32nd Annual International Conference of the IEEE EMBS
Buenos Aires, Argentina, August 31 - September 4, 2010

Combining EMD with ICA for Extracting Independent Sources from Single Channel and Two-Channel Data

Bogdan Mijović, Maarten De Vos, Ivan Gligorijević and Sabine Van Huffel

Abstract— **Blind Source Separation (BSS) techniques are frequently needed in the processing of biomedical signals. This need comes from the fact that these signals are often composed of many different sources, which are mixed in the measured signal. However, we are usually only interested in examining one or a limited set of sources of interest separately. A variety of algorithms exist for separating multichannel mixtures into its independent sources (e.g. different Independent Component Analysis (ICA) techniques). These techniques only work if the number of channels is larger than, or equal to the number of sources present in the signal. On the other hand, only a few algorithms have been reported for the analysis of single channel sources, or other mixtures where the number of sources is higher than the number of channels. In this work we show a new technique which combines Empirical Mode Decomposition (EMD) and Independent Component Analysis (ICA). We will show that this technique is capable in separating independent sources when the number of these sources is higher than the number of channels available. We show the performance in single channel and two-channel biosignal processing.**

I. INTRODUCTION

In the field of biomedical signal processing, independent sources are often mixed together in the measured signal. Our task is then to unmix contributing sources in order to have a closer look at the signal of interest. In multichannel recordings, such as Electroencephalogram (EEG), this problem is efficiently handled by means of *Blind Source Separation* (BSS) techniques which unmix the given signal into sources (see e.g. [1], [2]).

Independent Component Analysis (ICA) is a BSS technique that extracts statistically independent sources (called independent components) from a set of recorded signals. Standard algorithms tackle the problem when the number of electrodes (channels) is larger than or equal to the number of sources. On the other hand, there is a group of algorithms called "undetermined ICA", that can recover more signals than channels available (see [3] and references herein). However, these techniques require some minimal number of channels to work well. In the limit, the goal can be the extraction of independent sources from a single-channel or 2-channel recordings. Examples are cleaning the Electromyogram (EMG) signal contaminated by an Electrocardiogram (ECG) artifact [4], [5], especially at the electrodes placed on the left side of the back. The ECG artifact is very prominent in this case. An example of processing 2-channel signals is EEG signal processing in neonatal brain. In this case, mostly two channels are recorded, and there is a lot of brain activity that has to be revealed.

An adaptation of ICA to single-channel signals, called Single Channel ICA (SCICA), has already been proposed in the literature [6]. It has two major drawbacks: first, the algorithm assumes stationary sources, and second the sources are assumed to be disjoint in the frequency domain. These assumptions, however don't always hold in applications.

Another approach of decomposing a signal of interest into different sources is as follows. If a multichannel recording can be created from a single channel signal, ICA might be used to select the signal components which are independent from each other. One way to decompose a single-channel signal into multiple components is to decompose it into different spectral modes (e.g. using a Wavelet transform or Empirical Mode Decomposition (EMD) [7]).

A wavelet decomposition splits a signal at each step in a predetermined manner by means of predefined linear time-invariant filters, thereby precluding the possibility of adapting the decomposition to local variations of the oscillation. The Wavelet-ICA (WICA) technique has been proposed in the literature [8]. A wavelet transform is used to expand a 1D signal into 2D by dividing it into its frequency subbands. Then ICA is used to extract the sources. However, in the field of biomedical signal processing it has always been referred to as a multichannel signal processing technique, where the wavelet transform has been used only for denoising (e.g. for extracting ECG artifacts from the EMG signal, when the latter has also been recorded in parallel [9], or for separating the fetal ECG from mother ECG when 6 ECG signals were recorded in parallel [10]).

A new technique for single-channel signal analysis, that first decomposes the given signal into *spectrally independent* modes using the EMD algorithm, and then applies ICA to extract *statistically independent sources* is introduced in [11]. In that work it is shown, through different simulations and real-life examples, that this technique, so-called EMD-ICA outperforms the above-mentioned WICA and SCICA techniques.

In this paper we show the performance of EMD-ICA algorithm on single-channel data and extend the approach to 2-channel EMD-ICA. We will show the extraction capabilities of the algorithm on 2-channel EEG signals and we will discuss the benefits and improvements compared to single-channel EMD-ICA.

All the authors are with the Department of Electrical Engineering (ESAT), division SCD, Katholieke Universiteit Leuven, 3001 Leuven, Belgium, e-mail: bogdan.mijovic@esat.kuleuven.be.

II. METHODS

A. Single Channel EMD-ICA

The Empirical Mode Decomposition (EMD) [7] is a signal analysis tool which is able to decompose any time series into a set of spectrally independent oscillatory modes, called Intrinsic Mode Functions (IMF). IMFs are meant to be *monocomponent*, zero-mean oscillatory functions, which are orthogonal to each other and a set of IMF's should be complete. Here the term monocomponent means that all the IMF's contain only one frequency at a time, which is called the Instantaneous Frequency. The orthogonality property implies that different IMF's do not have similar frequency content. Oscillatory here means that they have the same number of local maxima and minima, and that all the maxima are positive, and the minima are negative.The advantage of EMD, compared to wavelets, is that the EMD is a data driven algorithm. This means that it decomposes a signal in a natural way without prior knowledge about the signal of interest embedded in the data series [12].

After the signal is decomposed with EMD into a set of IMFs, Independent Component Analysis (ICA) has been applied to extract independent sources. The ICA algorithm used in this study is FastICA [13], [14]. Other ICA algorithms can also be used, although FastICA and Infomax [15] show the best performance. After selecting the ICs of interest, the signal is reconstructed first by multiplying it by the mixing matrix to derive a new IMF set, in which only the IC of interest is present. Summing over the newly derived IMF set, the appearance of the desired source in the original signal is obtained.

B. 2-channel EMD-ICA

Extension of the EMD technique to 2-channel signals was not straight forward. The complex EMD algorithm [16] effectively applies real-valued EMD to the signals corresponding to the positive and negative frequency component of the spectrum of analytic signals. The rotation invariant complex EMD (RI-EMD) [17], extends the real-valued EMD to the complex (bivariate) domain in a generic way. Since all the operations are performed directly in the complex field, this method provides a single set of complex IMFs. An extension of RI-EMD [17], also termed the bivariate EMD [18], separates fast rotating components of a complex signal from slowly rotating ones. Envelope curves are obtained by projecting a bivariate signal in multiple directions and interpolating their extrema. The local mean is calculated by averaging the envelope curves, and is then subtracted from the original signal repeatedly to sift out rotating components within the signal.

The major advantage of Complex EMD is that it has the ability of extracting common oscillatory modes for both component signals and capture them in the same complex IMF set. This implies that the number of IMFs is the same for both components.

After complex EMD has been performed, our data is stored in a 3D matrix (Tensor). The dimensions are spatial (different channels), times time, times frequency (the set of IMFs). There are several ways how to proceed from this point, and perform a data reduction and independent component analysis. The fact that after Complex EMD, the set of IMFs contains common oscillations from both channels allows us to meaningfully perform the Singular Value Decomposition (SVD) , and thus to reduce the dimensionality in the third dimension (IMFs) [19]. This way we perform the dimension reduction, and still keep the highest variance information for each channel separately. When the data reduction has been performed, the data is stored in a new matrix, which is obtained by merging both sets of reduced IMFs (putting them on top of each other). After this, ICA is applied and a set of common independent sources has been extracted. The reader should note that all the processes are reversible, so the appearance of each source in each of the original channels can be back-reconstructed.

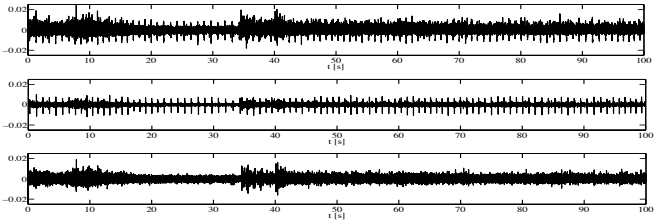


Fig. 1. EMG recording. Upper trace - original EMG signal, middle trace - extracted ECG artifact, lower trace - cleaned EMG signal.

C. Data

To illustrate the performance of our algorithm on real-life data we consider two applications. In the first, we will demonstrate how our new method performs in separating the ECG artifact from the EMG recordings. In the second, we apply the EMD-ICA algorithm to one of the EEG channels contaminated by muscle artifact, eye artifact, and seizure activity. We try to separate this information into different sources. In the EMD-ICA algorithm, the number of independent components to be extracted was set to 5 in the EMG example. The number of IC's to be extracted was set to 7 in the example of EEG recording, because we were interested in 3 different sources and not all of them would appear in the first 5 IC's.

In the upper trace of Fig. 1 we give a recording of 100 seconds of the EMG signal contaminated with the ECG artifact. This recording corresponds to the phase immediately after the muscle activity, when the muscle is still not in the completely relaxed state. The muscle is not in complete rest however, and therefore the power of the EMG signal is still fairly high.

In Fig. 2, 10 seconds ($250Hz$) of 21-channel scalp EEG recordings from a long-term epilepsy monitoring unit are shown. This recording contains ictal activity from a patient with *Mesial Temporal Lobe Epilepsy*, contaminated with eye blinks and muscle artifact. The epileptic seizure is constantly visible on the T2 lead. Although the single channel technique

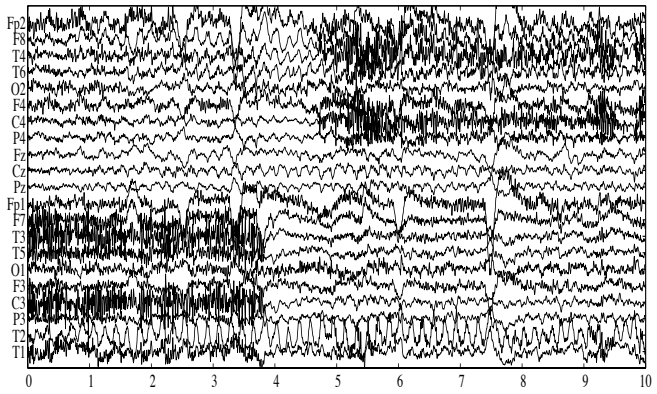


Fig. 2. 10 seconds of 21 channel EEG recordings

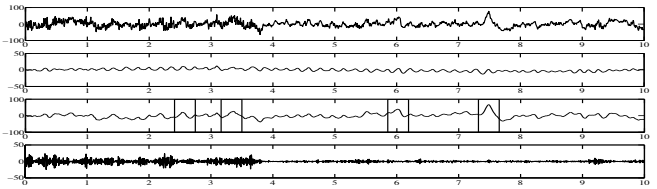


Fig. 3. EEG channel at T1 decomposed in the independent components. First row - original channel, second row - seizure event, third row - eye artifact (rectangles), fourth row - muscle activity

in this case is not necessary, we chose this example because lower spatial sampling can be required in wireless systems applications, and our technique might be of use. Also, it enables us to further validate our technique. Eye-artifacts can be observed around 2.5, 3.5, 6 and 7.5 seconds, and are most emphasized in the Fp1 and Fp2 channels, as expected. Muscle activity starts on one side of the head, and moves to another one. We extract all underlying signals from channel T1 by applying EMD-ICA, because their contributions are rather weak in this channel and therefore a successful extraction of these sources is most challenging.

After the extraction of independent sources from the T1 channel in EEG, we added another channel to observe what is the benefit of using 2-channel EMD-ICA. The second channel we used was F4. This lead is interesting to us because its position is at the right side of the head (contrary to T1) and the artifacts in this channel are still not clearly visible.

III. RESULTS

Fig. 1 shows the EMG signal contaminated by the ECG artifact. In the middle trace of this figure the extracted ECG artifact is shown, and the lower trace shows the cleaned EMG signal obtained by subtracting the extracted artifact from the original signal. We see that the ECG artifact is nicely removed from the signal without distorting the original shape and amplitude of the EMG signal. After applying the EMD-ICA algorithm, 4 sources were detected, only one of them containing the ECG artifact.

In Fig. 3, the T1 channel of the EEG recording from Fig. 2 is decomposed. After applying the EMD-ICA algorithm,

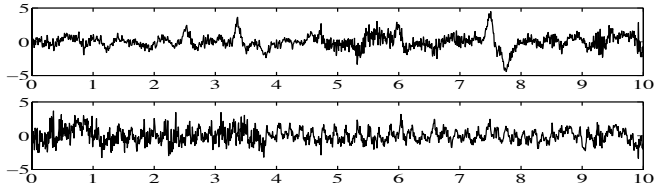


Fig. 4. The independent sources extracted by applying the FastICA algorithm to EEG channels T1 and F4

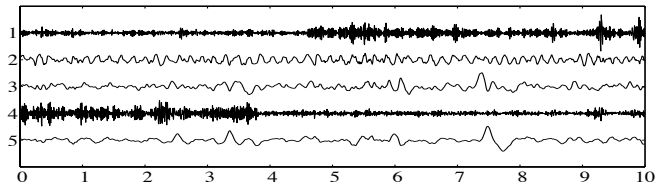


Fig. 5. The 5 out of 7 independent sources extracted from EEG channels T1 and F4 by applying 2-channel EMD-ICA

7 independent components were extracted among which we show 3 to back reconstruct the signals shown in the same figure, representing the seizure event, eye artifact and muscle activity respectively. We can clearly notice a successful separation of the important components. The extracted oscillatory activity is in phase with the seizure activity visible in the T2 channel. This confirms that this is related to the seizure. This is an interesting result, since channels T1 and T2 are located on the opposite sides of the head. This means that the seizure has been picked up from the averaged reference despite the very high noise-to-signal ratio. It is also clear that the eye artifact is more clearly (although not perfectly) expressed than in the original signal and that the muscle activity is well separated. We would like to stress here that these artifacts (e.g. eye-artifact) can sometimes be more clear in other leads (Fp1 and Fp2). However, in some cases (e.g. in neonates), EEG is recorded only with 8, and sometimes even with as few as 2 leads. In these cases it is very difficult to capture all the brain activity clearly, and our algorithm can be beneficial.

Fig. 4 shows the 2 independent components when only FastICA is applied to channels T1 and F4. It is obvious that there are more than 2 sources present and that FastICA was unable succesfully disjoint them, since we only gave two channels as an input. Then we show the result when 2-channel EMD-ICA algorithm is performed. EMD-ICA has extracted 7 components, 5 of which are shown in the Fig. 5.

IV. DISCUSSION

From Fig. 1 we see that the ECG artifact has been successfully cleaned from the EMG signal (trace 2). The shape and the amplitude of the EMG signal have been preserved. Please note here that by closer observation of the signal, it can be seen that no EMG activity has been removed. This is a very successful application of our algorithm, because otherwise the EMG activity here cannot be accurately studied.

Fig. 4 shows the decomposition of channel T1 into independent sources. The seizure activity and muscle activity are

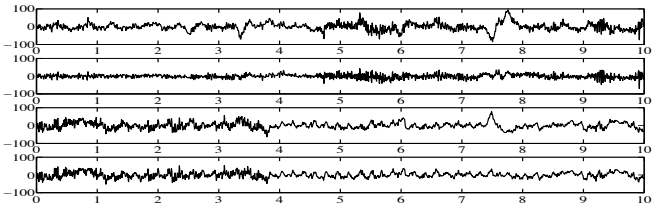


Fig. 6. Channels T1 and F4 (traces 1 and 3) with cleaned eye-artifacts (traces 2 and 4) using two-channel EMD-ICA. It is obvious that only eye-artifact is removed, but the seizure and muscle artifact are still present

nicely captured. The extracted oscillatory activity is in the phase with the channel T2. This proves that the captured activity is indeed the seizure. The eye artifacts are also more clearly (although not perfectly) isolated. However, this performance is enhanced by performing 2-channel EMD-ICA on channels T1 and F4 (see Fig. 5).

The performance of the 2-channel EMD-ICA is shown in Fig. 5. Channels T1 and F4 are given as inputs. It is clear that the eye artifacts (sources 3 and 5) are more clearly isolated than in the case of single-channel EMD-ICA (especially the artifacts appearing around 2.5 and 3.5 seconds). The eye-artifact channel is decomposed into two components, one of which (source 3) allows better reconstruction of the separate appearance of the artifact in the two observed channels. Good extraction of the eye-artifact is obvious from Fig. 6. Both channels are shown (channel F4 in trace 1, and channel T1 in trace 3). The corresponding channels after cleaning the eye artifact are shown in traces 2 and 4 respectively. It is obvious that other artifacts are not removed after cleaning the eye-artifact. Beside the eye-artifact, the muscle activity from the right side of the head (where the electrode F4 is positioned) is also nicely captured (trace 1 in Fig. 5). The sources which had already been extracted before by applying the single-channel EMD-ICA to the channel T1 (muscle artifact on the left side of the head and the seizure activity) are still present in 2-channel EMD-ICA. Therefore we can conclude that the 2-channel EMD-ICA makes better estimation of the common sources (eye-artifact), while the activity present in each channel separately is still preserved.

V. CONCLUSION

In this work we have shown the performance of the single channel EMD-ICA introduced in [11]. We have also introduced the extension to 2-channel signals and shown that it preserves the extraction of the independent sources from each channel separately, as well as it enhances the performance of common dynamics present in the both input channels. We have also shown that this method is capable of extracting more sources than channels recorded. This allows for analyzing signals in spatially under-sampled signals, which opens new possibilities in biomedical signal analysis.

VI. ACKNOWLEDGMENTS

Research supported by Research Council KUL: GOA-AMBioRICS, GOA MaNet, CoE EF/05/006, OPTEC, IDO 05/010 EEG-fMRI, IDO 08/013 Autism, IOF-KP06/11 FunCopt, by FWO: FWO G.0302.07 (SVM), G.0341.07, G.0427.10N, by (ICCoS, ANMMM); IWT: TBM070713-Accelero, TBM070706-IOTA3, TBM080658-MRI; by DWTC: IUAP P6/04 (DYSCO), by ESA: PRODEX No 90348, by EU: FAST (FP6-MC-RTN-035801), Neuromath (COST-BM0601)

The scientific responsibility is assumed by its authors.

REFERENCES

[1] T. Jung, S. Makeig, C. Humphries, T. Lee, M. McKeown, V. Iragui, and T. Sejnowski, "Removing electroencephalographic artifacts by blind source separation," *Psychophysiology*, vol. 37, pp. 163–178, 2000.

[2] M. De Vos, A. Vergult, L. De Lathauwer, W. De Clercq, S. Van Huffel, P. Dupont, A. Palmini, and W. Van Paesschen, "Canonical decomposition of ictal scalp EEG reliably detects the seizure onset zone," *NeuroImage*, vol. 37, no. 3, pp. 844–854, 2007.

[3] L. De Lathauwer, J. Castaing, and J.-F. Cardoso, "Fourth-order cumulant based underdetermined independent component analysis," *IEEE Transaction on Signal Processing*, vol. 55, pp. 2965–2973, 2007.

[4] P. Zhou and T. Kuiken, "Eliminating cardiac contamination from myoelectric control signals developed by targeted muscle reinnervation," *Physiological Measurement*, vol. 27, pp. 1311–1327, 2006.

[5] J. Taelman, A. Spaepen, and S. Van Huffel, "Wavelet-Independent Component Analysis to remove electrocardiography contamination in surface electromyography," in *Proceedings of the 29th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (IEEE/EMBC)*, 2007, pp. 682–685.

[6] M. E. Davies and C. J. James, "Source separation using single channel ICA," *Signal Processing*, vol. 87, no. 8, pp. 1819 – 1832, 2007, independent Component Analysis and Blind Source Separation.

[7] N. E. Huang, M. L. Wu, S. R. Long, S. S. Shen, W. D. Qu, P. Gloersen, and K. L. Fan, "The Empirical Mode Decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of Royal Society Lond*, vol. 454A, no. 1971, pp. 903–993, 1998.

[8] J. Lin and A. Zhang, "Fault feature separation using Wavelet-ICA filter," *NDT & E International*, vol. 38, no. 6, pp. 421 – 427, 2005.

[9] B. Azzerboni, G. Finocchio, M. Ipsale, F. L. Foresta, and F. Morabito, "A new approach to detection of muscle activation by Independent Component Analysis and Wavelet Transform," *Springer-Verlag, Lecture Notes in Computer Sciences*, vol. 2486, pp. 109–116, Sep. 2002.

[10] B. Azzerboni, F. L. Foresta, N. Mammone, and F. Morabito, "A new approach based on Wavelet-ICA algorithms for fetal electrocardiogram extraction," in *European Symposium on Artificial Neural Networks Bruges, ESANN-2005.*, Apr 2005, pp. 193–198.

[11] B. Mijović, M. De Vos, I. Gligorijević, J. Taelman, and S. Van Huffel, "Source separation from single-channel recordings by combining empirical mode decomposition and independent comnponent analysis," *to appear in IEEE Transaction on Biomedical Engineering*, 2010.

[12] P. Flandrin and P. Goncalves, "Empirical Mode Decomposition as data-driven Wavelet-like expansion," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 2, no. 4, pp. 477–496, 2004.

[13] A. Hyvarinen and E. Oja, "Independent Component Analysis: algorithms and applications," *Neural Networks*, vol. 13, no. 4-5, pp. 411 – 430, 2000.

[14] http://www.cis.hut.fi/projects/ica/fastica/code/dlcode.shtml, "Fastica graphical user interface."

[15] A. Bell and T. Sejnowski, "An Information-Maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[16] D. Mandić and V. Goh, "Complex-valued non-linear adaptive filters: Noncircularity, widely linear neural models," *Wiley*, 2006.

[17] M. Altaf, T. Gautama, T. Tanaka, and D. Mandić, "Rotation-invariant complex empirical mode decomposition," vol. 3, 2007, pp. 1009–1012.

[18] G. Rilling, P. Flandrin, P. Goncalves, and J. M. Lilly, "Bivariate empirical mode decomposition," *IEEE Signal Processing Letters*, vol. 14, pp. 936–939, 2007.

[19] L. De Lathauwer, B. De Moor, and J. Vandewalle, "A multilinear singular value decomposition," *SIAM journal on Matrix Analysis and Applications*, vol. 21, no. 4, pp. 1253–1378, 2000.