

SEPARATING SOURCES FROM SEQUENTIALLY ACQUIRED MIXTURES OF HEART SIGNALS

Fabio L. Hedayioglu^{*} Maria G. Jafari[†] Sandra S. Mattos[†] Mark D. Plumbley[‡] Miguel T. Coimbra^{*}

^{*} Department of Computer Science, University of Porto, Institute of Telecommunications,
{fheday,mcoimbra}@dcc.fc.up.pt

[†]Unit of Fetal and Pediatric Cardiology, Royal Portuguese Hospital, Recife, Brazil

[‡]Queen Mary University of London, Centre for Digital Music
{maria.jafari,mark.plumbley}@elec.qmul.ac.uk

ABSTRACT

In this paper, we consider the problem of separating a set of independent components when only one movable sensor is available to record the mixtures. We propose to exploit the quasi-periodicity of the heart signals to transform the signal from this one moving sensor, into a set of measurements, as if from a virtual array of sensors. We then use ICA to perform source separation. We show that this technique can be applied to heart sounds and to electrocardiograms.

Index Terms— Independent Component Analysis, Periodic Signals, FastICA, Blind Source Separation Application, Heart Sounds, Electrocardiogram, foetal Electrocardiogram.

1. INTRODUCTION

Independent component analysis (ICA) is a multidimensional signal processing technique used to separate signals arising from linear mixtures of several unobserved components or sources into distinct components. In biomedical signal processing, ICA can be used to analyse the electroencephalogram, to separate out the components arising from different brain sources [1], the electrocardiogram, for example in the context of extracting the foetal ECG from a recording taken from an expectant mother, and which contains the maternal heart and other interfering sources. ICA can also correlate different exams, such as MEG data and EEG [2]. Underlying the ICA generative model, is the assumption that a set of source signals, $\mathbf{s}(t) \in \mathbb{R}^m$, are transmitted simultaneously through a mixing medium, and a set of sensors simultaneously record a weighted linear combination of them, denoted as $\mathbf{x}(t) \in \mathbb{R}^n$. Typically, the basic mixing model assumes an instantaneous

and noiseless mixture of sources, that can be written as

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t) \quad (1)$$

where $\mathbf{A} \in \mathbb{R}^{n \times m}$ is the mixing matrix, and it is assumed that the source signals are mutually statistically independent, with at most one source having a Gaussian probability distribution [3]. For simplicity, we assume that there are as many sensors as sources $n = m$. To recover the original sources, an ICA algorithm such as Infomax [4], JADE [5], or FastICA [6], is used to estimate the unmixing matrix \mathbf{W} . The sources are finally estimated using:

$$\mathbf{y}(t) = \mathbf{W}\mathbf{x}(t). \quad (2)$$

where $\mathbf{y}(t)$ are the recovered sources. ICA suffers from a permutation and scaling ambiguity, which means that the ordering and scale of the estimated sources cannot be pre-determined [2].

In this paper, we address the problem of separating a set of independent components assuming that we have access to one moving sensor and we also assume that the underlying source signals that we wish to separate are quasi-periodic (this is a well-known characteristic of heart signals). Specifically, in our algorithm, the sensor is a stethoscope that is placed at particular locations in the thoracic (chest) region, during the routine listening of the heart sounds.

At first glance this might appear to be a single sensor separation problem such as discussed in [7]. However, the availability of a moving sensor and the cyclic-stationary nature of the signal means that we can simulate an array of n spatially separated sensors which acquire n mixtures by time-shifting to align the mixtures. Thus, $x_1(t)$ will be recorded during the time interval $0 < t < t_1$, $x_2(t)$ during $t_1 < t < t_2$, etc., with each mixture being zero outside the time interval.

As result, the signal from the moving sensor can be segmented to extract the mixture signals, which are then artificially synchronised and treated as if they were acquired simultaneously from an array of sensors. Then, any ICA algorithm can be used to extract the underlying sources.

This work was supported by: EU Framework 7 FET-Open project FP7-ICT-225913-SMALL: Sparse Models, Algorithms and Learning for Large-Scale data; Platform Grant: Centre for Digital Music (EP/E045235/1); Leadership Fellowship (EP/G007144/1) from the UK Engineering and Physical Sciences Research Council (EPSRC); and FCT (Portuguese Science and Technology Agency) grants: SFRH/BD/61655/2009 and PTDC/EIA-CCO/100844/2008

2. HEART SIGNALS

Heart signals include the electrocardiogram (ECG) signal, and heart sounds. The ECG is a recording of the difference in potential between two electrodes during the cardiac cycle, and provides important information about the performance of the heart. ECG signal analysis typical entails removal of noise and interference. A related problem is the elimination of maternal ECG (MECG) components from ECG signals recorded during pregnancy, also known as foetal ECG extraction. When risk factors are present during pregnancy, electrocardiograms, along with other measurement methods, may be of vital importance to both mother and child. This problem has been studied at length [8], and several ICA algorithms have been used to extract the foetal heartbeat. Because of this, we use our algorithm to address this problem, as a step to establish the validity of the proposed method. We do this by simulating a relocated sensor, then aligning the mixtures, and comparing the output of ICA applied on this, with the output of conventional ICA.

The act of hearing the heart sounds is known as *auscultation*, and is one of the cheapest and most useful techniques for the diagnosis of heart disease. Whenever a clinician is performing an auscultation he observes the main constituents of a cardiac cycle: The first heart sound (typically referred to as S1) and the second heart sound (S2), which has two components: A2 and P2, as shown in Figure 1(a). The A2 and P2 components are of interest to the clinician, because they can be used as an indicator of the presence and severity of a number of cardiac abnormalities [9].

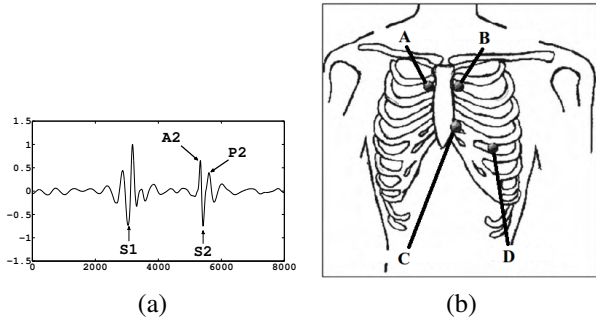


Fig. 1. a) The heart sound and b) Standard auscultation sites

3. SEPARATION VIA SEQUENTIAL SIGNAL SYNCHRONISATION

In this paper we aim to separate the A2 and P2 components. We do this in a way that is inconspicuous to the clinician, who remains free to perform a routine examination, including listening to the heart sounds sequentially at four standard sites, as shown in Figure 1(b). Since simultaneous recordings are not available, ICA cannot be applied directly. An alternative

approach is to take advantage of the quasi-periodic nature of the heart sound, and align the observed signals so that they are artificially synchronised. These signals are then presented to the ICA algorithm as if they were acquired simultaneously.

Algorithm 1 Separation via sequential signal synchronization

1. Segment the observed signal to extract the recordings at the different locations using:

$$\tilde{x}_i(t) = \begin{cases} \bar{x}_i(t) & \text{if } t_i < t < t_{i+1} \\ 0 & \text{otherwise} \end{cases}$$

where t_i represents the time at which recording at the next thoracic (chest) location begins.

2. Align the mixture signals using

$$\hat{x}_i(t) = \begin{cases} \tilde{x}_i(t - t_i) & \text{if } i = 1 \\ \tilde{x}_i(t - (t_i + \delta_i)) & \text{otherwise } (i > 1) \end{cases}$$

3. Form the mixture vector

$$\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \dots, \hat{x}_n(t)]^T$$

4. Apply FastICA.
-

In the scheme that we present, only one sensor is available, and it is firstly used to acquire a mixture of the sources from location *A* on the chest (see Figure 1(b)). The same sensor is then placed at the *B* location, and so on, until four observations are obtained. The sensor signal, $\bar{x}(t)$, contains the heart signal from the four locations, with periods of silence when the sensor is being relocated. Note that the timing between the source signals is not affected by the way the signals are acquired. In order to generate a mixture signal vector as in equation (1), we proposed to segment $\bar{x}(t)$ into four signals, so that the i -th heart signal is given by

$$\tilde{x}_i(t) = \begin{cases} \bar{x}_i(t) & \text{if } t_i < t < t_{i+1} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

where t_i represents the time at which recording at the next thoracic (chest) location begins (eg. the sensor is placed at *A* at time t_1 , at *B* at time t_2 , and so on). We align the signals, so that they can be presented to the ICA algorithm as if they were acquired simultaneously. In doing this, we exploit the quasi-periodicity of the heart cycle. We seek to align the peaks of the mixtures in $\tilde{x}_i(t)$, according to

$$\hat{x}_i(t) = \begin{cases} \tilde{x}_i(t - t_i) & \text{if } i = 1 \\ \tilde{x}_i(t - (t_i + \delta_i)) & \text{otherwise } (i > 1), \end{cases} \quad (4)$$

where δ_i is the relative time shift to get the peaks to align. It is evaluated using the cross-correlation function:

$$\delta_i = \sum_{n=0}^{N-m-1} \tilde{x}_i(m) \tilde{x}_i^*(n+m), \text{ for } i \neq j. \quad (5)$$

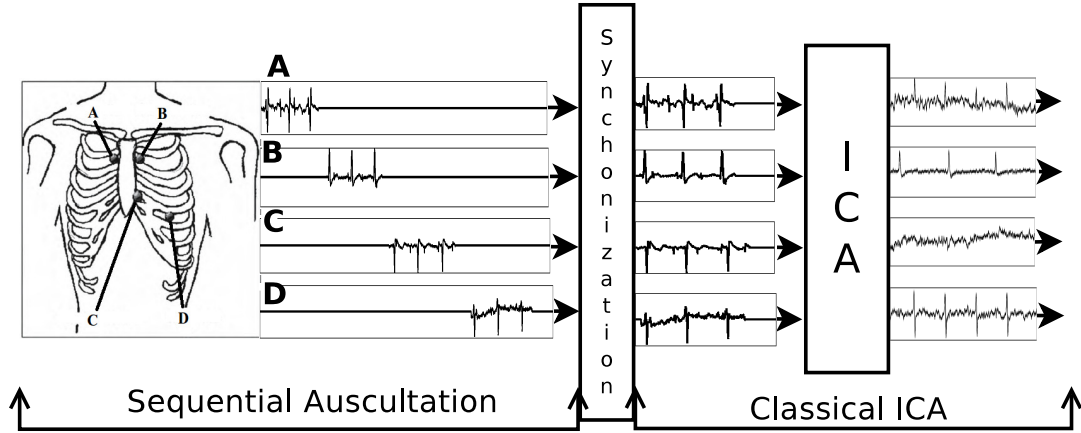


Fig. 2. The proposed method

Finally, the FastICA algorithm is used to separate the independent components from $\hat{\mathbf{x}}(t) = [\hat{x}_1(t), \dots, \hat{x}_n(t)]^T$. The proposed algorithm is summarized in Algorithm 1.

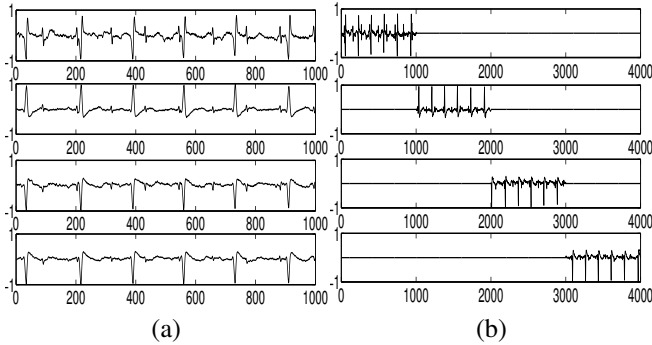


Fig. 3. Maternal ECG mixed with foetal ECG. The classical input to ICA a) and the input for the proposed method b).

4. VALIDATION ON ELECTROCARDIOGRAM - FOETAL ECG EXTRACTION

To test the validity of our method, we used it to address the foetal ECG extraction problem, on the signals described in [10]. To simulate a relocating sensor, we selected random sections of the recorded signals from each of four sensors. Then we aligned the sections as described previously and applied FastICA[6] to perform source separation. Figure 3(a) shows the mixture signals, while Figure 3(b) illustrates the simulated sequential signals, prior to alignment. The separated signals are illustrated in Figures 4(a) and (b), where the former shows the output of the FastICA algorithm using the classical approach (note the graph the first extracted signal has the higher heart rate, from the foetus), while the latter shows the output of the FastICA algorithm using the proposed method. Comparing the two figures, we can see that the proposed algorithm

successfully extract the foetal ECG, and generally recovers the same sources as conventional ICA. The scaling ambiguity of ICA causes the difference in the second source extracted by the two algorithms.

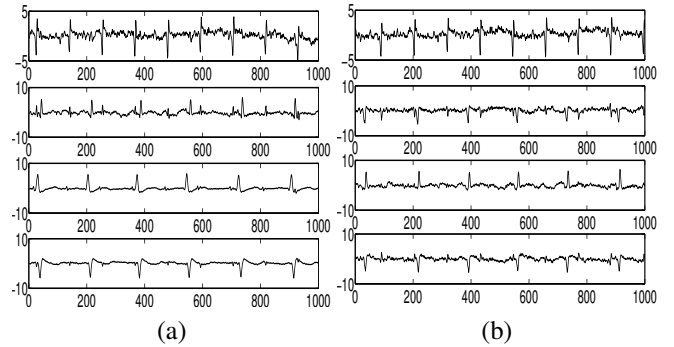


Fig. 4. Output of ECG signals using classical (a) and proposed method (b). In a) and b): the first signal is the foetal ECG. The second and forth signals are a mixture with maternal and foetal ECG, and the third signal is maternal ECG.

5. EXPERIMENTS ON HEART SOUNDS

In this section we apply our algorithm to heart sounds acquired sequentially at the locations shown in Figure 1, from a healthy adult. They were recorded using an electronic stethoscope from Welch Allyn model Elite, connected to a laptop to capture the signal. The sound was digitised at 44.1KHz, with 16 bits resolution.

Since most of the frequency content of heart sounds is found below about 1.5KHz, the audio signal was resampled down to 8KHz. The signal was also normalised so that the amplitude of the waveform lies between $[-1, +1]$. To remove undesired noise and to emphasise the heart sounds we performed wavelet threshold denoising [11].

The proposed algorithm was then applied to the signal. Figure 5 shows examples of one heart cycle taken from the four waveforms that have been aligned, and prior to input to ICA. They are labelled with the locations to which they correspond. The FastICA algorithm was subsequently applied to this set of synchronised mixture signals, and the following parameters were selected: non-linearity function $g(u) = \tanh(a_1 * u)$, fine tuning function $g(u) = u^3$, and epsilon 10^{-7} .

Figure 6 shows the separated components. In particular, the A2 component can be seen in the second output, at sample 5032, while the P2 component is observed in the fourth output, at sample 5119. These A2 and P2 components are similar to the components reported in the literature [12]. Moreover, their relative sizes are in agreement with what is expected in a healthy patient (A2 comes earlier than P2 and has a greater amplitude). These results were validated by a clinician (SSM).

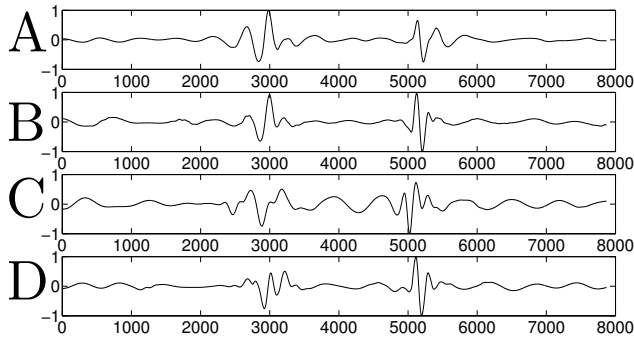


Fig. 5. An example of heart cycles acquired on each of the four sites depicted on Figure 1.

6. CONCLUSIONS

In this paper we introduced a new method for the separation of sources when only one sensor is available to record their mixtures. The proposed method segments the recorded signal into several mixtures, and exploits the periodic nature of heart signals to align these. The result is a mixture vector from which ICA can separate the underlying sources. The algorithm was validated on ECGs and subsequently applied to heart sounds, from which it was found to extract two components that are particularly interesting from the clinical point of view.

7. REFERENCES

[1] S. Makeig, Anthony J. B., T.P. Jung, and T.J. Sejnowski, "Independent Component Analysis Of Electroencephalographic Data," *Adv. in neural inf. proc. sys.*, pp. 145–151, 1996.

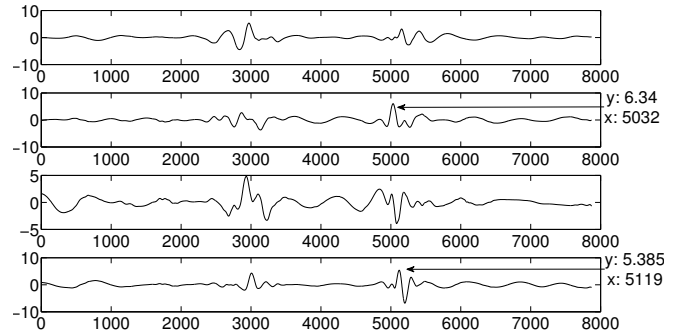


Fig. 6. FastICA output. The estimated A2 component is depicted on second graph, the estimated P2 component is depicted on fourth graph.

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

[3] A. Hyvarinen, "Survey on Independent Component Analysis," *Neural Computing Surveys*, vol. 2, pp. 94–128, 1999.

[4] A.J. Bell and T.J. Sejnowski, "An Information Maximization Approach To Blind Separation And Blind Deconvolution," *Neural computation*, vol. 7, no. 6, pp. 1129–1159, 1995.

[5] J.F. Cardoso and A. Souloumiac, "Blind Beamforming For Non-Gaussian Signals," *IEE Proc. F Radar And Signal Processing*, vol. 140, pp. 362–362, 1993.

[6] A. Hyvarinen, "Fast And Robust Fixed-Point Algorithms For Independent Component Analysis," *IEEE Trans. On Neural Networks*, vol. 10, no. 3, pp. 626–634, 1999.

[7] S.T. Roweis, "One microphone source separation," *Neural Inform. Process. Syst.*, pp. 793–799, 2000.

[8] M. Jafari, "Novel Sequential Algorithms For Blind Source Separation of Instantaneous Mixtures," September 2002.

[9] Morton E. Tavel, *Clinical Phonocardiography And External Pulse Recording*, Y. B. Medical Publishers, 1972.

[10] L. De Lathauwer, D. Callaerts, B. Moor, and J. Vandewalle, "Fetal Electrocardiogram Extraction By Source Subspace Separation," in *In Proc. HOS'95*, 1994.

[11] D.L. Donoho, "De-noising By Soft-thresholding," *IEEE trans. on Info. theory*, vol. 41, no. 3, pp. 613–627, 1995.

[12] V. Nigam and R. Priemer, "A Procedure To Extract The Aortic And The Pulmonary Sounds From The Phonocardiogram," in *IEEE 28th EMBS*, 2006, pp. 5715–5718.