Emotion Recognition System Based on Electroencephalography

Alma Areli Vicencio-Martínez
Instituto Politécnico Nacional
UPIITA
Mexico City, Mexico
alma.vicencio.m@gmail.com

Blanca Tovar-Corona
Instituto Politécnico Nacional
UPIITA
Mexico City, Mexico
bltovar@ipn.mx

Laura I. Garay-Jiménez
Instituto Politécnico Nacional
UPIITA
Mexico City, Mexico
0000-0001-9478-4835

***Abstract*— A system based on electroencephalography was designed and implemented to analyze the relation of the electroencephalography with the self-assessment manikin evaluation when sound stimuli that provoke emotions are applied. The analysis was made in volunteers between 18 and 40 years old, who were not deaf, to whom a protocol was applied that allowed them to evoke their emotions using sounds. The sounds used during the test were pre-sorted by each volunteer into four main groups, according to the emotional characteristics of activation and valence that correspond to the quadrats presented in the Russell model. A 7-channel electroencephalographic acquisition equipment was used to get the signals from positions T3, C3, F3, FP1, T4, C4 and F4 following the international system 10-20. Then a feature extraction was carried out on the signals for an emotion quadrant classification. Statistical difference was found in rations delta/beta of C3, C4, T3 and T4 positions. It was found that the best classification obtained was with neural networks that achieved an efficiency from 75% to 100%, depending on the participant, for the differences between F3 and F4 of the power spectral density, using the Yule Walker method, and the energy obtained with the Fourier Transform.**

***Keywords*— electroencephalography, emotions, valence, activation.**

I. INTRODUCTION

The emotion recognition has been done by facial features, voice tones and gestures, although these forms of recognition are efficient, they might be manipulated since the results could be altered at the user's will. For this reason, the most reliable methods for the detection of emotions are those based on imaging techniques such as positron emission tomography (PET) or functional magnetic resonance imaging (fMRI) [1]. However, these methods require equipment, infrastructure and specialized personnel, making their cost high. Some methods detect emotions through electroencephalographic (EEG) signals, being non-invasive, low cost and reliable due to advances in electronics.

The emotions are internal states of superior organisms that allow regulating the interactions with the environment, and they consist of a series of physical changes that occur in response to a given stimulus. [2].

On one hand, authors as Marosi et al. [3], Tortella et al.[4] and Ismali et al. [5], showed that there are changes in the delta and theta frequency bands of the electroencephalographic signals in the parietal, occipital and frontal zones associated to anger, sadness, happiness and surprise. Tortella found that emotional changes can be measured through the relations between the theta/beta and delta/beta bands, that shows changes in parietal and frontal regions for visual stimuli[4].

On the other hand, there is evidence that the two neurophysiological variables of the Russell model are orthogonal factors that can capture the entire spectrum of basic emotions, meaning that it can be represented in a two-dimensional plane as shown in Fig. 1. Valence goes from unpleasant to pleasant, meanwhile activation from activated to deactivated or calm condition, on a quantitative scale from 1 to 9 units. These attributes are useful since in this way emotions are not classified with isolated and discrete values but using a continuum in which each affective state is the result of a linear combination of the two independent factors mentioned above, that are interpreted as representing a particular emotion [1], [6] [7].

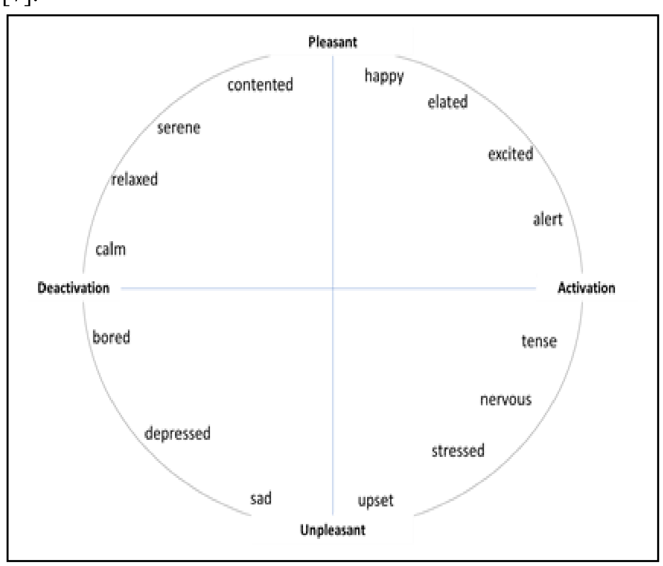

Fig. 1 Activation-Valence plane [6]

A useful tool to map the emotions in coordinates into the dimensional plane arousal-valence is the self-assessment manikin (SAM), which is a non-verbal graphical evaluation that allows directly measuring the pleasure, intensity and

978-1-7281-4840-3/19/$31.00 ©2019 IEEE

control associated with an affective reaction in people with a wide variety of stimuli.

In practice, users are asked to rate their emotions in the SAM evaluation by selecting an image for each line associated with a specific discrete level in each dimension [8], as shown in Fig. 2. The SAM was digitalized, and the participant was able to select any position into the range from 1 to 9. This work did not include the dominance in the analysis because the Russell model was followed, and there are several papers that report it as enough for the emotion classification [5-6].

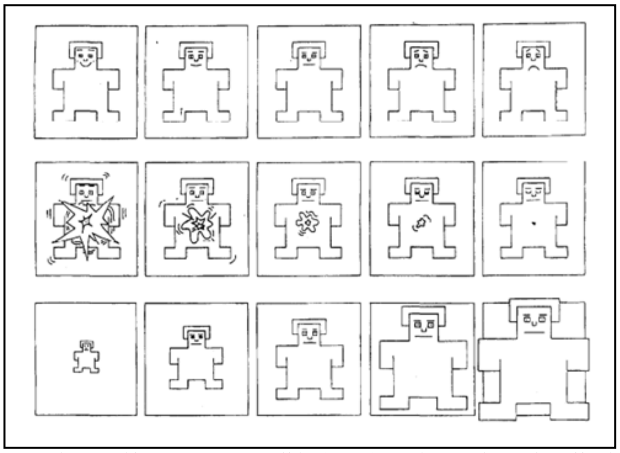

Fig. 2 Self-assessment manikin (SAM) used to evaluate the affective dimension of valence (top panel), activation (middle panel) and dominance (bottom panel) in paper format. [8].

There are several methods to evoke emotions mainly with visual or auditory stimuli, which can be images, videos or music. For this study, the International Affective Digital Sound (IADS) database was used as stimuli. This database has 167 sounds which were rated by at least 100 people of both genders, according to their valence and activation dimension using SAM [9].

This paper exposes the results of the analysis of the relation between the EEG parameters with the SAM results.

II. Materials and Methods

The process of emotion classification consists of several steps, as shown in Fig 3.

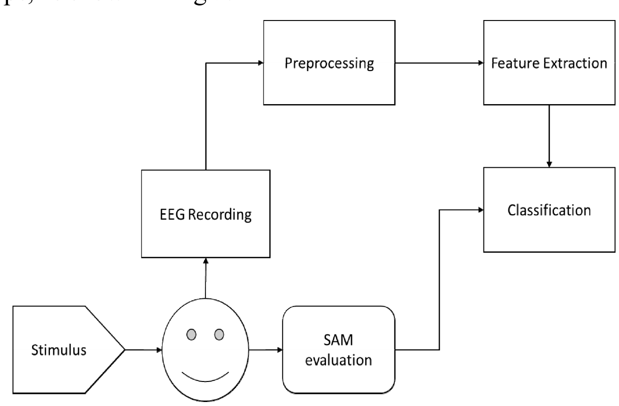

Fig. 3 Emotion classification process

A. Subjects

Twenty-five subjects participated in the study; they were between 18 and 40 years old, without gender distinction. Eight of them, who comply with the same characteristics of health, based on the SF-36 results, were selected for an EEG recording [11].

B. Stimulus

In order to elicit emotions, 40 sounds of six seconds duration were selected from the AIDS database. As shown in Table I, the ranges can be defined according to SAM.

TABLE I. SAM Ranges Definition

Range	1-4	5	6-9
Valence	Unpleasant	neutral	Pleasant
Activation	Calm	neutral	activated

Because of the stimuli set size, ten sounds of each quadrant were selected far away from the neutral state. The sounds selected are shown in Fig. 4.

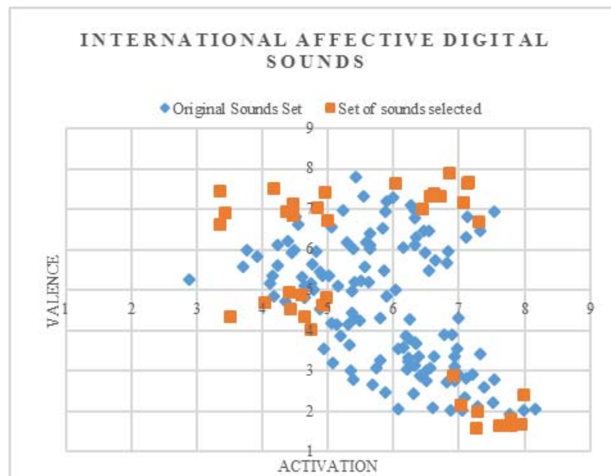

Fig. 4 Sounds selection

C. EEG Recording

A self-designed 7 channel equipment with an unipolar output was used to record the signals from the positions F3, T3, C3, FP1, T4, C4 and F4 with references in the mastoid, according to the standard 10-20. Commercial gold cup electrodes were used with Nuprep® gel to prepare the skin and Ten20 conductive gel to place them.

The equipment used was designed and implemented based on the design developed by A. Alberto Ramos [10], the system is composed by the steps shown in Fig. 5.

Each channel has a high-pass filter and a low-pass filter at 0.5 Hz and 50 Hz, respectively, using fourth order Butterworth type filters in Sallen-Key structure. The total gain is 110000. The operation voltage of each chanel is ±9V, and the energy consuming is in the range of 25.7 mA to 34.8 mA.

The system uses an acquisition board from National Instruments™ to digitalize the signals, 16 bits resolution, at 256 Hz sampling frequency and a program to record the signals designed in LabVIEW, that comprise a digital bandpass filter from 0.5 to 50Hz and a bandstop of 59 to 61 Hz, both of fourth order Butterwoth type.

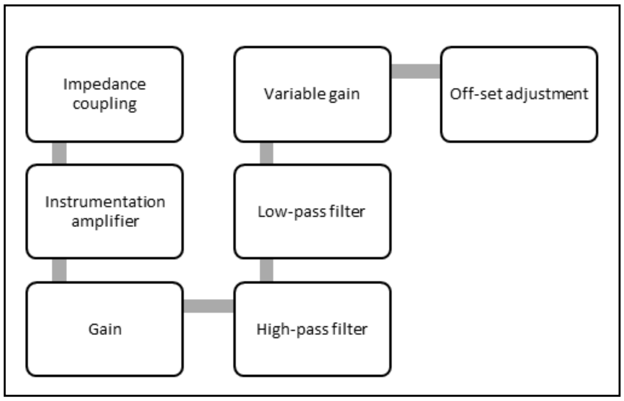

Fig. 5 Block diagram of the conditioning for the EEG signal

During a protocol to elicit the emotions, the EEG was recorded. First, the procedure was explained to the participant, and a consent letter was signed. Second, the participant was asked to answer the health survey SF-36, which allows measuring the perception of the participant about his health, if a threshold of 66.37 is exceeded, the participant data could be considered in the analysis. The use of this survey on the Mexican population was validated by Miguel A. Zuñiga [11]. Third, the placement of the electrodes was carried out. Fourth, the EEG signals recording initiates with a period of 60 seconds of silence to allow the participant being ready for the stimulus. After this period, the participant began to assess the set of sounds selected from the IADS after listening to each sound attentively. The sound was played continuously until the participant rates the sensation provoked using a digital SAM evaluation, and then a new one was randomly chosen, as it is shown in Fig 6. Finally, the recordings were saved for offline analysis.

Fig. 6 Answering digital self-assessment manikin questionnaire

D. Preprocessing and Feature Extraction

The electrical noise due to external sources is attenuated using fourth order bandpass IIR filters from 0.5Hz to 30Hz in a Butterworth configuration. This band was selected in order to limit the range to the EEG waves: δ, θ, α and β. Then the discrete wavelet transform (Db4) was used to decompose the signal up to level 5 in order to reduce the blinking artefact in the signal correspondent to FP1 position, eliminating the detail in level 5.

The signals were segmented according to the time required to answer the SAM questionnaire for each sound. This time differs in duration according to each participant performance because the next sound is played until each assessment was finished. In order to normalize the segment duration, two seconds of EEG signal was selected, centred for each sound.

Signal characteristics were extracted using two techniques: one non-parametric and one parametric. The first one was using a Hanning window, and the Fast Fourier transform (FFT). Then it was calculated its energy (En1), dominant frequency (Fn1) and the amplitude of the dominant frequency(An1) for the delta, theta, alpha and beta bands. The second proposal estimated the power spectral density with the Yule-Walker method (PSD), due to its minimum variation in low frequencies segments with short duration. The method was implemented with a resolution of 0.5 Hz using an order 20 as it was found the best compromise between resolution and accuracy due to the signal's characterisitics in frequency content and duration of the events due to the stimuli. It was used to extract characteristics such as dominant frequency (Fn2) and amplitudes of the dominant frequencies (An2) for the delta, theta, alpha and beta bands as [3] and [5] proposed using the FFT techique. The parametric method was proposed in order to campare its performance in the emotions recognition versus the FFT.

Ratios between the frequency bands delta/beta (δ/β) and theta/beta (θ/β) were calculated for each of the characteristics mentioned above, which are considered emotion monitors within visual stimuli [4].

Because of the number of parameters and looking for a reduction of dimensionality, a statistical analysis using Two Way ANOVA with post-test based on Tuckey test with $p=0.01$ was used as a reference to obtain the positions and characteristics with significant statistical differences. A post-test comparison was made using the Sidak´s test [12]. Each parameter for each channel was used as a seldom input of the classifiers, considering the subdivision in the 4 bands.

E. Classification

The classifiers selected were neural networks trained with the backpropagation algorithm and the hard C-means classifier [13][14]. The neural network consisted of two layers as it shows Fig. 7, one hidden layer with 7 neurons using log-sigmoid transfer function and an output layer with 4 neurons using lineal transfer function, one for each quadrant. It was trained with 70% of the data set, validated and tested with 15% and 15% respetively. The unsupervised method, a Hard C-means clustering was programmed to divide the set into four groups, that should correspond to each quadrant of the Russell model.

During the training, it was needed to expand the database, using machine learning methods [15], due to the displacement of the coordinate of some samples into another quadrant after the personalized SAM results. The expansion of the characteristics extracted from the EEG signals was done adding random noise no higher than the average standard deviation of the characteristic, this in order to balance the small data set. The imbalance compensation was applied to the parameter dataset previously to the classifier training.

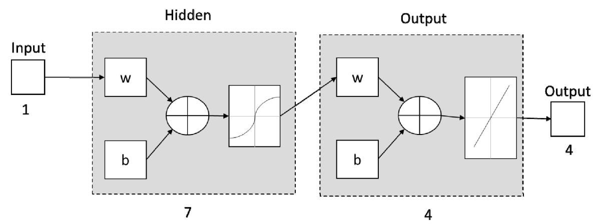

Fig. 7 Neural Network diagram

All preprocessing, feature extraction and classification algorithms were implemented in MATLAB® 2013a.

III. Results and Discussion

A. Stimulus

Even though the chosen sounds were carefully selected from the cases which were far away from the neutral condition (5,5) into the original distribution. It was observed that the third quadrant has less amount of sounds, and they are closer to the neutral condition. The original distributions of the chosen sound groups and the distribution after the tagged procedure, are observed in Fig.8.

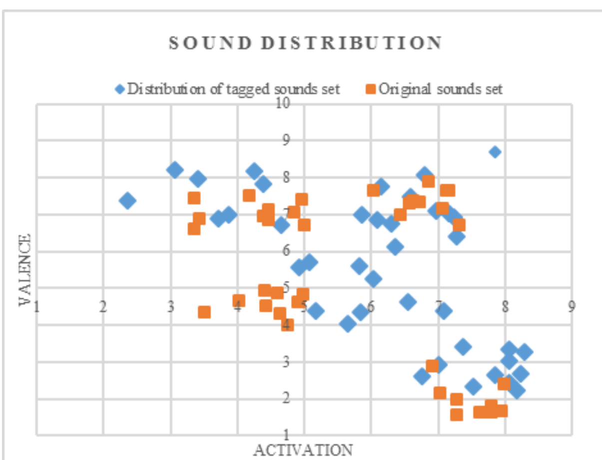

Fig. 8 Sound distribution.

Each selected sound set of the four quadrants had dispersion less than 0.647; however, when the 25 participants tagged them, the distribution of the sounds scatters towards the neutral condition, and the third quadrant almost disappeared. The mean of the cluster was moved towards the neutral condition. In a more in-depth analysis of each participant responses, it is observed that the third quadrant disappears, and the sound set moves to the fourth quadrant associated to unpleasant/activated or the first quadrant associated to the pleasant/activated condition in 63% of the cases. In Table II, the mean and standard deviation is summarized for each sound set.

B. Preprocessing and Feature Extraction

Considering that the method for the power spectral density computation (PSD), the energy representation is slightly different from the FFT method, both methods were computed looking for better performance. The parametric method estimates better with short length data than the Fourier method, but the last one is reported in several studies of emotion [3][5].

As it was mentioned before, the tendency of the data to cluster in fewer quadrants than the original assignment generates an imbalanced dataset. On the other hand, the number of features should be limited because of the curse of dimensionality, so these inconveniences were overcome using a preselection and an extension of the small datasets [16] [17].

The statistical analysis showed that there are electrode positions with significant statistical differences among pairs of quadrants summarized in Table III. The ratio δ/β (En) is calculated using the normalized energy, and the ratio δ/β (An) is computed with the normalized amplitude associated to the dominant frequency.

TABLE II. Statistical Information Of The Sounds distribution

Quadrant	Valence				Activation			
	Original set		*Tagged set*		*Original set*		*Tagged set*	
	μ	σ^2	x'	s^2	μ	σ^2	x'	s^2
1	7.376	0.358	7.700	1.068	6.787	0.391	7.263	1.133
2	1.947	0.429	3.145	1.482	7.549	0.395	7.861	1.118
3	4.581	0.289	4.338	0.935	4.517	0.457	4.389	1.001
4	7.054	0.311	7.667	1.065	4.243	0.647	3.151	1.338

TABLE III. Electrode Positions And Features With Significance Statistical Differences

Electrode position	Method	Characteristic	Quadrants
C4	PSD	δ/β (En)	4 vs 2
T4	PSD	δ/β (An)	4 vs 1
			4 vs 3
T3	PSD	δ/β (An)	4 vs 1
			4 vs 2
T3	FFT	δ/β (En)	4 vs 1
			4 vs 2
C3	FFT	δ/β (En)	3 vs 1
			3 vs 2
F3-F4	FFT	Delta Band En	1 vs 2
			2 vs 3
F3-F4	PSD	Delta Band An	4 vs 2

In the case of Tortella et al. study, the stimulus was visual, and it was studied the spontaneous emotion regulation [4]. They proposed the ratio δ/β (En), which is mainly reported into attention studies and found changes in parietal and frontal zones. This study was looking for the viability of these ratios with auditive stimuli, and the results show changes in the Delta/Beta ratio associated to temporal and central zones. The T3 and T4 electrodes were placed over the auditory associated zone as well as C3 in the principal motor zone [4]. However, the energy bands into the frontal zone presents no changes in these ratios.

Besides, in agreement with [3][5] and [18][19], an increase in the difference between F3-F4 energy or amplitude of dominant frequency was observed during the different sounds associated to each quadrant, even though they reported these changes with video stimuli and in constrast, for electroencephalogram monitoring associated to an attention test in the frontal zone reported by Putman et. al., the use of this ratios did not show difference in this spontaneous emotional response [20].

On the other hand, the results of SAM generated a tagged sound set for the eight participants, shown in Table IV.

As it is observed in Table II and IV, there is a difference in the IADS evaluation results, and the participant response, where a 62.5% of the sample shows a predominate unpleasant and activated condition, associated to stress, nervous and tense according to Russell model. 25% of the participants feel mainly calm or serene, and 12.5% alert to happy during the SAM test.

TABLE IV TAGS ACCORDING TO SAM RESULTS

Participant	*Q1*	*Q2*	*Q3*	*Q4*
1	**9**	**9**	**5**	**17**
2	**13**	**6**	**8**	**13**
3	**14**	**5**		**21**
4	**10**	**6**	**2**	**22**
5	**9**	**14**	**10**	**7**
6	**5**	**12**	**1**	**22**
7	**9**	**15**	**6**	**11**
8	**14**	**9**	**5**	**12**

C. Classification

The classifiers were trained and proved using each characteristic found in Table III and considering the imbalance problem observed in Table IV. Then, the final set after the ampliation is shown in Table V, The set with less than three elements was eliminated and just three quadrants considered. In the SAM responses with imbalance condition, in the case of large sets, it was randomly chosen the responses considered in the reduced set, and in contrast in the small sets, it was at least duplicated.

The difference between F3 and F4 was a constant in the studies reviewed, and it was also presented in the results of the group of participants considered for EEG recording (P1-P8), then it was used as input in the Neural Network classifier considering the SAM results as targets. Neural Network performance is presented in Table VI.

TABLE V. TAGS ACCORDING TO SAM RESULTS

Participant	*Q1*	*Q2*	*Q3*	*Q4*
P1	**18**	**18**	**10**	**17**
P2	**13**	**12**	**13**	**13**
P3	**10**	**10**		**11**
P4	**10**	**12**		**11**
P5	**9**	**10**	**10**	**10**
P6	**10**	**12**		**11**
P7	**9**	**15**	**12**	**11**
P8	**10**	**9**	**10**	**12**

For participants 3, 4 and 6, just three quadrants were considered, and the mean efficiency was from 75% to 100%, as shown in Table VI.

On the other hand, considering the clustering performance indicating the percentage of each clustered group (G) to be associated with the quadrant(Q) assigned by the participant, is not consistent, as it is shown in Table VII. In the hard C-means results, the quadrants were mainly reduced. It is observed that the quadrants cannot be identified properly with this method, it could be due to the amount of data, and because the sounds for the subjects analyzed tend to concentrate towards the centre.

TABLE VI. RESULTS OF THE NEURAL NETWORK CLASSIFIER.%EFFICIENCY

ELECTRODE POSITION	**METHOD**	**CHARACTERISTIC**	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**
F3-F4	FFT	En	88.88	62.5	100	100	83.33	80	85.71	71.42
F3-F4	FFT	An	77.7	62.5	80	100	66.66	80	85.71	71.42
F3-F4	PSD	An	77.77	100	80	100	83.33	100	71.42	85.71
MEAN EFICIENCY			81.45	75.00	86.67	100	77.77	86.67	80.95	76.18

TABLE VII. RESULTS OF THE HARD C-MEANS (PROPOSED/OBTAINED)

ELECTRODE POSITION	METHOD	CHARACTE-RISTIC	P1	P2	P3	P4	P5	P6	P7	P8
F3-F4	FFT	En	4/3	4/2	3/2	3/2	4/2	3/1	4/1	4/3
F3-F4	FFT	An	4/3	4/2	3/2	3/1	4/1	3/1	3/1	4/1
F3-F4	PSD	An	4/3	4/2	3/2	3/2	4/2	3/3	4/2	4/2

IV. CONCLUSIONS

In this preliminary study of emotion associated to sound stimuli, it is shown that the difference between the frontal electrodes is presented with sound stimuli. Besides, it was proved that the beta, theta and delta bands are involved in the process of emotions, as others reported it in video stimuli [3][4][18][19]. The ratios in T3 and C3 had not been reported as possible parametes in spontaneous emotion from sound stimuli and these results demonstrate that they could be used to detect the response associated to the sound emotional stimuli as well.

The best efficiency classification of the emotions associated to the quadrants of the Russell model varies according to each participant response in this study eventhough the statistical results show differences in delta/beta ratio into C3, C4, T3 and T4 for different emotion. For this preliminary study, the classification achieved 75% to 100% ranges using the neural network for the differences between F3 and F4.

It would be interesting to remake the studio with more sounds in order to get more data available for the training of the classifiers and achieve a better classification efficiency as well as propose a combination of them.

This EEG technique and proposed processing method could become an alternative to locating the emotional state of the person without being necessary to ask directly about the emotional state to the participant thus preventing subjectivity.

ACKNOWLEDGEMENT

Thanks to Center of Emotion and Attention (CSEA) at the University of Florida for allowing the use of their database of auditory stimuli.

Special thanks to Ana Georgina Guerrero for developing the digital platform for the SAM evaluation.

REFERENCES

[1] V. S. Ramachandran, "Encyclopedia of the Human Brain", San Diego, Elsevier Science, 2002, pp. 169-191.
[2] W. James, "What is an Emotion?", *Oxford Journals,* vol. 9, Abril 1884.
[3] E. Marosi Holczberger, D. M. B. Prieto Corona, M. G. Yáñez Téllez, M. A. Rodriguez Camacho, H. Rodrígez Camacho y V. Guerrero Juárez, "Quantitative spectral EEG assessments during affective states evoked by the presentation of the international affective pictures", *Journal of behaviour, heal social issues,* vol. 4, nº 2, Abril 2013.
[4] M. Tortella Feliu, A. Morillas Romero, B. M., J. Llabrés, X. Bornas y P. Putman, "Spontaneous EEG activity and spontaneous emotion regulation", International Journal of Psychophysiology, pp. 365-372, 2014.
[5] W. Ismail, M. Hanif, S. Mohamed, N. Hamah y Z. I. Rizman, "Human emotion detection via brain waves study by using electroencephalogram (EEG)", *International Journal on Advanced Science Engineering Information Technology,* vol. 6, nº 6, 2016.
[6] J. Posner, J. A. Russell y B. S. Peterson, "The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development and psychopathology", *Development and Psychopathology,* vol. 1, nº 17, pp. 715-734, 2005.
[7] A. J. Gerber, J. Posner, D. Gorman, T. Colibazzi, S. Yu, Z. Wang, A. Kangarlu, H. Zhu, J. Russell y B. S. Peterson, "An affective circumplex model of neural systems subserving valence, arousal, and cognitive overlay during the appraisal of emotional faces", *Neuropsychology,* nº 46, pp. 2129-2139, 2008.
[8] M. M. Bradley y P. J. Lang, "Measuring emotion: The selff-assessment manikin and the semantic differential", *Journal of Behavioral Therapy and Experimental Psychiatry,* vol. 25, nº 1, pp. 49-59, 1994.
[9] M. M. Bradley y L. P. J., The International Affective Digitized Sounds (2nd Edition; IADS-2): Affective Ratings of Sounds and Instruction Manual, Gainsville. Fl., 2007.
[10] A. Alberto Ramos, "Estudio de la señal electroenceflográfica bajo la intención de control para su uso en una interfaz cerebro computadora", Master Thesis, Instituto Politécnico Nacional - UPIITA, Ciudad de México, 2017.
[11] M. A. Zuñiga, C. J. G. T., P. J. Fos, B. Gandek y M. Medina Moreno, "Evaluación del estado de salud con la encuesta SF-36 : resultados preliminares en México", Salud pública de México, vol. 41, nº 2, 1999.
[12] Cardillo, G. (2006). Holm-Sidak t-test: a routine for multiple t-test comparisons. Available in: http://www. mathworks. de/matlabcentral/fileexchange/12786 (Consulted June 29th 2019).
[13] M. T. Hagan, Neural Network Design, 2nd edition. . ISBN-13: 978-0971732117
[14] T. J. Ross, Fuzzy logic with engineering applications, New Mexico: John Wiley & Sons, Ltd, 2004
[15] S. J. Raudys y A. K. Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners", IEEE Transactions on Pattern Analysis & Machine Intelligence, vol. 3, pp. 252-264., 1991.
[16] R. O. Duda, H. P. E. y S. D. G., Pattern Classification, Second ed., John Wiley & Sons Inc., 1973.
[17] M. A. Mazurowski, P. A. Habas, J. Zurada, J. Y. Lo, J. A. Baker and G. D. Tourassi, «Training neural networks clasifiers for medical decision making: The effects of imbalanced datasets on classification performance.,» Neural networks, vol. 21, nº 2-3, pp. 427-436, 2008
[18] A. Angelidis, W. van der Does, L. Schakel & P. Putman, "Frontal EEG theta/beta ratio as an electrophysiological marker for attentional control and its test-retest reliability", Biological psychology, vol.121, pp 49-52,2016.
[19] T.Valenzi, T. Islam, P. Jurica and A. Cichocki, A., "Individual S. classification of emotions using EEG", Journal of Biomedical Science and Engineering, vol.7, num. 8, pp 604,2014.
[20] P. Putman, B. Verkuil, E. Arias-Garcia, I. Pantazi, Ch. van Schie, "EEG theta/beta ratio as a potential biomarker for attentional control and resilience against deleterious effects of stress on attention", Cogn affect Behav Neurosci, 14(2), pp 782-791, 2016. DOI 10.3758/s13415-013-0238-7