A NEW METHOD FOR BLIND SOURCE SEPARATION OF NONSTATIONARY SIGNALS

Douglas L. Jones

Department of Electrical and Computer Engineering & Coordinated Science Laboratory
University of Illinois at Urbana-Champaign
dl-jones@uiuc.edu

ABSTRACT

Many algorithms for blind source separation have been introduced in the past few years, most of which assume statistically stationary sources. In many applications, such as separation of speech or fading communications signals, the sources are nonstationary. We present a new adaptive algorithm for blind source separation of nonstationary signals which relies only on the nonstationary nature of the sources to achieve separation. The algorithm is an efficient, on-line, stochastic gradient update based on minimizing the average squared cross-output-channel-correlations along with deviation from unity average energy in each output channel. Advantages of this algorithm over existing methods include increased computational efficiency, a simple on-line, adaptive implementation requiring only multiplications and additions, and the ability to blindly separate nonstationary sources regardless of their detailed statistical structure.

1. INTRODUCTION

The separation of multiple unknown sources from multi-sensor data has many applications, including the isolation of individual speech signals from a mixture of simultaneous speakers (as in video conferencing or the often-cited "cocktail party" environment), the elimination of cross-talk between horizontally and vertically polarized microwave communications transmissions, and the separation of multiple cellular telephone signals at a base station. In the past decade or so, a number of significant methods have been introduced for blind source separation, of which we review a few of the most popular here. One of the earliest and most effective methods is a constant-modulus-based method published in 1985 by Treichler and Larimore [1]. This method achieves simultaneous separation and equalization (i.e., separation of convolutive mixtures) by minimizing the deviation of the separated output magnitudes from a fixed gain. This method is very simple and efficient and works well

even for non-constant-modulus signals with a sub-Gaussian kurtosis (which includes most communications signals).

Jutten and Herault introduced one of the most popular methods [2]. This method works well in many applications, particularly cross-talk situations in which a relatively modest amount of mixing occurs. Methods for non-Gaussian sources have also been developed, including [3] and others[1]. A method by Belouchrani, *et al.* can separate stationary Gaussian sources with different autocorrelation statistics [4].

In many applications of blind source separation, the received signals are nonstationary. Nonstationarity may arise either from the source signals themselves (such as speech), or from channel impairments (such as fading in wireless communications channels). Most techniques for blind source separation assume stationarity of the signals and depend on reliable estimation of second-order or higher-order statistics. These methods may have difficulty when applied to nonstationary signals.

Several methods developed explicitly for the nonstationary source separation have been published recently. Belouchrani and Amin have developed a time-frequency extension of the method in [5] for nonstationary sources, and Parra, *et al.* have developed another method based on frequency decomposition of several successive blocks of time [6]. While these methods appear effective, and the latter can also separate convolutive mixtures, they are block-based methods requiring somewhat sophisticated and expensive processing. Matsuoka, *et al.* present an on-line, adaptive extension of the Jutten-Herault method which, somewhat like the method proposed here, attempts to minimize the average cross-correlation between separated channels while normalizing the output energy [7].

We propose a new method for blind source separation which requires only nonstationarity and independence of the sources to achieve separation. A very efficient on-line, LMS-like algorithm is derived which achieves separation while normalizing the average energy of each output channel. This algorithm offers advantages in terms of simplicity

[1] It should be noted here that the CMA-based method by Treichler and Larimore also depends on the non-Gaussianity of the sources.

or efficiency as compared to existing methods, along with tracking capability for time-varying mixtures. The optimization criterion is presented in the second section of this paper, an efficient adaptive algorithm is derived in the third section, and simulations which illustrate its performance are presented in the fourth section. Some perspectives on the results are discussed in the final section.

2. A NONSTATIONARY SOURCE SEPARATION CRITERION

The general source separation problem with instantaneous mixtures can be described as

$$\mathbf{x}(n) = \mathbf{A}\mathbf{s}(n), \tag{1}$$

where $\mathbf{s}(n)$ is a vector of M zero-mean, statistically independent source processes at time-sample n, $\mathbf{x}(n)$ is a vector of N sensor measurements, $N \geq M$, and $\mathbf{A}$ is a mixing matrix of rank equal to or greater than the number of sources. The goal of blind source separation is to determine a matrix $\mathbf{B}$ which, when applied to the received sensor data as in

$$\mathbf{y}(n) = \mathbf{B}\mathbf{x}(n), \tag{2}$$

recovers (separates) the individual sources up to an unknown permutation and unknown channel gains, which cannot be uniquely determined without additional information [7]. In a noise-free case, $\mathbf{B}$ will be a pseudo-inverse of the mixing matrix $\mathbf{A}$ up to unknown permutation, gain, and null-space components.

It has been observed in many papers on blind source separation that a necessary condition for the separation of zero-mean, statistically independent sources is that the cross-correlations of the output channels equal zero. However, this is not a sufficient condition, as is well known and easily illustrated by the following example. Consider the simple case of two sources and two sensors, where the sources are stationary processes with variances σ_1^2 and σ_2^2, and the mixing matrix is $\mathbf{A} = \mathbf{I}$. Clearly, the received signals are already separated, and $\mathbf{B} = \mathbf{I}$ is the desired solution. However, it is easily shown that the cross-correlation

$$E[y_1 y_2] = b_{11}b_{21}\sigma_1^2 + b_{12}b_{22}\sigma_2^2, \tag{3}$$

and, for given σ_i's, there are an infinite number of $\mathbf{B}$ matrices with zero cross-channel correlation which do not separate the sources. However, it is noted in [7] that this ambiguity exists only for sources with fixed σ_i's. While for any arbitrary pair of variances σ_1^2 and σ_2^2 there exist an infinite number of decorrelating but not separating $\mathbf{B}$'s, these classes are *different* for different source variances, and only a truly separating solution yields zero cross-channel correlation for all variance combinations. This is the key insight on which nonstationary blind source separation algorithms

are based. In effect, these methods take multiple snapshots of the short-time cross-correlation at different times, and by minimizing all of these simultaneously, exploit the changes in the relative channel variances to find a truly separating solution.

This paper uses the same basic insight, but proposes a new criterion for exploiting it which leads to a particularly efficient and convenient algorithm. We propose to minimize the following criterion:

$$\min_{\mathbf{B}} E\left[\sum_{i=1}^{M}\sum_{j \neq i}^{M} \hat{r}_{y_i y_j}^2 + \lambda \sum_{i=1}^{M}(\hat{r}_{y_i y_i} - 1)^2\right] \tag{4}$$

where

$$\hat{r}_{y_i y_j}(n) = \sum_k h(k)y_i(n-k)y_j(n-k) \tag{5}$$

where $h(k)$ is a lowpass averaging filter for computing a short-term estimate of the cross-correlation of output channels y_i and y_j at time n. The first term in the criterion is to minimize the average squared magnitude of the short-term cross-correlations of the output signals (which, as shown in [7], is only achieved for nonstationary signals by a separating solution), while the second term demands that the output signals in each channel have unit energy on average. In our experiments, we have always chosen λ equal to the ratio of the number of cross-correlation and auto-correlation terms, to equally weight in some sense the two components of the criterion.

3. ADAPTIVE ALGORITHM

There are many ways to construct a numerical algorithm based on the above criterion for blind nonstationary source separation, yielding different tradeoffs in terms of computational efficiency, convergence rate, block-based or adaptive forms, etc. However, in many applications, a computationally efficient, adaptive method which can track slow variations in the mixing parameters is desired. We derive here a stochastic gradient (LMS-like) algorithm which has these characteristics.

For the optimization of the demixing matrix, $\mathbf{B}$, a stochastic gradient update takes the form

$$\mathbf{B}_{n+1} = \mathbf{B}_n - \mu \nabla_n \tag{6}$$

where

$$\nabla_n = \left[\frac{\partial}{\partial b_{pq}}\left\{\sum_{i=1}^{M}\sum_{j \neq i}^{M} \hat{r}_{y_i y_j}^2 + \lambda \sum_{i=1}^{M}(\hat{r}_{y_i y_i} - 1)^2\right\}\right] \tag{7}$$

where p and q are the row and column indices of the gradient matrix. Note the use of the instantaneous value at time

n of the error function in (4) in the gradient computation. The (p, q)th element of the instantaneous gradient matrix can easily be shown to be

$$\nabla_{pq}(n) = 4 \sum_{i \neq p}^{M} \hat{r}_{y_p, y_i} \hat{r}_{y_i, x_q} + 2\lambda (\hat{r}_{y_p, y_p} - 1)(\hat{r}_{y_p, x_q}) \quad (8)$$

The first sum in the above equation can conveniently be expressed as

$$\sum_{i \neq p}^{M} \hat{r}_{y_p, y_i} \hat{r}_{y_i, x_q} = \sum_{i=1}^{M} \hat{r}_{y_p, y_i} \hat{r}_{y_i, x_q} - \hat{r}_{y_p, y_p} \hat{r}_{y_p, x_q} \quad (9)$$

With this observation, the gradient can be succinctly expressed in matrix form as

$$\nabla_n = 4(\hat{\mathbf{R}}_{yy} \hat{\mathbf{R}}_{yx} - \text{diag}(\hat{\mathbf{R}}_{yy})[11 \cdots 1] \otimes \hat{\mathbf{R}}_{yx})$$
$$+ 2\lambda((\hat{\mathbf{R}}_{yx} - 1) \otimes \hat{\mathbf{R}}_{yx}) \quad (10)$$

where $\otimes$ is a Kronecker (term-by-term) product, $\hat{\mathbf{R}}_{yy}$ is the matrix of averaged correlation products $\hat{r}_{y_i y_j}(n)$ defined in (5), $\hat{\mathbf{R}}_{yx}$ is the corresponding matrix of averaged cross-correlation products between the output and input vectors, and $\mathbf{1}$ is the all-ones matrix.

The above matrix expression for the stochastic gradient update yields an efficient and straightforward computation once the short-time correlations are available. We now derive efficient recursive updates for these components for a convenient form of the averaging filter. For computational efficiency, we select a first-order IIR averaging filter with impulse response

$$h(k) = \alpha^k u(k) \quad (11)$$

where $u(k)$ is the unit step function and $0 < \alpha < 1$. With this form, the elements of $(1 - \alpha)^{-1} \hat{\mathbf{R}}_{yy}$ can easily be updated recursively according to

$$\tilde{r}_{y_i y_j}(n) = \alpha \tilde{r}_{y_i y_j}(n) + y_i(n) y_j(n), \quad (12)$$

and similarly

$$\tilde{r}_{y_i x_j}(n) = \alpha \tilde{r}_{y_i x_j}(n) + y_i(n) x_j(n) \quad (13)$$

for $(1 - \alpha)^{-1} \hat{\mathbf{R}}_{yx}$. This yields the following simple recursive algorithm for nonstationary blind source separation.

Compute output:
$$\mathbf{y}_n = \mathbf{B}_n \mathbf{x}_n \quad (14)$$

Update short-time correlations:

$$\mathbf{R}_{yy_{n+1}} = \alpha \mathbf{R}_{yy_n} + \mathbf{y}_n \mathbf{y}_n^T \quad (15)$$

$$\mathbf{R}_{yx_{n+1}} = \alpha \mathbf{R}_{yx_n} + \mathbf{y}_n \mathbf{x}_n^T \quad (16)$$

Update separation matrix:

$$\mathbf{B}_{n+1} = \mathbf{B}_n - \mu(1 - \alpha)^2$$
$$(4(\mathbf{R}_{yy} \mathbf{R}_{yx} - \text{diag}(\mathbf{R}_{yy})[11 \cdots 1] \otimes \mathbf{R}_{yx})$$
$$+ 2\lambda(\mathbf{R}_{yx} - 1) \otimes \mathbf{R}_{yx}) \quad (17)$$

For M sources and N sensors, the cost of the output computation is MN multiplications and $M(N - 1)$ additions per sample time. The $\mathbf{R}_{yy}$ and $\mathbf{R}_{yx}$ updates require $M^2 + M$ multiplications and $(M^2 + M)/2$ additions, and $2MN$ multiplications and MN additions, respectively. The gradient update requires $M^2 N + 4MN$ multiplications and $M^2 N + 3MN$ additions, for a total operation count of $M^2 N + 5MN + M^2 + M$ multiplications and $M^2 N + 4MN + (M^2 - M)/2$ additions per vector sample.

4. SIMULATIONS

Several simulations have been performed to confirm the efficacy of the proposed method. For the following simulation with two sources and sensors, the mixing matrix is

$$\mathbf{A} = \begin{bmatrix} 2 & 1.5 \\ 0.5 & 1 \end{bmatrix} \quad (18)$$

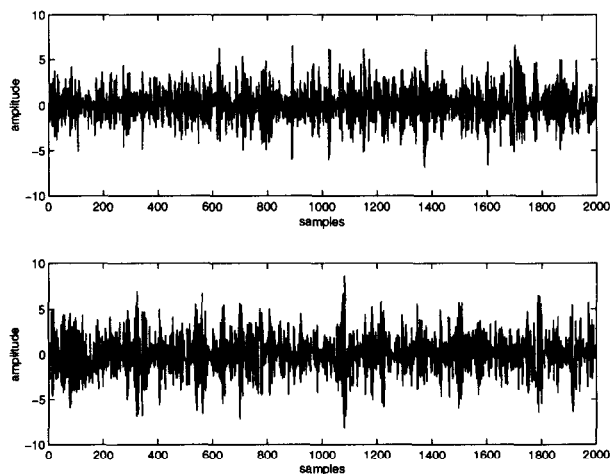as used in [7]. The sources, shown in Figure 1, are binary

Figure 1: Nonstationary sources used in the simulation

random signals multiplied by lowpass filtered Gaussian signals, and may be considered a crude approximation to communications signals undergoing fading. Figure 2 plots the adaptation of the four coefficients in the demixing matrix $\mathbf{B}$ with $\mu = 0.0001$ and $\alpha = 0.9$. Note the very rapid initial convergence to substantial separation, followed by much slower refinement to almost perfect separation. Figure 3 shows the separated outputs, which closely resemble
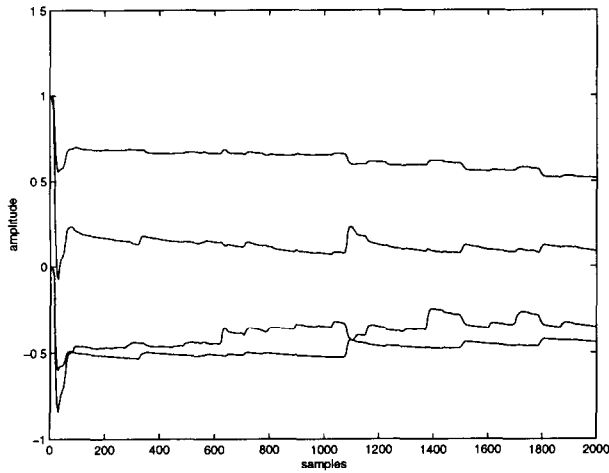
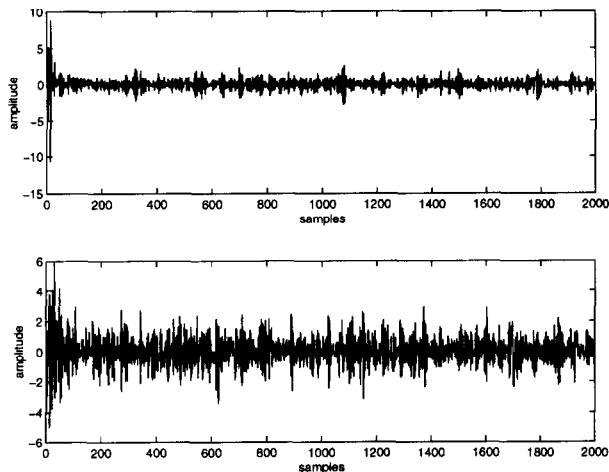Figure 2: Evolution of the demixing matrix coefficients during adaptation

Figure 3: Separated source outputs

scaled versions of the original sources after initial convergence. Experiments with more sources and sensors yield similar results.

5. CONCLUSIONS

Effective blind source separation can be achieved by exploiting nonstationarity of the sources. Nonstationary blind source separation algorithms appear particularly relevant for practical applications because many sources of interest, such as speech or fading signals, exhibit nonstationarity but may not otherwise present features (such as non-Gaussian statistics or different auto-correlation structure) required by other methods. In comparison with other nonstationary blind source separation algorithms, the method proposed here results in a simple and efficient on-line stochastic gradient algorithm requiring only multiplications and additions, which are efficiently implemented in signal processing hardware. It appears to exhibit the traditional characteristics of LMS-like algorithms including excellent robustness and numerical stability, the ability to track slow variations in the environment, and relatively slow convergence.

6. REFERENCES

[1] J. R. Treichler and M. G. Larimore, "New processing techniques based on the constant modulus adaptive algorithm," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, pp. 420–431, April 1985.

[2] C. Jutten and J. Herault, "Blind separation of sources, part I: An adaptive algorithm based on neuromatic architecture," *Signal Processing*, vol. 24, pp. 1–10, July 1991.

[3] J.-F. Cardoso, "Iterative techniques for blind separation using only fourth-order cumulants," in *Signal Processing IV - Theories and Applications, Proceedings fo EUSIPCO-92, Sixth European Signal Processing Conference*, vol. 2, pp. 739–742, 1992.

[4] A. Belouchrani, K. A. Meraim, J.-F. Cardoso, and E. Moulines, "A blind source separation technique using second order statistics," *IEEE Transactions on Signal Processing*, vol. 45, pp. 434–444, February 1997.

[5] A. Belouchrani and M. G. Amin, "Source separation based on the diagonalization of a combined set of spatial time-frequency distribution matrices," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, ICCASP - 97*, (Germany), April 1997.

[6] L. Parra, C. Spence, and B. de Vries, "Convolutive blind source separation based on multiple decorrelation," in *Proceedings of 1998 IEEE Workshop on Neural Networks for Signal Processing*, (Cambridge, UK), September 1998.

[7] K. Matsuoka, M. Ohya, and M. Kawamoto, "A neural net for blind separation of nonstationary signals," *Nueral Networks*, vol. 8, no. 3, pp. 411–419, 1995.