Multi-Science Journal 2015; 1(3):40-45

ARTIGO ORIGINAL

Blind source separation by multiresolution analysis using AMUSE algorithm

Bruno Rodrigues de Oliveira[1], Marco Aparecido Queiroz Duarte[2] & Jozué Vieira Filho[1]

ABSTRACT
Algorithms for blind source separation have been extensively studied in the last years. This paper proposes the use of multiresolution analysis in three decomposition levels of the wavelet transform, such as a preprocessing step, and the AMUSE algorithm to separate the source signals in distinct levels of resolution. Results show that there is an improvement in the estimation of the signals and in the mixing matrix even in noisy environment if compared to the use of AMUSE only.
Keywords: audio signal separation, wavelet transform, second order statistical.

Separação cega de fontes pela análise multirresolução utilizando o algoritmo AMUSE

RESUMO
Algoritmos para separação cega de fontes têm sido extensivamente estudados nos últimos anos. Este artigo propõe o uso da análise multirresolução pela transformada wavelet em três níveis de decomposição, na etapa de pré-processamento, e o algoritmo AMUSE para separar sinais fontes em distintos níveis de resolução. Os resultados mostram que há melhora na estimação dos sinais e da matriz de mistura, mesmo em ambiente ruidoso, se comparado ao uso do AMUSE somente.
Palavras-chave: separação de sinais de áudio, transformada wavelet, estatísticas de segunda ordem.

Autor para correspondência: Bruno Rodrigues de Oliveira
Universidade Estadual Paulista (UNESP), Ilha Solteira, SP, Brasil.
E-mail: bruno.ro.net@gmail.com
Recebido em: 04 jun. 2015
Aceito em: 08 jul. 2015
Editor responsável: Prof. Dr. Guilherme Malafaia

[1]Universidade Estadual Paulista (UNESP), Ilha Solteira, SP, Brasil.

[2]Universidade Estadual de Mato Grosso do Sul, MS, Brasil.

INTRODUCTION

Blind Source Separation (BSS) is encountered in various branches of applied mathematics: medical applications such as EEG, ECG, fetal ECG, MEG and fMRI; in telecommunications such as multiuser detection; as a tool for financial analysis, helping to minimize the risk in investment strategy; in audio separation; in feature extraction, which allows implementing pattern recognition systems (HYVÄRINEN *et. al*, 2001; COMON; JUTTEN, 2010).

Many researchers have considered the inclusion of preprocessing steps using the Discrete Wavelet Transform (DWT) (LÓ *et al.*, 2011; MISSAOUI *et al.*, 2011; MIJOVIC *et al.*, 2011; TALBI *et al.*, 2012; SHAYESTEH; FALLAHIAN, 2010). Some of them use the DWT to remove noise. Others exploit the decomposition into several frequency bands where the particularities of the signals can be emphasized, as in Huang *et al.* (2003) which proposes a parallel architecture to separate signal in low and high frequencies. Another approach is the recognition of images as in Leo *et al.* (2003) that uses DWT and BSS method for detecting ball in soccer game, aided by neural network.

In this work, it is proposed a BSS method that uses the Multiresolution Analysis (MRA) proportioned by the DWT as a preprocessing step. The advantage of the proposed method is that there is no necessity to return to time domain, since the observed signals are stored, and after the identification of the mixing matrix, in different resolution levels, they are used to obtain the separated signals. The proposed method consists in an update to AMUSE, which means the inclusion of the wavelet preprocessing step aiming the improvement of the estimated sources.

MATERIAL AND METHODS

The methods used in this research were the wavelet transform and AMUSE algorithm. As material, male and female speech signals recorded in common environment were used.

Blind Source Separation

BSS problems are characterized as MIMO (multiple-input-multiple-output) systems, where each output is the combination of multiple inputs (sources) with some noise. The inputs of this system and the subsystem that mixes the source signals are both unknown.

Let $\mathbf{x} = [\mathbf{x}_1(t) \ \ldots \ \mathbf{x}_n(t)]^T$ be the observation vector of the output of the system. The input vector is $\mathbf{s}_0 = [\mathbf{s}_1(t) \ \ldots \ \mathbf{s}_n(t)]^T$, composed by the source signals and $\mathbf{v} = [\mathbf{v}_1(t) \ \ldots \ \mathbf{v}_n(t)]^T$ is the additive noise vector, where $[\,\cdot\,]^T$ denotes the transposed vector and t is the time. $\mathbf{A}_0 \in \mathbb{R}^{m\times n}$ is the actual matrix that characterizes the mixing system. If we consider the case of instantaneous mixtures, the output of the system is seen as

$$\mathbf{x} = \mathbf{A}_0\mathbf{s}_0 + \mathbf{v} \qquad (1)$$

In order to solve the problem, $\mathbf{s}_0$ and A_0 must be estimate only through $\mathbf{x}$. The main difficulty is the lack of information. Therefore some assumptions must be made, so that a model can be proposed.

Definition 1

A BSS model is an ordered pair $(\mathbf{A}_0, \mathbf{s}_0)$ such that: equation (1) is valid; $rank(\mathbf{A}_0) = n$; $\mathbf{s}_0$ and $\mathbf{v}$ are zero-mean WSS (Wide Sense Stationary) vectors and the covariance matrix is $\mathbf{R}_\mathbf{v} = \sigma^2\mathbf{I}$, where σ^2 is the variance of the noisy process and $\{\mathbf{s}_i(t), i = 1, \ldots, n\}$ are uncorrelated soruces.

Theoretically, $\mathbf{A}_0$ can be any non-singular matrix. However, in practical situations the waveform of the source should be preserved, so that the estimated signals are intelligible, as well as the estimated sources are not of the same order or magnitude (YEREDOR, 2010). The following theorem establishes a relationship with this property.

Theorem 1

A relationship $\mathfrak{R}$ that preserves the waveform of source signals is an equivalence relation on BSS model defined as a couple of ordered pairs $(\mathbf{A}', \mathbf{s}')$ and $(\mathbf{A}_0, \mathbf{s}_0)$ such that:

$$\mathbf{A}' = \mathbf{A}_0\mathbf{\Lambda}^{-1}\mathbf{P}^T \qquad (2)$$

$$\mathbf{s}' = \mathbf{P}\mathbf{\Lambda}\mathbf{s}_0 \qquad (3)$$

where $\mathbf{\Lambda}$ is a scaling matrix and $\mathbf{P}$ a permutation matrix. The proof of this theorem can be found in Tong *et al.* (1991).

Definition 2

In order to solve a BSS model an ordered pair $(\mathbf{A}', \mathbf{s}')$ must be found such that $(\mathbf{A}', \mathbf{s}')\mathfrak{R}(\mathbf{A}_0, \mathbf{s}_0)$.

Therefore, it is not sufficient to estimate a mixing matrix and sources. These estimates should be related through $\mathfrak{R}$ as it preserves the waveform of the actual source signals and also the direction of the column vectors of $\mathbf{A}_0$, differing only by norm and/or permutation.

Algorithm to solve the BSS problem

In this work we use AMUSE (Algorithm for Multiple Unknown Signals Extraction) to solve the BSS problem. In an overview, AMUSE explores the temporal structure of sources projecting the observation vector in an orthogonal space. In this space, the n largest singular values of the covariance matrix of the observation vector are distinct (CHICHOCKI; AMARI, 2003).

Then, it is possible to find a solution to the BSS problem by the eigenvalues decomposition of such matrix.

Multiresolution Analysis

The Multiresolution Analysis (MRA) allows the observation of a signal at different scales, and in each scale it is possible to obtain more or less information.

It is defined as a sequence of closed subspaces $V_j \subset V_{j-1}$, j is an integer number, requiring $f(\cdot) \in V_j \Leftrightarrow f(2^j \ \cdot) \in V_0$ and exists a function $\phi(t) \in V_0$ so that $\{\phi_{0,k} = \phi(t-k); k \in \mathbb{Z}\}$ is an orthonormal basis for those subspaces; more generally $\{\phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k); j,k \in \mathbb{Z}\}$ is an orthonormal basis for V_j (DAUBECHIES, 1992). Since $\phi_{-1,k}$ is an orthonormal basis for V_{-1} and $V_0 \subset V_{-1}$, there exists a sequence $h[n]$ so that

$$\phi(t) = \sqrt{2}\sum_{n} h_n \, \phi(2t-n) \qquad (4)$$

where $h_n = \langle \phi, \phi_{-1,k} \rangle$ and $\sum_n |h_n|^2 < +\infty$.

The function in eq. (4) implements a lowpass filtering with coefficients $h[n]$ and also a downsampling operation by a factor of two, that can be seen in the matrix form in eq. (6).

For other resolution levels, the output is again filtered. When a signal is decomposed by the DWT a MRA process is performed (DAUBECHIES, 1992; STRANG; NGUYEN, 1997).

Proposed method

Our proposal is to insert a preprocessing step before using AMUSE algorithm. This step will act as a filtering process in the signal decomposed by the DWT, i.e., in the wavelet domain.

The separation is performed in each resolution level and it is not necessary to return to time domain, because the observation vector, in the jth resolution, is only used to calculate the mixing matrix and then, in order to estimate the sources, it is used to multiply the observation vector in the time domain. In the BSS model a filtering process can be implemented without changing its structure (HYVÄRINEN *et. al*, 2001), as shown in eq. (5), considering eq. (1):

$$\mathbf{x}F = \mathbf{A}_0\mathbf{s}_0F + \mathbf{v}F = \mathbf{A}_0\mathbf{s}_0^{\#} + \mathbf{v}^{\#} \qquad (5)$$

where F is a matrix used in the filtering process and $\mathbf{s}_0^{\#}$ and $\mathbf{v}^{\#}$ are the filtered signals. In the DWT decomposition four Daubechies wavelet coefficients are used:

$$h_0 = \frac{1+\sqrt{3}}{4\sqrt{2}}, h_1 = \frac{3+\sqrt{3}}{4\sqrt{2}}$$
$$h_2 = \frac{3-\sqrt{3}}{4\sqrt{2}}, h_3 = \frac{1-\sqrt{3}}{4\sqrt{2}}$$

In a matrix form, filtering operation F can be seen in eq. (6) (WEEKS, 2007):

$$F = \begin{bmatrix} h_3 & h_2 & h_1 & h_0 & 0 & 0 & 0 & 0 & \cdots & (\\ 0 & 0 & h_3 & h_2 & h_1 & h_0 & 0 & 0 & \cdots & (\\ 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 & h_0 & \cdots & (\\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & (\\ 0 & 0 & 0 & 0 & 0 & 0 & h_3 & h_2 & h_1 & h \end{bmatrix} \qquad (6)$$

Obviously the filter must operate in the signals such that for a given solution $(\mathbf{A}', \mathbf{s}')$, $(\mathbf{A}', \mathbf{s}')\Re(\mathbf{A}_0, \mathbf{s}_0)$, i.e., the filtering process does not change the waveform of the source signals. It is clear that is was expected to happen with the proposed method. Since the DWT implements only a matrix multiplication with downsampling by factor 2. However, as shown in Figure (1), the waveform of the signal is modified from some level of decomposition. Therefore the AMR should be used in some levels, as done in the following experiments.

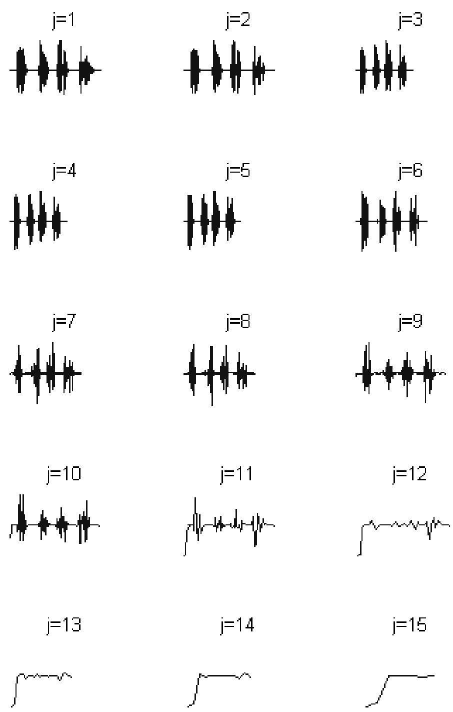

Figure 1. Changing the waveform at different decomposition levels

Performance measures

Among the several useful measures for evaluating BSS algorithms, one of the most used is ζ (Amari metric), proposed by Amari (KAWAGUCHI *et al.* 2012). It compares the elements of $\mathbf{Q} = \widehat{\mathbf{A}}_0^{-1}\mathbf{A}_0$, with the jth column of $\mathbf{Q}$ given by q_{ij}. The closer to zero it is the better is estimation of the mixing matrix. Such measure is defined in (7), its not-normalized version:

$$\zeta = \sum_{i=1}^{n}\left[\frac{\sum_{j=1}^{n}|q_{ij}|}{\max_{j}|q_{ij}|} - 1\right] + \sum_{j=1}^{n}\left[\frac{\sum_{i=1}^{n}|q_{ij}|}{\max_{i}|q_{ij}|} - 1\right] \quad (7)$$

The Source-to-Interference-Ratio - SIR, proposed by Radu *et al.* (2003) and defined in (8), is used to measure the interference among the sources in the BSS process. SIR values are given in dB (decibels) and the higher they are the lower is the interference among the sources.

$$SIR = 10log_{10}\left[\frac{1}{m}\frac{\sum_{i=1}^{m} max(q_j)^2}{q_j^T q_j - max(q_j)^2}\right] \quad (8)$$

RESULTS AND DISCUSSION

In this section two experiments are performed in order to evaluate the efficiency of the proposed method, based on the metrics presented in section 6. Results are obtained for AMUSE and for the proposed method. In the implementation of the proposed method, three wavelet decomposition levels are tested ($j = 1,2,3$).

Experiment One

In this experiment two Portuguese speech signals are considered, one with male voice and the other with female voice, both sampled at 16 kHz rate with 7s of duration.

First, sources are mixed without noise and results are presented in Table (1) and Figure (2). After, before the mixing process, both signals are corrupted by Gaussian white noise, now results are in Table (2) and Figure (3).

Measure	AMUSE Method	Proposed Method		
		$j = 1$	$j = 2$	$j = 3$
ζ	0.0176	0.0141	0.0143	0.0307
SIR	35.9570	38.5584	39.0679	30.4950

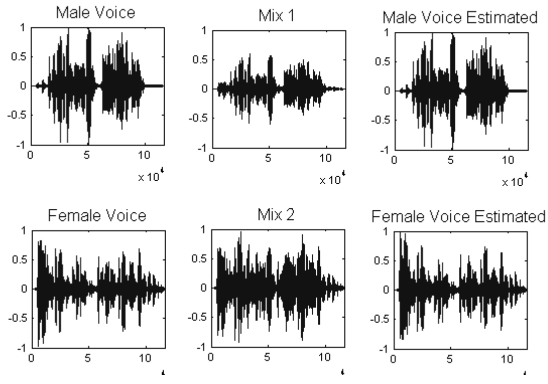

Figure 2. Waveform of the signals of Table 1 in $j = 2$

Measure	AMUSE Method	Proposed Method		
		$j = 1$	$j = 2$	$j = 3$
ζ	0.0835	0.0152	0.0142	0.0310
SIR	21.9108	37.5799	39.057	30.3961

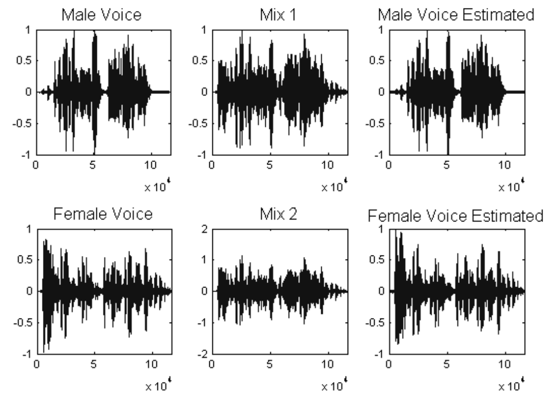

Figure 2. Waveform of the signals of Table 2 in j=2

Observing Tables (1) and (2) it is perceptible the efficiency of the proposed method over AMUSE. When noise-free signals are considered, the proposed method overcomes AMUSE in the first and second wavelet decomposition levels, with better ζ value for $j = 1$ and better SIR for $j = 2$.

When noisy signals are considered, the proposed method also overcomes AMUSE in the in the first and second wavelet decomposition levels, but now better ζ and SIR values are reached when $j = 2$. It is worth noting that for the two cases, the use of a third wavelet decomposition level is not so advantageous, since results are becoming worst from second level on.

Experiment Two

This experiment considers as source signals, for mixture, the Portuguese pronunciations of the letters "i" and "a" by a male speaker. Signals are sampled at 16kHz rate with 1s of duration. Again, mixtures of noise-free signals and noisy signals are tested, with noisy signals corrupted with white Gaussian noise. Table (3) and Figure (4) presents results when BSS is performed with noise-free signals and results for noisy signals are presented in Table (4) and Figure (5).

Measure	AMUSE Method	Proposed Method		
		$j=1$	$j=2$	$j=3$
ζ	0.0123	0.0021	0.0016	0.3425
SIR	38.6951	55.5826	56.4386	9.7450

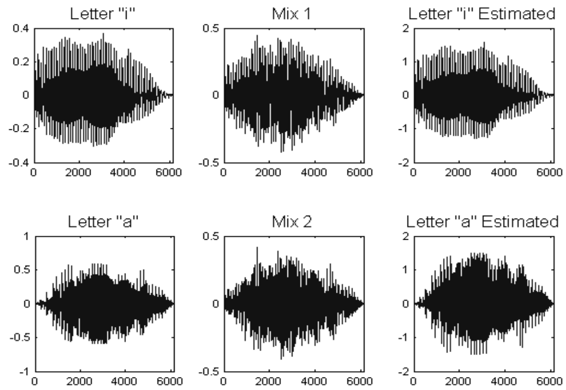

Figure 3. Waveform of the signals of Table 2 in j=2.

Measure	AMUSE Method	Proposed Method		
		$j=1$	$j=2$	$j=3$
ζ	0.0288	0.0021	0.0018	0.3361
SIR	31.9500	54.6894	55.7146	9.9048

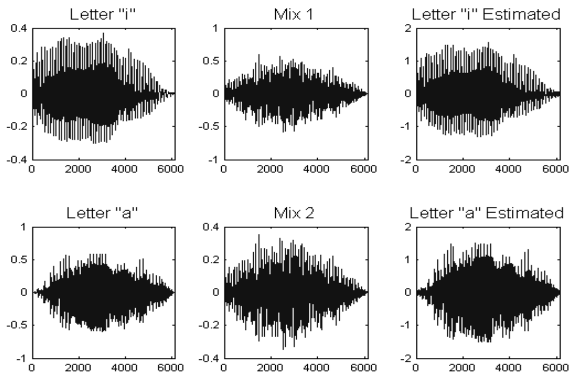

Figure 4. Waveform of the signals of Table 2 in j=2.

Observing Tables (3) and (4) it is again perceptible the efficiency of the proposed method over AMUSE and, in the two considered cases, better results are reached for $j=2$. When $j=3$, results become worst, under acceptable values, because there is few information and the waveform of the source signals change.

In general, analyzing results presented in Tables (1) to (4), a wavelet preprocessing step improves AMUSE, since results are superior than simple AMUSE. Considering average values for ζ and SIR it is worth saying that the ideal preprocessing wavelet decomposition level is $j=2$. After $j=2$, there is a clear deterioration in the sources.

Another fact that should be noted is that in a noisy environment the proposed method has better results than AMUSE only. Figure (6) shows, for different signal-to-noise-ratio (SNR), how this preprocessing step contributes for better estimation of the source signals, i.e., higher SIR.

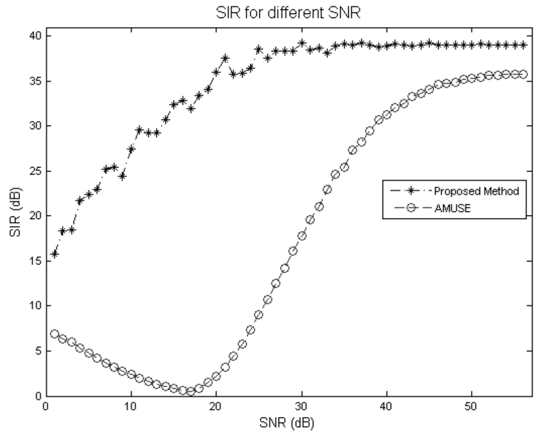

Figure 5. SIR for different SNR to AMUSE and Proposed Method

CONCLUSION

In this paper an improvement for AMUSE BSS algorithm was proposed, which consists in the wavelet decomposition of the mixture before the actual BSS process.

Results showed significant improvements in the estimation of the sources up to the second wavelet resolution level in all experiments, reaching approximately 18dB of improvement in average.

Experiments also showed that it is not worth using more than two wavelet resolution levels, since the use of a third resolution level deteriorates the sources. Such worsening in the estimation is due to the fact that there is not enough information to source separation, since the third wavelet level has only 1/8 of samples of the observation vector.

In noisy environments results were slightly worse, since the separation does not ignore noise, estimating it along with the sources. In fact each estimated source is a linear combination of the noise sources present in the sensors. Therefore, in order to attenuate noise, the structure to acquire sources should have more sensors than sources.

Further works should consider the use of other wavelet filters, trying to get the better (or optimal) wavelet filter for BSS process. The case of more sensors than sources should be also studied.

REFERENCES

CHICHOCKI, A.; AMARI, S. Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications. West Sussex: Wiley, 2003.

COMON, P.; JUTTEN, C. Handbook of Blind Source Separation: Independent Component Analysis and Applications. Oxford: Elsevier, 2010.

DAUBECHIES, I. Ten Lectures on Wavelets. CBMS-NSF regional conference series in applied mathematics, Philadelphia, 1992.

HYVÄRINEN, A.; KARHUNEN, J.; OJA, E. Independent Component Analysis. New York: Wiley, 2001.

HUANG, R.; CHEUNG, Y.; ZHU, S. A Parallel Architecture Using Discrete Wavelet Transform for Fast ICA Implementation. IEEE Int. Conf. Neural Networks & Signal Processing on, v. 14-17, p. 1358-1361, 2003.

KAWAGUCHI, A.; TRUONG, Y. K.; HUANG, X. Application of Polynomial Spline Independent Component Analysis to fMRI Data. In: NAIK, G. R. Independent Component Analysis for Audio and Biosignal Applications. Rijeka: InTech, 2012. p. 209-220.

LEO, M.; D'ORAZIO, T.; DISTANTE, A. Feature extraction for automatic ball recognition: comparison between wavelet and ICA preprocessing. Image and Signal Processing and Analysis, ISPA 2003. Proceedings of the 3rd International Symposium on, vol. 2, p. 587-592, 2003.

LÓ, P. M. G.; LOZANO, H. M.; SÁNCHEZ, F. L. P.; MORENO, L. N. O. Blind Source Separation of audio signals using independent component analysis and wavelets. Electrical Communications and Computers CONIELECOMP 21st International Conference on, p.152-157, 2011.

MISSAOUI, I.; LACHIRI, Z. Blind speech separation based on undecimated wavelet packet-perceptual filterbanks and independent component analysis. IJCSI International Journal of Computer Science Issues on, v. 8, 2011.

MIJOVIC, B.; VOS, M. D.; GLIGORIJEVIC, I.; TAELMAN, J.; HUFFEL, S. V. Using Wavelet Transformation in Blind Sources Separation of the Fetal Electrocardiogram. Majlesi Journal of Electrical Engineering on, v. 5, p. 2188-2196, 2011.

TALBI, M.; AICHA, A. B.; SALHI, L.; CHERIF, A. Bionic Wavelet Based Denoising Using Source Separation. Int J Comput Commun on, v. 7, p. 574-585, 2012.

RADU, M.; HULLE, M. M. V. A comparative survey on adaptive neural network algorithms for independent component analysis. Romanian Reports in Physics on, v. 55.1, p. 49-74, 2003.

SHAYESTEH, M.; FALLAHIAN, J. Source Separation From Single-Channel Recordings by Combining Empirical-Mode Decomposition and Independent Component Analysis. IEEE Transactions on Biomedical Enginnering on, v. 57, p. 33-37, 2010.

STRANG, G.; NGUYEN, T. Wavelets and Filter Banks. Cambridge: Wellesley, 1997.

TONG, L.; LIU, R.; SOON, V. C.; HUANG, Y. Indeterminacy and identifiability of blind identification. Circuits and Systems, IEEE Transactions on, v. 38, p. 499-509, 1991.

YEREDOR, A. Second-order methods based on color. In: COMON, P.; JUTTEN, C. Handbook of Blind Source Separation: Independent Component Analysis and Applications. Oxford: Elsevier, 2010. p. 227-278.

WEEKS , M. Digital Signal Processing using MATLAB and Wavelets. Massachusetts: Jones and Bartlett Publishers, 2007.