The Effect of the Whitening Matrix in Determining the Final Solution in Blind Source Separation of Biomedical Signals

Hasan Al-Nashash and Husein Abdul-Hamid
Email: {hnashash, hhamid}@aus.ac.ae
American University of Sharjah, P.O. Box 26666, Sharjah, UAE

Abstract: In this paper, independent component analysis (ICA) is used for blind source separation of biomedical signals. Visual and quantitative tests of the ability of ICA to separate signals were performed using a fast ICA algorithm. Results obtained from simulated and FECG signals show that the ICA performance using the whitening matrix of the mixed signals was superior to that of random initial weights.

I. INTRODUCTION

The problem of detecting and separating a desired biomedical signal corrupted by other periodic interference and random noise signals is of extreme importance in medicine. Examples include the Fetal Electrocardiograph (FECG) [1], His Purkinje System Electrogram (HPSE) [2], Ventricular Late Potentials (VLP) [3], and the Diaphragmatic Electromyogram (EMGdi) [4]. Electronic random noise and 50 Hz power line interference represent two major sources of undesired signals in all of the above applications. The random noise signal generated by muscular activities corrupts the FECG, HPSE and VLP signals. Also, the ECG is considered as an undesired signal in both the FECG and EMGdi. The FECG signal reflects the electrical activity of the fetal heart. It contains information on the health status of the fetus and therefore, an early diagnosis of any cardiac defects before delivery increases the effectiveness of the appropriate treatment [1]. The HPSE signal is recorded at the body surface as microvolt potentials that reflect the electrical activity of the specialized conduction system of the heart. It helps to further understand the physiology and identify various cardiac pathologies associated with the HPS and not with the myocardium. VLP are microvolt signals that are part of the terminal portion of the QRS complex and continue into the ST-T segment. They represent areas of delayed ventricular activation, which are manifestations of slowed conduction velocity. One of the most advantages of detecting VLP is the ability to predict the likelihood of sustained ventricular tachycardia, ventricular fibrillation, and sudden cardiac death in patients post myocardial infarction. The EMGdi signals are recorded during inspiration and expiration and used for respiration monitoring and respiratory control mechanisms. From the clinical point of view, the EMGdi can be used to assess the level of respiratory fatigue which is characterized by a progressive inability of respiratory muscles to maintain the work load demanded by the respiratory drive. In addition to separating signal from noise, the estimation, location and distribution of the electric current sources within the brain from the Electroencephalograph (EEG) signals are fundamental problems in the area of neurological monitoring and instrumentation [5].

For each of the above signals, several research groups have proposed different techniques to improve the signal to noise ratio (SNR) of the desired signal. Ensemble averaging, spatial averaging, cross-correlation, adaptive filtering and wavelet analysis were among several of these techniques [2-4, 6-7]. Despite the reported successes of these methods, they are still not used clinically at a large scale. In this research, we aim to investigate the use of independent component analysis technique for the detection of desired biomedical signals mixed with other types of undesired signals.

Technically, the above problem can be thought of as a set of desired and undesired signals linearly mixed to produce another set of body surface signals. It is assumed that these signals are non-Gaussian (except the random noise signal) and independent. ICA decomposes the mixed signals into as statistically independent components as possible. ICA has been used recently to detect FECG [1] and to process magnetoencephalogram (MEG) data [8]. Several ICA algorithms have been proposed. In this paper, the Fast ICA (FICA) algorithm, which was proposed by [9-11], is used.

Section II includes the theory of ICA followed by results obtained from computer simulations and biomedical signal analysis as shown in in section III. Finally, conclusions and future work are given in section IV.

II. THEORY

Assume that the set of desired and undesired signals, \( S \), have been mixed to produce an array of body surface signals, \( O \). This can be modeled by a linear latent vector, \( O \), as shown in the following equation:

\[
O = MS
\]
The Effect of the Whitening Matrix in Determining the Final Solution in Blind Source Separation of Biomedical Signals

<table>
<thead>
<tr>
<th>Title and Subtitle</th>
<th>Contract Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Effect of the Whitening Matrix in Determining the Final Solution in Blind Source Separation of Biomedical Signals</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Project Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performing Organization Name(s) and Address(es)</th>
<th>Project Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>American University of Sharjah, P.O. Box 26666, Sharjah, UAE</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sponsor/Monitor's Acronym(s)</th>
<th>Task Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sponsor/Monitor's Report Number(s)</th>
<th>Work Unit Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performing Organization Report Number</th>
<th>Task Number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution/Availability Statement</th>
<th>Task Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Approved for public release, distribution unlimited</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Abstract</th>
<th>Classification of Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td>Papers from the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 25-28 October 2001, held in Istanbul, Turkey. See also ADM001351 for entire conference on cd-rom., The original document contains color images.</td>
<td>unclassified</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject Terms</th>
<th>Limitation of Abstract</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>UU</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Number of Pages</th>
<th>Classification of this page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>unclassified</td>
</tr>
</tbody>
</table>
The choice of function $g$ is the derivative of the general contrast function $J$ that measures the non-normality of a zero-mean random variable $y$:

$$J_g(y) = |E_g(G(y)) - E_g(G(\mathbf{v}))|^p$$

(4)

where $\mathbf{v}$ is a standardized Gaussian random variable, and the exponent $p=1.2$ typically.

Despite the above good properties, logistic drawbacks have been noticed which affected the usefulness of the algorithm:

1) The order of the signals was changing after different runs. This prevented us from indexing the signals, which is very important for clinical applications. Even, in many applications, one does not need to estimate all the independent components. In the ideal case where the one-unit contrast functions are optimized globally, the independent components are obtained in the order of (descending) non-Gaussianity.

2) Although that the outputs signals resemble the input signals morphology, the gain factor is not unity. This may give false indications of the source magnitude values.

3) Sometimes the output signals are inverted in a random manner at different runs.

III. RESULTS

To investigate the usefulness of the FICA algorithm in separating desired from undesired signals, visual and quantitative experiments were conducted on simulated and biomedical signals. The data used for this study was contributed by Lieven De Lathauwer [12,13]. These signals were recorded from eight different skin electrodes located on different points of a pregnant thoracic region containing only MECG and five of these simultaneous signals containing MECG and thoracic region containing only MECG. Fig.1 shows eight

![Fig.1: Output of the FICA using random weights.](image-url)
signals obtained after applying the FICA algorithm where
the initial weights were selected randomly. It was found
that the random noise was estimated as the first
independent component while the desired FECG was
estimated as the seventh component. Furthermore, it was
also found that the order of the FECG signal was changing
after almost every run of FICA. This problem was solved
by using the whitening matrix computed using PCA to
initialize the weights. The whitening matrix represents the
eigenvectors of the covariance matrix of the observed
signals. Fig. 2 shows the same eight signals but the FECG
was estimated as the first independent component.

To further investigate this problem, three signals were
simulated: a periodic pulsating signal, a sinusoidal signal
and a Gaussian random signals which resemble the
desired, interference and noise signals respectively as
shown in Fig 3. Fig. 4 shows the signals after being mixed
with a mixing matrix $M$. Fig. 5 shows the output signals
extracted using the FICA algorithm were the whitening
matrix was again used to initialize the weights of the
FICA algorithm. It was also found that the order of the
estimated independent components was fixed after
performing several runs.

The importance of the whitening matrix stems from the
fact that PCA acts to orthogonolize and decorrelate the
mixed signals. It is clear that the covariance matrix of the
mixed signals has a great role in determining the
Guassianity of these signals. Therefore, the magnitude of
the eigenvalues helps to speed the FICA algorithm to
search for an estimate of the first non-Gaussian
independent component.

To quantify the performance of the FICA with random
initial weights against using the whitening matrix, the
percent root mean square difference was used. It is defined
as:

$$\text{PRD}_\% = \frac{\sum_{i=1}^{N} [\text{Source}(i) - \text{Estimated}(i)]^2}{\sum_{i=1}^{N} [\text{Source}(i)]^2} \times 100$$  \hspace{1cm} (4)$$

where $\text{Source}(i)$ is the $i^{th}$ sample of the original pre-mixed
desired signal, $\text{Estimated}(i)$ is the $i^{th}$ sample of the
estimated desired signal and $N$ is the total number of
source samples under test.

Fig. 6 shows the averaged PRD% of the desired pulsating
signal versus its SNR at the input of the FICA for both
random and whitening matrices. Results show that the
PRD% is almost zero when using the whitening matrix
while it ranges between 30%-50% when using random
initial weights. These results confirm the earlier findings.

IV. CONCLUSIONS

Independent component analysis can be used to solve the
problem of separating desired biomedical from undesired
biomedical, random and periodic noise signals. In this
paper, the FICA algorithm was used to detect the FECG. It
was found that it suffers from a number of practical
limitations. The effect of initial weights on the FICA
algorithm performance was investigated.
The authors would like to acknowledge Lieven De Lathauwer, lieven.delathauwer@esat.kuleuven.ac.be of Katholieke Universiteit Leuven for providing us with the FECG data.

REFERENCES


