Biomedical Signal processing by Blind source separation on ECG signals

Prema K N[1]
[1]*Assistant Professor, Department of ECE, JNNCE, Shivamogga, Karnataka,India*

Abstract- Signal processing is a vast research area, in that bio medical signal processing has become the challenging area. Bio medical signals such as EEG, ECG and EMG are required to analyze for the proper diagnosis. The direct signals consist of noise and artefacts which should be reduced. There are many methods for reducing the noise and artefacts; here blind source separation method is used for the denoising process along with the filters.
Keywords:Electrocardiography, Blind source separation, Independent Component Analysis,

I. INTRODUCTION

The ECG is a diagnostic tool which is commonly used in the checking the electrical and muscular functions of the heart. While it is a relatively simple test to perform, the interpretation of the ECG tracing requires significant amounts of training.In a healthy heart, the heartbeat will have an orderly progression of a wave of depolarization that is triggered by the cells in the Sino atrialnode, spreads out through the atrium, passes through "intrinsic conduction pathways" and then spreads all over the ventricles. It is detected as small variations in the voltages of two electrodes which are connected across either side of the heart which is displayed as a wavy line either on a screen or on paper. The display indicates the overall rhythm of the heart and weaknesses in different parts of the heart muscle.

The recorded electrocardiograms (ECGs)signals aresometimes corrupted by different types of artefacts and many efforts have been made to enhance their quality by reducing the noise or artefacts. Noise is defined to be part of the real signal that confuses analysis (e.g. muscle movements) and artefact is defined to be any distortion of the signal caused by the recording process, such as electrode movement. Many attempts have been made to detect and eliminate noise sources and artefacts from the actual electrocardiographic signals.

Analogue or digital filters are widely used to reduce the influence of interference superimposed on the ECG. Early work on noise and artefact reduction in the ECG used either temporal or spatial averaging techniques. The temporal averaging method requires a large number of time frames for effective noise reduction, while the main drawback of spatial averaging is the physical limitation of placing a large number of electrodes in the same region. Besides linear noise filtering, several adaptive filtering methods have been proposed for separation and identification of the component waves from noisy ECGs . The quasi periodic pattern of the cardiac signal has also been exploited by synchronizing the parameters of the filter with the period of the signal. Other proposed methods include subspace rotations, neural networks , and bi−spectral analysis.

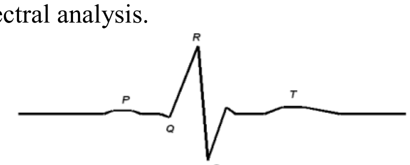

Fig1:Typical ECG waveform with the P, Q, R, S and T waves for one heart beat

ICA is a newly developed source separation method, and its application to biomedical signals is rapidly expanding. In the field of ECG analysis, Cardoso presented a good example of ICA decomposition for fetal and maternal ECGs recorded simultaneously from 8 electrodes placed on the mother’s chest and abdomen. Wisbeck et al. used ICA to isolate the breathing artefacts (large baseline shifts due to the physical movement of the electrodes in relation to the heart) from 4−channel ECG recordings.

The aim of biomedical data processing is to extract the medical signals clinically, biochemically or pharmaceutically relevant information (e.g. metabolite concentrations in the brain) in terms of parameters out of low quality measurements in order to enable an improved medical diagnosis .Since biomedical data are affected by large measurement errors, largely due to the noninvasive nature of the measurement process or the severe constraints to keep the input signal as low as possible for safety and bio-ethical reasons.

II. BIO SIGNALS

The physiological changes in the body affects on the measurement of the biomedical source signals which will be indicated by the function or malfunction of various physiological systems. To extract the relevant information for diagnosis and therapy, expert knowledge in medicine and engineering is also required. Biomedical source signals are usually weak, geostationary signals and distorted by noise and interference. Moreover, they are usually mutually superimposed. Besides classical signal analysis tools (such as adaptive supervised filtering, parametric or non parametric spectral estimation, time frequency analysis, and higher order statistics).

ICA (Independent component analysis)

The concept of independence can be defined by considering the two scalar valued random variables y_1 and y_2. Basically, the variables y_1 and y_2 are said to be independent if information on the value of y_1 does not give any information on the value of y_2, and vice versa. Above, we noted that this is the case with the variables s_1, s_2 but not with the mixture variables x_1,x_2.

Technically, independence can be defined by the probability densities. Let us denote by $p(y_1,y_2)$ the joint probability density function (pdf) of y_1 and y_2. Let us further denote by $p_1(y_1)$ the marginal pdf of y_1, i.e. the pdf of y_1 when it is considered alone:

$$p_1(y_1) = \int p(y_1,y_2)dy_2$$

Similarly for y_2. Then y_1 and y_2 are independent if and only if the joint pdf is factorizablein the following way

$$p(y_1,y_2) = p_1(y_1)p_2(y_2)$$

This definition extends naturally for any number n of random variables, in which case the joint density must be a product of n terms.

The definition can be used to derive a most important property of independent random variables. Given two functions, h_1 and h_2, we always have

$$E\{h_1(y_1)h_2(y_2)\} = E\{h_1(y_1)\}E\{h_2(y_2)\}$$

ICA can be applied, for example, for blind source separation, in which the observed values of x correspond to a realization of an m-dimensional discrete-time signal x(t), t = 1, 2,3,…. Then the components $S_i(t)$ are called source signals, which are usually original, uncorrupted signals or noise sources. Often such sources are statistically independent from each other, and thus the signals can be recovered from linear mixtures X_i by finding a transformation in which the transformed signals are as independent as possible, as in ICA. Another promising application is feature extraction, in which S_i is the coefficient of the i-th feature in the observed data vector X The use of ICA for feature extraction is motivated by results in neurosciences that suggest that the similar principle of redundancy reduction explains some aspects of the early processing of sensory data by the brain[1].

III. BLIND SOURCE SEPARATION

Blind source separation (BSS) is the method that decomposes the signal mixtures into the original sources. The technique of Independent component analysis (ICA) can be used to estimate the mixing block and the original sources based on the information of their independence's by ICA is important research field because of a lot of applications in biomedical signal processing, geophysical data processing, data mining, speech recognition and enhancement, wireless communication and so on.

Because ICA with same number of sources and mixtures has the square mixing matrix, the sources can be reconstructed almost perfectly by learning the inverse of mixing matrix with the independency of sources. But the underdetermined ICA, which has less mixture than sources, does not have the square mixing matrix, so we cannot learn the mixing matrix precisely.

The classical application of the ICA model is blind source separation. In blind source separation, the observed values of x correspond to a realization of an m-dimensional discrete-time signal X(t), t = 1; 2; :::. Then the independent components $S_i(t)$ are called source signals, which are usually original, uncorrupted signals or noise sources. A classical example of blind source separation is the cocktail party problem. Assume that several people are speaking simultaneously in the same room, as in a cocktail party. Then the problem is to separate the voices of the different speakers, using recordings of several microphones in the room[4].

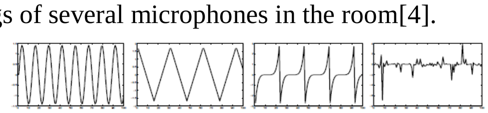

Fig2: This shows the 4 signals or independent components

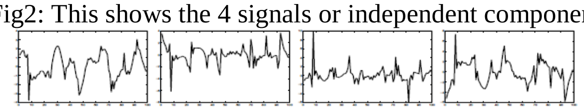

Fig3:The above signal due to some external circumstances

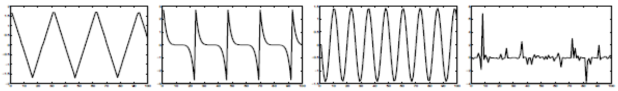

Fig4:After the application of Ica for the signals with noise

3.1 Statistical Independence

A key concept that constitutes the foundation of ICA is statistical independence. To simplify the above discussion consider the case of two different random variables s1 and s2. The random variable s1 is independent of s2, if the information about the value of s1 does not provide any information about the value of s2, and vice versa. Here s1 and s2 could be random signals originating from two different physical processes that are not related to each other.

3.2 Independence definition

Mathematically, statistical independence can be defined in terms of probability density of the signals. Consider the joint probability density function (pdf) of s1 and s2 be p(s1, s2). Let the marginal pdf of s1 and s2 be denoted by p(s1) and p(s2) respectively. s1 and s2 are said to be independent if and only if the joint pdf can be expressed as

$p(s1, s2) = p(s1)p(s2)$

Similarly, independence could be defined by replacing the pdf by the respective cumulative distributive functions as;

$E\{p(s1)p(s2)\} = E\{g1(s1)\}E\{g2(s2)\}$

Where E{.} is the expectation operator. In the following section the above properties are used to explain the relationship between uncorrelatedness and independence.

3.3Kurtosis

Kurtosis is the classical method of measuring non-Gaussianity. When data is preprocessed to have unit variance, kurtosis is equal to the fourth moment of the data. The Kurtosis of signal (s), denoted by kurt (s), is defined by

$kurt(s) = E\{s4\} - 3(E\{s2\})$

This is a basic definition of kurtosis using higher order (fourth order) cumulant, this simplification is based on the assumption that the signal has zero mean. To simplify things, we can further assume that (s) has been normalized so that its variance is equal to one: E {s2} = 1. Hence equation 2.7 can be further simplified to

$kurt(s) = E\{s4\} - 3$

For Gaussian signal, E{ s^4 } = 3(E{ s^4 })^2 and hence its kurtosis is zero. For most non Gaussian signals, the kurtosis is nonzero. Kurtosis can be both positive and negative. Random variables that have positive kurtosis are called as super-Gaussian or platykurtotic, and those with negative kurtosis are called as sub-Gaussian or leptokurtosis. Non-Gaussianity is measured using the absolute value of kurtosis or the square of kurtosis3.Kurtosis has been widely used as a measure of non-Gaussianity in ICA and related fields because of its computational simplicity. Theoretically, it has a linearity property such that

$kurt(s1 \pm s2) = kurt(s1) \pm kurt(s2)$

and

$kurt(\alpha s1) = \alpha 4 kurt(s1)$

Where α is a constant. Computationally kurtosis can be calculated using the fourth moment of the sample data, by keeping the variance of the signal constant. In an intuitive sense, kurtosis measured how "spikiness" of a distribution or the size of the tails. Kurtosis is extremely simple to calculate, however, it is very sensitive to outliers in the data set. Its values may be based on only a few values in the tails which means that its statistical significance is poor. Kurtosis is not robust enough for ICA. Hence a better measure of non-Gaussianity than kurtosis is required.

3.4 Negentropy

A second very important measure of nongaussianity is given by negentropy. Negentropy is based on the information theoretic quantity of (differential) entropy. Entropy is the basic concept of information theory. The entropy of a random variable can be interpreted as the degree of information that the observation of the variable gives. The more "random", i.e. unpredictable and unstructured the variable is, the larger its entropy. More rigorously, entropy is closely related to the coding length of the random variable, in fact, under some simplifying assumptions, entropy is the coding length of the random variable. For introductions on information theory. Entropy H is defined for a discrete random variable Y as

$H(Y) = -\sum_i P(Y=ai) \log p(y=ai)$

Where the aiare the possible values of Y. This very well-known definition can be generalized for continuous-valued random variables and vectors, in which case it is often called differential entropy. The differential entropy H of a random vector y with density f (y) is defined as

$H(y) = -\int f(y)\log f(y)dy$
A fundamental result of information theory is that a Gaussian variable has the largest entropy among all random variables of equal variance. This meansthat entropy could be used as a measure of nongaussianity. In fact, this shows that the Gaussian distribution isthe "most random" or the least structured of all distributions. Entropy is small for distributions that are clearlyconcentrated on certain values, i.e., when the variable is clearly clustered, or has a pdf that is very "spiky".To obtain a measure of nongaussianity that is zero for a Gaussian variable and always nonnegative, one oftenuses a slightly modified version of the definition of differential entropy, called negentropy. NegentropyJ is defined as follows
$J(y) = H(y_{gauss}) - H(y)$
Where ygauss is a Gaussian random variable of the same covariance matrix as y. Due to the above-mentioned properties, negentropy is always non-negative, and it is zero if and only if y has a Gaussian distribution. Negentropy has the additional interesting property that it is invariant for invertible linear transformations.
The advantage of using negentropy, or, equivalently, differential entropy, as a measure of nongaussianity is that it is well justified by statistical theory. In fact, negentropy is in some sense the optimal estimator of nongaussianity, as far as statistical properties are concerned. The problem in using negentropy is, however, that it is computationally very difficult. Estimating negentropy using the definition would require an estimate (possibly nonparametric) of the pdf. Therefore, simpler approximations of negentropy are very useful, as will be discussed next.

IV. PREPROCESSING

Before examining specific ICA algorithms, it is essential to discuss preprocessing steps that are generally carried out before ICA.

4.1 Centering

A simple preprocessing step that is commonly performed is to "center" the observation vector x by subtracting its mean vector m = E{x}. That is then the centered observation vector, xc is obtained and is as follows:
$x_c = x - m$
This step simplifies ICA algorithms where in a zero mean is assumed. Once the unmixing matrix has been estimated using the centered data, the actual estimates of the independent components are obtained and is as follows:
$\hat{s}(t) = A^{-1}(x_c + m)$
From this point on, all observation vectors will be assumed centered. The mixing matrix, on the other hand, remains the same after this preprocessing, so this can be done always without affecting the estimation of the mixing matrix[3].

4.2 Whitening

Another step which is very useful in practice is to pre-whiten the observation vector x. Whitening involves linearly transforming the observation vector such that its components are uncorrelated and have unit variance. Let xw denote the whitened vector, then it satisfies the following equation.
$E\{x_w x_w^T\} = I$
where $E\{x_w x_w^T\}$ is the covariance matrix of xw. Also, since the ICA framework is insensitive to the variances of the independent components, it can be assumed that without loss of generality the source vector, s, is white, i.e. $E\{ss^T\} = I$
A simple method to perform the whitening transformation is to use the Eigen Value Decomposition (EVD) of x. The covariance matrix of x is decomposed as
$E\{xx^T\} = VDV^T$
where V is the matrix of Eigenvectors of $E\{xx^T\}$, and D is the diagonal matrix of Eigen values, i.e. D = diag{λ1, λ2, ..., λn}. The observation vector can be whitened by the following transformation:
$X_w = VD^{-1/2}V^T x$
Where the matrix $D^{-1/2}$ is obtained by a simple component wise operation
$D^{-1/2} = \mathrm{diag}\{\lambda_1^{-1/2}, \lambda_2^{-1/2}, \ldots, \lambda_n^{-1/2}\}$
Whitening transforms the mixing matrix into a new one, which is orthogonal
$vx_w = VD^{-1/2}V^T A s = A_w s$
hence, $E\{x_w x_w^T\} = A_w E\{ss^T\} A_w^T$
$= AA_w^T$
$= I$
Whitening thus reduces the number of parameters to be estimated. Instead of having to estimate the n^2 elements of the original matrix A, only the new orthogonal mixing matrix need to be estimated, where an orthogonal matrix has

n(n−1)/2 degrees of freedom. One can say that whitening solves half of the ICA problem. This is a very useful step as whitening is a simple and efficient process that significantly reduces the computational complexity of ICA.

V. ADAPTIVE FILTER

The Adaptive filters are used because some of the ECG signal may be overlapped with noise spectrum. In order to extract this hidden signal in noise we go for Adaptive filter. Adaptive filter uses LMS algorithm.An adaptive filter is a filter that self-adjusts its transfer function according to an optimizing algorithm. Because of the complexity of the optimizing algorithms, most adaptive filters are digital filters that perform digital signal processing and adapt their performance based on the input signal. By way of contrast, a non-adaptive filter has static filter coefficients (which collectively form the transfer function).Generally speaking, the adapting process involves the use of a cost function, which is a criterion for optimum performance of the filter (for example, minimizing the noise component of the input), to feed an algorithm, which determines how to modify the filter coefficients to minimize the cost on the next iteration [5].

VI. PSEUDO CODE FOR ECG SIGNAL PROCESSING

Take four mixed signals
Plot the mixed signals //ECG signals mixed with noise
Call band pass chebyshev filter for pre processing
Plot preprocessed signals
Call FASTICA
Plot separated signals
Call adaptive filter for post processing
Plot filtered signals
Determine SNR
Display SNR

VII. SIMULATION RESULT OF ECG SIGNAL

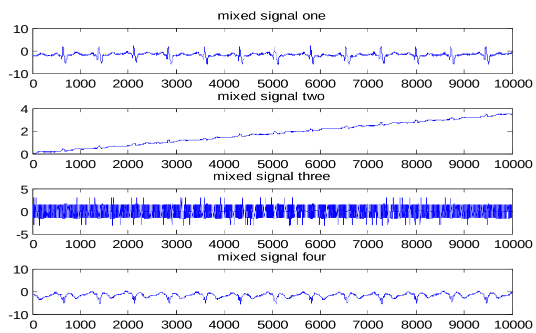

Fig5:Mixed ECG signal

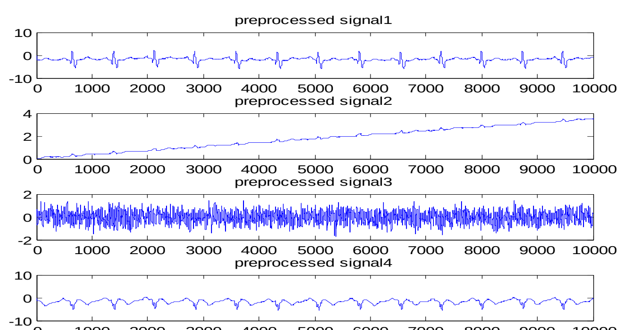

Fig6: Preprocessed signal using chebyshev band pass filter

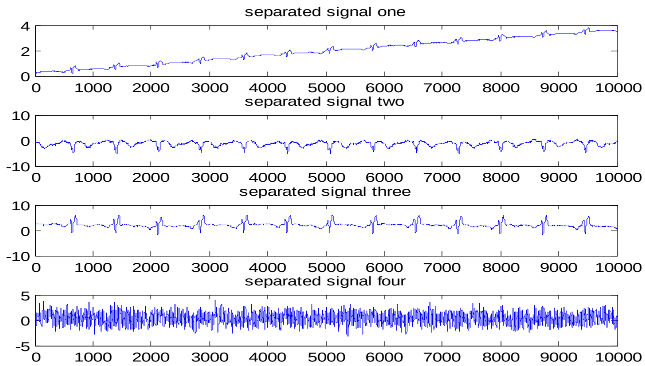

Fig7:Separated signals after ICA

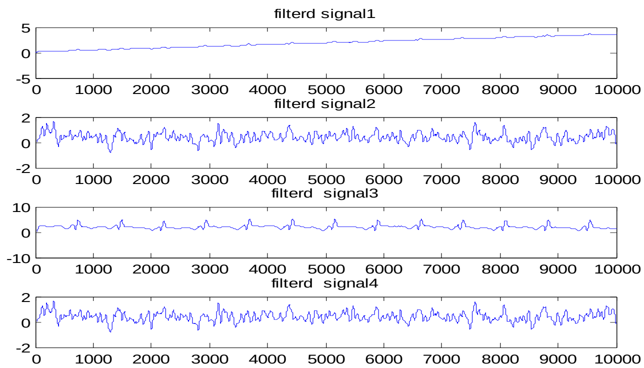

Fig8: Post Processed Signals using adaptive filter

VIII. CONCLUSION AND FUTURE ENHANCEMENT

Biomedical signals are collective electrical signal acquired from any organ that represents a physical variable of interest. These signals are complexed in nature as it contains various artifacts with it, the main criterion is to extract and analyze these signals independently. The algorithm used here i.e. Independent Component Analysis (ICA) is a member of a class of BSS(Blind Source Separation). Detection of ECG signals with powerful and advance methodologies is becoming a very important requirement in biomedical engineering. The main reason for the interest in ECG or s is in clinical diagnosis and biomedical applications.The result demonstrate that signal of ECG signals are effectively extracted from their respective recordings and isolating it from artifacts associated with it at the time of recording. In this workChebyshev band pass filter is used as preprocessing filter and adaptive filter as post processing filter which removed the noise.

Thus it is observed that using of FASTICA algorithm, Chebyshev band pass filter and adaptive filter efficiently isolate ECG signal from unwanted noise.In this work, 4 ECG recorded signal are processed. So, this work can be enhanced byprocessing real time biomedical signal such as EEG, ECG, EMG and MEG.In case processing ECG signal, QRS complex can detected by applying QRS detection algorithm. Increasing the order of the signal i.e. by considering more number of inputs. This need higher order matrices.The processing of biomedical signals can be implemented in to hardware.

IX. REFERENCES

[1] Taigang He, Gari Clifford and Lionel Tarassenko,"Application of ICA in Removing Artefacts from the ECG",Signal Processing & NeuralNetworks Research Group,Department of Engineering ScienceUniversity of Oxford, Oxford, UK,2006.

[2] AapoHyvärinen Helsinki,"Fast and Robust Fixed-Point Algorithms for Independent Component Analysis", University of TechnologyLaboratory of Computer and Information Science P.O. Box 5400,FIN-02015.

[3] HUT, Finland), "Independent component analysis, a new concept?",IEEETrans. on Neural Networks, 10(3):626-634.1999.Comon, P.(1994,Signal Processing 36(3), 287–314. 13.

[4] Seungjin Choi, AndrzejCichocki, Hyung-Min Park and Soo-Young Lee,"Blind Source Separation and Independent Component Analysis: AReview", Neural Information Processing - Letters and Reviews, Vol.6,No.1, January 2005.

[5] Scott C. Douglas, "Fixed-point algorithms for the blind separation of complex-valued mixtures",Department of Electrical EngineeringSouthern Methodist UniversityDallas, Texas 75275 USA, June 2006.