COMPENSATING FREQUENCY RESPONSE OF THE SENSORS IN BLIND SEPARATION OF SOURCES

Massoud BABAIE-ZADEH[1,2], Christian JUTTEN[1], and Kambiz NAYEBI[2]

[1] Institut National Polytéchnique de Grenoble (INPG), Laboratoire des Images et des Signaux (LIS), 46 Avenue Félix Viallet, Grenoble, France

[2] Department of Electrical Engineering, Sharif University of Technology, Tehran, Iran.

mbzadeh@yahoo.com

Christian.Jutten@inpg.fr

knayebi@sina.sharif.ac.ir

ABSTRACT

Blind Source Separation is a basic problem in signal processing. For general convolutive mixtures, the separation criterion is complex because it must apply to stochastic processes. In this paper, we deal with a special case of convolutive mixtures in which the channel creates a linear instantaneous mixture and the frequency response of sensors are modeled by LTI filters. For this case, we show the independence criterion of random variables is sufficient for source separation.

1. INTRODUCTION

Blind Source Separation (BSS) is an important problem in signal processing and has been considered intensively in the last decade. In linear instantaneous case, the mixture is assumed to be of the form:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

where $\mathbf{s}$ is the source vector, $\mathbf{x}$ is the observation vector, and $\mathbf{A}$ is the (constant) mixing matrix. To separate sources, a separation matrix $\mathbf{B}$ must be estimated such that:

$$\mathbf{y} = \mathbf{B}\mathbf{x} \tag{2}$$

In this case, it can be shown [1] that the independence of the components of y, is sufficient to achieve separation (up to a scale and a permutation indeterminacy).

In the convolutive case, the mixing and separating systems are matrices of Linear Time Invariant (LTI) filters [2, 3, 4]. For these mixtures too, it has been shown [5] that the output independence is sufficient for signal separation (up to a permutation and a filtering indeterminacy). However, it must be noted that in this case we deal with the stochastic processes, not with random variables as in the instantaneous case. Therefore, the independence of say y_1 and y_2, cannot be deducted from the sole independence of $y_1(n)$ and $y_2(n)$. Instead, $y_1(n)$ and $y_1(n-m)$ must be independent for all pairs (n, m).

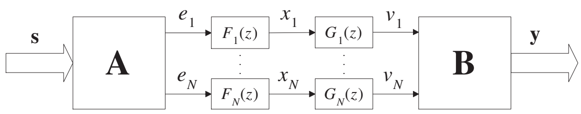

Figure 1: Mixing-Separating system

In this paper, we have addressed a special case of convolutive mixtures where a linear instantaneous mixture is created in the channel, but the frequency response of the sensors and corresponding amplifiers are modeled by filters. We will show that for this special case, independence at instant n, is sufficient for separation. This leads to a very simple algorithm to separate these mixtures.

2. PRELIMINARY ISSUES

For the general convolutive mixtures, it is possible that $y_1(n)$ and $y_2(n)$ be independent, when y_1 and y_2 are not independent [4].

Consider now the system of Fig. 1, where the sources are first mixed by a linear instantaneous system, and then the i-th mixtures is filtered by $F_i(z)$, which corresponds to the frequency response of the sensors or the amplifiers following them. We represent the observations by x_i, $i = 1, \ldots, N$. For separating the sources, one must first compensate for the filtering effect (by estimating the filters G_i), and then estimate the separation matrix. The outputs of filters G_i are represented by v_i, $i = 1, \ldots, N$. For simplicity, we represent the total filter of each branch by:

$$H_i(z) = G_i(z)F_i(z) \tag{3}$$

It is obvious that if:

$$H_i(z) = c_{ij}H_j(z) \qquad \forall i, j \tag{4}$$

and $\mathbf{BA}$ is the product of a permutation and a diagonal matrix, then the outputs will be independent. Therefore, as in the convolutive mixtures, sources can be obtained only up to a filtering indeterminacy.

Relating to this special convolutive mixture, we can state the following theorems, where their proofs are given in the appendix.

Theorem 1 *Let the sources s_1 and s_2 be iid with at most one Gaussian source, and the mixing matrix $\mathbf{A} = \begin{pmatrix} 1 & a_1 \\ a_2 & 1 \end{pmatrix}$ is a nonsingular matrix and $a_1 a_2 \neq -1$. Also suppose that there is an m such that $h_1(m)h_2(m) \neq 0$. Then, if $y_1(n)$ and $y_2(n)$ are independent, $y_1(n)$ and $y_2(n-m)$ will be independent for all m.*

Theorem 2 *Under the assumptions of theorem 1, and if the matrix $\mathbf{A}$ satisfies the condition: $a_1 \neq 0$ or $a_2 \neq 0$, then the independence of $y_1(n)$ and $y_2(n)$ implies that $H_1(z) = cH_2(z)$, where c is an arbitrary constant.*

The first theorem indicates that for this special mixtures, the independence of $y_1(n)$ and $y_2(n)$ is a necessary and sufficient condition for the independence of stochastic processes y_1 and y_2. This theorem can easily be extended to the case $N \times N$, by assuming pair wise independence.

The second theorem shows that if the sources are 'really mixed', then the global filters on each branch are the same, up to a scale factor.

3. ESTIMATING EQUATIONS

For developing estimating equations, we assume that:

1. The compensating filters can be FIR causal filters of order M.
2. The filters $F_i(z)$ are causal.
3. $f_i[0] \neq 0 \quad \forall i$.

By these assumptions, the compensating filters can be chosen of the form:

$$G_i(z) = 1 + \sum_{j=1}^{M} g_{ij} z^{-j} \tag{5}$$

From causality of the filters, and from $f_i[0] \neq 0$ and $g_i[0] \neq 0$, we can deduce $h_i[0] \neq 0$, and hence the theorem 1 can be applied.

Let $\mathbf{y} = (y_1, y_2, \ldots, y_N)^T$ be the output random vector. Then, as in the instantaneous mixtures, we use the mutual information of the outputs, $I(\mathbf{y})$, as the separation criterion. Recall that:

$$I(\mathbf{y}) = \sum_{k=1}^{N} H(y_k) - H(\mathbf{y}) \tag{6}$$

where $H(\cdot) = -E\{\ln p_{(\cdot)}(\cdot)\}$ denotes the Shannon's entropy.

3.1. Instantaneous stage

From:

$$p_{\mathbf{y}}(\mathbf{y}) = \frac{p_{\mathbf{v}}(\mathbf{v})}{|\det(\mathbf{B})|} \tag{7}$$

we have:

$$H(\mathbf{y}) = H(\mathbf{v}) + \ln|\det(\mathbf{B})| \tag{8}$$

Hence, the gradient of $H(\mathbf{y})$ with respect to $\mathbf{B}$ is:

$$\frac{\partial H(\mathbf{y})}{\partial \mathbf{B}} = \frac{\partial}{\partial \mathbf{B}} |\det(\mathbf{B})| = \mathbf{B}^{-T} \tag{9}$$

On the other hand, from the definition of entropy and $y_j = \sum_{k=1}^{N} b_{jk} v_k$, it can be easily shown that:

$$\frac{\partial}{\partial b_{ij}} \left(\sum_{k=1}^{N} H(y_k) \right) = -E\{\psi_i(y_i) v_j\} \tag{10}$$

where $\psi_i(\cdot) = \frac{p'_{y_i}(\cdot)}{p_{y_i}(\cdot)}$ is the score function of y_i.

By combining (9) and (10), we will have:

$$\frac{\partial I(\mathbf{y})}{\partial \mathbf{B}} = -E\left\{\boldsymbol{\psi}_{\mathbf{y}}(\mathbf{y}) \mathbf{v}^T\right\} - \mathbf{B}^{-T} \tag{11}$$

where $\boldsymbol{\psi}_{\mathbf{y}}(\mathbf{y}) = (\psi_1(y_1), \ldots, \psi_N(y_N))^T$.

3.2. Filtering stage

From (8), we have:

$$\begin{aligned} \frac{\partial H(\mathbf{y})}{\partial g_{ij}} = \frac{\partial H(\mathbf{v})}{\partial g_{ij}} &= -E\left\{\frac{\partial}{\partial g_{ij}} \ln p_{\mathbf{v}}(\mathbf{v})\right\} \\ &= -E\left\{\frac{\partial \ln p_{\mathbf{v}}(\mathbf{v})}{\partial v_i} . \frac{\partial v_i}{\partial g_{ij}}\right\} \end{aligned} \tag{12}$$

Thus:

$$\frac{\partial H(\mathbf{y}(n))}{\partial g_{ij}} = -E\{\varphi_{\mathbf{v},i}(\mathbf{v}(n))\, x_i(n-j)\} \tag{13}$$

where $\varphi_{\mathbf{v},i}(\mathbf{x})$ is the i-th component of the joint score function [4] of $\mathbf{v}$, defined by:

$$\begin{aligned} \boldsymbol{\varphi}_{\mathbf{v}}(\mathbf{x}) &= \nabla_{\mathbf{x}} \ln p_{\mathbf{v}}(\mathbf{x}) \\ &= \frac{1}{p_{\mathbf{v}}(\mathbf{x})} \left(\frac{\partial p_{\mathbf{v}}(\mathbf{x})}{\partial x_1}, \cdots, \frac{\partial p_{\mathbf{v}}(\mathbf{x})}{\partial x_N} \right)^T \end{aligned} \tag{14}$$

For calculating the derivative of $\sum H(y_k)$ with respect to g_{ij} we write:

$$\begin{aligned} \frac{\partial H(y_k)}{\partial g_{ij}} &= -E\left\{\frac{\partial}{\partial g_{ij}} \ln p_{y_k}(y_k)\right\} \\ &= -E\left\{\psi_k(y_k) \cdot \frac{\partial y_k}{\partial g_{ij}}\right\} \\ &= -E\left\{\psi_k(y_k) \frac{\partial}{\partial g_{ij}} \left(\sum_{l=1}^{N} b_{kl} v_l\right)\right\} \\ &= -E\left\{b_{ki} \psi_k(y_k) \frac{\partial v_i}{\partial g_{ij}}\right\} \end{aligned} \tag{15}$$

Thus:

$$\frac{\partial}{\partial g_{ij}} \sum_{k=1}^{N} H(y_k) = -E\left\{ \left(\sum_{k=1}^{N} b_{ki}\psi_k\left(y_k(n)\right) \right) x_i(n-j) \right\} \tag{16}$$

By defining the $N \times M$ matrices $\mathbf{G} = (g_{ij})$ and:

$$\mathbf{X}(n) = (\mathbf{x}(n-1), \mathbf{x}(n-2), \cdots, \mathbf{x}(n-M)) \tag{17}$$

and the vector:

$$\boldsymbol{\psi}_{\mathit{diff}}(n) = \boldsymbol{\varphi}_{\mathbf{v}}(\mathbf{v}(n)) - \mathbf{B}^T \boldsymbol{\psi}_{\mathbf{y}}(\mathbf{y}(n)) \tag{18}$$

and applying (13) and (16), we will have:

$$\frac{\partial I(\mathbf{y})}{\partial \mathbf{G}} = E\left\{ diag\left(\boldsymbol{\psi}_{\mathit{diff}}(n) \right) \mathbf{X}(n) \right\} \tag{19}$$

4. THE ALGORITHM

Implementation of the estimation equations, requires estimation of marginal and joint score functions, and we have used kernel estimators [6] to estimate them.

In the instantaneous stage, as it can be seen from (19), the estimating equation is exactly the same as instantaneous mixtures. Therefore, it leads to the following algorithm for estimating $\mathbf{B}$:

$$\mathbf{B}_{n+1} = (I + \mu_1 \mathbf{D})\mathbf{B}_n \tag{20}$$

where:

$$d_{ij} = \begin{cases} 1 - y_i^2 & \text{if } i = j \\ \hat{E}\left\{\psi_i(y_i) y_j\right\} & \text{if } i \neq j \end{cases} \tag{21}$$

and from (19), and by applying the steepest descent algorithm, the filters are estimated by the equation:

$$\mathbf{G}_{n+1} = \mathbf{G}_n - \mu_2 \hat{E}\left\{ diag\left(\boldsymbol{\psi}_{\mathit{diff}}(n) \right) \mathbf{X}(n) \right\} \tag{22}$$

5. EXPERIMENTAL RESULTS

For justifying our method, we have chosen $F_i(z)$ as first order low-pass filters:

$$F_i(z) = \frac{1}{1 - \alpha_i z^{-1}} \tag{23}$$

and the compensating filters of the form:

$$F_i(z) = 1 + g_i z^{-1} \tag{24}$$

Note that by these choices, if the signals are really mixed, we do not deal with the filtering indeterminacy.

In the first experiment, 100 samples of two iid sources with uniform distribution are mixed by matrix:

$$A = \begin{pmatrix} 1 & 0.5 \\ 0.3 & 1 \end{pmatrix}$$

The coefficients of the filters are $\alpha_1 = 0.6$ and $\alpha_2 = 0.8$. We used $\mu_1 = 0.1$ and $\mu_2 = 0.05$. The separation results with the proposed algorithm were averaged through 50 experiments. In average, the estimated values for g_1 and g_2 were -0.5954 and -0.7942, with variances 0.0023 and 0.0008, respectively.

In the second experiment, we have chosen an ill-conditioned mixing matrix:

$$A = \begin{pmatrix} 1 & 5 \\ 1 & 4 \end{pmatrix}$$

In this case the values $\mu_1 = 0.1$ and $\mu_2 = 0.0005$ were used, and the same experiment was repeated. The resulting estimated values for g_1 and g_2 were, in average, 0.6004 and 0.8001, with the variances 7.1×10^{-5} and 3.7×10^{-5}, respectively.

It can be seen that the result of filter estimation is better in the case of hard mixtures. This is because, in this case, algorithm (20) is more sensitive to the distortion introduced by the filters.

Theorem 1 has stated and proved only for iid sources. It is interesting to test the robustness of the algorithm against this hypothesis. Hence, in the third experiment we have mixed two non-iid sources: a sinusoidal and a triangular signal. We have used 100 samples, and the mixing matrix:

$$A = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$$

and step sizes are $\mu_1 = 0.1$ and $\mu_2 = 0.01$. Figure 2 shows the output Signal to Noise Ratios (SNR's) versus iteration. SNR is defined by the relation (assuming no permutation):

$$\text{SNR}_i = 10 \log_{10} \frac{E\left\{y_i^2(n)\right\}}{E\left\{ \left. y_i^2(n) \right|_{s_i(n)=0} \right\}} \tag{25}$$

The figure, clearly shows the ability of the algorithm to separate non-iid sources as well as iid sources.

6. CONCLUSION

In this paper, we have presented an algorithm for separating a special case of convolutive mixtures. For this type of mixtures, instantaneous independence of the outputs is sufficient to estimate the sources up to a filtering indeterminacy. Currently, we are working on its generalization to nonlinear case, where there is a nonlinearity between the mixing matrix and each filter, which models the sensor nonlinearities.

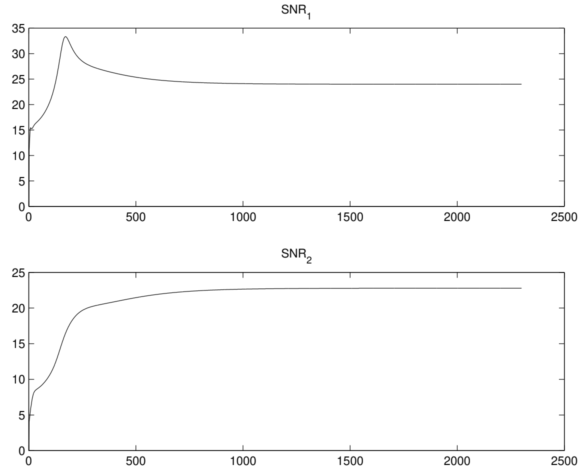

Figure 2: Output SNR's in separating two non-iid sources.

A. APPENDIX

Proof of the theorem 1: By using the separation matrix $B = \begin{pmatrix} 1 & b_1 \\ b_2 & 1 \end{pmatrix}$, we will have:

$$\begin{aligned} y_1(n) = &\sum_{k=-\infty}^{\infty} \{h_1(k) + a_2 b_1 h_2(k)\}\, s_1(n-k) \\ &+ \sum_{k=-\infty}^{\infty} \{a_1 h_1(k) + b_1 h_2(k)\}\, s_2(n-k) \end{aligned} \tag{26}$$

$$\begin{aligned} y_2(n) = &\sum_{k=-\infty}^{\infty} \{b_2 h_1(k) + a_2 h_2(k)\}\, s_1(n-k) \\ &+ \sum_{k=-\infty}^{\infty} \{a_1 b_2 h_1(k) + h_2(k)\}\, s_2(n-k) \end{aligned} \tag{27}$$

Suppose that $y_1(n)$ and $y_2(n)$ are independent. Applying Darmois' theorem [1] for the random variables $s_1(n-k)$ and $s_2(n-k)$, we will have:

$$\begin{cases} b_2 h_1(k) + a_2 h_2(k) = 0 \\ a_1 h_1(k) + b_1 h_2(k) = 0 \end{cases} \tag{28}$$

or:

$$\begin{cases} h_1(k) + a_2 b_1 h_2(k) = 0 \\ a_1 b_2 h_1(k) + h_2(k) = 0 \end{cases} \tag{29}$$

By the assumptions: $\exists m$ such that $h_1(m)h_2(m) \neq 0$. Now suppose that for $k = m$, equation (28) holds. Hence:

$$b_1 = -a_1 \frac{h_1(m)}{h_2(m)} \tag{30}$$

$$b_2 = -a_2 \frac{h_2(m)}{h_1(m)} \tag{31}$$

Let $c = \frac{h_1(m)}{h_2(m)}$ and suppose that at least one of the values of a_1 or a_2 is nonzero. Inserting the above values in (28), we have $h_1(k) = ch_2(k)$, and inserting them in (29) we obtain:

$$\begin{pmatrix} h_2(m) & -a_1 a_2 h_1(m) \\ -a_1 a_2 h_2(m) & h_1(m) \end{pmatrix} \begin{pmatrix} h_1(k) \\ h_2(k) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{32}$$

The determinant of the coefficient matrix of this equation is $\left(1 - a_1^2 a_2^2\right) h_1(m) h_2(m)$ and cannot vanish by the assumptions. Hence, its unique solution is $h_1(k) = h_2(k) = 0$, which is the solution of (28), too. Therefore, (28) is satisfied for all k, and we have always $h_1(k) = ch_2(k)$. Also, we have:

$$y_1(n) = \sum_{k=-\infty}^{\infty} \{h_1(k) + a_2 b_1 h_2(k)\}\, s_1(n-k) \tag{33}$$

$$y_2(n) = \sum_{k=-\infty}^{\infty} \{a_1 b_2 h_1(k) + h_2(k)\}\, s_2(n-k) \tag{34}$$

Hence, it is clear that $y_1(n)$ and $y_2(n-m)$ are independent for all m. The same result can be obtained from the assumption that equation (29) holds for $k = m$. □

The proof of theorem 2 results directly from the above proof. In fact, if $a_1 \neq 0$ or $a_2 \neq 0$, then $\forall k$, $h_1(k) = ch_2(k)$, where c is a real constant. Conversely, if $a_1 = 0$ and $a_2 = 0$, then the sources are not mixed and the outputs will be independent for all filters. □

REFERENCES

[1] P. Comon, "Independent component analysis, a new concept?," *Signal Processing*, vol. 36, no. 3, pp. 287–314, 1994.

[2] H.L. Nguyen Thi and C. Jutten, "Blind sources separation for convolutive mixtures," *Signal Processing*, vol. 45, pp. 209–229, 1995.

[3] C. Simon, *Séparation aveugle des sources en mélange convolutif*, Ph.D. thesis, l'université de Marne la Vallée, Novembre 1999, (in French).

[4] M. Babaie-Zadeh, C. Jutten, and K. Nayebi, "Separating convolutive mixtures by mutual information minimization," in *Proceedings of IWANN'2001*, Granada, Spain, Juin 2001, pp. 834–842.

[5] D. Yellin and E. Weinstein, "Criteria for multichannel signal separation," *IEEE Trans. Signal Processing*, pp. 2158–2168, August 1994.

[6] W. Härdle, *Smoothing techniques with implementation in S*, Springer-Verlag, 1991.