Noise removal of functional Near Infrared Spectroscopy signals using Emperical Mode Decomposition and Independent Component Analysis

Pham Thi Kim Chi[1], Vo Nhut Tuan[1], Nguyen Hoai Thuong[1], Ho Thi Kieu Khanh[1], Hui Yu[2], Nguyen Duc Thang[1*]

[1] Biomedical Engineering Department, International University- Vietnam National University, Ho Chi Minh City, Vietnam
[2] School of Creative Technologies, University of Portsmouth, United Kingdom

Abstract— **Currently, researchers are getting more interests in discovering brain activities by non-invasive methods of using functional near-infrared spectroscopy (fNIRS). However, fNIRS collected signals usually contain noises which significantly affect the measurement of fNIRS experiments. There have been available methods proposed to remove artifacts of fNIRS signals. Among those approaches, adaptive filers are effective to mitigate physiological noises measured by extra sensors. However, the use of sensors attached on the human subjects during fNIRS measurement is uncomfortable for a user and is complicated for setup. Therefore, the method to extract the physiological signals automatically from fNIRS signals without the needs of other sensors is getting more attention from research community.**

In this work, we propose the combination of emperical mode decomposition (EMD) method and independent component analysis (ICA) to extract the heart rate signal. EMD is the fundamental part of Hilbert – Huang transform which is used to decompose signal into intrinsic mode functions that are not set analytically and are instead determined by an analyzed sequence alone. ICA uses Hyvarinen's fixed-point algorithm to estimate the independent components from given multidimensional signals. Our proposed approach is able to extract the heart rate signal from multiple fNIRS channels with the accuracy of 80% to 90% compared with the one measured from the real device. Our further work will integrate this result with noise attenuation using adaptive filters to mitigate the global inference of physiological activities to fNIRS measurement.

Keywords— **Near infrared spectroscopy, Emperical mode decomposition, Independent component analysis, Noise attenuation, Physiological signal.**

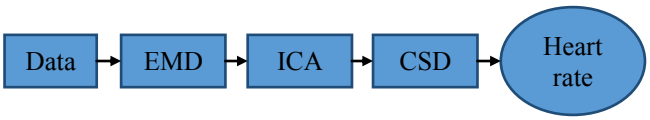
I. INTRODUCTION

fNIRS is a noninvasive method which bases on the study of near infrared light to detect the changes in the concentration of oxygenated (oxyHb) and deoxygenated (deoxyHb) hemoglobin molecules in the blood. This is an effective optical neuroimaging method to monitor hemodynamic response to brain activation.

The measured data includes not only the chromophore mobilization but also noises such as instrument noise, motion artifact, global interference including biological noise. As a result, several filtering methods is apply to remove those noses which include band pass filter, smoothing filter, principle component analysis (PCA), Kalman filter, EMD, ICA and so on.

During fNIRS measurement, the collected data always contains physiological lowfrequency oscillations such as Mayer's wave, or high frequency oscillations caused by instrument noise. In this case, band-pass filtering is a suitable candidate for lessening those unwanted signals. Smoothing is a processing method capturing important patterns in the data by averaging the data points with their neighbors in series of time. Smoothing is able to suppress the high-frequency signals and enhance low-frequency ones in the data. Beisdes, principal component analysis is a statistical procedure using an orthogonal transformation to convert a set of observations of correlated variables into a set of values of linearly uncorrelated variables called principal components whose number is not greater than that of the original variables.

In order to detect signals with known frequency range, the combination of EMD and ICA method is well considered. With the application of Hilbert – Huang transform (HHT) in EMD method, fNIRS signal can be decomposed into intrinsic mode functions which are determined by an analyzed sequqence alone. Then the independent components from these functions are estimated by Hyvarinen's fixed-point algorithm of ICA method. The dominant frequency of every component signal is then detected by Correntropy Spectral Density (CSD) method by which the heart rate of subjects can be estimated to be eliminated from the collected data by further methods.



II. FILTERING METHODS

A. EMD

EMD is the fundamental part of Hilbert – Huang transform (HHT) by which signal is decompose into intrinsic mode functions (IMFs) that are not set analytically and are instead determined by an analyzed sequence alone. The IMF must satisfy the following conditions:

- In the whole data set, the number of extrema and that of zero crossings must either equal or differ at most by one.

- At any data point, the mean value of the envelope defined by local maxima and the envelope defined by the local minima is zero. [1]

[2] With the input x(t), the mean of upper and lower envelopes m_1, the difference between x(t) and m_1 is the first protomode, h_1.

$$h_1 = x(t) - m_1 \quad (1)$$

Then the sifting process is occured to eliminate background waves which the IMF is riding on and to make the wave profiles more symmetric. The sifting process will be repeated to make the extracted signal satisfy the condition of an IMF. The proto – IMF, h_1, is treated as the data in the next iteration:

$$h_{11} = h_1 - m_{11} \quad (2)$$

$$h_{1k} = h_{1(k-1)} - m_{1k} \quad (3)$$

The stoppage criterion determines the number of sifting steps to produce an IMF when

$$SD_k = \frac{\sum_{t=0}^{T} |h_{k-1}(t) - h_k(t)|^2}{\sum_{t=0}^{T} h_{k-1}^2(t)} \quad (4)$$

is smaller than a predetermined value.

When the approximate local envelope symmetry condition is satisfied, and h_{1k} becomes the IMF c_1 which is the first IMF component. This first IMF should contain the finest scale or the shortest-period oscillation in the signal, which can be extracted from the data by

$$x(t) - c_1 = r_1 \quad (5)$$

The residue, r_1, still contains longer-period variations. The procedure is repeated to obtain an IMF of lower frequency

$$r_1 - c_2 = r_2 \quad (6)$$

$$r_{n-1} - c_n = r_n \quad (7)$$

The decomposition process finally stops when the residue, r_n, becomes a monotonic function or a function with only one extremum from which no more IMF can be extracted. In summary:

$$x(t) = \sum_{j=1}^{n} c_j + r_n \quad (8)$$

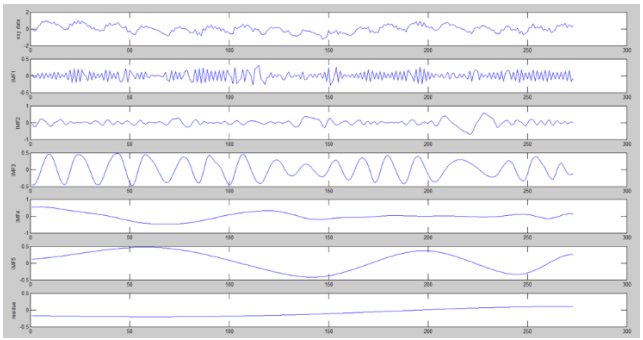


Fig 1. A 15-second interval of oxygenated fNIRS data after being decomposed into components by EMD method.

B. ICA

[3] Fast ICA uses Hyvarinen's fixed-point algorithm to estimate the independent components from given multidimensional signals. Each row of the input matrix is one observed signal. [4] In general, the stabllized fixed – point algorithm for one unit with non sphered data is defined as:

$$w^+ = w - \frac{\mu[C^{-1}E\{xg(w^Tx)\} - \beta w]}{E\{g'(w^Tx)\} - \beta} \quad (9)$$

$$w^* = \frac{w^+}{\sqrt{(w^+)^T C w^+}} \quad (10)$$

Where $\beta = E\{w^T x g(w^T x)\}$, μ is a step size parameter that may change with the iteration count and $C = E\{xx^T\}$ is the covariance matrix of the data.

To estimate n independent components, this algorithm is run n times. To prevent different neurons from converging to the same maxima we must decorrelate the outputs $w_1^T x, \ldots, w_n^T x$ after every iteration. A simple iterative algorithm is a deflation scheme based on a Gram-Schmidt-like decorrelation. This means that we estimate the independent components one by one. When we have estimated p independent components, or p vectors $w_1,\ldots,w_p$, we run the one-unit fixed-point algorithm for w_{p+1}, and after every iteration step subtract from w_{p+1} the 'projections' $w^T_{p+1} w_j w_j$, j = 1,..., p of the previously estimated p vectors, and then renormalize w_{p+1}:

$$w_{p+1} = w_{p+1} - \sum_{j=1}^{p} w_{p+1}^T C w_j w_j \quad (11)$$

$$w_{p+1} = \frac{w_{p+1}}{\sqrt{w_{p+1}^T C w_{p+1}}} \quad (12)$$

C. CSD

[5] The CSD is a generalization of the conventional power spectral density. It is based on the Fourier transform of the centered correntropy function used to convert signal from time domain to frequency domain. Firstly, a sliding time window of about 15 seconds is used to segment signal

into segments assumed to be stationary and suitable for spectral analysis. Secondly, the CSD is applied to the signal segments. Thirdly, the Heart rate is estimated by detecting the maximum frequency peak within the cardiac frequency band (0.5Hz – 3Hz).

IMF3 on the figure 1 is determined as heart component as its peak frequency is 1.4Hz which lies in the cadiac frequency band.

III. ASSESSMENT AND RESULT

fNIRS signal and ECG data is conducted concurrently to record the cardial signal at corresponding time, by which the estimated Heart rate from fNIRS data is compared with the real measured heart rate for accuracy assessment. Measured ECG signal is converted into approximate PPG signal by nomalization, smoothing and lowpass filter with cutcoff frequency is 2.5Hz to obtain Heart rate.

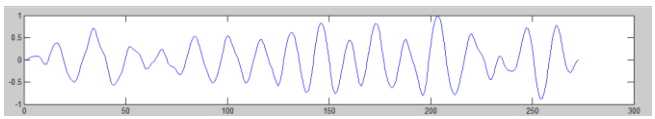


Fig 2. Heart rate smoothed from measured ECG signal

Next, smoothed ECG signal and estimated Heart rate are both converted to frequency domain by PSD algorithm which describes the distribution of power into frequency components composing that signal.

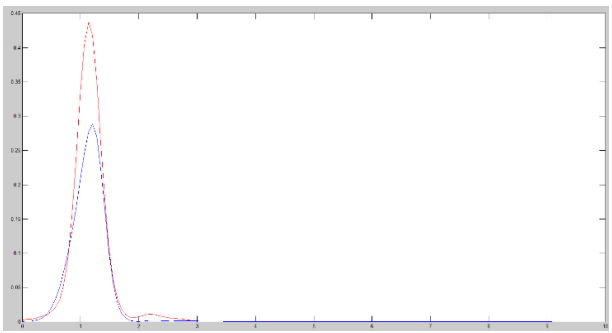


Fig 3. PSD comparision of estimated Heart component and PPG signal in frequency domain. The 'red' lines display the Heart component and the 'blue' ones illustrate the PPG signal.

$$correlation = 1 - \frac{|f_1 - f_2|}{mean(f_1, f_2)} \qquad (13)$$

Note:

f_1: frequency of estimated Heart component

f_2: frequency of measured PPG signal

correlation: the similarity of f_1 and f_2 calculated by the ratio of the frequencies difference and the mean of f_1 and f_2.

The table 1 illustrates the frequencies and correlations of estimated Heart component and PPG signal of Subject 1 and Subject 2 in two minutes with 15-second time intervals

Table 1.

Time interval (second)	Subject 1			Subject 2		
	f_1(Hz)	f_2(Hz)	correlation	f_1(Hz)	f_2(Hz)	corelation
0-15	1.14	1.21	0.94	1.42	1.35	0.95
15-30	1.21	1.14	0.94	1.35	1.35	1
30-45	1.21	1.21	1	1.27	1.27	1
60-75	1.14	1.14	1	1.27	1.21	0.94
75-90	1.21	1.14	0.94	1.27	1.27	1
90-105	1.21	1.06	0.88	1.27	1.27	0.95
105-120	1.14	1.14	1	1.35	1.35	1
120-135	1.21	1.06	0.88	1.35	0.99	0.71

IV. CONCLUSIONS

The assessment on the result of EMD and ICA method has shown high average correlation of measured Heart rate and the estimated one which is 80% to 90%. That means the Heart component detected by these methods has no significant differences from the measured one.

As a result, the combination of EMD and ICA can be well applied to extract biological noise (heart rate) from fNIRS signal with significantly high accuracy.

REFERENCES

1. Metatrader 5 - Statistics and Analysis: Introduction to Emperical Mode Decomposition
2. Norden E.Huang and Zhaohua Wu: A review on Hilbert-Huang transform: Method and its applications to geophysical studies.
3. The Neurophysiological Biomarker Toolbox (NBT): ICA - Independent Component Analysis
4. Helsinki University of Technology – Laboratory of Computer Science:
   A Fast Fixed-Point Algorithm for Independent Component Analysis
5. Ainara Garde, Walter Karlen, J. Mark Ansermino, Guy A. Dumont: Estimating Respiratory and Heart Rates from the Correntropy Spectral Density of the Photoplethysmogram