Automatic Selection of Frequency Bands for Electroencephalographic Source Localization

Pablo Andrés Muñoz
Universidad del Quindío
Email: pamunoz@uniquindio.edu.co

Maximiliano Bueno López
Universidad de la Salle
Email: maxbueno@unisalle.edu.co

Eduardo Giraldo
Universidad Tecnológica de Pereira
Email: egiraldos@utp.edu.co

Marta Molinas
Norwegian University of Science and Technology
Email: marta.molinas@ntnu.no

***Abstract*—This paper shows a method to locate actives sources from pre-processed electroencephalographic signals. These signals are processed using multivariate empirical mode decomposition (MEMD). The intrinsic mode functions are analyzed through the Hilbert-Huang spectral entropy. A cost function is proposed to automatically select the intrinsic mode functions associated with the lowest spectral entropy values and they are used to reconstruct the neural activity generated by the active sources. Multiple sparse priors are used to locate the active sources with and without multivariate empirical mode decomposition and the performance is estimated using the Wasserstein metric. The results were obtained for conditions with high noise (Signal-to-Noise-Ratio of -5dB), where the estimated location, for five sources, was better for multiple sparse prior with Multivariate Empirical Mode Decomposition, and with low noise (Signal-to-Noise-Ratio of 20dB), where the estimated location, for three sources, was better for multiple sparse prior without MEMD.**

I. Introduction

Electroencephalographic (EEG) Source Localization (ESL) has been widely used in different medical fields (neuroscience studies or clinical applications) for its high temporal resolution that allows to measure the changes of neural activity in time intervals of the order of milliseconds. The main drawback of ESL is to solve the neuromagnetic inverse problem which is ill-posed and it does not have an unique solution. Therefore, to obtain an approximated locations of neural current sources from EEG, it is necessary to solve the inverse problem using some *a priori* information or applying some constraints over the source space [1], [2]. Nowadays, spatio-temporal constraints have been used in different works, in [1] was included, to improve the spatial resolution, a basis set for smoothing the source space (localized areas that could be potentially active brain regions) and based on a Markovian assumption applied at each sample time to estimate the brain activity, the time resolution was improved. Another spatio-temporal constraints were incorporated as a small and locally patches to reconstruct sparse brain activity; to smooth the solution over the time, temporal constraint was imposed for penalizing the difference between consecutive time points [2].

Currently, some research have focused their studies to analyze the neural activity in frequency bands, in this way, they have found it e.g. some theta-band activities of low-amplitude desynchronised were associated to visual areas when they were compared among motion stimuli and static stimuli [3]. Besides, in [4], the authors focused the research in alpha-band oscillations because they are the main frequency components, associated to neural activity, present in EEG signals. Recently, some works have proposed a method whose structure is based on data-driven analysis. An example for this kind of analysis is the empirical mode decomposition (EMD) and some applications in brain activity reconstruction are shown in [5], [6]. One of the results shown in [6] was the way how the neural activity was split in frequency bands which can be seen in the intrinsic mode functions (IMFs). Similar results can be seen in [5], but these results were analyzing according to the retained energy and the amount of entropy in each IMFs. Despite of the relevant results, some issues associated to EMD method were regarded in the full reconstruction, namely, mode mixing and mode splitting.

Solutions for reducing the mode mixing have been highlighted in [7] e.g. Noise aided EMD computation (EEMD) and multivariate empirical mode decomposition (MEMD). In this paper is presented a method, based on data-driven analysis using MEMD, to improve the localization of the actives sources in the brain and for reducing the mode mixing problem. To separate the frequency bands is used MEMD method and the relevant IMFs are chosen from marginal Hilbert-Huang spectrum (MHHS) and entropy analysis. A cost function based on entropy is proposed to dismiss the IMFs with Hilbert-Huang Spectral Entropy (HHSE) greater than the estimated HHSE threshold, and those IMFs with lower HHSE are chosen to located the actives sources. The proposed method is evaluated by comparing ESL using multiple sparse prior (MSP) with and without MEMD, and the performance is measured with Wasserstein metric on simulated brain activity.

II. Material and Methods

A. Multivariate Empirical Mode Decomposition (MEMD)

Signals represented in multivariate form should have a coherent treatment to obtain a suitable time-frequency estimation, because these signals contain generalized oscillations (joint rotational modes). Therefore, it is important to remark that when the single EMD is applied (channel by channel) to multichannel signals, this approach is obstructed by [7]:

- Nonuniformity. Each channel would not be decomposed with the same number of IMFs using standard EMD.
- Scale alignment. It is possible that the scales across data channels do not have the same-index.
- Nature of IMFs. It is not convenient to enforce the same number of IMFs for each channel, because the t-f estimation could be affected, as such IMFs are typically not monocomponent.

Common mode alignment and nonuniqueness have been the greatest obstacles for application of the EMD in studies where is necessary same-index IMFs within of the same scale for the corresponding information (synchrony, causality, data/image fusion), being a problem in applications data/image fusion [7]. For multivariate signals, the local *maxima* and *minima* can not be calculated directly and, the notion of ”oscillatory modes” to define an IMF is confuse in this case [8]. This method proposes to take a signal projections along of multiple directions that have been distributed in a uniform way within of a n-dimensional space to obtain multiple envelopes which are averaged and then, interpolated (using cubic spline) their extrema to estimate the local n-dimensional mean. Especial attention is required to choose a suitable set of directions from the signal projections taken in the n-dimensional space [7].

The following algorithm summarizes how the MEMD works [7]:

1) Using the Hammersley sequence, as a uniformly sampling a n-dimensional sphere, generate a P-point.
2) Projections $q_{\theta_p}(t_k)$ of the signal $y(t_k)$ must be calculated in the same direction vector $\mathbf{x}_{\theta_p}$, for $p = 1, ..., P$ and then to obtain a set of projections $\{q_{\theta_p}(t_k)\}_{p=1}^{P}$.
3) Find the time instants $\{t^i_{\theta_p}\}_{p=1}^{P}$ that correspond to the maxima of the set of projections of signals $\{q_{\theta_p}(t_k)\}_{p=1}^{P}$.
4) Interpolate $[t^i_{\theta_p}, s(t^i_{\theta_p})]$ to obtain the envelope curves $\{e_{\theta_p}(t_k)\}_{p=1}^{P}$.
5) Calculate the mean of the P multidimensional envelopes

$$\boldsymbol{m}(t_k) = \frac{1}{P}\sum_{p=1}^{P} \boldsymbol{e}_{\theta_p}(t_k) \tag{1}$$

6) Extract the “detail” $d(t_k) = s(t_k) - m(t_k)$. If $d(t_k)$ fulfills the stoppage criterion for a multivariate IMF, apply the above procedure to $s(t_k) - d(t_k)$, else repeat for $d(t_k)$.

B. Hilbert-Huang Spectral Entropy

Spectral entropy can be defined as a measure of the amount of disorder and this definition is based on the spectrum of a signal. The *Hilbert-Huang Epectral Entropy* (HHSE), for non-stationary signals, is calculated from Hilbert spectrum following these steps [9]:

1) The signal $\boldsymbol{x}(t)$ is decomposed into a series of IMFs (IMF_j).
2) The hilbert transform is applied to IMF_j $(1 \leq j \leq n)$ to obtain Y_{IMF_j}
3) The analytical signal is calculated for each IMF_j:

$$\boldsymbol{Z}_{IMF_j}(t) = \boldsymbol{IMF}_j(t) + i\mathbf{Y}_{IMF_j}(t) = \boldsymbol{a}_j(t)e^{i\theta_j(t)} \tag{2}$$

where

$$\boldsymbol{a}_j(t) = [\boldsymbol{IMF}_j^2(t) + \mathbf{Y}_{IMF_j}^2(t)]^{\frac{1}{2}} \tag{3}$$

and

$$\boldsymbol{\theta}_j(t) = \arctan(\frac{\boldsymbol{Y}_{IMF_j}}{\boldsymbol{IMF}_j(t)}) \tag{4}$$

4) The instantaneous frequency is calculated for IMF_j $(1 \leq j \leq n)$:

$$\omega_j(t) = \frac{d\theta_j(t)}{dt}, \tag{5}$$

The time series is expressed as:

$$\boldsymbol{x}(t) = \sum_{j=1}^{n} \boldsymbol{a}_j(t) exp(i \int \boldsymbol{\omega}_j dt) \tag{6}$$

The equation 6 represents, as function on time, the amplitude and the instantaneous frequency, therefore, this equation corresponds to the Hilbert Transform $\boldsymbol{H(\omega, t)}$. The Hilbert spectrum is the energy-time-frequency distribution over the signal $\boldsymbol{x}(t)$ y HHSE es calculated using the *frequency marginal* by integrating the Hilbert spectrum over the time-axis.

C. Neuromagnetic Inverse Problem

The neural activity can be generated through the following model of EEG generation:

$$\boldsymbol{y}(t_k) = \boldsymbol{M}\boldsymbol{x}(t_k) + \epsilon(t_k) \tag{7}$$

being the EEG at sample time t_k termed $\boldsymbol{y}(t_k) \in R^{d\times 1}$, the lead field matrix $\boldsymbol{M} \in R^{d\times n}$ and the neural activity $\boldsymbol{x}(t_k) \in R^{n\times 1}$. The forward problem indicated in (7), allows to define that the estimation of the neural activity can be obtained by solving the inverse problem based on the EEG measurements $\boldsymbol{y}(t_k)$ and the knowledge of the lead field matrix $\boldsymbol{M}$. Besides, to get an unique solution, it is necessary to consider some spatio-temporal dynamics of EEG signals, which can improve the approximated location of the active sources [1]. The MSP method was proposed by [10] and this method apply a hierarchical or empirical Bayes model as spatio-temporal constraints to reconstruct the inverse problem in a distributed way, and multiple cortical sources with a spatial support, specified in terms of empirical priors, are automatically selected.

III. Experimental Setup

Studies in neuroscience have set five frequency bands, namely: delta-band (0-4 Hz), theta-band (4-8 Hz), Alpha-band (8-14 Hz), beta-band (14-30 Hz) and gamma-band (30-150Hz) [3]. The aim was to simulate brain activity for three sources and five sources, these sources were randomly located in three (delta, alpha and beta bands) and five (delta, theta, alpha, beta and gamma bands) different frequency bands, they were also located randomly in different areas in the brain. The activity in each source was simulated using the following expression:

$$x_i(t_k) = e^{-\frac{1}{2}\left(\frac{t_k - c_i}{\sigma}\right)^2} \sin(2\pi f_i t_k), \tag{8}$$

c_i being the center of the windowed signal in seconds (1, 3 and 5 seconds for three sources and 1, 2, 3, 4 and 5 seconds for five sources), the frequency of the signal (f_i) was chose randomly within of the ranges according with the frequency bands mentioned above and $\sigma = 0.2$. In this work were simulated 30 trials for Signal-to-Noise-Ratio (SNR) of 20dB, 10dB, 0dB and -5dB using the model of generation (7).

After applying the HHSE to each trial and each noise level, It was possible to find that the lowest spectral entropy values were associated to the IMFs where the simulated activity was observed in the frequency bands. For this reason, the subset of of IMFs whose entropy was under a threshold τ_e were chosen to locate the active sources.

The proposed entropy function is the following:

$$e_j = -\sum_k \|\boldsymbol{IMF}_j(t)\|_2^2 \log(\|\boldsymbol{IMF}_j(t)\|_2^2) \tag{9}$$

It is applied over each IMF $\boldsymbol{IMF}_j(t)$ where e_j is the entropy of each IMF, and $\boldsymbol{e} = [e_1 \dots e_N]$. The estimated EEG signal $\tilde{\boldsymbol{y}}(t)$ from IMFs with highest entropy (chosen automatically) is rebuilt according to the measured entropy e_i.

$$\tilde{\boldsymbol{y}}(t) = \sum_{i \in O} \boldsymbol{IMF}_j(t) \tag{10}$$

Access to a standard EEG database is important because it is necessary to know the underlying source activity to evaluate the methods for solving the inverse problem. We used a model with $n = 8,196$ sources and 32 electrodes for simulation, as described by ([1]).

IV. Results

After analyzing all the trials with four noise levels, it could be found that the most suitable threshold to choose the relevant IMFs for locating the active sources, was the IMFs with lowest spectral entropy and the chosen IMFs were those whose sum did not exceed 40 percent of the normalized HHSE for all IMFs. Two simulations were carried out under controlled conditions to show the results of this work, especially with respect to the location of the active sources, which were located for a clear visualization.

The first one was simulated for three active sources with $f_1 = 2Hz$, $f_2 = 9Hz$ and $f_3 = 22Hz$, the SNR was of $20dB$. In fig. 1 are shown three of the six IMFs chosen by entropy cost function; the simulated EEG fig. (1A), IMF2 Fig. (1B) associated to frequency beta-band ($f_3 = 22Hz$), IMF5 fig. (1C) associated to frequency alpha-band ($f_2 = 9Hz$) and IMF8 fig. (1D) associated to frequency delta-band ($f_1 = 2Hz$).

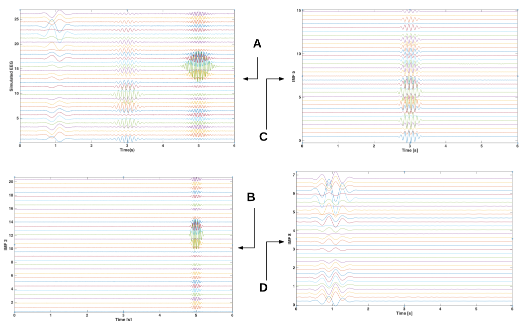

Fig. 1. Selected IMFs for 3 sources with SNR 20dB

Each IMF used to locate the active source can be seen in fig. 2 whose sum allows to obtain the full location for the three active sources fig. 2B. The Wasserstein metric for this estimation was the 3.1467 and the location without MEMD was 3.2313 fig. 2C, this measurements compared with the ground truth fig. 2A.

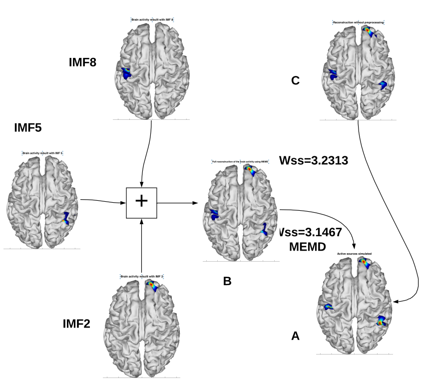

Fig. 2. Wasserstein metric with and without MEMD for 3 active sources located with SNR of 20dB

The second simulation was done for 5 sources with $f_1 = 1.5Hz$, $f_2 = 4Hz$, $f_3 = 9Hz$, $f_4 = 20Hz$ and $f_5 = 45Hz$, the SNR was of $-5dB$. The high level of noise can be seen in fig. 3A and the another figures are shown 5 of the 6 IMFs chosen. The advantage by using the MEMD is to be able to separate the activity in different bands of frequency e.g., in fig. 3B corresponds to gamma-band ($f_5 = 45Hz$) and was decomposed in the IMF2 with some noise. In the IMF4 (fig. 3C) was located the frequency associated to beta-band ($f_4 = 20Hz$) and the same way, it can be seen in fig. 3D the IMF5 with the frequency in alpha-band ($f_3 = 9Hz$), in fig. 3E the IMF7 with the frequency in theta-band ($f_2 = 4Hz$) and in fig. 3F the IMF9 with the frequency in delta-band ($f_1 = 1.5Hz$).

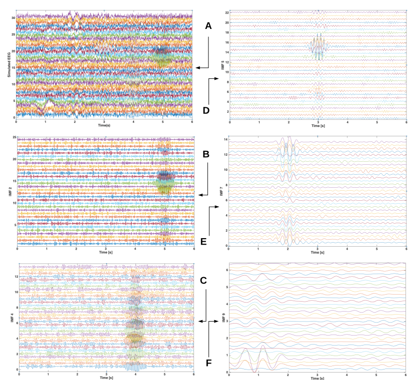

Fig. 3. Selected IMFs for 5 sources with SNR -5dB

In the fig. 4 is presented that the Wesserstien metric for MSP with MEMD (fig. 4B) was lower than the metric for MSP without MEMD (fig. 4C), compared with the ground truth fig. 4A.

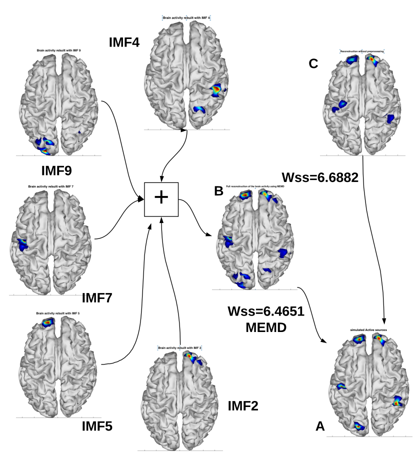

Fig. 4. Wasserstein metric with and without MEMD for 5 active sources located with SNR of -5dB

V. Conclusion

A method based on data-driven analysis, for improving the accuracy for EEG source localization (ESL), was evaluated. The MEMD was used in order to decompose the EEG signal in its main modes and separate the noisy components in order to locate the active sources with a minimum noise. It could also be seen that the EEG signal was decomposed in IMFs within different frequency bands and to each IMF was associated a specific spectral entropy value. Those IMFs with incorporated frequency band or source activity allow to reconstruct the brain activity of that source. The cost function of entropy was proposed for choosing the IMFs with lowest spectral entropy (calculated by using HHSE) and up to a maximum of 40 percent, with this cost function, all the active sources were located. The performance of MSP with MEMD, according to the Wesserstein metric, was better under SNR of -5db while for SNR of 20dB the performance of MSP was better without MEMD. In both cases, the method for choosing the IMFs took in account additional IMFs, because of the mode splitting generated by the MEMD method.

Acknowledgment

This work was carried out under the funding of COLCIENCIAS. Research project: 111077757982: Sistema de identificacin de fuentes epileptognicas basado en medidas de conectividad funcional usando registros electroencefalogrficos e imgenes de resonancia magntica en pacientes con epilepsia refractaria: apoyo a la ciruga resectiva.

References

[1] E. Giraldo-Suarez, J. D. Martinez-Vargas, and G. Castellanos-Dominguez, "Reconstruction of neural activity from eeg data using dynamic spatiotemporal constraints," *International Journal of Neural Systems*, vol. 26, no. 07, p. 1650026, 2016, pMID: 27354190. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/S012906571650026X

[2] J. D. Martínez-Vargas, F. M. Grisales-Franco, and G. Castellanos-Dominguez, "Estimation of m/eeg non-stationary brain activity using spatio-temporal sparse constraints," in *Artificial Computation in Biology and Medicine*, J. M. Ferrández Vicente, J. R. Álvarez-Sánchez, F. de la Paz López, F. J. Toledo-Moreo, and H. Adeli, Eds. Cham: Springer International Publishing, 2015, pp. 429–438.

[3] S. B. Agyei, F. R. van der Weel, and A. L. van der Meer, "Longitudinal study of preterm and full-term infants: High-density eeg analyses of cortical activity in response to visual motion," *Neuropsychologia*, vol. 84, pp. 89 – 104, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0028393216300379

[4] K. Mahjoory, V. V. Nikulin, L. Botrel, K. Linkenkaer-Hansen, M. M. Fato, and S. Haufe, "Consistency of EEG source localization and connectivity estimates," *NeuroImage*, vol. 152, no. March, pp. 590–601, 2017. [Online]. Available: http://dx.doi.org/10.1016/j.neuroimage.2017.02.076

[5] P. Muoz-Gutirrez, M. Molinas, E. Giraldo, and M. Bueno, "Localizing the focal origin of epileptic activity using eeg brain mapping based on empirical mode decomposition," in *Proceedings. (ITISE 2018). (International conference on Time Series and Forecasting, 2018.*, September 2018.

[6] M. Bueno-Lopez, E. Giraldo, and M. Molinas, "Analysis of neural activity from EEG data based on EMD frequency bands," in *24th IEEE International Conference on Electronics, Circuits and Systems (ICECS)*, vol. 1. Batumi, Georgia: IEEE, December 2017, pp. 1–5.

[7] D. P. Mandic, N. Ur Rehman, Z. Wu, and N. E. Huang, "Empirical mode decomposition-based time-frequency analysis of multivariate signals: The power of adaptive data analysis," *IEEE Signal Processing Magazine*, vol. 30, no. 6, pp. 74–86, 2013.

[8] N. Rehman and D. P. Mandic, "Multivariate empirical mode decomposition," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 466, no. 2117, pp. 1291–1302, 2010.

[9] A. Humeau-Heurtier, C. W. Wu, S. D. Wu, G. Mahe, and P. Abraham, "Refined Multiscale Hilbert-Huang Spectral Entropy and Its Application to Central and Peripheral Cardiovascular Data," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 11, pp. 2405–2415, 2016.

[10] K. Friston, L. Harrison, J. Daunizeau, S. Kiebel, C. Phillips, N. Trujillo-Barreto, R. Henson, G. Flandin, and J. Mattout, "Multiple sparse priors for the M/EEG inverse problem," *NeuroImage*, vol. 39, no. 3, pp. 1104 – 1120, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1053811907008786