Blind Source Separation with Filtered Time Delay Decorrelation

Ryszard Szupiluk, Krzysztof Siwek, Sławomir Jabłonowski

Abstract — **In this paper we propose an algorithm for blind signals separation in the presence of additive noise. Our method is based on second order statistics and spatio-temporal whitening. We propose using filtered time delay covariance matrix for noisy signals. The computer simulation experiment confirms validity of our conception.**

I. Introduction

BLIND signal separation (BSS) is a fundamental problem that is encountered in many practical applications such as telecommunications, array signal processing, image processing, speech processing, multiple sensor biomedical signals. The BSS task is to estimate the source without having special information about the sources and additive noise mixed with the original source. Many methods and tools for solving the BSS problem have been developed e.g. neural networks, higher order statistics, Kalman filters [1,7].

The main property we explore in this paper, addresses to nonstationary structure of sources. Mainly we are interested in the second-order nonstationarity (in the sense that sources have time-varying variance) which leads to second-order BSS algorithms [2,3,5,6,11].

The method presented in this paper can be an treaded as extension of AMUSE algorithm [7,12]. Especially we focus on robust properties to additive noise influence. To resolve that problem we need a generative model for observed signals. We make an assumption that the m-dimensional observation vector x(t) is generated by

$$\boldsymbol{x}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{v}(t) \tag{1}$$

where:

1. $\boldsymbol{A} \in \Re^{m \times n}$ is the unknown full column rank mixing matrix with $m \geq n$,
2. $\boldsymbol{s}(t)$ is the n-dimensional source vector, $\boldsymbol{v}(t)$ is the additive noise vector that is assumed to be statistically independent of $\boldsymbol{s}(t)$.
3. Sources are spatially uncorrelated with different autocorrelation functions,
4. Additive noises $\boldsymbol{v}(t)$ are spatially correlated but temporally white $E[\boldsymbol{v}(\mathrm{t})] = \boldsymbol{0}$ and $E[\boldsymbol{v}(r)\boldsymbol{v}(s)^T] = \boldsymbol{R_v}\delta_{rs}$ where: δ_{rs} is the Kronecker symbol, $\boldsymbol{R_v}$ is a diagonal noise correlation matrix, $E[.]$ is expectation operator.

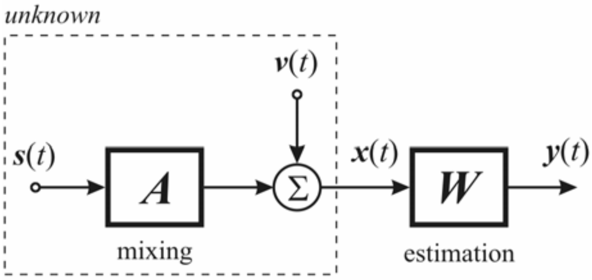

Fig. 1. Blind separation model: mixing, adding noise and estimation.

The task of BSS is to estimate the demixing matrix $\boldsymbol{W}$ in order to estimate original source signals. The basis difficulty is lack of knowledge both $\boldsymbol{A}$ and $\boldsymbol{s}$ in (1). This fact introduces two ambiguities in the solution. It is impossible to recover original variance and order of $\boldsymbol{s}$. So, we accept estimated signals rescaled and reordered in comparison with original signals [7]. Let $\boldsymbol{W}$ describe demixing matrix and

$$\boldsymbol{y}(t) = \boldsymbol{W}\boldsymbol{x}(t) = \boldsymbol{W}\boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{W}\boldsymbol{v}(t) \tag{2}$$

Our aim is to find such matrix $\boldsymbol{W}$ that

$$\boldsymbol{G} = \boldsymbol{W}\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D} \tag{3}$$

where $\boldsymbol{G}$ is the global transformation matrix which combines the mixing and separating system ($\boldsymbol{G}$ is called the generalized permutation matrix), $\boldsymbol{P}$ is some permutation matrix, $\boldsymbol{D}$ is some scaling nonsingular diagonal matrix [1,6,7]. After (3) is achieved separated signals can be covered by random noise witch can be eliminated by classical filtering [8,10].

II. Second-order statistics and diagonalisation

To present our method we start with second order statistic properties and diagonalization processes for matrix separation in noiseless case. Let's define time delay correlation matrix [9]

$$\boldsymbol{R_s}(p) = \mathrm{E}[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{T}}(t\text{-}p)] \tag{4}$$

Ryszard Szupiluk and Sławomir Jabłonowski are with Warsaw School of Economics, Al. Niepodleglosci 162, 02-554 Warsaw, Poland, Ryszard Szupiluk is also with Polska Telefonia Cyfrowa Ltd., Al. Jerozolimskie 181, 02-222 Warsaw, Poland. Krzysztof Siwek is with the Institute of Theory of Electrical Engineering, Measurement and Information Systems, Warsaw University of Technology, ul. Koszykowa 75, 00-662 Warsaw, Poland, e-mail: *rszupiluk@era.pl*

and to obtain matrix with symmetric properties we formulate [6]

$$\widetilde{\boldsymbol{R}}_{\boldsymbol{s}}(p)=\frac{1}{2}\left\{\boldsymbol{R}_{\boldsymbol{s}}(p)+\boldsymbol{R}_{\boldsymbol{s}}^{T}(p)\right\} \tag{5}$$

For simplicity we assume that $E[\boldsymbol{s}(t)]=0$. Notice that if we set p=0 we have standard correlation matrix $\boldsymbol{R}_{\boldsymbol{s}}(0)=E[\boldsymbol{s}(t)\boldsymbol{s}^{\mathrm{T}}(t)]$. In the same way we build correlation matrices for observed and separated signals denoted as $\boldsymbol{R}_{\boldsymbol{x}}(p)$, $\widetilde{\boldsymbol{R}}_{x}(p)$, $\boldsymbol{R}_{\boldsymbol{y}}(p)$, $\widetilde{\boldsymbol{R}}_{y}(p)$. With diagonal correlation matrix of sources $\boldsymbol{R}_{s}(0)$, and with stationary white additive noise ($\boldsymbol{R}_{v}(p)=0$, $p\neq 0$), we get:

$$\boldsymbol{R}_{\boldsymbol{x}}(0)=\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{s}}(0)\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{R}_{\boldsymbol{v}}(0) \tag{6}$$

$$\boldsymbol{R}_{\boldsymbol{x}}(p)=\boldsymbol{A}\boldsymbol{R}_{\boldsymbol{s}}(p)\boldsymbol{A}^{\mathrm{T}} \tag{7}$$

The following theorem will be useful our considerations [6,7].

Theorem 1

Let $\boldsymbol{\Lambda}_1,\boldsymbol{\Lambda}_2,\boldsymbol{D}_1,\boldsymbol{D}_2\in\Re^{\mathrm{n\times n}}$ be diagonal matrices with nonzero diagonal entries. Suppose that $\boldsymbol{G}=\Re^{\mathrm{n\times n}}$ satisfies the following decompositions:

$$\boldsymbol{D}_1=\boldsymbol{G}\boldsymbol{\Lambda}_1\boldsymbol{G}^{\mathrm{T}}, \tag{8}$$

$$\boldsymbol{D}_2=\boldsymbol{G}\boldsymbol{\Lambda}_2\boldsymbol{G}^{\mathrm{T}}, \tag{9}$$

where: $\boldsymbol{D}_1^{-1}\boldsymbol{D}_2$ and $\boldsymbol{\Lambda}_1^{-1}\boldsymbol{\Lambda}_2$ have distinct diagonal entries then the matrix $\boldsymbol{G}$ is the generalized permutation matrix.

Correlation matrices of observed signals $\widetilde{\boldsymbol{R}}_{\boldsymbol{x}}(p_1)$ and $\widetilde{\boldsymbol{R}}_{\boldsymbol{x}}(p_2)$, where $p_1\neq p_2\neq 0$ allow to use the Theorem 1. First we make transformation

$$\boldsymbol{z}(t)=\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_1^{T}\boldsymbol{x}(t) \tag{10}$$

where $\boldsymbol{D}_1$, $\boldsymbol{U}_1$ are from eigenvalue decomposition

$$\widetilde{\boldsymbol{R}}_{\boldsymbol{x}}(p_1)=\boldsymbol{U}_1\boldsymbol{D}_1\boldsymbol{U}_1^{T} \tag{11}$$

The correlation of transformed signals is

$$\widetilde{\boldsymbol{R}}_{z}(p_1)=\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_1^{T}\widetilde{\boldsymbol{R}}_{x}(p_1)\boldsymbol{U}_1\boldsymbol{D}_1^{-1/2}=\boldsymbol{I} \tag{12}$$

Performing EVD decomposition of $\widetilde{\boldsymbol{R}}_{z}(p_2)$ for chosen p_2 we have

$$\widetilde{\boldsymbol{R}}_{z}(p_2)=\boldsymbol{U}_2\boldsymbol{D}_2\boldsymbol{U}_2^{T} \tag{13}$$

Transformation of the form

$$\boldsymbol{y}(t)=\boldsymbol{U}_2^{T}\boldsymbol{z}(t) \tag{14}$$

gives us

$$\widetilde{\boldsymbol{R}}_{y}(p_2)=\boldsymbol{U}_2^{T}\widetilde{\boldsymbol{R}}_{z}(p_2)\boldsymbol{U}_2=\boldsymbol{D}_3 \tag{15}$$

where $\boldsymbol{D}_3$ is a diagonal matrix. With assumption that sources have diagonal correlation matrices $\boldsymbol{R}_{s}(p)$ we can write

$$\boldsymbol{U}_2^{T}\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_1^{T}\boldsymbol{A}\widetilde{\boldsymbol{R}}_{s}(p_1)\boldsymbol{A}^{T}\boldsymbol{U}_1\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_2=\boldsymbol{I} \tag{16}$$

$$\boldsymbol{U}_2^{T}\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_1^{T}\boldsymbol{A}\widetilde{\boldsymbol{R}}_{s}(p_2)\boldsymbol{A}^{T}\boldsymbol{U}_1\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_2=\boldsymbol{D}_3 \tag{17}$$

Based upon Theorem 1 matrix $\boldsymbol{G}=\boldsymbol{U}_2^{T}\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_1^{T}\boldsymbol{A}$ is the generalized permutation matrix and $\boldsymbol{W}=\boldsymbol{U}_2^{T}\boldsymbol{D}_1^{-1/2}\boldsymbol{U}_1^{T}$ is a demixing matrix (separating matrix). The properties described above provide to efficient algorithms. But one of main problem is how to reduce influence of additive noise.

To increase robust properties for additive noise we propose filtered time-delayed correlation matrices defined as

$$\overline{\boldsymbol{R}}_{\boldsymbol{x}}(p)=\mathrm{E}[\boldsymbol{x}(t)\overline{\boldsymbol{x}}^{\mathrm{T}}(t\text{-}p)] \tag{18}$$

where $\overline{\boldsymbol{x}}(t)$ is filtered version of $\boldsymbol{x}(t)$. It can be simple FIR filtration described as

$$\overline{\boldsymbol{x}}(t)=\sum_{k=1}^{K}b(k)\boldsymbol{x}(t-k) \tag{19}$$

where $b(k)$ are filter coefficients the same for all signals in $\boldsymbol{x}$. Similar to previous analysis we define symmetric filtered time delay correlation matrix

$$\breve{\boldsymbol{R}}_{\boldsymbol{x}}(p)=\frac{1}{2}\left\{\overline{\boldsymbol{R}}_{\boldsymbol{x}}(p)+\overline{\boldsymbol{R}}_{x}^{T}(p)\right\} \tag{20}$$

The exploration of above formula leads us to iterative separation algorithm which for simplicity will be described for equal number source and observed signals.

III. ALGORITHM OUTLINE: FILTERED MULTISTAGE TIME DELAY DECORRELATION (FMTDD)

We propose a new algorithm, called Filtered Multistage Time Delay Decorrelation (FMTDD).

1. Let $\boldsymbol{z}(t)=\boldsymbol{x}(t)$, p=0.

2. Estimate the correlation matrix of observed signals with mixing matrix.

$$\widetilde{\boldsymbol{R}}_z(p) = \mathrm{E}\left[z(t)\bar{z}^T(t-p)\right] \approx \frac{1}{N}\sum_{t=1}^{N} z(t)\bar{z}^T(t-p) \quad (21)$$

where

$$\bar{z}(t) = \sum_{k=1}^{K} b(k)z(t-k) \quad (22)$$

$\bar{z}(t)$ is filtered version of $z(t)$, it can be simply smoothing process with not large order.

3. Compute the symmetric matrix

$$\breve{\boldsymbol{R}}_z(p) = \frac{1}{2}\left\{\overline{\boldsymbol{R}}_z(p) + \overline{\boldsymbol{R}}_z^T(p)\right\}$$

4. Apply the eigenvalue decomposition of matrix $\boldsymbol{R}_z(p)$

$$\boldsymbol{R}_z(p) = \boldsymbol{U}\boldsymbol{\Sigma}\,\boldsymbol{U}^T \quad (23)$$

where $\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2, ..., \boldsymbol{u}_m] = \Re^{m \times m}$ is orthogonal matrix of eigenvectors, $\boldsymbol{\Sigma} = diag\{\sigma_1, ..., \sigma_m\}$ is diagonal matrix of eigenvalues.

5. Perform decorrelation for set delay

$$\boldsymbol{y}(t) = \boldsymbol{\Sigma}^{-1/2}\boldsymbol{U}^T z(t) \quad (24)$$

6. Let $z(t) = \boldsymbol{y}(t)$, p=p+1 and we repeat steps 2-6

Typical number of iterations is not large (between 2 and 20). FMTDD algorithm explore the same idea as SOBI or AMUSE [3,12], but estimates a mixing matrix from a filtered time delayed correlation matrix. Due to filtration process in correlation matrix calculation FMTDD is robust to additive noise in significant level.

In the case without additive noise we can use standard time delay covariance matrix (against filtered matrix). Note that such algorithm is vary fast, by its simplicity, and gives good results even if there are many sources.

For more sensors than sources $n < m$ the number of sources can be detected by inspecting the dominant singular values $\mathbf{R}_x(0)$ [7].

IV. Computer simulation

We consider problem of blind signal separation in the presence of the large additive noise. Noise level is given by signal to noise ratio *SNR*, defined as

$$SNR_i = 10\log_{10}\left(\mathrm{var}(\hat{x}_i)/\mathrm{var}(v_i)\right) \quad (25)$$

where $\mathrm{var}(\hat{x}_i)$ is a variance of mixed sources without the additive noise, $\mathrm{var}(v_i)$ variance of additive noises. We generate five source signals and mixing matrix **A**. All the elements of mixing matrix are drawn form uniform distribution (-1,1). Both signals and matrix are assumed to be unknown.

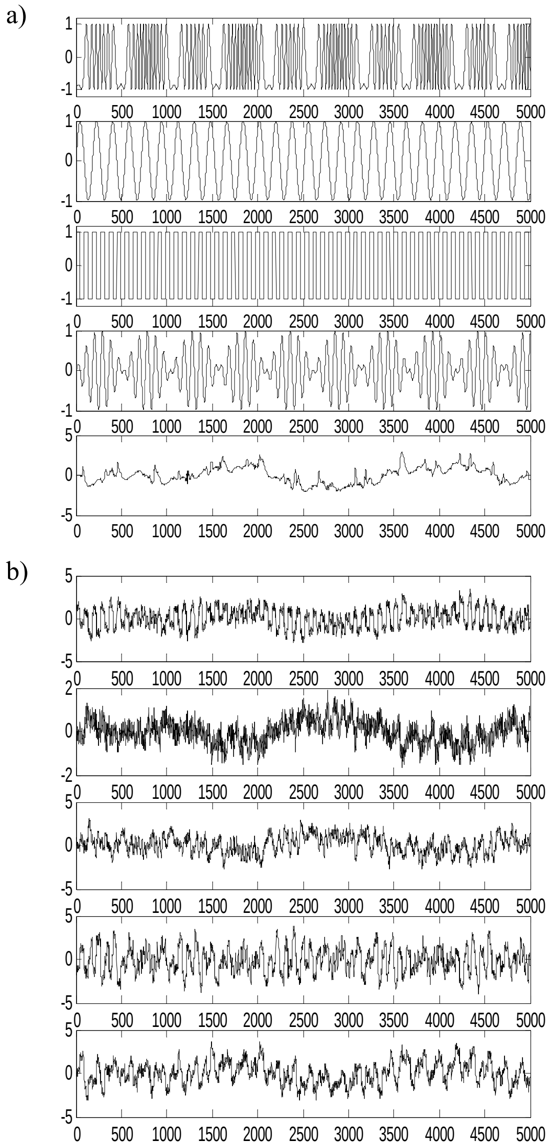

Fig. 2. Experiment results a) source signals, b) mixed signals

Source signals are mixed and disturbed by additive noise, Fig 2. The noise level in particular channels is SNR1=0.9875, SNR2=-6.6863, SNR3=-0.1585, SNR4=3.0856, SNR5=2.0073. To measure the performance of our method we use performance index defined as

$$\mathrm{PI} = \frac{1}{n(n-1)}\sum_{i=1}^{n}\left\{\left(\sum_{k=1}^{n}\frac{|g_{ik}|}{\max_j |g_{ij}|} - 1\right) + \left(\sum_{k=1}^{n}\frac{|g_{ki}|}{\max_j |g_{ji}|} - 1\right)\right\} \quad (26)$$

where g_{ij} is the (i,j) element of the global system matrix $\boldsymbol{G}$=$\boldsymbol{WA}$. If the performance index is zero, then the perfect estimation is achieved.

Signals after separation and next after denoising filtration are presented in Fig. 3. We can show the efficiency of the algorithm by presentation diagram of global transformation matrix G which is close to global permutation matrix. In computation was used algorithm with delays $p=\{0,1,...,14\}$ and smoothing filtering. In this case we obtain PI=0.08.

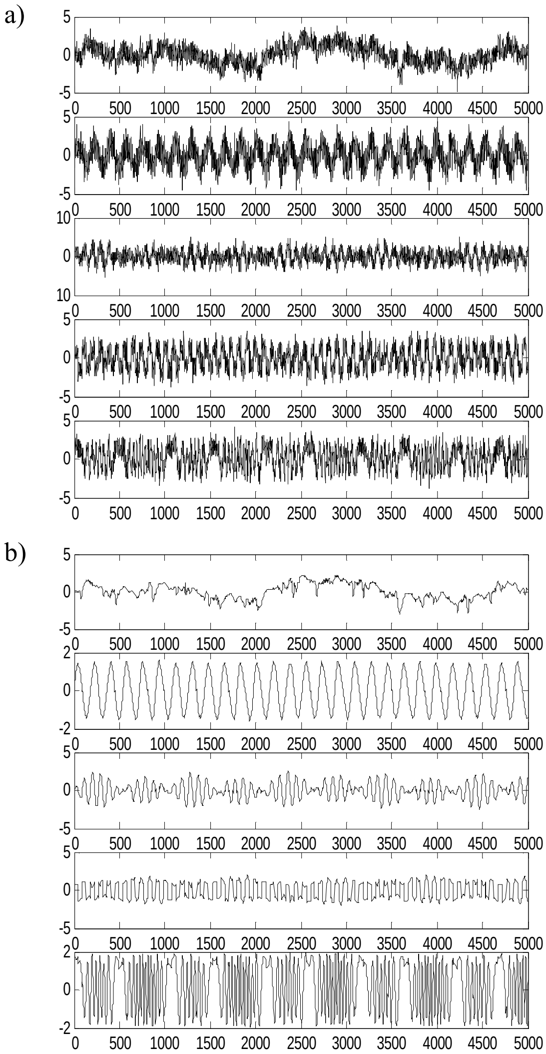

Fig. 3. Experiment results a) separated signals, b) signals after noise cancellation

V. Conclusions

We have presented a new method of BSS in the framework of filtered time-delayed correlation analysis. Presented methods gives good results in the presence of medium level of the additive noise. In case of estimating mixed signals with additive noise, the better results are obtained when the number of observed signals is bigger than number of source signals. In noiseless case we can left out filtration process time delayed covariance matrix what leads us to simple efficient algorithm where only two correlation matrices for p=0 and p=1.

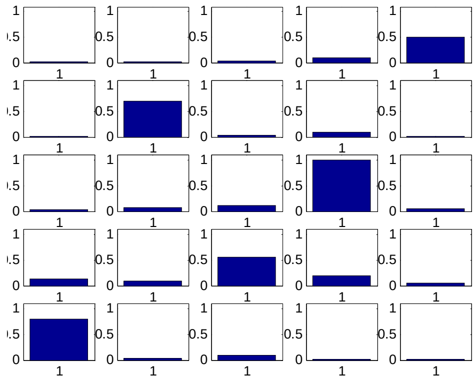

Fig. 4. Global transformation matrix (**G**) diagram

In the noisy case we typically need more matrices but it is open question how to determine a priori optimal number of them. The second question is what type of filtration should be used. We not recommend complex filtration to not disturb statistical structure processed signals. The good performance of the proposed method was confirmed by numerical experiments.

References

[1]. Amari S., Cichocki A.: „Adaptive blind signal processing-neural network approaches", Proc. of the IEEE, Special Issue on Blind Identification and Estimation, vol. 86, no. 10, pp. 2026-2048, Oct. 1998.

[2]. Belouchrani A., Cichocki A.: Robust whitening procedure in blind source separation context, Electronics Letters, Vol. 36, No. 24, pp. 2050-2053, 2000.

[3]. Belouchrani A., Meraim A.K., Cardoso J.F., Moulines E.: A blind source separation technique using second order statistics, IEEE Trans. on Signal Processing, Vol. 45, no 2, pp. 434-44, Feb. 1997.

[4]. Belouchrani A., Meraim A.K., Cardoso J.F., Moulines E.: Second-order blind separation of correlated sources, Proc. Int. Conf. on Digital Sig. Proc., pp. 346-351, Cyprus, 1993.

[5]. Choi S., Cichocki A.: „Blind separation of nonstationary sources in noisy mixtures", Electronics Letters, vol.36, pp. 848-849, Apr. 2000.

[6]. Choi S., Cichocki A., Belouchrani A.: Second order nonstationary source separation, Journal of VLSI Signal Processing, vol. 32, no. 1-2, pp. 93-104, Aug. 2002.

[7]. Cichocki A., Amari S.: Adaptive Blind Signal and Image Processing, John Wiley, Chichester, 2002.

[8]. Haykin S.: Adaptive filter theory, 3rd ed., Upper Saddle River, Prentice-Hall, 1996.

[9]. Molgedey L., Schuster H.: Separation of a mixture of independent signals using time delayed correlations, Phisical Review Letters, no. 72, vol. 23, 1994.

[10]. Oppenheim A.V., Willsky A.S., Nawab S.H.: Signals and systems, 2nd ed., Upper Saddle River, Prentice-Hall, 1997.

[11]. Pham D.-T., Cardoso J.F.: Blind separation of instantaneous mixtures of non stationary sources, IEEE Trans. Signal Processing, pp. 1837-1848, vol. 49, no 9, 2001.

[12]. Tong L., Liu R., Huang Y.-F.: Indeteminacy and Identifiability of Blind Identification, IEEE Trans. on Circuits and Systems, vol. 38, no. 5, May 1991.