BLIND SEPARATION OF CONVOLVED SOURCES BASED ON INFORMATION MAXIMIZATION

Kari Torkkola

Motorola, Inc., Phoenix Corporate Research Laboratories
2100 E. Elliot Rd, MD EL508, Tempe AZ 85284, USA
tel: (602)413-4129, fax: (602)413-5934, email: A540AA@email.mot.com

ABSTRACT

Blind separation of independent sources from their convolutive mixtures is a problem in many real world multi-sensor applications. In this paper we present a solution to this problem based on the information maximization principle, which was recently proposed by Bell and Sejnowski for the case of blind separation of instantaneous mixtures. We present a feedback network architecture capable of coping with convolutive mixtures, and we derive the adaptation equations for the adaptive filters in the network by maximizing the information transferred through the network. Examples using speech signals are presented to illustrate the algorithm.

1 INTRODUCTION

Blind source separation (BSS) denotes observing mixtures of *independent* sources, and by making use of these mixture signals only and nothing else, recovering the original signals. Due to a number of interesting applications in communications and speech and medical signal processing BSS has received recently lots of attention. Most of the work on BSS so far addresses the case of instantaneous mixtures, where a linear mixture model is assumed: $x(t) = As(t)$, where $s(t)$ is the vector of sources at instant t, A is the mixing matrix, and $x(t)$ is the observed vector of mixtures (ignoring noise). The sources can be separated by finding a matrix W with the condition $WA = PD$, where P is a permutation matrix and D is a diagonal matrix. Thus the recovered sources $u(t) = Wx(t) = PDs(t)$ are the original ones up to permutation and scaling. This instantaneous mixture case is also called Independent Component Analysis (ICA) [7].

To learn W by observing the mixtures only requires making use of higher order statistics [8, 7, 2]. Existing algorithms can be roughly divided into two categories, those that take some aspects of higher order statistics into account implicitly through some nonlinear functions (neural network-like algorithms) [14, 3, 22, 16, 9, 2, 5], and those that do it explicitly [12, 7, 27, 1]. Also, generalized cross-correlations can be used [26, 23]. A tutorial review to methods in the first category is given in [15].

As pointed out in [23, 2], there are several obstacles to applying these algorithms to real world situations. These include the effect of noise to successful learning of the separating solution, possibly unknown number of sources (especially noise sources), and the assumption that the source signals are stationary. Most notable is the assumption of the simultaneous mixing of the sources. In any real-world recording, where the propagation of the signals through the medium is not instantaneous (like sound through air or water), there will be phase differences between the sources in the mixtures. Current BSS algorithms cannot tolerate these delays. A related problem is the multipath arrival of the signals. In general, the sensor is not observing just a clean copy of the source, but a sum of multipath copies distorted by the environment, for example, by the impulse response of a room. These can be modeled as convolutive mixtures.

Some of the literature addresses these problems using various different methods. As independence can be defined in terms of vanishing higher-order cross statistics, many papers learn filters for separation by either implicitly or explicitly minimizing some function of some high-order cross statistics with a range of time lags, either combined or without using second order statistics [22, 12, 11, 27, 20, 24, 17, 28, 6]. Another approach is to return to the instantaneous mixture case by means of array processing [18], by moving to the frequency domain [4, 19], or by means of multivariate linear prediction [10, 13]. The whole problem of BSS can also be casted as maximum likelihood density estimation problem [21].

Only in [28] separation experiments done with real acoustic signals in a real reverberant environment are reported. No general and rigorous solution seems to exist so far, unless the approach in [21] will turn out to be one.

Bell and Sejnowski have presented an appealing approach of using mutual information as the separating criterion for instantaneous mixtures [2]. In a previous paper we extended Bell's information-theoretic approach to the separation of delayed signals [25]. In this paper will show how the problem of convolutive mixtures is also solvable using the same approach. We will first review the algorithm of Bell and Sejnowski. We then extend the information maximization approach to convolutive mixtures and, finally, we will present examples using speech signals to illustrate the performance of the algorithm.

2 INFORMATION MAXIMIZATION

Bell and Sejnowski proposed to learn the separating matrix W by minimizing the mutual information between components of $y(t) = g(u(t))$, where g is a nonlinear function approximating the cumulative density function (cdf) of the sources (see Fig. 1 for a two-source network) [2]. They showed that for positively kurtotic signals (like speech) minimizing the mutual information between components of y is equal to maximizing the entropy of y, which can be written as $H(y) = -E[ln(f_y(y))]$, where $f_y(y)$ denotes the probability

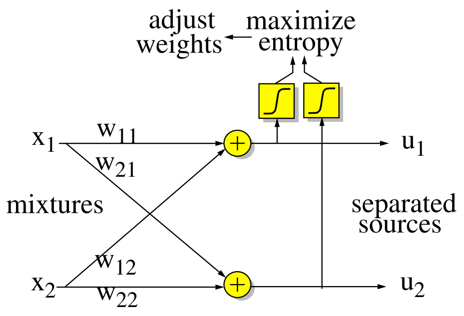


Figure 1: The separation network of Bell and Sejnowski for instantaneous mixtures of two sources (ignoring the bias weights).

density function of y. Denoting the determinant of the Jacobian of the whole net by $|J|$, $f_y(y)$ can be written as $f_x(x)/|J|$ (Jacobian is a matrix with entries of $\partial y_i/\partial x_j$.) Now, maximizing the entropy of the output leads to maximizing $E[ln|J|]$, which in turn can be developed into the following stochastic gradient ascent rule using instances of $x(t)$ and $y(t)$, instead of using the expectation:

$$\Delta W \quad \propto \quad (1 - 2y(t))x(t)^T + [W^T]^{-1} \tag{1}$$

To derive the adaptation rule above, $g(u) = 1/(1+e^{-u})$ was used to approximate the cdf. Bell and Sejnowski demonstrated the efficiency of this rule in the separation of as many as ten artificially mixed audio sources.

The algorithm can be made more efficient and independent of the conditioning of the mixing matrix by using the so called natural gradient instead of the absolute gradient [1].

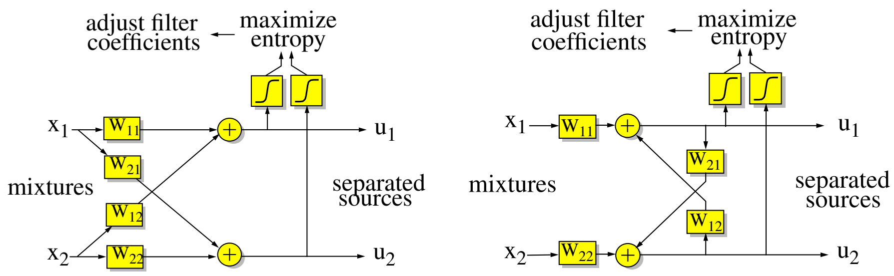
3 CONVOLUTIVE MIXTURES

Real audio signals present one example where the instantaneous mixing assumption does not hold. The acoustic environment (like a room) imposes a different impulse response between each source and microphone pair. Moreover, the microphones may have different characteristics, or at least their frequency responses may differ for sources in different directions. This kind of situations can be modeled as convolved mixtures. Let us look at two sources in the z-transform domain for simplicity; this can be generalized to any number of sources. For the present, we ignore the noise.

$$\begin{aligned} X_1(z) &= A_{11}(z)S_1(z) + A_{12}(z)S_2(z)\ , \\ X_2(z) &= A_{21}(z)S_1(z) + A_{22}(z)S_2(z), \end{aligned} \tag{2}$$

where A_{ij} are the z-transforms of any kind of filters. We can solve the sources S in terms of the mixtures X:

$$\begin{array}{lll} S_1(z) = (& A_{22}(z)X_1(z) & -A_{12}(z)X_2(z))/G(z) \\ S_2(z) = (- & A_{21}(z)X_1(z) & +A_{11}(z)X_2(z))/G(z) \end{array} \tag{3}$$



a) A feedforward network with adaptive filters for the separation of convolved mixtures.

b) A feedback network with adaptive filters for the separation of convolved mixtures.

Figure 2: Blind separation network architectures for two-source mixtures.

where $G(z) = A_{11}(z)A_{22}(z) - A_{12}(z)A_{21}(z)$. This suggests a feedforward architecture for separation. The network is illustrated in Fig. 2 a), where $W_{11}(z) = A_{22}(z)/G(z)$, $W_{12}(z) = -A_{12}(z)/G(z)$, and so on, at the ideal separating solution. The existence of a separating solution requires that the stable inverse of $G(z)$ exists (this implies that it is possible to construct mixtures impossible to separate).

This network is the combined deconvolution and separation network mentioned in [2]. The adaptation rules for the filter coefficients are easy to derive directly in the time domain, again by maximizing the entropy at the output (See Appendix A).

However, entropy maximization with this network does not result in the solution of (3). In the instantaneous mixture case we had the indeterminacy of the recovered sources up to permutation and scaling. With convolutive mixtures the indeterminacy is up to permutation and arbitrary filtering. Entropy maximization has the side effect of temporally whitening the outputs, which may not be desirable in many applications, for example, in those involving speech. The effect is caused by $W_{11}(z)$ and $W_{22}(z)$, which tend to remove all redundancies in time within each signal. A separating solution without the whitening effect exists, though. By rearranging (3), so that the direct filters $W_{11}(z)$ and $W_{22}(z)$ become unity we get

$$\begin{aligned} G(z)A_{22}(z)^{-1}S_1(z) &= X_1(z) - A_{12}(z)A_{22}(z)^{-1}X_2(z) \\ G(z)A_{11}(z)^{-1}S_2(z) &= X_2(z) - A_{21}(z)A_{11}(z)^{-1}X_1(z) \end{aligned} \tag{4}$$

Unfortunately the cross-filters $W_{12}(z)$ and $W_{21}(z)$ will not converge to these solutions when the adaptation of Appendix A is used. Rather, they will become mere temporal whitening filters. Obviously, in this case, temporal whitening will contribute more to the total entropy increase than mere separation.

The whitening effect can be avoided by using a different network architecture, namely the feedback structure depicted in Fig. 2 b). In this architecture the outputs before nonlinearities are:

$$\begin{aligned} U_1(z) &= W_{11}(z)X_1(z) + W_{12}(z)U_2(z) \ , \\ U_2(z) &= W_{22}(z)X_2(z) + W_{21}(z)U_1(z). \end{aligned} \tag{5}$$

The ideal separating solution in terms of A's and X's will of course be exactly the same as in the feedforward case, but now with different filters W_{ij} because of the different architecture. Using (2) and (5) we can easily derive a solution for perfect separation and deconvolution:

$$\begin{aligned} &W_{11}(z) = A_{11}(z)^{-1}, \quad W_{12}(z) = -A_{12}(z)A_{11}(z)^{-1} \\ &W_{22}(z) = A_{22}(z)^{-1}, \quad W_{21}(z) = -A_{21}(z)A_{22}(z)^{-1} \end{aligned} \tag{6}$$

Again, maximizing the entropy at the output will result W_{11} and W_{22} not only inverting A_{11} and A_{22}, but also whitening the sources. This can be avoided by forcing $W_{11}(z)$ and $W_{22}(z)$ to mere scaling coefficients. In the ideal case W_{12} and W_{21} will have the following solution:

$$\begin{aligned} &W_{11}(z) = 1, \quad W_{12}(z) = -A_{12}(z)A_{22}(z)^{-1} \\ &W_{22}(z) = 1, \quad W_{21}(z) = -A_{21}(z)A_{11}(z)^{-1} \end{aligned} \tag{7}$$

Why is whitening avoided with this configuration? Since the output of a cross filter is summed to a *different* branch than where its input comes from it has no chance of removing redundancies from its input signal. Instead, it will learn to remove the redundancies *between* the signals. Thus no deconvolution is performed at all, only separation. We must be then content with $U_1(z) = A_{11}(z)S_1(Z)$ and $U_2(z) = A_{22}(z)S_2(Z)$, which is what each sensor would observe in the absence of the interfering sources without any other distorting effects. The existence of this solution requires $A_{11}(z)$ and $A_{22}(z)$ have stable inverses, in addition to $G(z)$ in (3) also having a stable inverse. If these are not fulfilled, the network is unable to achieve separation. Derivation of the adaptation rules for this architecture [1] is presented in Appendix B.

4 EXPERIMENTS WITH SPEECH SIGNALS

In this section we present an example of separation of two speech signals mixed using the filters in (8). The filters were chosen so that the correctness of the solution would be easy to verify by eye.

$$\begin{aligned} A_{11}(z) &= 1 - 0.4z^{-25} + 0.2z^{-45}, \\ A_{12}(z) &= 0.4z^{-20} - 0.2z^{-28} + 0.1z^{-36}, \\ A_{21}(z) &= 0.5z^{-10} + 0.3z^{-22} + 0.1z^{-34}, \\ A_{22}(z) &= 1 - 0.3z^{-20} + 0.2z^{-38}. \end{aligned} \tag{8}$$

We used the feedback architecture with FIR-filters (see Fig. 2 b). W_{11} and W_{22} were reduced just to adaptive scaling coefficients to avoid the temporal whitening effect. W_{12} and W_{21} had 100 taps each. Figure 3 depicts the learned coefficients of each filter.

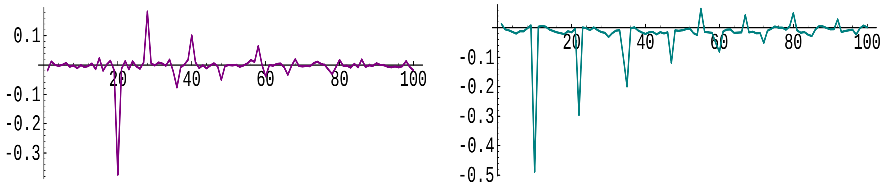


Figure 3: Coefficients of separating FIR-filters W_{12} (left) and W_{21} (right).

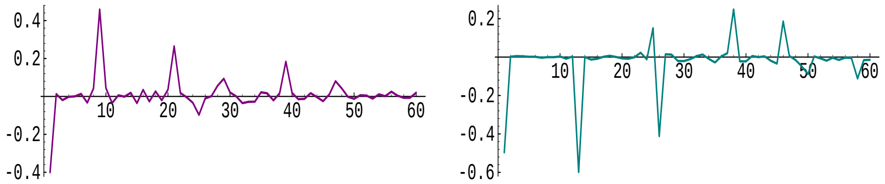


Figure 4: Coefficients of separating IIR-filters W_{12} (left) and W_{21} (right).

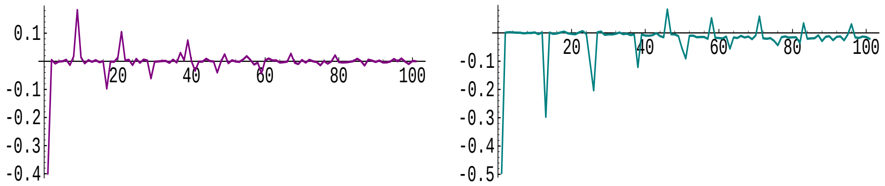


Figure 5: Impulse responses of the IIR-filters W_{12} (left) and W_{21} (right).

In this case, W_{12} and W_{21} should follow (7). For example, looking at W_{21}, it is easy to locate the three peaks corresponding to $-A_{21}$. The rest of the peaks correspond to the convolution of $-A_{21}$ with the $1/A_{11}$ (which has an infinite length as an FIR-filter, ideally). We should see the first of these peaks at delays 35 (10+25), 47 (22+25), 55 (10+45), 59 (34+25), and 67 (22+45). This is the case indeed. The audible separation was nearly perfect.

In some cases the FIR-filters need to be very long to be able to model the inverses of the A_{ii}'s. Shorter IIR-filters suffice for the same effect. Figure 4 presents the coefficients of two IIR-filters of the direct form of length 60. The delay to the first nonzero tap was also learned by the delay adaptation rule [25]. This way we can avoid wasting resources to taps with zero coefficients in the beginning of the filter. The learning rate for the delays needs to be larger than the learning rate for the weights to make the delays "win" first.

The impulse responses of the filters are depicted in Fig. 5 up to 100 samples. These are identical to the FIR-filters in Fig. 3, which should be the case. Notice that in Fig. 5 the impulse responses begin at the first peaks of Fig. 3, because the initial delays were learned in this case, too.

In addition to artificial convolutive mixtures we have also experimented with

[1] It is also possible to switch the order of W_{ii} and W_{ij} filters.

real world recordings. For example, with two simultaneous speakers in a small hard wall conference room using two omnidirectional microphones on a table in the middle of the room, we were able to achieve a 12 dB attenuation of the unwanted source component.

5 CONCLUSION

Convolutive mixtures of independent sources arise in many real world multi-sensor applications. We have presented a feedback network architecture of adaptive filters for blind separation of such sources. We have derived adaptation equations for the filters based on the information maximization principle presented by Bell and Sejnowski [2]. Experiments with speech sources demonstrate the efficiency of the approach. Though all presented examples concern only two sources, the method is easily extendable to multiple sources as mentioned in Appendix B. Future work includes investigating the use of noncausal filters in the separation network, quantifying the effect of independent noise to the adaptation, and the analysis of the statistical stability and convergence of the algorithm.

Appendix A. Adaptation rules for the separation of mixtures of convolved sources using a feedforward network architecture.

Referring to Fig. 2 a) we derive the adaptation equations in the time domain using mixtures of two sources. g is the logistic function $y = g(u) = 1/(1+e^{-u})$. Following Bell and Sejnowski, we can minimize the mutual information between outputs y_1 and y_2 by maximizing the entropy at the output, which is equals maximizing $E[ln|J|]$.

Assuming causal FIR-filters for W_{ij} the network carries out the following:

$$
\begin{aligned}
u_1(t) &= \sum_{k=0}^{L_{11}} w_{1k1} x_1(t-k) + \sum_{k=0}^{L_{12}} w_{1k2} x_2(t-k), \\
u_2(t) &= \sum_{k=0}^{L_{22}} w_{2k2} x_2(t-k) + \sum_{k=0}^{L_{21}} w_{2k1} x_1(t-k).
\end{aligned}
\tag{9}
$$

w_{ikj} denotes the weight associated with delay k in the filter from mixture j to output i.

The determinant of the Jacobian of the network is now

$$
\begin{aligned}
|J| &= \frac{\partial y_1}{\partial x_1}\frac{\partial y_2}{\partial x_2} - \frac{\partial y_1}{\partial x_2}\frac{\partial y_2}{\partial x_1} = y_1' y_2' D \quad \textit{and} \\
ln|J| &= ln(y_1') + ln(y_2') + ln(D).
\end{aligned}
\tag{10}
$$

where

$$
\begin{aligned}
D &= \left(\frac{\partial u_1}{\partial x_1}\frac{\partial u_2}{\partial x_2} - \frac{\partial u_1}{\partial x_2}\frac{\partial u_2}{\partial x_1}\right) = w_{101}w_{202} - w_{102}w_{201}, \\
y_1' &= \frac{\partial y_1}{\partial u_1}, \quad y_2' = \frac{\partial y_2}{\partial u_2}.
\end{aligned}
$$

The adaptation rule for each parameter of the network can now be derived by computing the gradient of $ln|J|$ with respect to that parameter. For w_{101} we get

$$\Delta w_{101} \propto \frac{\partial ln|J|}{\partial w_{101}} = \frac{1}{y_1'}\frac{\partial y_1'}{\partial w_{101}} + \frac{1}{y_2'}\frac{\partial y_2'}{\partial w_{101}} + \frac{1}{D}\frac{\partial D}{\partial w_{101}} \tag{11}$$

For the logistic function $\frac{\partial y_i'}{\partial y_i} = 1 - 2y_i$. Thus we can write for the partial derivatives for w_{101}:

$$\begin{aligned} \frac{\partial y_1'}{\partial w_{101}} &= \frac{\partial y_1'}{\partial y_1}\frac{\partial y_1}{\partial u_1}\frac{\partial u_1}{\partial w_{101}} = (1-2y_1)y_1'x_1, \\ \frac{\partial y_2'}{\partial w_{101}} &= \frac{\partial y_2'}{\partial y_2}\frac{\partial y_2}{\partial u_2}\frac{\partial u_2}{\partial w_{101}} = 0, \\ \frac{\partial D}{\partial w_{101}} &= w_{202}. \end{aligned} \tag{12}$$

Thus the adaptation rules for zero delay weights become the following:

$$\begin{aligned} &\Delta w_{101} \propto (1-2y_1)x_1 + w_{202}/D, \quad \Delta w_{102} \propto (1-2y_1)x_2 - w_{201}/D, \\ &\Delta w_{201} \propto (1-2y_2)x_1 - w_{102}/D, \quad \Delta w_{202} \propto (1-2y_2)x_2 + w_{101}/D. \end{aligned} \tag{13}$$

If written in matrix form this is exactly the same as (1). The zero delay weights thus follow the solution of Bell and Sejnowski, as could be expected.

The adaptation rules for weights other than zero delay can be derived exactly in the same fashion. Only one term will be nonzero in the partial derivatives, and we arrive to the decorrelation rule

$$\Delta w_{ikj} \quad \propto \quad (1-2y_i)x_j(t-k). \tag{14}$$

Where time indices have been left out, t is assumed.

Appendix B. Adaptation rules for the separation of mixtures of convolved sources using a feedback network architecture.

Referring to Fig. 2 b) we derive the adaptation equations using mixtures of two sources only. Assuming causal FIR-filters for W_{ij}, the network carries out the following in the time domain

$$\begin{aligned} u_1(t) &= \sum_{k=0}^{L_{11}} w_{1k1}x_1(t-k) + \sum_{k=1}^{L_{12}} w_{1k2}u_2(t-k), \\ u_2(t) &= \sum_{k=0}^{L_{22}} w_{2k2}x_2(t-k) + \sum_{k=1}^{L_{21}} w_{2k1}u_1(t-k). \end{aligned} \tag{15}$$

For the Jacobian, we get

$$ln|J| = ln(y_1') + ln(y_2') + ln(D) = ln(y_1') + ln(y_2') + ln(w_{101}w_{202}). \tag{16}$$

There will now be three different cases: zero delay weights in direct filters, other weights in direct filters, and weights in feedback cross-filters. Following the steps in Appendix A and in [25] for all these cases, we get:

$$\begin{aligned}\Delta w_{i0i} &\propto (1-2y_i)x_i + 1/w_{i0i}, \\ \Delta w_{iki} &\propto (1-2y_i)x_i(t-k), \\ \Delta w_{ikj} &\propto (1-2y_i)u_j(t-k). \end{aligned} \tag{17}$$

The zero delay weights scale the data to maximize the information passed through the sigmoid, other weights in the direct branches of the network decorrelate each output from the corresponding input mixture (temporal whitening), and the weights of the feedback branches decorrelate each output y_i from all the other sources u_j at every time instant within the scope of the filters (separation). The derivation for IIR-filters or filters of any structure follows the same steps. Since the rules are local, generalization to multiple sources is trivial.

References

[1] S. Amari, A. Cichocki, and H. H. Yang. A new learning algorithm for blind signal separation. In *Advances in Neural Information Processing Systems 8*. MIT Press, 1996.

[2] A. J. Bell and T. J. Sejnowski. An information-maximisation approach to blind separation and blind deconvolution. *Neural Computation*, 7(6):1129–1159, 1995.

[3] G. Burel. Blind separation of sources: A nonlinear neural algorithm. *Neural Networks*, 5(6):937–947, 1992.

[4] V. Capdevielle, C. Serviere, and J. Lacoume. Blind separation of wide-band sources in the frequency domain. In *Proc. ICASSP*, pages 2080–2083, Detroit, MI, May 9-12 1995.

[5] J.-F. Cardoso. The invariant approach to source separation. In *Proceedings of the International Symposium on Nonlinear Theory and Applications (NOLTA)*, Las Vegas, Nevada, USA, December 10-14 1995.

[6] D. B. Chan, P. J. W. Rayner, and S. J. Godsill. Multi-channel signal separation. In *Proc. ICASSP*, Atlanta, GA, May 7-10 1996.

[7] P. Comon. Independent component analysis, a new concept? *Signal Processing*, 36(3):287–314, 1994.

[8] P. Comon, C. Jutten, and J. Herault. Blind separation of sources, part II: Problems statement. *Signal Processing*, 24(1):11–20, 1991.

[9] N. Delfosse and P. Loubaton. Adaptive separation of independent sources: A deflation approach. In *Proc. ICASSP*, volume IV, pages 41–44, Adelaide, Australia, April 19-22 1994.

[10] N. Delfosse and P. Loubaton. Adaptive blind separation of convolutive mixtures. In *Proc. ICASSP*, Atlanta, GA, May 7-10 1996.

[11] A. Dinc and Y. Bar-Ness. A forward/backward bootstrapped structure for blind separation of signals in a multi-channel dispersive environment. In *Proc. ICASSP*, pages 376–379, Minneapolis, MN, USA, April 27-30 1993.

[12] A. M. Engebretson. Acoustic signal separation of statistically independent sources using multiple microphones. In *Proc. ICASSP*, volume II, pages 343–346, Minneapolis, MN, USA, April 27-30 1993.

[13] S. Icart and R. Gautier. Blind separation of convolutive mixtures using second and fourth order moments. In *Proc. ICASSP*, Atlanta, GA, May 7-10 1996.

[14] C. Jutten and J. Herault. Blind separation of sources, part I: An adaptive algorithm based on neuromimetic architecture. *Signal Processing*, 24(1):1–10, 1991.

[15] J. Karhunen. Neural approaches to independent component analysis and source separation. In *Proc. ESANN'96 (4th European Symposium on Artificial Neural Networks)*, Bruges, Belgium, April 24-26 1996.

[16] J. Karhunen, L. Wang, and R. Vigario. Nonlinear PCA type approaches for source separation and independent component analysis. In *Proc. ICNN*, pages 995–1000, Perth, Western Australia, November 27 - December 1 1995.

[17] R. H. Lambert. A new method for source separation. In *Proc. ICASSP*, pages 2116–2119, Detroit, MI, May 9-12 1995.

[18] S. Li and T. J. Sejnowski. Adaptive separation of mixed broad-band sound sources with delays by a beamforming Hérault-Jutten network. *IEEE Journal of Oceanic Engineering*, 20(1):73–79, January 1995.

[19] K. Matsuoka and M. Kawamoto. Blind signal separation based on mutual information criterion. In *Proceedings of the International Symposium on Nonlinear Theory and Applications (NOLTA)*, pages 85–90, Las Vegas, Nevada, USA, December 10-14 1995.

[20] M. Nájar, M. A. Lagunas, and I. Bonet. Blind wideband source separation. In *Proc. ICASSP*, Vol IV, pages 65–68, Adelaide, Australia, April 19-22 1994.

[21] B. A. Pearlmutter and L. C. Parra. A context-sensitive generalization of ICA. In *International Conference on Neural Information Processing*, Hong Kong, Sept. 24–27 1996. Springer. In press.

[22] J. C. Platt and F. Faggin. Networks for the separation of sources that are superimposed and delayed. In J. Moody, S. Hanson, R. Lippmann, eds., *Advances in Neural Information Processing Systems 4*. Morgan-Kaufmann, 1992.

[23] K. Pope and R. Bogner. Blind separation of speech signals. In *Proceedings of the Australian International Conference on Speech Science and Technology*, Perth, Western Australia, December 6-8 1994.

[24] H.-L. N. Thi and C. Jutten. Blind source separation for convolutive mixtures. *Signal Processing*, 45(2), 1995.

[25] K. Torkkola. Blind separation of delayed sources based on information maximization. In *Proc. ICASSP*, Atlanta, GA, May 7-10 1996.

[26] E. Weinstein, M.Feder, and A. Oppenheim. Multi-channel signal separation by decorrelation. *IEEE Trans. on Speech and Audio Processing*, 1(4):405–413, 1993.

[27] D. Yellin and E. Weinstein. Criteria for multichannel signal separation. *IEEE Trans. on Signal Processing*, 42(8):2158–2168, August 1994.

[28] D. Yellin and E. Weinstein. Multichannel signal separation: Methods and analysis. *IEEE Trans. on Signal Processing*, 44(1):106–118, January 1996.