2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

A comparison between the Hilbert-Huang and Discrete Wavelet Transforms to recognize emotions from electroencephalographic signals

Camilo E. Valderrama [1]

Abstract— **Recent studies have attempted to recognize emotions by extracting features from electroencephalographic (EEG) signals using either linear and stationary, or linear and non-stationary transformations. However, as EEG signals are non-linear and non-stationary, it seems that a non-linear and non-stationary transformation may be more suitable. Despite the attractiveness of this hypothesis, until now, little studies have used such transformation. The current work presents a comparison between an approach to recognize positive and negative emotions using a non-linear and non-stationary transformation (Hilbert-Huang Transformation) with an approach using linear and non-stationary transformation (Discrete Wavelet Transform). The two approaches were compared using 200 EEG signals recorded from 10 subjects. The comparison indicated that an approach using the Hilbert-Huang Transformation statistically significantly classified emotions more accurately than a Wavelet-based approach ($P <$ 0.02). This result implies that Hilbert-Huang Transformation is a promising tool to increase the prediction of emotional states, thereby helping to designing and developing more robust emotion recognition approaches.**

Clinical relevance— **This remarks the potential of the Hilbert-Huang transform to enhance EEG-based emotion recognition systems, which can potentially help to diagnose and treat mental diseases, such as autism and depression.**

I. Introduction

The Affective Computing (AC) community has remarked on the relevance of emotion recognition to improve human-machine interactions [1]. The most common techniques to recognize emotions in the AC community are neuroimaging techniques, with electroencephalography (EEG) being the most used. The advantage of EEG over other techniques, such as magnetoencephalogram, positron emission tomography, and functional resonance imaging, is that EEG is less intrusive and has a better time resolution [2].

Emotion recognition approaches commonly extract spectral features from the EEG signals to recognize emotions [3], [4] by either using linear and stationary transformations or by using linear and non-stationary transformations. However, since EEG signals are non-linear and non-stationary, the commonly used transformations, such as the Fourier transformation or the discrete Wavelet transformation (DWT), may be unable to precisely describe the nature of such signals.

One transformation that could overcome this limitation is the Hilbert-Huang transform (HHT), which was introduced in the 90s to decompose real-world non-linear, non-stationary, and stochastic processes [5]. Although HHT was introduced 30 years ago, Jenke et al. (2014) [6], in a review of emotion recognition from EEG, reported that only one out of 33 works had used this transformation for extracting features.

Hadjidimitriou et al. (2012) [7] pointed out the advantages of HHT to recognize emotional states by reporting that the Hilbert-Huang spectrum method is less affected by noise than methods based on a spectrogram (STFT) and the Zhao-Atlas-Marks. Recent works have also reported the relevance of applying HHT to EEG signals to classify pleasant and arousal emotional values induced by musical and speech stimuli [8]–[10]

The empirical advantage of HHT over linear transformations has been analytically discussed in [11], [12]. Huang et al. (2008) [11] reported that the capacity of the HHT to obtain local and instantaneous frequency from the signals allows extracting a more accurate frequency and time resolution than the DWT. This adaptive property was further analyzed by Bueno-Lopez et al. (2017) [12] by comparing the capacity of HHT and DWT to detect instantaneous frequency from EEG signals. The authors reported that the HHT could describe the EGG oscillations, whereas the DWT resulted in no interpretable components.

This work presents a simple experiment to assess the benefit of the HHT over the DWT for emotion recognition from EEG signals. The work first introduces an HHT-based approach to recognize emotions, and then compares this approach with a previously published DWT-based approach [13].

II. Methods

A. Emotion induction

Positive emotions and negative emotions were induced using images from the International Affective Picture System (IAPS) [14]. The IAPS provides the level of arousal (or intensity) and valence (pleasantness/unpleasantness) for 956 color photographs. According to Lang (1995) [15], in the IAPS system, positive emotions have high arousal and valence values. In contrast, negative emotions have high arousal values and low valence values.

In this work, 20 IAPS images were used to evoke emotions. These images were selected based on their arousal and valence values. The ten images with the highest arousal and valence values were selected as positive emotion inducers, whereas the ten pictures with the highest arousal and the lowest valence values were selected as negative emotional stimuli.

The sequence was built by locating a negative emotion picture followed by a positive one, repeating this pattern

[1]Camilo E. Valderrama is with Data Intelligence for Health Lab, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada. camilo.valderramacua@ucalgary.ca

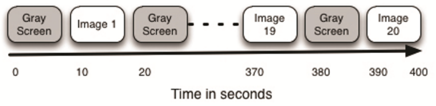

Fig. 1. Image sequence for evoking negative and positive emotions. Adapted from [29]

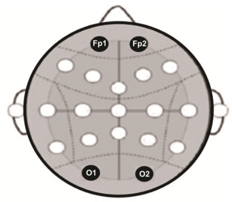

Fig. 2. Electrode montage to acquire EEG signals. Adapted from [29]

until all the pictures were placed. In order to neutralize the emotional state during the transition between two pictures, a gray screen was placed between the negative and positive stimuli. Thus, the pattern gray screen, unpleasant picture, gray screen, and pleasant picture composed a run. Each picture of this run was shown for five seconds. Fig. 1 shows the built image sequence.

B. EEG recording

The EEG signals extraction procedure was approved by the Institutional Review Board. The extraction procedure was performed in a laboratory environment with proper temperature and illumination. The EEG signals were captured using a device ML 4818 PowerLab T15 from AD-Instrument and the software LabChart on a computer (Intel Core 2 Duo, 2.33 GHz). The signals were recorded at a sampling frequency of 1000 Hz using four electrodes of the International 10/20 System: O_1, O_2, F_{p1}, and F_{p2}. The reason for choosing these electrodes was because O_1 and O_2 are located in the occipital lobe, which is the brain area initiating the visual process, perceiving the shape, movement, and color of the observed object [16]. With regard to F_{p1} and F_{p2}, these electrodes are located in the frontal lobe, which is a zone participating in the processing of the emotions [17].

The approach used a bipolar montage, and therefore the brain activity was measured in both hemispheres using the two brain channels (F_{p1}, O_1), and (F_{p2}, O_2). Fig. 2 shows the used montage, using the International 10/20 System as a reference.

The EEG signals were extracted from ten mentally healthy subjects (half for each gender) whose age ranged from 18 to 28 years old. Before reproducing the image sequence, each participant was informed about the experiment protocol. To avoid noise and artifacts provoked by blinking or muscular movements, each subject was warned to remain as still as possible during the projection of the sequence. Each subject was sat in front of a computer screen, and an electroencephalograph cap, using wet electrodes, was placed onto his/her head. The subject's EEG signals were recorded while observing the pictures. When the sequence ended, each participant filled out a survey, indicating for each picture whether he or she had felt a positive or negative emotion. Table I shows the self-evaluation responses of participated subjects.

TABLE I
SELF-EVALUATION OF EMOTIONAL INDUCTION OF THE PARTICIPATED SUBJECTS

Subject answer	**IAPS Category**	
	Positive emotion	**Negative emotion**
Positive emotion	83.75	2.50
Negative emotion	6.25	85.00
Do not know	10.00	12.25

After recording the EEG signals from all the subjects, a total of 400 signals were obtained for each subject. Half of the recorded EEG signals corresponded to each brain hemisphere. Likewise, half of the signals were induced for each type of emotion (positive or negative).

III. Emotion Recognition Approaches

Two different EEG emotion recognition approaches were used. Both approaches shared the same block diagram (preprocessing, feature extraction, and classification stages), but they differed in the transformation used for extracting the features. One approach extracted features using the intrinsic mode functions (IMF) obtained with the HHT, whereas the second one used the detail coefficients obtained with the DWT.

Both approaches also extracted fractal features from the EEG signals since it was previously reported that the combination of spectral and fractal features improved emotion recognition in the EEG signals used in this experiment [13].

A. Preprocesing

Since EEG signals are prone to artifacts and noise, soft-thresholding was applied to the recorded EEG signals following steps presented in [18].

B. Spectral features

1) HHT-based approach: HHT decomposes a signal into a set of functions called intrinsic mode functions (IMFs). IMFs satisfy two properties: *i*) The number of extrema maxima and minima and the number of zero crossings must be equal or differ at most by 1; *ii*) The mean value between the envelope of the local maxima and the envelope of the local minima must be zero at any point. To obtain the IMFs, EMD uses an intuitive algorithm called 'sifting procedure'. It is an iterative procedure, which finds all the IMFs of the signal until the difference between output and the input of the sifting procedure becomes a monotonic function. More details of the method can be found in [19].

For each IMFs, the logarithmic power (LP), the logarithmic energy (LE), the entropy (H), and the absolute logarithmic energy efficiency (ALREE) were calculated as presented in [20].

The mean, standard deviation, minimum and maximum of the LP, LE, H, and ALREE over the IMFs were taken as the

TABLE II
FREQUENCY DECOMPOSITION FOR THE DWT

Level	Frequency range (Hz)	Coefficients	Sub-band
1	250.00-500.00	d1	
2	125.00-250.00	d2	
3	62.50-125.00	d3	
4	31.25-62.50	d4	Gamma
5	15.63-31.25	d5	Beta
6	7.81-15.63	d6	Alpha
7	3.91-7.81	d7	Theta
7	0.00-3.91	a7	Delta

final spectral features, having a total of 16 spectral values for each EEG signal.

2) DWT-based approach: The DWT-based approach was previously introduced in [13]. This approach obtained spectral features by applying statistical and energy functions to the Wavelet detail coefficients extracted from EEG signals using the DWT.

The signal was decomposed into seven levels to obtain coefficients located in the frequency range (2-64 Hz) (see Table II). The detail coefficient corresponded to beta, alpha, and theta were selected to extract spectral features. The coefficients from other frequency bands were discarded because frequencies below 4 Hz are associated with eye movements and blinking, whereas those below 1.2 Hz are related to cardiac movements [21]. Moreover, frequencies above 30 Hz are related to muscular reflexes.

The DWT-based approach calculated 187 spectral features corresponding to the logarithmic power, logarithmic energy, ALREE, entropy energy, entropy power, energy, absolute logarithmic power efficiency, mean, standard deviation, maximum, minimum, and the square root of the eigenvalues of the coefficient matrix. This feature vector was reduced to six values using principal component analysis. More details of the DWT-based approach can be found in [13].

C. Fractal features

Two different fractal techniques were used to extract features from the EEG signals. Firstly, the signal singularity spectrum of the EEG was calculated following the method proposed in [22], choosig the the five exponents with the highest spectrum values as features. Secondly, the fractal dimension was calculated using Higuchi's algorithm [23].

D. Classification

The final dataset for classification was obtained by concatenating the features of the two channels corresponding to the same visual stimuli. The final dataset thus comprised 200 samples, 20 per subject, were half corresponding to each type of emotion (positive and negative).

Leave-out-one cross-validation (LOOCV) was used over the 200 samples. At each iteration, the training samples (199) were used to train a binary support vector machine (SVM) with a Gaussian kernel function. The C and σ parameters were selected using nested cross-validation in the search space 0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100, 300, 1000, 3000.

The nested cross-validation accuracy rates obtained through the LOOCV were used as a first indicator to compare the two approaches' performance.

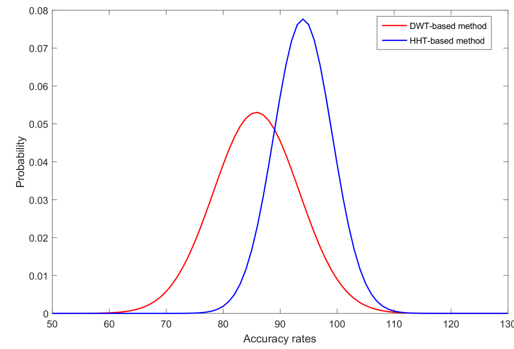

Fig. 3. Probability density distribution for the training sets during the LOOCV

The comparison of the approaches was performed using the prediction of the held-out sample of the LOOCV procedure. Each iteration of the LOOCV resulted in a prediction error of either 0% or 100% since the classifier was binary and only one sample was in the test set. The 200 estimations were used to compute the accuracy, sensitivity, specificity, precision, and F1-score were calculated.

Finally, the metrics of both approaches were compared with a right-sided two-sample t-test with the null hypothesis that the corresponding metric of HHT-based approach was lower than or equal to the accuracy of the DWT-based approach ($H_o : \pi_{HHT} \leq \pi_{DWT}$; $H_a : \pi_{HHT} > \pi_{DWT}$).

IV. RESULTS

A. Training performance

Fig. 3 shows the probability distribution for the nested cross-validation accuracy rates obtained through the LOOCV. The HHT-based approach probability distribution was centered further right than that of the DWT-based approach. The HHT-based approach also achieved a lower dispersion around the center, thus having a lower variance over the LOOCV iterations.

B. Testing performance

Both approaches equally well classified positive emotions, yielding the same sensitivity (Table III). However, the DWT-based approach showed a lower capacity to detect negative emotions, obtaining a specificity lower by more than 15 percent. The better capacity of the HHT-based approach to detect both types of emotion resulted in statistically significant differences for the accuracy, specificity, precision, and F1-score metrics (right-sided two-sample t-test; $P < 0.02$).

V. DISCUSSION AND CONCLUSION

The results presented in this work indicate that an HHT-based approach can recognize human emotion from EEG signals more accurately than a DWT-based approach.

The higher capacity of the HHT-based approach to discriminate between positive and negative emotions suggests that HHT transformation was able to extract the non-stationary and non-linear properties of EEG signals. This

TABLE III
ACCURACY, SENSITIVITY, SPECIFICITY, PRECISION AND F1 SCORE OF THE HHT-BASED AND WAVELET-BASED APPROACHES FOR THE LOOCV. A † INDICATES A STATISTICALLY SIGNIFICANT DIFFERENCE BETWEEN THE TWO APPROACHES (RIGHT-SIDED TWO-SAMPLE T-TEST; $\alpha = 0.05$)

Method	Accuracy	Sensitivity	Specificity	Precision	F1 score
HHT-based method	94.5 %†	94.0 %	95.0 %†	94.9 %†	94.5 %†
DWT-based method	86.0 %	94.0 %	78.0 %	81.0 %	87.0 %

result is consistent to [11], [12], in which by analytical comparison, reported that the HHT was more accurate than the DWT to describe oscillations and extract instantaneous frequencies from EEG signals.

The statistically significant difference in specificity for the two approaches suggests that negative emotions might be correlated with non-linearity properties of the EEG signals since the DWT, which is a linear transformation, achieved a lower specificity. However, further research is needed to validate this statement.

A limitation of this work is that it only compared the HHT transformation with the DWT, and no other transformations were considered. However, the obtained results suggest that the accuracy of the EEG emotion recognition approach can be increased using HHT. Thus, future research should compare the HHT-based approach with different transformations.

Another limitation is that only two emotional states were considered for the comparison. Although providing approaches able to classify different emotions is relevant for the AC community, the main purpose of the current work was to empirically assess whether HHT could improve the classification of emotions.

Moreover, this work only used an SVM to compare the approaches, and no advanced models, such as deep learning methods or random forest, were considered. However, the capacity of HHT for extracting non-linear properties goes beyond the selected classifier, and similar results may be expected with different classification techniques.

The presented result suggests that HHT is a promising transformation for designing and developing more robust emotion recognition approaches.

REFERENCES

[1] R. W. Picard, "Affective computing: challenges," *International Journal of Human-Computer Studies*, vol. 59, no. 1, pp. 55–64, 2003.

[2] H. Gunes and M. Pantic, "Automatic, dimensional and continuous emotion recognition," *International Journal of Synthetic Emotions*, vol. 1, no. 1, pp. 68–99, January 2010. [Online]. Available: http://doc.utwente.nl/75888/

[3] C. E. Valderrama and G. Ulloa, "Análisis espectral de parámetros fisiológicos para la detección de emociones," *Sistemas & Telemática*, vol. 10, no. 20, pp. 27–49, 2012.

[4] ——, "Spectral analysis of physiological parameters for emotion detection," in *Image, Signal Processing, and Artificial Vision (STSIVA), 2012 XVII Symposium of*, 2012, pp. 275–280.

[5] N. E. Huang and S. S. P. Shen, *Hilbert-Huang transform and its applications*, 2nd ed. Toh Tuck Link; Singapore: World Scientific, 2014.

[6] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Transactions on Affective Computing*, vol. 5, no. 3, pp. 327–339, 2014.

[7] S. K. Hadjidimitriou and L. J. Hadjileontiadis, "Toward an eeg-based recognition of music liking using time-frequency analysis," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 12, pp. 3498–3510, 2012.

[8] C. Shahnaz, S. S. Hasan *et al.*, "Emotion recognition based on wavelet analysis of empirical mode decomposed EEG signals responsive to music videos," in *2016 IEEE Region 10 Conference (TENCON)*. IEEE, 2016, pp. 424–427.

[9] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from eeg signals using multidimensional information in emd domain," *BioMed Research International*, vol. 2017, 2017.

[10] J. Chen, H. Li, L. Ma, H. Bo, and X. Gao, "Application of eemd-hht method on eeg analysis for speech evoked emotion recognition," in *2020 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*. IEEE, 2020, pp. 376–381.

[11] M. Huang, P. Wu, Y. Liu, L. Bi, and H. Chen, "Application and contrast in brain-computer interface between Hilbert-huang transform and wavelet transform," in *2008 The 9th International Conference for Young Computer Scientists*. IEEE, 2008, pp. 1706–1710.

[12] M. Bueno-Lopez, M. Molinas, and G. Kulia, "Understanding instantaneous frequency detection: a discussion of Hilbert-huang transform versus Wavelet transform," in *International Work-Conference on Time Series Analysis-ITISE*, vol. 1. University of Granada Granada, 2017, pp. 474–486.

[13] C. E. Valderrama and G. Ulloa, "Combining spectral and fractal features for emotion recognition on electroencephalographic signals," *WSEAS Transactions on Signal Processing*, vol. 10, pp. 481–496, 2014.

[14] P. J. Lang, M. M. Bradley, and B. N. Cuthbert, "International affective picture system (IAPS): Affective ratings of pictures and instruction manual," The Center for Research in Psychophysiology, University of Florida, Gainesville, FL, Tech. Rep. A-8, 2008. [Online]. Available: http://csea.phhp.ufl.edu/Media.html

[15] P. J. Lang, "The emotion probe: Studies of motivation and attention." *American psychologist*, vol. 50, no. 5, p. 372, 1995.

[16] B. Kolb and I. Q. Whishaw, *Fundamentals of Human Neuropsychology*, 5th ed. New York; NY: Worth Publishers, 2015.

[17] F. Dolcos and R. Cabeza, "Event-related potentials of emotional memory: Encoding pleasant, unpleasant, and neutral pictures," *Cognitive, Affective, & Behavioral Neuroscience*, vol. 2, no. 3, pp. 252–263, 2002.

[18] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 613–627, 1995.

[19] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," in *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971. The Royal Society, 1998, pp. 903–995.

[20] S. Jerritta, M. Murugappan, W. Khairunizam, and S. Yaacob, "Electrocardiogram-based emotion recognition system using empirical mode decomposition and discrete fourier transform." *Expert Systems: The Journal of Knowledge Engineering*, vol. 31, no. 2, pp. 110–120, 2014.

[21] P. C. Petrantonakis and L. J. Hadjileontiadis, "Emotion recognition from brain signals using hybrid adaptive filtering and higher order crossings analysis," *IEEE Transactions on Affective Computing*, vol. 1, no. 2, pp. 81–97, 2010.

[22] P. Manimaran, P. K. Panigrahi, and J. C. Parikh, "On Estimation of Hurst Scaling Exponent and Fractal Behavior through Discrete Wavelets," Tech. Rep. physics/0604004, Apr 2006. [Online]. Available: http://cds.cern.ch/record/940057

[23] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena*, vol. 31, no. 2, pp. 277–283, 1988.