimization proceeds is shown in Figures 6 and 7, where the entire population and the final area of convergence are shown for a joint optimization of amplitude modulation correlation and the flatness criterion.

Table 2 shows the maximum degree of permutation correction that was achieved with each combination and single objective optimization run.

Table 2: Percentage of permutation correction with single and multiobjective optimization

	Cumulants	Amcor	Flatness	Correlation
Cumulants	75%	93.8%	84.4%	81.2%
Amcor		93.8%	87.5%	90.6%
Flatness			62.5%	78.2%
Correlation				68.8%

As can be seen from Table 2, optimization by amplitude modulation correlation has led to the best results for the set of test criteria. The set of solutions found in multiobjective optimization of amplitude modulation correlation together with the flatness criterion show, how this impacts the set of solutions found by multiobjective optimization. It can be seen from Figures 6 and 7, how the best solutions always lie at the lowest criterion values of amplitude modulation correlation, while the value of the other criterion is less significant regarding the degree of attainable permutation correction.

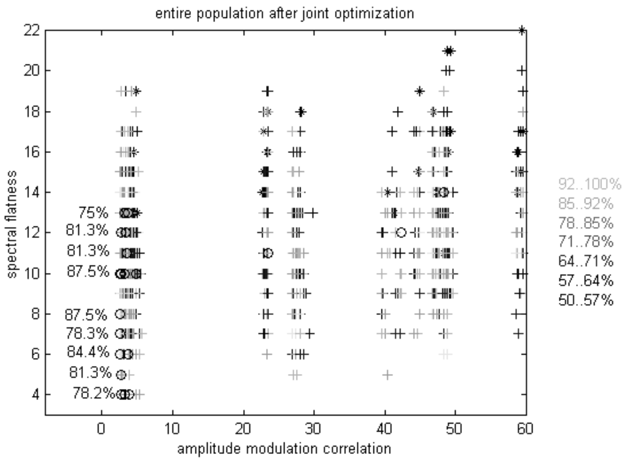

Fig 6. Entire population after joint optimization of spectral flatness with amplitude modulation correlation.

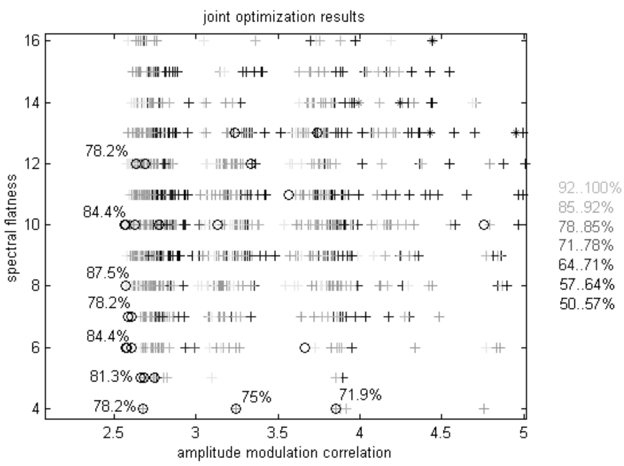

Fig 7. Results after joint optimization of spectral flatness with amplitude modulation correlation.

A similar effect can be observed when amplitude modulation correlation is optimized jointly with the cross-cumulant criterion or the cross-correlation criterion.

Thus, it can be seen that the three criteria tested here do not positively influence the outcome of optimization when used together with amplitude modulation correlation.

However, when two criteria are used for multiobjective optimization, which complement each other well, significant improvements can be attained over single objective optimization. This can be seen for the case of the flatness criterion together with cross cumulants, which had a very low correlation of 0.0181. In this case, multi-objective optimization delivered a result of 84.4%, which is a significant improvement over the 62.5% and 75.0%, respectively, when single objective optimization was applied. Also, optimizing cross-cumulants together with cross-correlation improves the results from 75.0% and 68.8%, respectively, in single objective optimization to 81.2% in joint optimization.

Figure 8 shows the results of the former case, where the final population is marked with circles while the initial population is marked by stars.

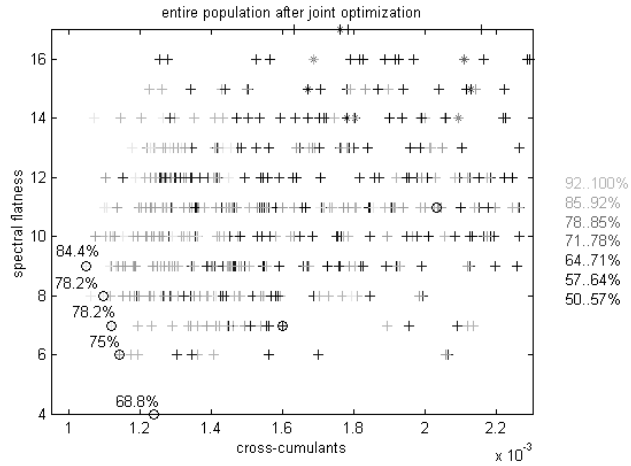

Fig 8. Entire set of populations after joint optimization of cross-cumulants with spectral flatness.

From these sets of results, in a next step it would be interesting to compute a criterion that captures all criteria with a suitable weighting. Based on the assumption that the optimal combination lies in the improvement of a weighted sum of all criteria, the appropriate weighting of the criteria can be calculated by taking the best point of the optimization as an ideal one. If this point is to be the optimum of single-objective weighted sum optimization, its weighted sum objective value

$$\sum_i w_i \cdot c_{i,opt} \qquad (6)$$

with w_i as the weight of the i^{th} criterion c_i, and $c_{i,opt}$ as the value of the i^{th} criterion at the optimal point, must be smaller than the objective value of any other Pareto-optimal point k, which leads to the requirement

$$\forall_k : \sum_i w_i \cdot c_{i,opt} \le \sum_i w_i \cdot c_{i,k} \; . \qquad (7)$$

This is in effect a standard linear programming problem, which allows the weights to be calculated by standard algorithms, subject to the constraint

$$\sum_i w_i = 1. \qquad (8)$$

With such a weighted sum value function, local optima should become smoother, leading to a quality function which integrates all aspects of optimality into one criterion and still lends itself well to other optimization algorithms, which perform faster than genetic algorithms but may not handle local optima as well.

5. CONCLUSIONS

Genetic optimization has been used to perform permutation correction for convolved speech signals in accordance with four different criteria. By using multi-objective optimization, it was possible to take more than one criterion into account simultaneously, which has allowed an evaluation of relative and joint merits of criteria.

Among the criteria that were tested, amplitude modulation correlation led to the best results, in single as well as multi-objective optimization.

However, multiobjective optimization has an advantage, when the criteria of interest complement each other well, in which case greatly improved results are attainable. Also, a multiobjective view of the final separation results can give a good overview of the relative merit of criteria, and may be used in a further step for deriving single objective quality functions that integrate more than one aspect of optimal source separation.

The four employed criteria were meant as a test case for the algorithm itself, and is expected that further multiobjective optimization of other possible criteria – e.g. feature signals in the time or frequency domain – will lead to further insight into the relationships between different criteria and ultimately to a suitable combination criterion.

REFERENCES

[1] A. J. Bell and T. J. Sejnowski (1995). An information-maximisation approach to blind separation and blind deconvolution. Neural Computation 7(6), pp. 1004-1034.

[2] A. Hyvärinen and E. Oja: A Fast Fixed-Point Algorithm for Independent Component Analysis. Neural Computation 9, pp. 1483-1492, 1997.

[3] J.-F. Cardoso and A. Souloumiac: Blind Beamforming for non Gaussian Signals. IEE Proceedings 140(6), pp. 362-370, 1993.

[4] T.-W. Lee, A. Ziehe. R. Orglmeister and T.J. Sejnowski: Combining Time Delayed Decorrelation and ICA: Towards Solving the Cocktail Party Problem. Proc. ICASSP, Vol. 2, pp. 1249-1252, 1998.

[5] L. Parra, C. Spence and B. De Vries: Convolutive Blind Source Separation Based on Multiple Decorrelation. Proc. IEEE NNSP workshop, pp. 23-32, Cambridge, UK, 1998.

[6] P. Smaragdis (1998). Blind Separation of Convolved Mixtures in the Frequency Domain. *Proc. International Workshop on Independence & Artificial Neural Networks*, Tenerife, Spain.

[7] W. Baumann: Implementierung and Analyse von Verfahren der Blinden Quellentrennung für Schallsignalaufnahmen in realen Umgebungen. Masters Thesis, Technische Universität Berlin, 1999.

[8] N. Murata, S. Ikeda and A. Ziehe: An Approach to Blind Source Separation Based on Temporal Structure of Speech Signals. BSIS Tech. Report, http://www.bsis.brain.riken.go.jp/, 1998.

[9] C. Mejuto, A. Dapena and L. Castedo: Frequency-Domain Infomax for Blind Separation of Convolutive Mixtures. Proc. ICA 2000, pp. 315-320, Helsinki, 2000.

[10] Z. Michalewicz: Genetic Algorithms + Data Structures = Evolution Programs. Springer Verlag, 1996.

[11] C. M. Fonseca and P. J. Fleming: An Overview of Evolutionary Algorithms in Multiobjective Optimization. Evolutionary Computation 3(1). pp. 1-16, 1995.

[12] D.E. Goldberg: Genetic Algorithms in Search, Optimization and Machine Learning, Addison-Wesley, 1989.

[13] C.M. Fonseca and P.J. Fleming: Multiobjective Genetic Algorithms Made Easy: Selection, Sharing and Mating Restriction, Proc. GALESIA 1995, pp. 45-52, 1995.

[14] D. Kolossa and G. Grübel: Evolutionary Computation and Nonlinear Programming in Multi-Model Robust Control Design, Lecture Notes in Computer Science, Vol. 1803, pp. 147-157, Springer Verlag, 2000.

[15] J. Anemüller and B. Kollmeier: Amplitude Modulation Decorrelation for Convolutive Blind Source Separation. Proc. ICA 2000, Helsinki, pp. 215-220.