Hybrid BSS techniques for foetal ECG extraction using framework for stress-testing extraction algorithms

Técnicas BSS híbridas para extracción del ECG fetal usando un marco de referencia para pruebas de esfuerzo de algoritmos de extracción

Luis Omar Sarmiento-Álvarez
M.Sc. Potencia Eléctrica
Unidades Tecnológicas de Santander
Bucaramanga, Colombia
lsarmiento@correo.uts.edu.co

José Millet-Roig
Ph. D. Ingeniería industrial
Universidad Politécnica de Valencia
Valencia, España
jmillet@eln.upv.es

1. INTRODUCTION

The foetal electrocardiogram (FECG) is a viable alternative to Cardiotocography (CTG) to assessing foetal cardiac activity. The foetal electrocardiogram (FECG) can be recorded invasively using a needle-like electrode attached to the foetal scalp, but this technique, have associated risk of infec-

Keywords – FECG, ICA, PCA, stress-testing
tion and it can be used during labour only. The non-invasive FECG (NI-FECG), makes use of surface electrodes placed onto the maternal abdomen. The challenge is to extract the foetal signal from the abdominal ECG (AECG) mixture since FECG overlaps both in frequency and time domain with maternal interfering source ECG (MECG), and various muscular noise sources.

Several FECG extraction approaches have produced interesting results: methods based on blind or semi-blind source separation techniques, independent component analysis (ICA), [1],[2], principal component analysis (PCA) or singular value decomposition (SVD) [3], [4], [5]; average MECG subtraction [6], [7]; different variants of adaptive filters [8],[9],[10]; wavelet decomposition[11].

BSS is the most widely used technique for extracting the FECG, since it has been shown that BSS methods are better than the adaptive filters [12], and nonlinear transformations have greater computational load and require some parameters to be set empirically.

In this paper we introduce COMBI [13] and MULTICOMBI [14] algorithms based on hybrid Blind Source Separation (BSS) for extracting FECG from AECG. Aiming at accurate detection of fetal QRS complex [FQRS], we compare the performance of COMBI, MULTICOMBI, ICA, JADE [15], PCA, WASOBI [16], and EFICA[17] using a standardized evaluation procedure for NI-FECG signal processing algorithms[18].

2. METHODS

2.1 Blind Source Separation

In the context of NI-FECG extraction, BSS aims to separate the FECG the others underlying statistically independent sources, MECG and noise. To solve this problem different methods have been proposed, Principal Component Analysis (PCA) and Independent Component Analysis (ICA). PCA assume that sources x have a Gaussian distribution function completely determined by its second-order statistics, while ICA assumes that sources x are statistically independent and are not Gaussian and that the distribution function is determined by its higher order statistics.

One way to implement PCA is the SOBI algorithm (Second Order Blind Identification)[19] and its improved version WASOBI (Weights-Adjusted Second Order Blind Identification)[20]. In turn, one way to implement ICA is the FASTICA algorithm [21] and its improved version EFICA [22].

ICA-based methods are the most used for the analysis of AECG because it is considered that the sources are predominantly non-Gaussian and statistically independent of one another. But, since real-world signals are both Gaussian and non-Gaussian components, it is interesting to test methods that incorporate both approaches. In that sense, algorithms have been developed which combine both WASOBI and EFICA algorithms. These algorithms are called COMBI and MULTICOMBI. Recently these algorithms have been applied in the task of extracting the FECG from the AECG using a proprietary semi synthetic database [23]. However, due to the lack of annotated public databases and defined protocols for assessing these algorithms, available studies may be biased and of questionable reproducibility[18].

2.2 Combi and Multicombi algorithms

Under certain conditions, WASOBI and EFICA are asymptotically optimal. WASOBI only take advantage of time-structure, disregarding the statistical distributions of the sources, whereas EFICA can only take advantage of non-Gaussianity of the sources, ignoring any time-structure. However, realistic mixtures are many times compound of sources which present both diverse time-structure and non-Gaussianity, rendering WASOBI and EFICA severely suboptimal. COMBI and MULTI-COMBI Algorithms offer novel schemes for combining WASOBI and EFICA, enabling exploit the strengths of both techniques [24].

To verify a good degree of separation, $G=WA$ is defined as the gain matrix, with $A$ the original mixing matrix. For a perfectly estimated de-mixing matrix, $W$, $G$ is equal to its identity matrix. The performance of blind-source separation algorithms is usually measured by the interference over signal ratio matrix, $ISR_{kl} = G_{kl}^2/G_{kk}^2$, $k,l = 1,2,...,d$, where $k$ and $l$ denote the observed and estimated sources, and $d$ is sources number. However, the
original mixing matrix, A, is not generally known for real data sets.

EFICA requires a user-defined choice of a set of nonlinear functions \( g_i(\cdot) \), for extracting each of the sources. Then, ISR matrix for the EFICA algorithm can be approximated by

\[
\text{ISR}_{ii}^{\text{EF}} \approx \frac{1}{N} \frac{\gamma_i(\gamma_i + \tau_i)}{\gamma_i + \tau_i}
\]

where \( \gamma_i = \beta_i - \mu_i \), \( \mu_i = E[S_i g_i(S_i)] \), \( \tau_i = |\mu_i - \rho_i| \), and \( \rho_i = E[S_i g_i(S_i)] \). The expectation operator and \( g_i(\cdot) \) denotes the derivative of \( g_i(\cdot) \), and \( S_i \) is the kth observed signals of \( S \) [14].

WASOBI is based on approximate joint diagonalization of several (say M) time-lagged estimated correlation matrices, \( R_{kl} = \frac{1}{N} \sum_{n=1}^{N} x_{ik} x_{jl} \tau_k \) for extracting each of the sources, \( \{a_n\}_{n=0}^{\infty} \) are the auto-regression coefficients of the lth source, and \( R_{n} \) is the autocorrelation of the \( k \)th source at time lag \( m \) [13].

In COMBI, the ISR matrices are obviously unknown (nor can they be consistently estimated from the data). However, it is possible to substitute these with the mean ISR, \( \overline{\text{ISR}}^{\text{M4}} \) and \( \overline{\text{ISR}}^{\text{OPT}} \). COMBI applies both EFICA and WASOBI to \( \overline{\text{ISR}}^{\text{M4}} \) and \( \overline{\text{ISR}}^{\text{OPT}} \) select for each source the reconstructed version that has the best total ISR of the two. This basic selection approach can then be turned into a successive scheme, such that in each iteration only the “best” separated sources are “accepted,” and the remaining signals (which are still weakly separated mixtures of the remaining sources) are subjected to an additional iteration of separation and selection [13].

MULTICOMBI uses a clustering technique based on “multidimensional component”. A multidimensional component is a cluster of signal components that can together be well separated from the other components in the mixture, yet are difficult to separate from one another. For EFICA, only components that have (nearly) Gaussian distributions might form such a cluster, hence at most one such cluster may exist. For WASOBI, any components sharing similar correlation structures (i.e., power spectra) are hardly separable from one another, but may be easily separated as a cluster, hence several such clusters might coexist [14]. MULTICOMBI uses this clustering technique in which both algorithms, EFICA and WASOBI, are run on the set of unseparated sources \( \overline{\text{X}} \) and their \( \overline{\text{ISR}}^{\text{M4}} \) and \( \overline{\text{ISR}}^{\text{OPT}} \), in (1) and (2), are estimated. The signals are then clustered depending on whether their specific ISR is lower for the EFICA or WASOBI case. Then, the process is repeated until all clusters are singletons, i.e. only contain one signal per cluster, and the signals are hence optimally separated [25].

2.3 Framework for stress-testing NI-FECG algorithms

In this work, we use a standardized methodology for stress testing NI-FECG algorithms [18]. The framework include a large database (FECGSYNDB) of realistic artificial signals of multichannel data absolute data, as well as extraction and evaluation routines. This dataset include several non-stationary events (e.g. foetal movements, uterine contractions and heart rate fluctuations) for evaluating extraction routines [26]. FECGSYNDB is the result of a simulator representing maternal and foetal hearts as punctual dipoles with different magnitudes and spatial positions. The database contains foetal–maternal mixtures by treating each abdominal signal component (e.g. foetal/maternal ECG or noise signals) as an individual source, whose signal is propagated onto the observational points (“electrodes”). This framework is useful to investigate the output of NI-FECG processing BSS methods in the presence of such non-stationarities. The modeled volume conductor, sampled from 34 channels (32 abdominal and 2 maternal ECG reference channels), is shown in Fig. 1.
2.4 Dataset and preprocessing.

In this work, we use the Sub01 dataset in FE-CGSYND (available in https://physionet.org/physiobank/database/fecgsyndb/). In this dataset, seven different physiological events were considered as described in Table I. Besides this, there were five different levels of additive noise (0, 3, 6, 9 and 12 dB) and 5 repetitions for each combination of settings. Overall a total of 7 (cases) x 5 (SNR levels) x 5 (repetitions) = 175 synthetic signals were used. Each signal had a duration of 5 minutes, and was sampled at 250 Hz with a 16-bit resolution. The SNR levels are 0, 3, 6, 9 and 12 dB.

<table>
<thead>
<tr>
<th>Case</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>Abdominal mixture (no noise or events)</td>
</tr>
<tr>
<td>0</td>
<td>Baseline (no events) + noise</td>
</tr>
<tr>
<td>1</td>
<td>Foetal movement + noise</td>
</tr>
<tr>
<td>2</td>
<td>MHR /FHR acceleration / decelerations + noise</td>
</tr>
<tr>
<td>3</td>
<td>Uterine contraction + noise</td>
</tr>
<tr>
<td>4</td>
<td>Ectopic beats (for both foetus and mother) + noise</td>
</tr>
<tr>
<td>5</td>
<td>Additional NI-FECG (twin pregnancy) + noise</td>
</tr>
</tbody>
</table>

Source: the authors

2.5 Performance parameters

In order to assess the accuracy of FQRS detection, we use the F1-measure[18]:

\[ F_1 = \frac{2TP}{2TP + FN + FP} \]  

where true positive (TP) denotes correctly detected peaks, false negative (FN) existing peaks which were not detected, and false positive (FP) nonexistent peaks that were falsely detected, as defined in ANSI/ AAMI EC57. It is defined a ± 50 ms acceptance interval between detection and the closest reference annotation to account for the higher expected fetal heart rate. The F1 measure well adopted by many participants in the Physionet Challenge 2013 as well as used for the Challenge 2014 [27].

3. EXPERIMENTS AND RESULTS

Two experiments were carried out. One to determine a WASOBI algorithm input parameter, also necessary in COMBI, and other to compare performance using F1 measure.

3.1 AR order

COMBI and MULTICOMBI algorithms combine WASOBI and EFICA algorithms. In WASOBI, an asymptotically optimal weight matrix must be obtained for the cases of Auto-Regressive (AR) Gaussian sources.

Therefore, the algorithm requires as input the AR order, the first experiment focused in this preliminary consideration relevant to the usage of COMBI in the following experiments. Sub01 dataset was
processed with COMBI, AR order between 2 and 10. In order to determine the highest achievable accuracy by each AR order, the foetal R peak detection results were obtained using the F1 measure. The Sub 01 dataset was pre-processed with 0.5-100 Hz Bandpass filter. Each signal has 32 channels available, as depicted in Fig. 1, we use eight (1, 8, 11, 14, 19, 22, 25 and 32).

3.2 COMBI and MULTICOMBI comparing

This experiment consisted of comparing COMBI and MULTICOMBI with ICA (deflation approach, symmetric method, and JADEICA), PCA, EFICA and WASOBI BSS based NI-FECG extraction techniques, in stationary and non-stationary scenarios by means of F1 measure. The experiment aims at obtaining an overview of how the presence of non-stationary mixtures affect extraction methods. The AR order used in this analysis was decided considering results from the previous experiment. The same channels and bandwidth in first experiment were used.

3.3 Results

Results of a previous study showed that COMBI produces better results than MULTICOMBI at the work removing the FECG from AECG [28]. For this reason, COMBI was selected to represent hybrid BSS techniques in experiment 1. Table II shows the F1 results of COMBI, AR order between 2 and 17. AR order = 8 and AR order = 10 have the same result. AR order = 10 was selected for its lower dispersion. Cases with noise or events were taken into account only, i.e. case 0 to case 5.

<table>
<thead>
<tr>
<th>COMBI AR order</th>
<th>F1 (%)</th>
<th>Case 0</th>
<th>Case 1</th>
<th>Case 2</th>
<th>Case 3</th>
<th>Case 4</th>
<th>Case 5</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>98.9 ± 2.1</td>
<td>95.8 ± 5</td>
<td>97.6 ± 6.2</td>
<td>83.4 ± 15.5</td>
<td>85.2 ± 2.1</td>
<td>91.4 ± 15.8</td>
<td>92.1 ± 6.5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>98.1 ± 6.9</td>
<td>94.4 ± 7.4</td>
<td>98.8 ± 3.1</td>
<td>84.6 ± 15.3</td>
<td>85.9 ± 2</td>
<td>91.1 ± 17.1</td>
<td>92.2 ± 5.5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>97.4 ± 10.6</td>
<td>97.1 ± 4.3</td>
<td>99.9 ± 0.2</td>
<td>87.9 ± 15.3</td>
<td>85.3 ± 2.1</td>
<td>93.4 ± 15.7</td>
<td>93.5 ± 5.3</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>97.5 ± 4</td>
<td>95.4 ± 6.3</td>
<td>99.8 ± 0.3</td>
<td>93.9 ± 8.1</td>
<td>85.1 ± 2.1</td>
<td>90.4 ± 17.8</td>
<td>93.7 ± 5.1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>99.5 ± 0.9</td>
<td>97.4 ± 4.1</td>
<td>97.6 ± 4.1</td>
<td>91.9 ± 10.5</td>
<td>82.1 ± 7.4</td>
<td>94.1 ± 10.2</td>
<td>93.7 ± 10.7</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>89.6 ± 14.2</td>
<td>90.7 ± 11.8</td>
<td>89.9 ± 13.9</td>
<td>85.1 ± 11.1</td>
<td>64.7 ± 17.6</td>
<td>91.8 ± 13.7</td>
<td>85.3 ± 13.8</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>98.5 ± 4.2</td>
<td>97.2 ± 5.1</td>
<td>98 ± 3.7</td>
<td>93.5 ± 8.1</td>
<td>82.6 ± 7.4</td>
<td>95.7 ± 10</td>
<td>94.3 ± 10.1</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>92.8 ± 13.8</td>
<td>96.9 ± 4.9</td>
<td>96.2 ± 9.7</td>
<td>92.2 ± 10.3</td>
<td>81.5 ± 9.3</td>
<td>93 ± 15.8</td>
<td>92.1 ± 8.5</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>98.6 ± 4.3</td>
<td>97.2 ± 5.1</td>
<td>98 ± 3.7</td>
<td>93.5 ± 8.1</td>
<td>82.6 ± 7.4</td>
<td>95.7 ± 10</td>
<td>94.3 ± 6.0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>98.4 ± 4.4</td>
<td>97.2 ± 4.8</td>
<td>97.5 ± 4.2</td>
<td>92.1 ± 10</td>
<td>83.3 ± 9.1</td>
<td>94.6 ± 10.4</td>
<td>93.9 ± 5.6</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>98.4 ± 4.4</td>
<td>97.1 ± 4.6</td>
<td>97.9 ± 4.3</td>
<td>91.8 ± 9.9</td>
<td>82.6 ± 9.7</td>
<td>94.1 ± 10.8</td>
<td>93.7 ± 5.9</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>98.3 ± 3.3</td>
<td>95.7 ± 6.2</td>
<td>97.8 ± 5</td>
<td>94.1 ± 9.3</td>
<td>81.4 ± 6.8</td>
<td>92.9 ± 12.5</td>
<td>93.4 ± 6.2</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>99.3 ± 1.6</td>
<td>97.5 ± 4.3</td>
<td>97.6 ± 4.1</td>
<td>91 ± 10.7</td>
<td>82.7 ± 8.9</td>
<td>92.9 ± 15.1</td>
<td>93.5 ± 6.2</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>99 ± 1.9</td>
<td>97.5 ± 4.3</td>
<td>97.4 ± 4</td>
<td>91.8 ± 11</td>
<td>83 ± 9</td>
<td>95.3 ± 10.1</td>
<td>94.0 ± 5.9</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>99 ± 1.9</td>
<td>97.3 ± 4.3</td>
<td>98.1 ± 3.5</td>
<td>93 ± 9.6</td>
<td>83.9 ± 5</td>
<td>93.8 ± 12.1</td>
<td>94.2 ± 5.56</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>99.3 ± 1.5</td>
<td>96.6 ± 5.6</td>
<td>98.1 ± 3.4</td>
<td>93.3 ± 9.2</td>
<td>83.4 ± 5.1</td>
<td>93 ± 12.8</td>
<td>94.0 ± 5.8</td>
<td></td>
</tr>
</tbody>
</table>

Source: the authors
Box plots may also have (whiskers) indicating

The boxplot in Fig. 2 shows F1 results for second experiment. The variability outside the upper and lower quartiles is indicating by lines extending vertically from the boxes, and the point inside the box is the second quartile (the median). Here, Met1= ICAsym, Met2=ICAdef, Met3=JADEICA, Met4=PCA, Met5=EFICA, Met6= WASOBI, Met7= COMBI and Met8= MULTICOMBI. COMBI has better results in Case 2, MHR /FHR acceleration / decelerations + noise and Case 5, additional NI-FECG (twin pregnancy) + noise. On the other hand, COMBI has similar results to higher performance algorithms in Case 0, abdominal mixture + noise and Case 3, uterine contraction + noise. Finally, COMBI has slightly lower results in Case 1, Foetal movement + noise and Case 4, ectopic beats (for both foetus and mother) + noise.

In Fig. 3, the overall median FQRS detection accuracies (i.e. considering all no stationary events) are obtained. For the best performing algorithms in each group were 96.4% for COMBI, 95.8% for EFICA, 95.2% for JADEICA, 94.6% for PCA, 93.6% for ICAdef, 92.9% for MULTICOMBI, 92.7% for ICAsym, and 89.7% for WASOBI.

4. CONCLUSIONS

In this work, we evaluate eight BSS based algorithms, including two hybrid algorithms that combining both WASOBI and EFICA. We use a standard framework for benchmarking and regulatory testing of NI-FECG extraction algorithms. Thus, our results can be checked and verified by other researchers. We find that the best AR order required as input for COMBI is 10. Benchmarking showed that COMBI have overall median FQRS detection performance. In turn, WASOBI exhibits the worst performance, but combining both WASOBI and EFICA...
in COMBI and MULTICOMBI, the strengths of both techniques are exploited exhibiting better performance. From these results we can conclude that FECG signals have both Gaussian and non-Gaussian components, Gaussian lesser degree, as also was demonstrated in [23],[24] using a semi-synthetic database. However, this statement must be verified in real databases. MULTICOMBI presents more dispersed results than COMBI. In that sense, the clustering scheme needs to be optimized for the task of separating the sources in AECG, or new methods that combine ICA and PCA and exploit characteristics present in AECG can be developed.

REFERENCES


