October 2004

Decorrelation: sufficient for convolutive blind source separation?

Jiangtao Xi
University of Wollongong, jiangtao@uow.edu.au

T. Mei
Dalian University of Technology, China, tiemin@uow.edu.au

Joe F. Chicharo
University of Wollongong, chicharo@uow.edu.au

F. Yin
Dalian University of Technology, China

Decorrelation: sufficient for convolutive blind source separation?

Abstract

This paper considers blind separation of signal sources in a convolutive mixing environment. It tries to show that decorrelation is sufficient for separation of convolutively mixed sources. Two algorithms are also proposed and tested by computer simulations.

Keywords

blind source separation, convolution, decorrelation

Disciplines

Physical Sciences and Mathematics

Publication Details

This paper originally appeared as: Xi, J, Mei, T, Chicharo, JF & Yin, F, Decorrelation: sufficient for convolutive blind source separation?, Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 20-22 October 2004, 475-478. Copyright IEEE 2004.

DECORRELATION: SUFFICNET FOR CONVOLUTIVE BLIND SOURCE SEPARATION?

Jiangtao Xi[1], Tiemin Mei[2], Joe Chicharo[1] and Fuliang Yin[2]

[1]School of Electrical,Computer and Telecommunications Engineering,
The University of Wollongong, NSW 2522, Australia, email: jiangtao@uow.edu.au

[2]School of Electronic and Information Engineering
Dalian University of Technology, Dalian, Liaoning, 116023 P. R. China

ABSTRACT

This paper considers blind separation of signal sources in convolutive mixing environment. It tries to show that decorrelation is sufficient for separation of convolutively mixed sources. Two algorithms are also proposed and tested by computer simulations.

1. INTRODUCTION

Blind source separation (BSS) has been an important area of signal processing due to its many potential applications, such as communication, speech and image processing, and various biomedical signal processing problems, etc. [1] [2]. There have been many different approaches for the BSS, including those based on non-linear neural networks [3], high order statistics (HOS) or polyspectra [4], and others. These algorithms are generally very complicated which require extensive computation. A simpler statistical approach is the one based on Second-order Statistics (SOS) or output decorrelation, e.g. [5-7]. Given the fact that SOS-based approaches are much simpler than those HOS-based approaches, we try to show that SOS is sufficient for BSS.

This paper is organized as follows. Section 2 gives the system model for the BSS problem. In Section 3 the condition of output de-correlation is analyzed, with the results that under certain conditions, output decorrelation is a sufficient condition for BSS. New SOS based algorithms are proposed in Section 4. Computer simulation results are given in Section 5. Finally Section 6 concludes the paper.

2. PROBLEM STATEMENTS

Consider the well-known BSS problem in convolutive environment: Two signal sources $s_1(t)$ and $s_2(t)$ are of zero mean and statistically uncorrelated stochastic processes. The two signal sources are mixed together via convolutive channels and yield two mixed signal measurements: $x_1(t)$ and $x_2(t)$. We assume that the channels are modelled by FIR filters with equal lengths. The relations between source signals and observations can be expressed mathematically as follows

$$x_1(t) = a_{11}(t) * s_1(t) + a_{12}(t) * s_2(t) \quad (1)$$

$$x_2(t) = a_{21}(t) * s_1(t) + a_{22}(t) * s_2(t) \quad (2)$$

The purpose is to recover the signal sources from those two measurements using a separation system with the similar structure:

$$u_1(t) = w_{11}(t) * x_1(t) + w_{12}(t) * x_2(t) \quad (3)$$

$$u_2(t) = w_{21}(i) * x_1(t) + w_{22}(t) * x_2(t) \quad (4)$$

Substituting (1)(2) into (3)(4) gives:

$$u_1(t) = h_{11}(t)*s_1(t)+h_{12}(t)*s_2(t) = v_1(t)+v_2(t) \quad (5)$$

$$u_2(t) = h_{21}(t)*s_1(t)+h_{22}(t)*s_2(t) = v_3(t)+v_4(t) \quad (6)$$

where $v_1(t)=h_{11}(t)*s_1(t)$, $v_2(t)=h_{12}(t)*s_2(t)$, $v_3(t)=h_{21}(t)*s_1(t)$ and $v_4(t)=h_{22}(t)*s_2(t)$, and

$$h_{11}(t)=a_{11}(t)*w_{11}(t)+a_{21}(t)*w_{12}(t) \quad (7)$$

$$h_{22}(t)=a_{12}(t)*w_{21}(t)+a_{22}(t)*w_{22}(t) \quad (8)$$

$$h_{12}(t)=a_{12}(t)*w_{11}(t)+a_{22}(t)*w_{12}(t) \quad (9)$$

$$h_{21}(t)=a_{11}(t)*w_{21}(t)+a_{21}(t)*w_{22}(t) \quad (10)$$

Note that $v_1(t)$ and $v_3(t)$ originate from $s_1(t)$ only, and $v_2(t)$ and $v_4(t)$ originate from $s_2(t)$ only.

3. DECORRELATION CONDITION

In this section, we will see what happens if we chose the separate filters $H_{ij}(z)$ (for $i,j=1,2$) so that $u_1(t)$ and $u_2(t)$ are uncorrelated. In other words, we want to show that BBS can be achieved under this condition.

The cross-correlation function of $u_1(t)$ and $u_2(t)$ is defined as follows:

$$r_{u_1u_2}(\tau) = E[u_1(t)u_2(t-\tau)] \quad (11)$$

Making use of the assumption of un-correlation between the two signal sources, and also assuming that both the mixing and separation channels do not change their mean value, we have:

$$r_{u_1u_2}(\tau) = r_{v_1v_3}(\tau) + r_{v_2v_4}(\tau) \quad (12)$$

We want to investigate what happens when the above cross-correlation function are set to zero:

$$r_{u_1u_2}(\tau) = r_{v_1v_3}(\tau) + r_{v_2v_4}(\tau) = 0 \quad (13)$$

Under the constrain that the $u_i(t) \neq 0$ (for i=1,2). There are two possible situations that make $r_{u_1u_2}(\tau)$ to be zero. One is that neither $r_{v_1v_3}(\tau)$ nor $r_{v_2v_4}(\tau)$ is zero. Hence $r_{u_1u_2}(\tau)=0$ when and only when $r_{v_1v_3}(\tau)=-r_{v_2u_4}(\tau)$, i.e. they are inversely symmetrical to each other. However, this is almost impossible as the length of mixing filters and the two sources can be very different in nature. Therefore we can assume that $r_{v_1v_3}(\tau)\neq -r_{v_2u_4}(\tau)$. In this case we must have:

$$r_{v_1v_3}(\tau)=0 \text{ and } r_{v_2u_4}(\tau)=0 \tag{14}$$

Let see $r_{v_1v_3}(\tau)=0$ first.

$$\begin{aligned} r_{v_1v_3}(\tau) &= E[v_1(t)v_3(t-\tau)] \\ &= E[(h_{11}(t)*s_1(t))(h_{21}(t-\tau)*s_1(t-\tau))] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h_{11}(\phi)h_{21}(\varphi)r_{s_1s_1}(\phi-\varphi-\tau)d\phi d\varphi \end{aligned} \tag{15}$$

where $r_{s_1s_s}(\tau)=E[s_1(t)s_s(t-\tau)]$ is the autocorrelation function of signal source $s_1(t)$, which is non-zero for at least some values of τ. Hence $r_{v_1v_3}(\tau)=0$ for all τ implies that

$$h_{21}(t)=0 \text{ or } h_{11}(t)=0 \tag{16}$$

Similarly for $r_{v_2v_4}(\tau)=0$, we can also have:

$$h_{12}(t)=0 \text{ or } h_{22}(t)=0 \tag{17}$$

There are four possible situations corresponding to (16) and (17):

Case 1: $h_{21}(t)=0$ and $h_{11}(t)\neq 0$ and $h_{12}(t)=0$ and $h_{22}(t)\neq 0$;

Case 2: $h_{21}(t)=0$ and $h_{11}(t)\neq 0$ and $h_{12}(t)\neq 0$ and $h_{22}(t)=0$;

Case 3: $h_{21}(t)\neq 0$ and $h_{11}(t)=0$ and $h_{12}(t)=0$ and $h_{22}(t)\neq 0$;

Case 4: $h_{21}(t)\neq 0$ and $h_{11}(t)=0$ and $h_{12}(t)\neq 0$ and $h_{22}(t)=0$;

From (5) and (6) we can see that Case 2 corresponds to $u_2(t)=0$ and Case 3 to $u_1(t)=0$ which are contrary to the constrain $u_i(t)\neq 0$ (for i=1,2). Therefore only Case 1 and Case 4 are possible.

Let go back to Case 1 again, from (5) and (6) we have

$$u_1(t)=h_{11}(t)*s_1(t) \tag{19}$$

$$u_2(t)=h_{22}(t)*s_2(t) \tag{20}$$

Obviously BSS is achieved up to an unknown filtering factor.

Let's consider $h_{12}(t)=0$, and taking the Z-transform gives:

$$A_{12}(z)W_{11}(z)+A_{22}(z)W_{12}(z)=0 \tag{21}$$

Assuming that $A_{12}(z)$ and $A_{22}(z)$ do not have common zeros, we must have [8]:

$$W_{11}(z)=\pm A_{22}(z) \text{ and } W_{12}(z)=\mp A_{12}(z) \tag{22}$$

Similarly, if $A_{11}(z)$ and $A_{21}(z)$ do not have any common zero, $h_{21}(t)=0$ leads to:

$$W_{22}(z)=\pm A_{11}(z) \text{ and } W_{21}(z)=\mp A_{21}(z) \tag{23}$$

Equations (22) and (23) imply that the unknown mixing channels can be identified as soon as BSS is achieved. In this case we have:

$$H_{22}(z)=H_{11}(z)=\pm W_{11}(t)W_{22}(z)\mp W_{21}(z)W_{12}(z) \tag{24}$$

Similar analysis can be done for Case 4 with the following results:

$$u_1(t)=h_{12}(t)*s_2(t) \tag{25}$$

$$u_2(t)=h_{21}(t)*s_1(t) \tag{26}$$

and

$$H_{12}(z)=H_{21}(z)=\pm W_{11}(t)W_{22}(z)\mp W_{21}(z)W_{12}(z) \tag{27}$$

Obviously no matter in Case 1 or Case 4, BSS can always be achieved and the output signals are filtered version of the original sources. Also it is interesting to note that the filter for both output and both cases are the same:

$$H(z)=\pm W_{11}(t)W_{22}(z)\mp W_{21}(z)W_{12}(z) \tag{28}$$

Original sources can be restored from $u_1(t)$ and $u_2(t)$ by de-convolution, using the following system:

$$E(z)=\frac{1}{H(z)}=\frac{1}{\pm W_{11}(t)W_{22}(z)\mp W_{21}(z)W_{12}(z)} \tag{29}$$

In other words, source signals can be restored by:

$$\hat{S}_i(t)=y_j(t)=u_j(t)*e(t) \quad \text{for } i,j=1,2 \tag{30}$$

Hence up to now we have proved that de-correlation of output signals might be a sufficient condition for BSS. Also the mixing channels can be identified as soon as BSS is achieved, based on which a de-convolution filter can be designed for restoring the original signal sources.

4. A SIMPLE MIXING CASE

In this section we consider a simple case when $a_{11}(t)=a_{22}(t)=\delta(t)$, that is

$$x_1(t)=s_1(t)+\sum_{i=0}^{n}a_{21}(i)s_2(t-i) \tag{31}$$

$$x_2(t)=s_2(t)+\sum_{i=0}^{n}a_{12}(i)s_1(t-i) \tag{32}$$

for the separation systems we also have we have $w_{11}(t)=w_{22}(t)=\delta(t)$. The outputs are

$$u_1(t)=v_1(t)+v_2(t) \tag{33}$$

$$u_2(t) = v_3(t) + v_4(t) \quad (34)$$

where

$$v_1(t) = s_1(t) - \sum_{k=0}^{n}\sum_{i=0}^{n} c(i)b(k-i)s_1(t-k) \quad (35)$$

$$v_2(t) = \sum_{k=0}^{n}[a(k) - c(k)]s_2(t-k) \quad (36)$$

$$v_3(t) = \sum_{k=0}^{n}[b(k) - d(k)]s_1(t-k) \quad (37)$$

$$v_4(t) = s_2(t) - \sum_{k=0}^{n}\sum_{i=0}^{n} d(i)a(k-i)s_2(t-k) \quad (38)$$

Now let us investigate the behaviours of the cross-correlation functions $r_{v_1v_3}(\tau)$ and $r_{v_2v_4}(\tau)$. Obviously that they all attenuate to zero when $|\tau| \to \infty$. However, from the above equations we see that the attenuating trends of $r_{v_1v_3}(\tau)$ and $r_{v_2v_4}(\tau)$ are very different: $r_{v_1v_3}(\tau)$ attenuates to zero more quickly when $\tau < 0$ than that when $\tau > 0$. On the contrary, $r_{v_2v_4}(\tau)$ attenuates to zero more slowly when $\tau < 0$ than that when $\tau > 0$. Therefore they are definitely not inversely symmetrical to each other. Hence the unique possible solution for $r_{u_1u_2}(\tau) = r_{v_1v_3}(\tau) + r_{v_2u_4}(\tau) = 0$ is that $r_{v_1v_3}(\tau) = 0$ and $r_{v_2v_4}(\tau) = 0$.

The mixing network and the de-convolution stage can be implemented together as recursive structure given by [6]

$$y_1(t) = x_1(t) - \sum_{i=0}^{n} c(i)y_2(t-i) \quad (39)$$

$$y_2(t) = x_2(t) - \sum_{i=0}^{n} d(i)y_1(t-i) \quad (40)$$

It is obvious that de-correlation of $u_1(t)$ and $u_2(t)$ is equivalent to the de-correlation of $y_1(t)$ and $y_2(t)$. Hence separation algorithms can be developed based on the de-correlation of $y_1(t)$ and $y_2(t)$. Without the loss of generality, we assume that $a_{12}(0) = 0, a_{21}(0) = 0$, $c(0) = 0$ and $d(0) = 0$, then $y_1(t)$ is only related to the past value of $y_2(t)$ and $y_2(t)$ related to the past value of $y_1(t)$. Such a system is strict causal coupling system [11].

Let

$$Y_1(t) = [y_1(t-1), \ldots, y_1(t-n)]^T \quad (41)$$

$$Y_2(t) = [y_2(t-1), \ldots, y_2(t-n)]^T \quad (42)$$

and

$$W_1 = [c(1), c(2), \ldots, c(n)]^T \quad (43)$$

$$W_2 = [d(1), d(2), \ldots, d(n)]^T \quad (44)$$

Under decorrelation condition, we can obtain two coupled Wiener-Hopf equations,

$$R_{Y_2Y_2}W_1 = E[x_1(t)Y_2(t)] \quad (45)$$

$$R_{Y_1Y_1}W_2 = E[x_2(t)Y_1(t)] \quad (46)$$

where $R_{Y_iY_i} = E[Y_i(t)Y_i^T(t)], \quad (i = 1,2)$.

The two coupled Wiener-Hopf equations (40) and (41) can be solved with the standard LMS or RLS algorithm. We name them as the Double-LMS and -RLS algorithms.

Double-LMS algorithms:

$$\begin{aligned} W_1(t) &= W_1(t-1) + \mu_1 y_1(t)Y_2(t) \\ W_2(t) &= W_2(t-1) + \mu_2 y_2(t)Y_1(t) \end{aligned} \quad (47)$$

Double- RLS algorithms:

$$\begin{aligned} W_1(t) &= W_1(t-1) + y_1(t)K_1(t) \\ W_2(t) &= W_2(t-1) + y_2(t)K_2(t) \end{aligned} \quad (48)$$

where

$$K_1(t) = \frac{P_1(t-1)Y_2(t)}{\lambda_1 + Y_2^T(t)P_1(t-1)Y_2(t)} \quad (49)$$

$$K_2(t) = \frac{P_2(t-1)Y_1(t)}{\lambda_2 + Y_1^T(t)P_2(t-1)Y_1(t)} \quad (50)$$

$$P_1(t) = \frac{1}{\lambda_1}\left[P_1(t-1) - K_1(t)Y_2^T(t)P_1(t-1)\right] \quad (51)$$

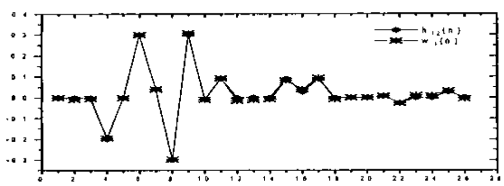
$$P_2(t) = \frac{1}{\lambda_2}\left[P_2(t-1) - K_2(t)Y_1^T(t)P_2(t-1)\right] \quad (52)$$

Note that the Double-LMS and Double-RLS algorithms require that the convolutive mixing filters should be that $a_{12}(0) = 0, a_{21}(0) = 0$. These constraints are not strict in practice. In fact, many practical systems can meet these requirements provided that sensors are placed far away enough from each other so that signals reach the sensors at different time, and the time-delay is longer than one sample period.

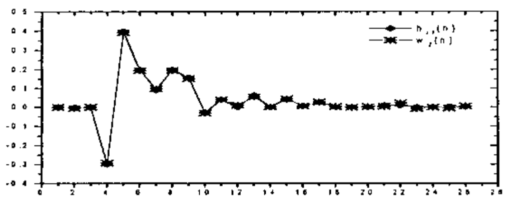
5. NUMERICAL EXPERIMENTS

Computer simulations are performed with the algorithms proposed in Section 4. Different convolutive mixing models are used in the simulations. If the two sources are white noises, then the separation filters will approximate the mixing filters. Contrarily, if the two sources are speeches, then the separation filters usually do not approximate the mixing filters, but the separation filters will approximate the mixing filters in frequency domain in the nonzero frequency bands of sources. Experimental results are pres below.

Fig. 1 shows the separation result when the sources are two white Gaussian noises. It is seen that the separation filter closely matches that of the mixing filter. Fig. 2 is the separation result when the sources are two speeches, from which we can see that there is not a match between the separation filters and mixing filters.

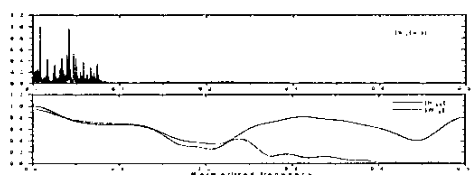


(a)

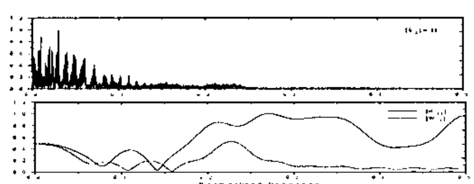


(b)

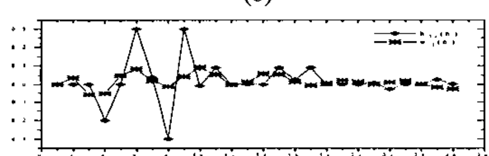
Fig.1. The separation of mixed white noises. (a): The impulse response of mixing and separation filters $a_{12}(n)$ and $w_1(n)$; (b): The impulse response of mixing and separation filters $a_{21}(n)$ and $w_2(n)$.



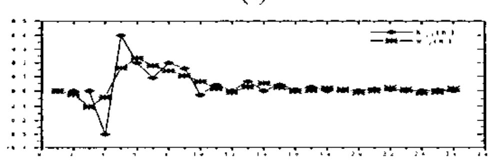
(a)



(b)



(c)



(d)

Fig.2. The separation results of two convolutely mixed speeches. (a): The spectrum of $s_1(t)$ (top), and the spectrum of $a_{12}(n)$ and $w_1(n)$ (bottom); (b): The spectrum of $s_2(t)$ (top), and the spectrum of $a_{21}(n)$ and $w_2(n)$ (bottom). (c): The impulse responses of mixing and separation filters $a_{12}(n)$ and $w_1(n)$; (d): The impulse responses of mixing and separation filters $a_{21}(n)$ and $w_2(n)$.

6. CONCLUSIONS

In this paper, we proved in time domain that, under some constraints, convolutely mixed sources can be separated by utilizing only their second-order statistics (SOS) without the aid of any other information. Also for the simple mixing structure where $a_{11}(t)=a_{22}(t)=\delta(t)$, blind separation is equivalent to the optimum filtering problem. We propose the Double-LMS and -RLS algorithms based on conventional standard LMS and RLS algorithms. These algorithms are very simple in implementation. They are independent on the statistical properties of source signals. Numerical experiments show that these algorithms are valid.

REFERENCE

[1] C. Jutten and J. Herault, "Blind separation of sources, Part I: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, pp. 1-10, 1991.

[2] P. Comon, "Independent component analysis, A new concept?," *Signal Processing*, vol. 36, pp. 287-314, 1994.

[3] S. Amari and A. Cichocki, "Adaptive blind signal processing-neural network approaches," *Proc. of the IEEE*, vol. 86, no. 10, pp. 2026-2048, 1998.

[4] J. F. Cardoso, "Blind signal separation: Statistical Principles," *Proc. of the IEEE*, vol.86, no.10, pp.2009-2025, 1998

[5] D. Yellin and E. Weinstein, "Multichannel signal separation: methods and analysis," *IEEE Trans. Signal Processing*, vol. 44, no. 1, pp. 106--118, 1996.

[6] E. Weinstein, M. Feder and A. V. Oppenheim, "Multichannel signal separation by decorrelation," *IEEE Trans. Speech and Audio Processing*, vol. 1, no. 4, pp. 405-413, 1993.

[7] U. A. Lindgren and H. Broman, "Source separation using a criterion based on second-order statistics," *IEEE Trans. Signal Processing*, vol. 46, no. 7, pp. 837-1850, 1998.

[8] H. Liu, G. Xu and L. Tang, "A deterministic approach to blind identification of multi-channel FIR filters," *ICASSP'94*, pp.IV581-584, 1994