Combining spectral and fractal features for emotion recognition on Electroencephalographic signals

CAMILO E. VALDERRAMA, GONZALO ULLOA
Information and telecommunications technologies research group
Universidad Icesi
Calle 18 # 122-135, Cali, Valle del Cauca
COLOMBIA
cevalderrama@icesi.edu.co, gulloa@icesi.edu.co

Abstract: - Recent studies have attempted to recognize emotions by extracting spectral and fractal features from electroencephalographic signals; however, up to now none of them have combined these two features to recognize emotions. This paper aims at providing a comparison between an accuracy rate of an approach that recognizes emotions by extracting both spectral and fractal features with that of those that extract only one of these features. To this end, we designed and implemented a procedure that recognizes positive and negative emotions by extracting spectral, fractal, or both features. Next, using this procedure, we built three different approaches to recognize positive and negative emotions; the first one extracted both spectral and fractal features, whereas the other two extracted each type of feature separately. Then, the accuracy rate of the approaches was calculated and compared among them. The comparison showed that the spectral-fractal approach recognizes emotions more accurately than the spectral and fractal approaches in 96% and 79% of the time, respectively. This suggests that it is possible to develop a more effective emotion recognition method by extracting both spectral and fractal features than extracting only one type of them.

Key-Words: - Affective Computing; Discrete Wavelet Transform; Electroencephalogram; Emotion recognition; Multifractal Analysis; Support Vector Machine; Pattern recognition.

1 Introduction

Emotion is a key aspect of human beings, which determines how a human reacts when interacting with another being or a specific situation. Nowadays, with the current proliferation of computers in the human life, humans perform many of their daily tasks using computers. However, these devices are unable to perceive human emotions, disabling the emotional aspect in the human and computer interaction. Being conscious of this, recent studies in Human-Computer interaction have attempted to endow machines with the ability to recognize human emotions. The importance of accomplishing this late objective is that it will give to machines the capacity of properly responding to the users' emotional state, improving thus the human-machine interactions and making the experience more pleasant for the users [1].

To recognize emotions, researchers have used many methods. One first method is through facial gestures and voice ([2, 3]), but although this has obtained high accuracy rates, it is unreliable because an individual could easily imitate both facial gestures and voice, manipulating the experiment's output. Another method is analyzing physiological signals from the Automatic Nervous System (ANS); this claims that this kind of physiological signals reflect the emotional changes that occur inside the body. Thus, skin conductance, pupil dilation and heart beat have been used for this purpose ([4, 5]). However, the disadvantage of this method is that other physiological processes affect the ANS signals, adding to them physical activity that is different from emotional experience.

A third method to recognize emotions is the analysis of brain activity using neuroimaging techniques. These techniques overcome some of the limitations that arise with the other two methods since it captures the emotional experience occurring in the central nervous system, which is less influenced by physiological processes and harder to manipulate by a person. Regarding the neuroimaging techniques, the technique that offers more advantages is Electroencephalography (EEG), which is less intrusive and has better resolution time than other techniques such as Magnetoencephalogram (MEG), Positron Emission Tomography (PET) and functional Magnetic

Resonance Imaging (fMRI). Consequently, conscious of the advantages offered by electroencephalography, recent studies have been focusing in the analysis and process of electroencephalographic signals to recognize emotions.

The way to recognize emotions from electroencephalographic signals is by developing a procedure that processes the signals, extracts feature from the signals, and classifies these features into a type of emotion. Thus, Petrantonakis and Hadileotiadis [6] developed a filtering procedure to decompose the electroencephalographic signals and to extract features from them. Then, they used four different classification methods to identify six emotions, obtaining an accuracy rate of 85.17%. Murugappan *et al.* [7] used the Discrete Wavelet Transform (DWT) for signal processing and feature extraction, and then, they classified five emotions using Support Vector Machines (SVM). Schaaff and Schultz [8] extracted features from the alpha frequency band to recognize three emotions induced by images.

Generally, most of the previous works have recognized emotions by extracting spectral features from frequency bands of the signals (e.g., [9], [10], [11], [12], [13] and [14]). However, some works have started to recognize emotional states by extracting features through fractal analysis. Thus, Liu *et al.*[15] extracted fractal features from electroencephalographic signals to detect the valence and arousal levels induced by musical tracks on subjects. Likewise, Sourina and Liu [16] developed a serious game that recognizes the player's emotional states by applying fractal techniques.

Up to now, little work has extracted spectral and fractal features jointly to recognize emotions. Nonetheless, Brodu suggested that combining spectral and fractal features allows recognizing neural patterns in better manner than using only spectral features [17]. Based on this idea and on the fact that it is necessary to continue conducting research to improve emotion recognition rates and develop more effective feature extraction methods [6], a comparison between an approach that extracts both spectral and fractal with those that extract only one type of these features would be of considerable interest, particularly with regard to recognize emotions in a more accurate manner.

In this paper, therefore, we compare the accuracy rate of an approach that recognizes emotions by extracting both spectral and fractal features with that of approaches that extract only spectral or fractal features independently. To this end, we designed and implemented a procedure that recognizes positive and negative emotions by extracting spectral, fractal, or both features. Next, using this procedure, we built three different approaches to recognize positive and negative emotions. The first approach extracted both spectral and fractal features, and the other two extracted each type of feature separately. Then, we calculated the accuracy rate of each approach, and finally we compared the rates among them. Our obtained results suggest that extracting both spectral and fractal features allows a better accuracy rate than extracting only one type of these features.

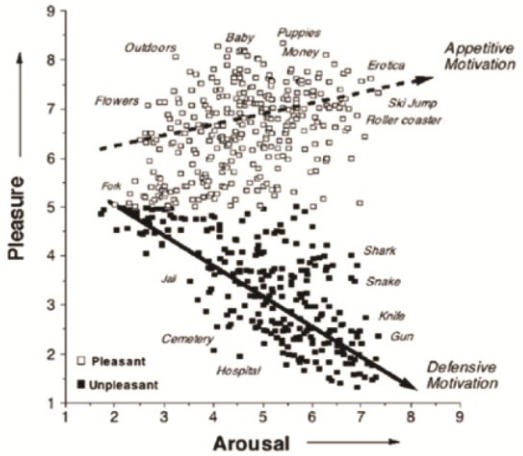
The remainder of the article is organized as follows: Section 2 describes the experiment that we used to induce emotions on subjects. Section 3 presents the procedure we designed and implemented to detect emotions. Section 4 shows the results, and finally, Section 5 presents the discussions and conclusions of the article.

2 Experiment design

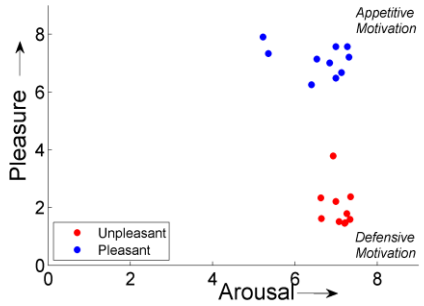
In this paper, we compare the accuracy rates of emotion recognition approaches that extract spectral and fractal features separately and jointly. For this purpose, we first designed an experiment to induce the emotions used in the comparison process. Thus, our experiment induced positive and negative emotions using pictures from the International Affective Picture System [18], which is a large set of colored pictures for eliciting emotions. Since this set contains a lot of pictures, we had to carefully select the appropriate pictures to induce the desired emotions. To this end, we used the valence-arousal model, which quantitatively measures emotions using two dimensions; the valence axis that varies between pleasure and displeasure, and the arousal axis that varies between calm and excitement.

Since we wanted to induce positive and negative emotions, which are a general type of emotions, before selecting the pictures we established positive emotions as pleasant stimuli and negative emotions as unpleasant stimuli. According to Lang [19], the pleasant stimuli are located in the high arousal and high valence subspace of the valence-arousal model supplied by the International Affective Picture System (Fig. 1a), whereas unpleasant ones are located in the high arousal and low valence subspace. Thus, to select the pleasant and

unpleasant pictures, we first ordered the images in a descending manner, taking arousal as the first discriminating value and valence as the second characteristic. Next, we chose ten images with high arousal and valence values as pleasant stimuli, and ten with high arousal and low valence values as unpleasant stimuli. Figure 1b shows the valence-arousal diagram of the selected images, and Table 1 shows the set of images selected.



(a) Arousal-Valence Model of the IAPS pictures [20].



(b) Arousal-Valence Model of the selected pictures.

Fig. 1: Arousal-Valence Model of the pictures.

Table 1. Selected pictures.

UNPLEASANT PICTURES			
Id	Title	Valence value	Arousal value
1120	Snake	3,79	6,93
3000	Mutilation	1,59	7.34
3001	Headless Body	1,62	6,64
3080	Mutilation	1,48	7,22
3170	Baby Tumor	1,46	7,21
3500	Attack	2,21	6,99
6230	Aimed Gun	2,37	7,35
9410	Soldier	1,51	7,07
9908	Car Accident	2,34	6,63
3010	Mutilation	1,79	7,26
	Mean	2,02	7,06
	Standard Deviation	0,72	0,27
PLEASANT PICTURES			
Id	Title	Valence value	Arousal value
4668	Erotic Couple	6,67	7,13
5621	Sky Divers	7,57	6,99
8030	Skier	7,33	5,35
8179	Bungee	6,48	6,99
8185	Skydivers	7,57	7,27
8186	Sky surfer	7,01	6,84
8341	Wing walker	6,25	6,40
8492	Rollercoaster	7,21	7,31
8501	Money	7,91	5,22
8163	Parachute	7,14	6,53
	Mean	7,11	6,93
	Standard Deviation	0,52	0,36

Using these 20 selected pictures, we built a video sequence. The sequence was built locating an unpleasant picture followed by a pleasant one, repeating this pattern until all the pictures were placed. In order to neutralize the emotional state of the viewer during the transition between two pictures, a gray screen was placed in the middle of the two. Every picture and every gray screen was projected for five seconds, and as result, the sequence last a total of 200 seconds. Fig. 2 shows the built video sequence.

Fig. 2 Sequence of images

The recording of the electroencephalographic signals was performed in a laboratory environment with proper temperature and illumination. We used a device ML 4818 PowerLab T15 to capture the signals, and the software LabChart to record these signals on a computer (Intel Core 2 Duo, 2.33 GHz); both the capture device and the software are from the AD-Instrument company. The signals were recorded using four electrodes of the International 10/20 System at a rate of 1000 Hz.

The four selected electrodes were O_1, O_2, F_{p1} and F_{p2}. The electrodes O_1 and O_2 were chosen because they are located in the occipital lobe, which is the brain zone that initiates the visual process by perceiving the shape, movement and color of the observed object [21], and the electrodes F_{p1} and F_{p2} were chosen because

they are located in the frontal lobe, which is a zone that participates in the processing of the emotions in the brain [22]. Using these four electrodes, we established two brain channels, (F_{p1}, O_1) and (F_{p2}, O_2) to measure the brain activity in both right and left hemispheres. Fig. 3 shows our montage using the International 10/20 System as reference.

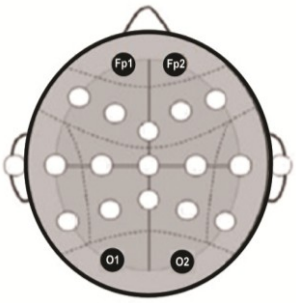


Fig. 3 Electrodes' montage.

To carry the experiment out, we selected ten mentally sane persons (five males and five females), aged between 18 and 28. The experiment was individually performed with each participant. Before the experiment started, each participant was informed about the protocol and details of the recording process. When the instructions were clear to the participant, we sat him/her in front of a computer screen, and then, we placed an electroencephalograph's cap onto his/her head. In order to minimize noise and artifacts in the electroencephalographic signals, the participant was warned to remain as still as possible during the projection of the video sequence. Additionally, the participant was also warned to avoid blinking during this phase.

Once the participant was ready to start the experiment, we projected the video sequence. The video sequence started with a gray screen whose function was to neutralize the emotional state of the participant. Next, an unpleasant picture was showed, followed by a new gray screen, and then, it was followed by a pleasant picture. These four pictures (the two gray screens, the unpleasant picture and the pleasant picture) composed a run, and each picture lasted five seconds. The video sequence was composed by 10 runs, having a total time of 200 seconds. While the participant was observing the pictures, we captured his/her electroencephalographic signals using the brain channels (F_{p1}, O_1) and (F_{p2}, O_2).

When the sequence ended, the participant filled out a survey, indicating for each pleasant and unpleasant picture whether it generated positive or negative emotions to him/her. Table 2 shows the results of the answers of all the participants. In these results, it is observed that in the majority of cases the participant perceived the proper emotions. Consequently, the selected pictures reliably induced positive and negative emotions on the subjects.

Table 2. Self-Assessment of the positive and negative emotions induction.

SUBJECTS' ANSWER	IAPS CATEGORIZATION	
	Pleasant stimuli	Unpleasant stimuli
Positive emotions	83,75%	2,50%
Negative emotions	6,25%	85,00%
Don't know	10,00%	12,50%

3 Procedure

Once positive and negative emotions were induced on subjects, the next step was to design and to implement a procedure for recognizing these emotions. To this end, we based on a previous literature review of emotion recognition made by the authors [23]. In this section, we introduce the procedure (Fig. 4) used to build all the approaches that are compared in this article. The only stage of the procedure that differs among all the approaches was the feature extraction stage, which in the first approach extracted both spectral and fractal features whereas in the other two ones extracted each feature separately. In the following subsections, we explain how each of the stage of the procedure was made.

3.1 Raw signal

After we project the video sequence to each participant, a total of 20 electroencephalographic signals were obtained. Since we used the brain channels (F_{p1}, O_1) and (F_{p2}, O_2) to record the signal, the half of these recorded signals were from the right hemisphere and the other half were from the left hemisphere.

Later, we created a sample set using the electroencephalographic signals. To this end, we first created a sample set for each signal by dividing the signal into 40 samples of five seconds. Each of the 40 samples corresponded to one of the pictures of the video sequence; thus, we obtained 20 samples

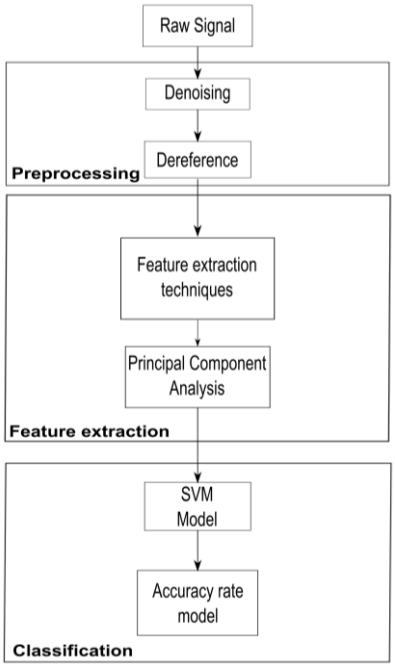


Fig. 4 The Proposed Scheme

for the gray screens, 10 samples for the unpleasant pictures and the 10 samples for the pleasant pictures. Then, we discarded all the samples corresponding to the gray screens, keeping only the samples corresponding to the pictures. Once the sample set was created for each signal, we combined all the individual sets (40 samples /participant) into a final set of 400 samples. This later set was used as an input for the emotion recognition procedure.

3.2 Pre-processing Stage

In this stage, the samples were pre-processed removing noises and other brain activities caused by the experiment and the reference points (ground and Cz electrode).

3.2.1 Noise Elimination

In order to remove noise and artifacts caused in the experiment, we applied threshold techniques based on the Discrete Wavelet Transform (DWT). The procedure was applied to each sample as follows: first, the sample was decomposed into 12 levels using the Daubechies-4 as Mother Wavelet. Then, the threshold value was calculated and normalized using the Eq. 1,

$$\lambda = \sqrt{2\log N} \times \frac{median\left(\left|d_k^1\right|\right)}{0.6745}, \qquad (1)$$

where $\left(\left|d_k^1\right|\right)$ is the set of wavelet detail coefficients of the first decomposition level, and N is the number of points of the sample.

Later, using the calculated threshold value, a soft thresholding was performed on the wavelet detail coefficients applying the following equation:

$$\forall_{i,k}\ \hat{d}_k^i = sgn(d_k^i)(|d_k^i| - \lambda), \qquad (2)$$

where i is the i-decomposition level; k, $1 < k < 2^{-i}N$, specifies the k-coefficient of the i decomposition level; and sgn is the sign function. Lastly, the sample without noise was obtained applying the Inverse Discrete Wavelet Transform to the modified coefficients. As an illustration, Fig. 5 shows what happens after removing the noise on one sample obtained in the experiment.

3.2.2 De-referencing

Besides the noise, the recorded electroencephalographic signals are also affected by the activity of the reference points used to capture the signals. Electroencephalography compares voltage measurements between an electrode and a reference point. Therefore, the obtained electric is a combination of both the electrode and the reference activity.

In our study, the activity caused by the reference points Cz and ground was removed applying the Current Source Density (CSD) technique. This technique defines the brain activity at each electrode as the voltage difference between the electrode and the weighted average of the surrounding electrodes. Thus, the CSD is estimated with the following equation:

$$x_n = x - \frac{1}{N_{neighbors}} \sum_{i=1}^{N_{neighbors}} X_i(t), \qquad (3)$$

where $N_{neighbors}$ are neighbor electrodes located around the electrode of interest.

Since our approach uses only two neural channels, one for each brain hemisphere, the CSD was performed as follows:

$$\hat{a}[n] = a[n] - b[n] \ \wedge\ \hat{b}[n] = b[n] - a[n],$$

where a is the sample obtained in the right hemisphere, b is the sample obtained in the left hemisphere, and n, $1 < n < N$, is the n-point of each sample.

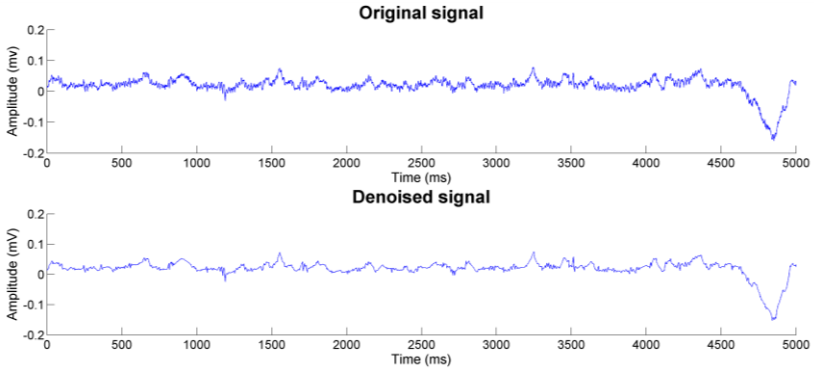


Fig. 5 Denoising Process

3.3 Extraction Features Stage

This stage is responsible for extracting features from the input samples. Since each of the approaches used a different set of features to recognize the induced emotions, this stage was different for all the approaches. Thus, the first approach extracted both spectral and fractal features, the second one extracted spectral features, and the third one extracted fractal features. At the end of this stage, the extracted features were consolidated using Principal Component Analysis. In this subsection, we explain the methods used to extract spectral and fractal features from the input samples and the process for consolidating such features.

3.3.1 Spectral Features Extraction

Spectral features are obtained decomposing a signal into its frequency components and calculating statistical and energy functions over them. In this study, we obtained the frequency components of the samples using the Discrete Wavelet Transform. This transform decomposes a signal in both time and frequency resolutions, providing at what time the frequency components occur. To this end, the signal is decomposed into levels using successive pairs of low-pass and high-pass filters (a signal of N points can obtain as maximum $log_2 N$ levels). At the end of each level, the pair of filters halves the signal's frequency and outputs a set of coefficients, which are called detail coefficients for the high-pass filter and approximation coefficients for the low-pass filter. The number of wavelet coefficients obtained at the j level is $2^{-j}N$, where $j = 1, \dots, log_2 N$, and N is the number of points of the signal. This process is repeated until every decomposition level is performed, taking the approximation coefficients of the level j as input of the $j + 1$ level.

In our work, we established the number of decomposition levels applying the Nyquist's Theorem, which states that the greatest frequency component of a signal is equal to half of the sample rate. Therefore, because the samples were recorded at 1000 Hz, the greatest frequency component of the samples was 500 Hz. Consequently, the samples were decomposed into 7 levels using the Daubechies-4 Mother Wavelet (Table 3).

Table 3. Frequency Decomposition.

Level	Frequency Range (Hz)	Coefficients	Corresponding Subband
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	0-3.91	a7	Delta

Once the sample is decomposed into its wavelet coefficients, the next step is to properly choose the coefficients that contain the most useful information, which, according to wavelet theory, are those with the highest amplitude. To this end, we applied the Discrete Wavelet to two special samples; the first one was the average of all the samples recorded on the left hemisphere, and the second one was the average of all the samples recorded on the right hemisphere.

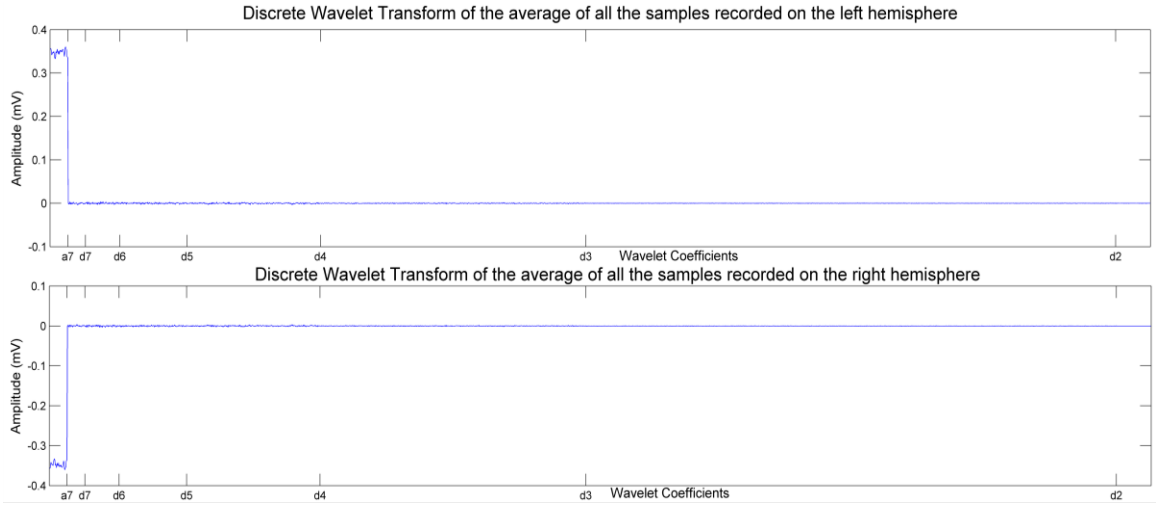


Fig. 6 Discrete Wavelet Transform of the samples

Figure 6 shows the obtained wavelet coefficients of these special samples. The coefficients with the most amplitude were located between 0 and 62.5Hz (a7-d4). However, although a high correlation was presented in low frequencies, we discarded these coefficients because frequencies below 4 Hz are associated with eye movements and blinking and those below 1.2 Hz are related to cardiac movements. Likewise, frequencies above 30 Hz were also discarded because they are related to muscular reflexes [6]. Therefore, we only considered the detail coefficients d5, d6 and d7, which correspond to the ranges 3.91-7.81Hz, 7.81-15.63Hz and 15.63-31.25Hz, respectively.

Therefore, we extracted spectral features from each of sample as follows: we first applied the Discrete Wavelet Transform to the sample, and then we used the detail coefficients of levels five through seven to calculate the 11 measures shown in Table 4. Finally, once all the measures were calculated, they were concatenated into a vector feature, constituting, hence, a final vector of dimension 187.

Table 4. Spectral features.

Feature	Description	Equation	Dimension Length
Logarithmic Potency	Measures the logarithmic potency of the detail coefficients at j level.	$LP_j = log\left(\frac{1}{N}\sum_{i=1}^{N}\left(d_i^j\right)^2\right)$ N: the number of coefficients at j level	3 values, one per each frequency band.
Logarithmic Energy	Measures the logarithmic energy of the detail coefficients at j level.	$LE_j = log\left(\sum_{i=1}^{N}\left(d_i^j\right)^2\right)$ N: the number of coefficients at j level	3 values, one per each frequency band.
Energy	Measures the energy of the detail coefficients at j level.	$E_j = \sum_{i=1}^{N}\left(d_i^j\right)^2$ N: number of coefficients at j level	3 values, one per each frequency band.

Continue in the next page

Table 4 - *Continuation of the previous page*

Feature	Description	Equation	Dimension Length
Entropy	Measures the entropy of the signal. It represents signal's nonlinearity.	$H = -\sum_{i=1}^{M}\left(\frac{E_i}{E_t}\right) log\left(\frac{E_i}{E_t}\right)$ M: number of decomposed levels E_i: Energy at i level E_t: Sum of the energy of the three selected frequency bands	1 value for the entire sample.
Alree Energy	Measures the absolute logarithmic value of the energy efficiency at j level.	$ALREE_j = \left\lvert log\left(\frac{E_j}{E_t}\right)\right\rvert$ E_j: Energy at level j E_t: Sum of the energy of the three frequency bands selected	3 values, one per each frequency band.
Alrpe Potency	Measures the absolute logarithmic value of the potency efficiency at j level.	$ALRPE_j = \left\lvert log\left(\frac{P_j}{P_t}\right)\right\rvert$ P_j: Energy at level i P_t: Sum of the potency of the three frequency bands selected	3 values, one per each frequency band.
Mean	Measures the mean of the detail coefficients at j level.	$M_j = \frac{1}{N}\sum_{i=1}^{N} d_i^j$ N: number of coefficients at j level	3 values, one per each frequency band.
Standard Deviation	Measures the standard deviation of the detail coefficients at j level.	$SD_j = \frac{1}{N-1}\sum_{i=1}^{N}\left(M_j - d_i^j\right)^2$ M: mean at j level N: number of coefficients at j level	3 values, one per each frequency band.
Maximum	Determines the maximum detail coefficient at j level.	$Max_j = max\left(d_1^j, d_2^j, \dots, d_N^j\right)$ N: number of coefficients at j level	3 values, one per each frequency band.
Minimum	Determines the minimum detail coefficient at j level.	$Min_j = min\left(d_1^j, d_2^j, \dots, d_N^j\right)$ N: number of coefficients at j level	3 values, one per each frequency band.
SVD	Measures the square root of the C^TC matrix's eigenvalues.	$SVD = \sqrt{\Delta(C^TC)}$ $C = \begin{bmatrix} d_5 \\ d_6 \\ d_7 \end{bmatrix}$ where d_i is the set of coefficients at i level	159 values for the entire signal.

3.3.2 Fractal Features Extraction

Fractal features is a suitable indicator to analyze the irregular and scaling properties of physiological signals [24, 25]. As consequence, in this work we extracted fractal features to properly analyze the irregular and scaling behavior of the obtained electroencephalographic signals. We extracted the fractal measures using two fractal techniques: the multifractal analysis and the fractal dimension. The first technique evaluates a repartition of the different irregularities present in a signal, providing a global distribution of the signal's irregularities, and the second technique measures the self-similarity of the entire signal.

To calculate the multifractal analysis, the wavelet transform is a very attractive technique because it can properly quantify the irregularities distribution of the signal [26]. Therefore, in this

work the multifractal analysis was calculated using the Discrete Wavelet Transform following the steps proposed by Manimaran et al. [27]. This procedure was applied to each sample obtained after the pre-processing stage.

At the beginning of the procedure, the profile signal of the input sample was determined using the following equation:

$$F(i) = \sum_{t=1}^{i} [x_t - \bar{x}], i = 1, \dots, N, \quad (4)$$

where x_t is the $t-th$ element of the input sample x, $\bar{x}$ is the average of input sample, and N is the number of points in the input sample x. Next, this profile signal was decomposed into 10 decomposition levels using the Discrete Wavelet Transform.

Next, for each of the 10 decomposition levels, the fluctuation segments of the level were determined. To accomplish this, first the polynomial behavior of the level was calculated by applying the inverse discrete wavelet transform on the approximation coefficients of the level. Then, the fluctuation of the level was calculated by subtracting the Profile signal with the polynomial behavior of the level. Subsequently, the fluctuation of the level was divided into non-overlapping $2M_s$ segments; here M_s equals$\lfloor N/s \rfloor$, where N is the size of the sample and $s = 2^{L-1}W$ is the scale of the level, which depends of the level(L) and the cardinality of mother wavelet's filters (W). The first M_s segments were determined splitting the sample from start to finish, and the last M_s segments were determined splitting the sample from finish to start. Table 5 shows the number of fluctuation segments that was obtained per decomposition level.

Table 5. Number of fluctuation segments per level.

Level	Scale	Number of fluction segments
1	4	2500
2	8	1250
3	16	624
4	32	312
5	64	156
6	128	78
7	256	38
8	512	18
9	1024	8
10	2048	4

Later, the q^{th} order fluctuation function of each scale was calculated. To this end, first the variance of each segment of each scale was calculated, and then, q^{th} order fluctuation function was calculated as follows:

$$F_q(s) = \left\{ \frac{1}{2M_s} \sum_{b=1}^{2M_s} [F^2(b,s)\,]^{q/2} \right\}^{1/q}, \quad (5)$$

where the variance of the fluctuation segment $b, 1 < b < 2M_s$, of the scale s is represented by the notation $F^2(b,s)$. The variable q is the order of moments, and it is defined for an interval of integers. In our approach, the values of q ranged between -41 and 41. Once the q^{th} order fluctuation function was calculated for each scale s and for each order moment q, we obtained the following matrix:

$$\begin{bmatrix} F_{-41}(4) & F_{-41}(8) & \cdots & F_{-41}(2048) \\ F_{-40}(4) & F_{-40}(8) & \cdots & F_{-40}(2048) \\ \vdots & \vdots & \ddots & \vdots \\ F_{41}(4) & F_{41}(8) & \cdots & F_{41}(2048) \end{bmatrix}.$$

In the obtained matrix, each row was a fluctuation function of the q order. Using each fluctuation function, the scale exponent $(h(q))$ of the order moment q was calculated by the relation $F_q(s) = s^{h(q)}$. Thus, the scale exponent $h(q)$ was determined by calculating the slope of the logarithmic-logarithmic curve $F_q(s)$ vs. the scales s. When the scale exponent was calculated for every order moment q, we obtained the following vector:

$$[h(-41) \quad h(-40) \quad \cdots \quad h(40) \quad h(41)]^T.$$

Then, the values of this vector were used to calculate the singularity spectrumf(α). First, the multifractal scale exponents were determined by the relation $\tau(q) = qh(q) - 1$, and finally the singularity spectrum was determined using the Legendre Transform, $\alpha = \tau'(q)$ and $f(\alpha) = q\alpha - \tau(q)$.

To show the reader what was obtained after applying multifractal analysis to a sample and how was extracted the fractal features, we performed the multifractal analysis to two special samples; the first sample was the average of the samples induced by unpleasant stimuli, and the second one was the average of the samples induced by pleasant stimuli. Fig. 7 shows the singularity spectrum obtained. As we see at the figure, the singular spectrum is a curve that has a maximum value of one.

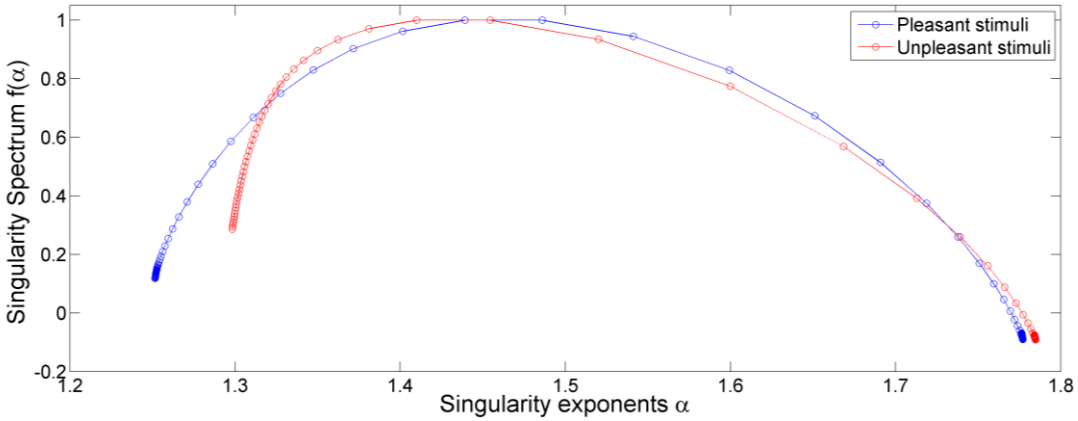


Fig. 7. Singularity Spectrum of the Signals.

We extracted the fractal features from the singularity spectrum as follows:

1. The maximum value of the singularity spectrum was identified.
2. The singularity exponent (value on the x axis), which corresponds to the maximum value of the singularity spectrum, was identified.
3. The two left points that adjoin with the selected singularity exponent at step 2 was taken.
4. The two right points that adjoin with the selected singularity exponent at step 2 was taken.
5. A vector with the five selected singularity exponents was built.

Thus, for the Fig. 7, the singularity exponents was $[1.4015, 1.4392, 1.4861, 1.5413, 1.5993]$ for the average of the samples induced by unpleasant stimuli (blue curve), whereas they was $[1.3810, 1.4100, 1.4545, 1.5200, 1.5999]$ for the average of the samples induced by pleasant stimuli (red curve).

On other hand, we also used the fractal dimensional to extract fractal features from the input samples. The fractal dimension was calculated using the Higuchi's algorithm. This algorithm determines the fractal dimension calculating the average length of the curve using a segment of k samples.

We calculated the fractal dimension to every sample obtained after the pre-processing stage. We proceeded as follows: firstly, the time interval k was fixed in 5. Next, a double nested loop was performed; the outer loop iterated k times, and for each iteration i of the outer loop $(1 \leq i \leq k = 5)$, the inner loop iterated i times. At each step of the inner loop, a new time series was calculated using the iteration indexes of both loops as follows:

$$X_i^m: X(m), X(m+i), X, \ldots, X\left(m + \left\lfloor \frac{(N-m)}{i} \right\rfloor \cdot i\right), \qquad (6)$$

where i is the index of the outer loop, m is the index of the inner loop $(1 \leq m \leq i)$, and N is the size of the sample. For example, for i=5, m=1, and N=5000, we obtained the time series: $X_5^1: X(1), X(6), X(11), \ldots, X(4996)$. Then, the length $L_m(i)$ of this new time series was calculated using the following equation:

$$L_m(i) = \left\{ \left(\sum_{j=1}^{\left\lfloor \frac{N-m}{k} \right\rfloor} |X(m+jk) - X(m+(j-1)k)| \right) \frac{N-1}{\left\lfloor \frac{N-m}{k} \right\rfloor i} \right\} (i)^{-1}, \qquad (7)$$

where i and m are the indexes of the loops, and N is the size of the sample.

At the end of outer loop, we obtained the following matrix:

$$\begin{bmatrix} L_1(1) & & & & \\ L_1(2) & L_2(2) & & & \\ L_1(3) & L_2(3) & L_3(3) & & \\ L_1(4) & L_2(4) & L_3(4) & L_4(4) & \\ L_1(5) & L_2(5) & L_3(5) & L_4(5) & L_5(5) \end{bmatrix}.$$

Using the values of the matrix, the average length value of each row i $(1 \leq i \leq 5)$ was calculated as follows:

$$\langle L(i) \rangle = \frac{1}{i} \sum_{m=1}^{i} L_m(i). \qquad (8)$$

Finally, the obtained average length values were used to calculate the fractal dimension through the relation $\langle L(k) \rangle \propto k^{-D}$. Therefore, the fractal dimension was determined by calculating the slope of the logarithmic-logarithmic graph of k vs. $\langle L(k) \rangle$.

At the end of the fractal features extraction process, a feature vector of dimension six was built. The first five values of this vector were the features that were extracted with the multifractal analysis, and the last one value was the value that was extracted with fractal dimension.

3.3.3 Principal Components Analysis

As mentioned before, this work compares the accuracy rate of an approach that recognizes emotions by extracting both spectral and fractal features with that of approaches that extract only spectral or fractal features independently. Since these approaches used a different set of features to recognize the induced emotions, the vector feature of each approach was built in a different manner. Thus, the first approach built the feature vector concatenating spectral and fractal features, whereas the second one and third one used each type of these features separately.

Moreover, as two electroencephalography signals were recorded for each participant, two samples and two vector features were obtained for each visual stimulus. Consequently, to create the final feature vector of a visual stimulus, the two feature vectors corresponding to the same visual stimulus were concatenated, obtaining, thus, a final feature vector set of 200 vectors.

Column 2 of Table 8 shows the number of features and the dimension of the vector features for each approach. Since these feature vectors had a considerable dimension, we consolidated the more useful information of these vectors using Principal Components Analysis. This technique creates a new sub-space from a vector feature keeping only the information that is most pertinent. Column 3 of Table 6 shows the new dimension of the feature vector after applying Principal Components Analysis. These 200 reduced feature vectors were used as an input in the classification stage.

Table 6. Number of fluctuation segments per level.

Approach	Dimension (number of features) of the feature vector	Dimension of the feature vector after PCA
Spectral-Fractal	386 (26)	6
Spectral	374 (22)	4
Fractal	12 (4)	3

3.4 Classification Stage

In this stage, a set of Support Vector Machine (SVM) was used to classify the input feature vectors between positive and negative emotions. This kind of classifier was chosen because according with previous works it has a better accuracy rate than other classifiers [28].

In our approach, the SVMs were built using the Gaussian function RBF as a kernel function:

$$k(x_i, x_j) = exp\left(\frac{-|x_i - x_j|^2}{2\sigma^2}\right). \quad (9)$$

We determined the success rate of each SVM using the Bootstrapping technique. This technique requires testing the SVM with every input feature vector, and therefore, each SVM was tested iterating over all the 200 feature vectors. Therefore, at each iteration step, one of the vectors was selected as test set, and the remaining 199 vectors were used as training set. Next, we built three different SVMs, whose parameters (C and sigma) were selected by performing a k-cross-validation onto the training set. Then, the success rates of these SVMs were calculated using the test set. Since the test set was compound by only one sample, the classifier's success rate was of 0% (failure) or 100% (success). This procedure was repeated until we took every feature vector out as a test set. At the end of the 200 iterations, a total of 200 individual success rates were obtained for each SVM. Finally, the final success rate of each SVM was calculated averaging its 200 individual success rates.

As mentioned above, we applied k-cross-validation onto the training set of each iteration step. To this end, the k-cross-validation process was performed fixing k value in 10. Hence, the input set was divided into 10 samples. Next, a loop iterated over all subgroups, selecting at each step the best pair of parameters that classified the subgroup. To select the best pair of parameters (C and sigma), the selected subgroup was used as validate set, and the other nine subgroups were used as training set. Then, the set {0.01, 0.03, 0.1, 0.3, 1, 3, 10, 30, 100, 300, 1000, 3000} was fixed as the set of values that the parameters C and sigma could take. Using all the 114 possible combinations of the parameter set, 114 temporal SVMs were built using the training set, and then, each SVM was tested with the validate set. Finally, the best pair of parameters was the one that classified better the validate set. At the end of the k-cross-validation process, we obtained a total of 10 pairs of best parameters (one for every subgroup).

Finally, the three different SVMs of the classification stage were built using the 10 pairs of best parameters of the k-cross-validation process. To this end, we followed the procedure proposed by Anguita et al. [29]. Table 7 describes the three models developed.

Table 7. SVM Models.

Model	Description
SMV meanParameters	The 10 pairs of parameters were averaged. The final SVM was built with the average of C and the average of sigma: $(\bar{C}, \bar{\sigma})$.
SVM bestParameters	The pair of parameters with the best success rate in the cross-validation process was selected. The final SVM was built with these parameters: (C, σ).
SVM bestScaledParameters	The best C parameter was scaled as follows: $C' = \frac{10C}{(10-1)}$ The sigma parameter was ignored because it depended only on the input size, which was always the same.

4 RESULT

Based on the fact that extracting spectral and fractal features allows to recognize neural patterns more accurate than extracting only spectral features [17], the aim of this paper was to compare the success rate of an approach that recognizes emotions by extracting both spectral and fractal features with that of those that extract only one of these features. To accomplish this, we first selected a set of pictures to induce positive and negative emotions, and next we projected these pictures to ten individuals. While the individuals were observing the pictures, we recorded their electroencephalographic signals. Once all individuals observed the sequence, we obtained a total of 20 electroencephalographic signals, 10 per each brain hemisphere.

Later on, using the method explained in the section 3, we built three approaches to recognize the induced emotions. Each one of these approaches extracted a different feature set to recognize emotions; thus, the first approach extracted both spectral and fractal features, and the other two extracted each type of feature separately. To calculate the success rate of each approach, we used the bootstrapping technique. Therefore, we left out one feature vector as testing set, and then we trained a SVM using the 199 remaining vectors. We repeated this procedure until we took every feature vector out as a test set.

In this section, we present and compare the success rates of the approaches. To this end, firstly, we reported the results of each approach. Subsequently, we used two-sample proportion tests to significantly compare the success rate of the approach that recognized emotions by extracting both spectral and fractal features with the success rates of the other two ones.

4.1 Spectral-Fractal Approach

In this approach, after denoising the recorded electroencephalographic signals, we extracted both spectral and fractal features. As previously explained in section 3, the spectral features were extracted by applying the Wavelet Discrete Transform to the samples, while, the fractal features were extracted by performing the multifractal analysis and the fractal dimension. At the end of extraction process, we obtained 200 feature vectors, whose dimension was reduced from 386 to 6 using Principal Component Analysis.

Later on, we used the three SVMs explained in 3.4 to recognize the induced emotions. Thus, we calculated the success rate of each SVM using bootstrapping technique. To this end, we left out one feature vector, and then we trained the SVMs using the remaining vectors. We repeated this procedure for each feature vector, having a total of 200 individual success rates at the end. Table 8 shows the three SVMS' accuracy rates discriminated for each type of stimulus. As can be seen, approximately in eight out of ten cases, the three SVMs satisfactorily predicted each type of stimulus. Furthermore, the average accuracy rates of all the SVMs to predict the unpleasant and pleasant stimuli were of 86.33% and 89.33%, respectively. Consequently, the extraction of spectral and fractal features allowed properly recognize the two types of stimuli.

Finally, comparing the SVMs between them, it is observed that classifiers SVM_bestParameters and SVM_bestScaledParameters slightly predict better

than SMV_meanParameters, however these difference are statistically not significant.

4.2 Separate-Features Approach

Once we calculated the accuracy rate of the approach that extracts both features, we repeated the emotion recognition process using other two different approaches. Each one of these new approaches only extracted one type of features. Therefore, we built one approach that recognized the induced emotions by extracting spectral features, and another approach that extracted fractal features for the same purpose. As the spectral-fractal approach, we calculated the accuracy rate of these new approaches following the procedure explained in 3.4.

Tables 9 and 10 present the accuracy rates of the approaches. It can be observed from these tables that for the spectral approach the average accuracy rates of all the SVMs to predict unpleasant and pleasant stimuli were of 77.33% and 85.33% respectively, while for fractal approach were of 87.66% and 82.33% respectively. The majority of these rates were lower than the average accuracy rate of the spectral-fractal approach; only the average accuracy rate to predict unpleasant stimuli of the fractal approach was slightly higher than that of the spectral-fractal approach. However, to determine if those differences were statistically significance, we used two-sample proportion test, whose results are presented in the next subsection.

Table 8. Model Accuracy.

	SMV meanParameters		SVM bestParameters		SVM bestScaledParameters	
Classifier's prediction \ Stimuli	Unpleasant	Pleasant	Unpleasant	Pleasant	Unpleasant	Pleasant
Unpleasant	**85%**	15%	**87%**	13%	**87%**	13%
Pleasant	10%	**90%**	11%	**89%**	11%	**89%**

Table 9. Model Accuracy of spectral-approach

	SMV meanParameters		SVM bestParameters		SVM bestScaledParameters	
Classifier's prediction \ Stimuli	Unpleasant	Pleasant	Unpleasant	Pleasant	Unpleasant	Pleasant
Unpleasant	**76%**	26%	**78%**	22%	**78%**	22%
Pleasant	14%	**86%**	15%	**85%**	15%	**85%**

Table 10. Model Accuracy of fractal-approach

	SMV meanParameters		SVM bestParameters		SVM bestScaledParameters	
Classifier's prediction \ Stimuli	Unpleasant	Pleasant	Unpleasant	Pleasant	Unpleasant	Pleasant
Unpleasant	**89%**	11%	**87%**	13%	**87%**	13%
Pleasant	17%	**83%**	18%	**82%**	18%	**82%**

4.3 Two-sample proportion test

We proved if the differences between the obtained accuracy rates were statistically significance using the equations of the two-sample proportion test presented by Navidi [28]. Previous to perform the tests, we first calculated the global accuracy rate of each approach averaging all its individual accuracy rates of its SVMs (Table 11).

Table 11. Global accuracy rate of the approaches

Approach	Global accuracy rate
Spectral-fractal	87.83%
Spectral	81.33%
Fractal	85.00%

Next, we performed two tests, the first one compared the global accuracy rates of the spectral-fractal and spectral approaches, and the second one compared the global accuracy rates of the spectral-fractal and fractal approaches. Since in each approach we worked with 200 feature vectors, we fixed the population of two-sample tests in 200. Thus, we compared the spectral-fractal and spectral approaches stating the following hypothesis:

- Null hypothesis: $\pi_{spectral-fractal} \leq \pi_{spectral}$
- Alternative hypothesis: $\pi_{spectral-fractal} > \pi_{spectral}$

The p-value of this test was 0.0359. This indicates that exists a statistically significance of 0.04. Consequently, in the majority of the times (96%), the spectral-fractal approach predicts better the emotions than the spectral approach.

Likewise, we compared the spectral-fractal and fractal approaches stating the following hypothesis:

- Null hypothesis: $\pi_{spectral-fractal} \leq \pi_{fractal}$
- Alternative hypothesis: $\pi_{spectral-fractal} > \pi_{fractal}$

The p-value of this test was 0.2044. This indicates that exists a statistically significance of 0.21. Consequently, in a substantial number of times (79%), the spectral-fractal approach predicts better than the fractal approach.

Our results confirm the previous findings of Brodu [17]. Thus, we recognized emotions with a better accuracy rate by extracting both spectral and fractal features than extracting only one of them. It should, however, be noted that we proved this hypothesis using only our introduced procedure (section 3), and also that we detected only two kinds of emotions (positive and negative emotions). Therefore, future work should prove this hypothesis for different procedures of pattern recognition and for recognizing many kinds of emotions. Nevertheless, our results suggest that combining spectral and fractal features allow recognizing emotion in a more reliable manner.

5 DISCUSSIONS AND CONCLUSIONS

In recent years, many studies have been attempted to endow machines with the capacity of recognizing emotions through the analysis of electroencephalographic signals. Up to now, those studies have obtained effective results, even though there are still some aspects that can be improved such as the feature extraction methods. Generally, studies have been recognizing emotions by extracting spectral features to recognize emotions; however, Brodu [17] suggested that extracting both spectral and fractal features allows recognizing neural patterns in a more accurately manner.

In this study, we presented a comparison between an approach that extracts both spectral and fractal with two approaches that extract only one type of these features. To this end, we developed a method that can recognize emotions extracting spectral, fractal, or both features. Next, using this method, we built three different approaches to recognize positive and negative emotions that were induced with pictures from the International Affective Picture System on ten individual. The first approach extracted both spectral and fractal features, and the other two extracted each type of feature separately. Finally, we calculated and compared the accuracy rate of those approaches.

We found that the accuracy rate of an approach that combined spectral and fractal features was better than that of approaches that used only one of them. Thus, our results indicate, on one hand, that in the majority of the times (96%) the spectral-fractal approach predicts better the emotions than the spectral approach, and on the other hand, that in a substantial number of times (79%) the spectral-fractal approach predicts better than the fractal approach. Our results extend the findings of Brodu, showing that, as in neural-pattern recognition, the combination of spectral and fractal features allows increasing the accuracy rate of human-emotion recognition.

Therefore, our study suggest that it is possible to develop a more effective emotion recognition method by extracting both spectral and fractal features than extracting only one type of them. However, although our study statistically supported

this last fact, we only proved it using our introduced emotion recognition procedure. Nevertheless, our results are encouraging and should be validated. Hence, future work in emotion recognition should extract both spectral and fractal features to evaluate whether this combination of features allows properly recognizing human emotions in different contexts.

Finally, our results could be applied to develop brain-computer applications capable of detecting emotions in a suitable manner. This would improve the interaction between humans and machines because computers could detect and respond to the emotional state of a user, which makes the experience more pleasant for him/her.

6. ACKNOWLEDGEMENTS

This work was supported by the Colombian Administrative Department of Science, Technology and Innovation (COLCIENCIAS).

References:

[1] T. Bradberry, J. Greaves, *Emotional Intelligence 2.0.*, Publishers Group West, 2009.

[2] H. Bung, S. Furui, Automatic recognition and understanding of spoken languages: a first step toward natural human machine communication, in: *Procedings of the IEEE*, 2000, pp. 1142-1165.

[3] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votss, S. Kollias, W. Fellenz, J. Taylor, Emotion recognition in human-computer interaction, *IEEE Signal Processing Magazine*, Vol. 18, No. 1, 2001, pp. 32-80.

[4] T. Partala, M. Jokiniemi, V. Surakka, Pupillary responses to emotionally provocative stimul, in: *Proceedings of the 2000 Symposium on Eye tracking research & applications*, 2000, pp. 123-129.

[5] K. Takahashi, Remarks on emotions recognition from biopotential signals, in: *2nd International Conference on Autonomous Robots and Agents*, 2004, 186-191.

[6] P.C. Petrantonakis, 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, 2010, pp. 81-97.

[7] M. Murugappan, R. Nagarajan, S. Yaacob, Combining Spatial Filtering and Wavelet Transform for Classifying Human Emotions Using EEG Signals. *Journal of Medical and Biological Engineering*, Vol. 31, No. 1, 2010, pp. 45-51

[8] K.S. Schaaff, T. Schultz, Towards emotion recognition from electroencephalographic signals, in: *3rd IEEE International Conference on Affective Computing and Intelligent Interaction and Workshops*, 2009, pp. 1-6.

[9] M. Murugappan, R. Nagarajan, S. Yaacob, Classification of human emotion from EEG using discrete wavelet transform, *Journal Biomedical Science and Engineering*, Vol. 2, No. 4, 2010, pp. 390-396.

[10] P.C. Petrantonakis, L.J. Hadjileontiadis, Emotion recognition from EEG using higher order crossings, *IEEE Transactions on Information Technology in Biomedicine*, 2010, Vol. 14, No. 2, pp. 186-197.

[11] S.A. Hosseini, M.A. Khalilzadeh, M.B. Naghibi-Sistani, V. Niazmand, Higher order spectra analysis of EEG signals in emotional stress states, in: *Second IEEE International Conference on Information Technology and Computer Science*, 2010 pp. 60-63.

[12] L. Brown, B. Grundlehner, J. Penders, Towards wireless emotional valence detection from EEG, in: *IEEE Annual International Conference of the Engineering in Medicine and Biology Society*, 2011, pp. 2188-2191.

[13] P.C. Petrantonakis, L.J. Hadjileontiadis, A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition., *IEEE Transactions on Information Technology in Biomedicine*, 2011, Vol. 15, No. 5, pp 737-746.

[14] E.I. Konstantinidis, C.A. Frantzidis, C. Pappas, P.D. Bamidis, Real time emotion aware applications: A case study employing emotion evocative pictures and neuro-physiological sensing enhanced by Graphic Processor Units, *Computer methods and programs in biomedicine*, Vol. 107, No. 1, 2012, pp. 16-27.

[15] Y. Liu, O. Sourina, M.K. Nguyen, Real-time EEG-based Emotion Recognition and its Applications, in: M. Gavrilova and C.J.K. Tan (Eds.), *Transactions on Computational Science XII*, Springer-Verlag, Berlin, 2011, pp. 256-277.

[16] O, Sourina, Y. Liu, A Fractal-based Algorithm of Emotion Recognition from EEG using Arousal-Valence Model, in: *Proceedings of the International Confernece on Bio-Inspired Systems & Signals Processing*, 2011 pp. 209-214.

[17] N. Brodu, Multifractal Feature Vectors for Brain-Computer Interfaces, in: *IEEE International Joint Conference on Neural Networks*, 2008 pp. 2883-2890.

[18] P.J. Lang, M.M. Bradley, B.N. Cuthbert, International affective picture system (IAPS): Affective ratings of pictures and instruction manual, *Technical Report A-8*, University of Florida, Gainesville, FL, 2008.

[19] P.J. Lang, The emotion probe: Studies of motivation and attention, *Amercian Psychologist*, Vol. 50, No. 5, 1995, pp. 372–385.

[20] M.M Bradley, P.J. Lang, Emotion and motivation, in: J.T. Cacioppo, L.G. Tassinary, and G. Berntson (Eds.), *Handbook of Phychophysiology*, Cambridge University Press, New York, 2007, pp. 581-607.

[21] B. Kolb, I.Q. Whishaw, *Fundamentals of Human Neuropsychology*, Worth Publishers, 2008.

[22] F. Dolcos, R. Cabeza, Event-related potentials of emotional memory: encoding pleasant, unpleasant, and neutral pictures, *Cognitive, Affective & Behavioral Neuroscience*, Vol. 2, No. 3, 2002, pp. 252-263.

[23] Valderama C.E., Ulloa G, Spectral, Spectral Analysis of physiological parameters for emotion detection, in: *XVII Symposium of Image, Signal Processing, and Artificial Vision (STSIVA)*, 2013, pp. 275-280.

[24] V. Kullish, A. Sourin, O. Sourina, Human electroencephalograms seen as fractal time series: Mathematical analysis and visualization, *Computers in Biology and Medicine*, Vol. 36, No. 3, 2006, pp. 291-302.

[25] V. Kullish, A. Sourin, O. Sourina, Analysis and visualization of human electroencephalograms seen as fractal time series, *Journal of Mechanics in Medicine and Biology*, Vo. 6, No. 2, 2006, pp. 175-188.

[26] P. Ivanov, L. Amaral, A. Goldberger, S. Havlin, M. Rosenblum, Z. Struzik, H. Stanley, Multifractality in human heartbeat dynamics, *Nature*, Vo. 399, No. 6735, 1999, pp. 461-465.

[27] P. Manimaran, P.K. Panigrahi, J.C. Parikh, On Estimation of Hurst Scaling Exponent and Fractal Behavior through Discrete Wavelets, *ArXiv Physics/0604004 e-prints*, 2006.

[28] C.M Bishop, M.N. Nasser, *Pattern Recognition and Machine Learning*, Springer Science, 2006.

[29] Anguita, D., A. Ghio, S. Ridella, D. Sterpi, K-Fold Cross Validation for Error Rate Estimate in Support Vector Machine, in: *Proceedings of the 5th International Conference on Data Mining*, 2009, pp. 291-297.

[30] Navidi, W.C., Statistics for engineers and scientists, McGraw-Hill Higher Education, 2008.