



Music-Induced Emotion Classification from the Prefrontal Hemodynamics

Pallabi Samanta¹, Diptendu Bhattacharya², Amiyangshu De¹,
Lidia Ghosh¹  , and Amit Konar¹

¹ Electronics and Tele-communication Engineering, Jadavpur University,
Kolkata, India

pallabi.samanta18@gmail.com, amiyangshu_de@yahoo.com,
lidiaghosh.bits@gmail.com, konaramit@yahoo.co.in

² Computer Science and Engineering, National Institute of Technology Agartala,
Agartala, India
diptendu1@gmail.com

Abstract. Most of the traditional works on emotion recognition utilize manifestation of emotion in face, voice, gesture/posture and bio-potential signals of the subjects. However, these modalities of emotion recognition cannot totally justify its significance because of wide variations in these parameters due to habitat and culture. The paper aims at recognizing emotion of people directly from their brain response to infrared signal using music as the stimulus. A type-2 fuzzy classifier has been used to eliminate the effect of intra and inter-personal variations in the feature-space, extracted from the infrared response of the brain. A comparative analysis reveals that the proposed interval type-2 fuzzy classifier outperforms its competitors by classification accuracy as the metric.

Keywords: Emotion classification · Functional near-infrared spectroscopy · Interval type-2 fuzzy set classifier · Evolution algorithm

1 Introduction

In recent times, brain computer interface (BCI) has earned immense popularity for its inherent advantages in understanding the biological basis of cognitive processes, involving perception, memory, emotion and sensory-motor coordination. Emotion is regarded as the conscious experience concerning the pleasure, neutrality and displeasure levels of cognitive reaction of the brain in response to external stimulation.

Among different types of brain stimulation, music is one of the common modalities [4] of emotion arousal. Although no scientific justification on the role of music to emotion arousal is known till date, it is noticed that slow tempos, minor harmonies and a fixed pitch of a music is responsible of arousal of sadness [5]. Similarly, fast tempos, cacophonous and a wide range of pitch and dynamics have correlation with arousal of fear. Similar characterizations of stimuli for arousal of specific emotion has been ascertained by different research groups [6].

Most of the existing techniques attempt to classify emotion experienced by a subject based on the external manifestation of his/her emotion, such as change in facial expressions, voice qualities and physiological characteristics (like body temperature and skin conductance). A few well-known works that need special mention in this regard are listed below.

Das *et al.* [2] and Halder *et al.* [1] report novel techniques of emotion recognition using facial expression, voice and EEG analysis [9]. Wu *et al.* reported the change in brain wave synchronization during listening to music [3]. Furthermore, Mathur *et al.* [5] demonstrate the role of Indian classical music structure in emotional arousal. Banerjee *et al.* [6] studied the effect of music in human brain and body.

However understanding the cognitive underpinnings of emotional arousal needs more studies based on brain-imaging and cellular neuro-dynamics. We, here, employ functional Near Infra-Red Spectroscopy (fNIRS 1100) device to capture the pre-frontal brain response during emotion arousal with an aim to recognize the emotion from the brain response along with interpreting the involved brain regions [7]. The main focus of the present study is to model intra-personal and inter-personal uncertainty in the feature space of the fNIRS data obtained from the subjects, when the subjects experience musical stimuli. The uncertainty captured by the model is used later by an interval type-2 fuzzy set induced pattern classifier to recognize emotion. Besides, the results of emotion classification for given fNIRS features obtained from one or more voxels in the pre-frontal regions explains the engagement of the brain modules in the arousal process of emotion.

The paper is divided into four sections. Section 2 presents all the main steps, including normalization and pre-processing of the fNIRS signals, feature extraction, feature selection and the proposed Interval Type-2 fuzzy classifier for emotion recognition. Section 3 deals with experiments on feature selection and classifier performance. Conclusions are listed in Sect. 4.

2 Principles and Methodologies

This section reports all the required tools and techniques to resolve the proposed problem. The following steps are performed to classify hemodynamics responses associate with emotional arousal from music: (1) Scaling of the raw data, (2) Processing of the raw data and artifact removal, (3) Feature extraction from the oxyhemoglobin (HbO) and de-oxyhemoglobin (HbR) data obtained from fNIRS signals, (4) Feature selection based on evolutionary algorithm and (5) Classification of hemodynamic features using Interval Type-2 Fuzzy classifier.

2.1 Scaling of the Raw Data and Artifact Removal

The scaling of the raw data is performed using a max-min technique adopted from the protocol of De *et al.* [4]. Such transformation returns normalized HbO and HbR in $[0, 1]$.

Different physiological and environmental artifacts are removed by means of elliptical IIR low-pass filter with cut-off frequency of 0.5 Hz [7, 8].

2.2 Feature Extraction

We use an fNIRS system having four sources and ten detectors that forms (4 sources) × (4 detectors) = 16 channels. The temporal hemodynamics in 16 channels is represented by (i) HbO and (ii) HbR absorption curve.

In the present scheme we used music to induce emotions, and these measured HbO(t) and HbR(t) at $t = kT$, where T is the samples interval, and $k = 0, 1, 2, \dots$. Here we have taken the HbO(t) and HbR(t) responses for 90 s, which are divided into 6 time-windows of 15 s each. We select a sampling rate of 2 Hz, i.e., 2 samples/second. Thus for 15 s duration, we have $(15 \times 2) = 30$ samples denoted by HbO(t) and HbR(t) for $k = 0$ to 29; and $T = 0.5$ s.

Features: We have taken difference $d(t) = \text{HbO}(t) - \text{HbR}(t)$ per window and obtain the Static Features [7] Mean(m), Variance(v), Skewness(sk), Kurtosis(ku) and Average Energy(E) from the standard definitions. To obtain the dynamic behavior, we compute the change in m, var, sk, ku and E over the transitions between each consecutive window of 15 s in a time frame of 90 s. For 5 transitions of windows we have $(5 \times 5) = 25$ features. Coinciding Static and Dynamic features together, we have as many as $(30 + 25) = 55$ features. Thus for 16 voxels we have $(16 \times 55) = 880$ features.

2.3 Feature Selection

We adopt an evolutionary algorithm based feature selection [10] to reduce 880 features into 200 features. The algorithm used for feature selection attempts to maximize inter-class separation and minimize intra-class separation. The following two objectives are designed to attain the above requirements. It is given that $\vec{a}_i^x = \vec{a}_{i,1}^x, \dots, \vec{a}_{i,R}^x$ is the i -th feature vector having R numbers of components falling in class x , b_j^x and b_j^y denote the component of class centroids for class x and y respectively, the two objective functions are as

$$J_1 = \sum_{c=1}^P \sum_{i=1}^Q \sum_{j=1}^R |a_{i,j}^x - b_j^x| \quad (1) \quad J_2 = \sum_{y=1, y \neq c}^P \sum_{c=1}^P \sum_{j=1}^R |b_j^x - b_j^y| \quad (2)$$

where P represents the number of classes, Q the number of data points, and R is the number of features. We have attempt to optimize $J = J_1 - J_2$ using evolutionary algorithm.

2.4 Fuzzy Classification

For the present emotion recognition problem, we considered 30 subjects and have performed repeated experiments of 10 times per subject. Let f_i be a feature. Then its ten instances are $f_i^1, f_i^2, \dots, f_i^{10}$. We take the mean and variance of $f_i^1, f_i^2, \dots, f_i^{10}$ and denote them by m_i and σ_i . We adopt Gaussian membership function to represent the membership of feature f_i . Now, for 30 experimental subjects we have 30 such Gaussian membership functions (MF) for a given feature f_i . We take the maximum and minimum of these 30 MFs to obtain the Interval Type- 2 Fuzzy set (IT2FS) [1].

Classifier Rule:

Rule_i: If f_1 is closed to its center, f_2 is closed to its center,....., and f_n is closed to its center, then class = *emotion_i*.

To resolve the classification problem, we determine the degree of firing strength of all the classifier rules. The rule having the highest firing strength represents the true emotion class.

Firing strength Computation: Let f_1, f_2, \dots, f_n be the measurement points. We obtain

$$LFS_i = \text{Min}(\mu_{\tilde{A}_1}(f_1), \mu_{\tilde{A}_2}(f_2), \dots, \mu_{\tilde{A}_n}(f_n)) \tag{3}$$

$$UFS_i = \text{Min}(\bar{\mu}_{\tilde{A}_1}(f_1), \bar{\mu}_{\tilde{A}_2}(f_2), \dots, \bar{\mu}_{\tilde{A}_n}(f_n)) \tag{4}$$

where LFS_i and UFS_i denote the lower and upper firing strength of the i -th rule. We take the average of LFS_i and UFS_i to denote the average strength of firing of the rule i . We denote this by FS_i .

Thus for n classes, we need to fire n rules with the same measurements and determine their firing strength $FS_i, i = 1$ to n . Let the i -th rule have the different firing strength, i.e.; $FS_j > FS_k$ for all $k \neq i$. Then we declare the j -th emotion class as the solution for the classification task (Fig. 1).

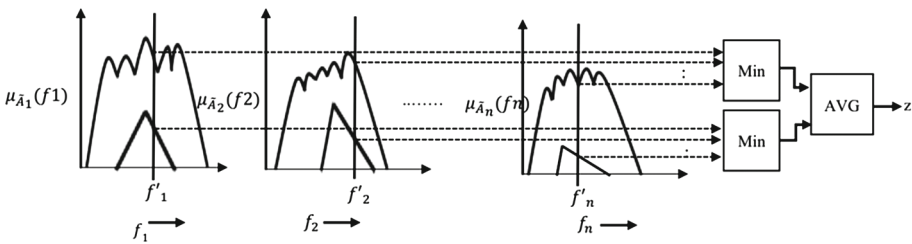


Fig. 1. Computing firing strength of an activated IT2FS induced rule

3 Experiments and Results

This section reports experimental protocols, the results from experimental instances and the inference derived from the experimental analysis.

3.1 Experimental Setup

30 right handed student volunteers, whose ages are between 20 to 27 took part in this experiment. Musical stimulus (Indian classical music) is presented by the head-phones mounted over the ears over a period of 90 s. Each participant undergoes 10 trials for each kind of music and a total set of six songs. 15 s interval, which generates maximum emotional depth, is considered as classification window. The hemodynamic data is recorded after removal of the base line.

3.2 Biological Inference of Hemodynamics in Emotion

Experimental analysis reports different emotion induced activations in prefrontal brain region. We adopt voxel plot approach from De *et al.* [11] (Fig. 3) using MATLAB 2015b to detect the spatial brain activation considering the mean HbO concentration during the specified 15 sec window of emotional activation. Here, we observe the least activation of the DorsoLateral Pre-frontal Cortex (DLPFC) [4] in happiness. The activation tends to rise in processing sadness and becomes highest in fear. Orbito-Frontal Cortex (OFC) shows a similar trend. The voxel plot for three emotions: (a) happiness, (b) sadness and (c) fear, is presented in Fig. 3 (due to lack of space the voxel plot of disgust is omitted). It also helps us to classify the different spatial pattern of brain activation due to emotional exposure from music.

3.3 Experiment 1: Feature Selection by Evolutionary Algorithm

Here, the feature dimension is reduced by using Evolutionary Algorithm. The best 200 features are selected among total 880 features. Figure 2 shows the discrimination of the selected 200 features by their relative amplitudes for each of the four classes: c_1 describing Happiness, c_2 describing Sadness, c_3 defining Fear and c_4 representing Disgust.

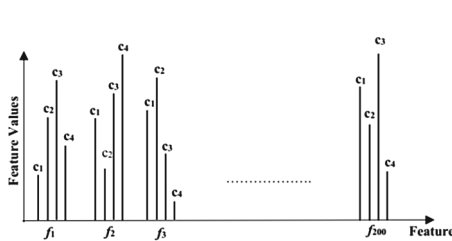


Fig. 2. Testing discrimination level of selected features

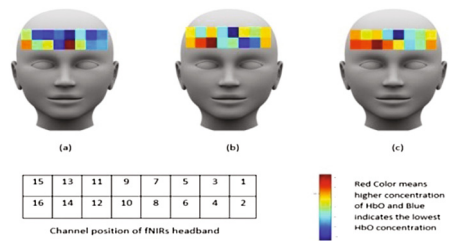


Fig. 3. Voxel plot of average HbO concentration

3.4 Experiment 2: Classifier Performance Analysis

To study the performance analysis, we compare the relative performance of the proposed IT2FS algorithm with three traditional emotion classification algorithms like Type-1 Fuzzy Classifier, multiclass Support Vector Machine (m-SVM) Classifier, Multi-layer Perceptron (MLP) algorithm. Table-I reveals that the final measure of the classification accuracy is the highest for the IT2FS method.

Table 1. Average Ranking of IT2FS over Three Traditional Classifiers according to their mean classifier accuracy

Emotion	IT2FS	Type-1 Fuzzy classifier	m-SVM	MLP
Happiness	83%	72%	67%	80.12%
Sadness	92%	80.67%	70%	82.12%
Fear	84%	75%	68.34%	81.45%
Disgust	78.89%	70%	65%	78.22%
Average ranking by Friedman test	1	3	4	2

3.5 Experiment 3: Statistical Comparison of Classifier Performance Using Friedman Test

To validate the importance of the work, we examine the performance of the classifier algorithms (IT2FS, Type-1 Fuzzy classifier, m-SVM and MLP) on four different databases using Friedman test. The Friedman Statistic used in this test has the standard definition as mentioned in [1]. The Friedman Statistic score is computed as $\chi_F^2 = 11.095$ with $N = 4$ and $k = 4$, which is greater than $\chi_{3,0.05}^2 = 7.815$. Thus, the null hypothesis is rejected at 3 degree of freedom with 78% accuracy suggesting that the classifier performance can be ranked according to their mean accuracy percentages. The average ranking of the classifiers computed using Friedman test is given in Table 1.

4 Conclusion

The paper introduces a novel approach to recognize music-induced emotion of subjects from their pre-frontal IR response. It is evident from the experimental results that the evolutionary algorithm based feature selection and the IT2FS induced classification approach together outperforms other conventional techniques. The IT2FS-induced classifier classifies the reduced 200 dimensional feature vectors into 4 emotion classes: happiness, sadness, fear and disgust with classification accuracies 83%, 92%, 84% and 78.89% respectively. The justification of IT2FS induced classifier is apparent due to its significance in intra-class and inter-class feature variation in the fNIRS signals. Experimental results further reveal that the DLPFC and OFC region of the brain is least activated during stimulation of happiness; the activation grows in sadness to disgust and with the highest response in fear.

References

1. Halder, A., Konar, A., Mandal, R., Chakraborty, A., Bhowmik, P., Pal, N.R., Nagar, A.K.: General and interval type-2 fuzzy face-space approach to emotion recognition. *IEEE Trans. Syst. Man Cybern. Syst.* **43**(3), 587–605 (2013)
2. Das, S., Halder, A., Bhowmik, P., Chakraborty, A., Konar, A., Janarthanan, R.: A support vector machine classifier of emotion from voice and facial expression data. In: *World Congress on Nature and Biologically Inspired Computing: (NaBIC 2009)* (2009)
3. Wu, J., Zhang, J., Ding, X., Li, R., Zhou, C.: The effects of music on brain functional networks: a network analysis. *Neuroscience* **250**, 49–59 (2013)
4. Schmidt, L.A., Trainor, L.J.: Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. *Cogn. Emot.* **15**(4), 487–500 (2001)
5. Mathur, A., Vijayakumar, S.H., Chakrabarti, B., Singh, N.C.: Emotional responses to Hindustani raga music: the role of musical structure. *Front. Psychol.* **30**, 1–11 (2015)
6. Banerjee, A., Sanyal, S., Sengupta, R., Ghosh, D.: Music and its effect on body, brain/mind: a study on indian perspective by neuro-physical approach. *iMedPub J.* **1**(1:2), 1–11 (2015)
7. Naseer, N., Hong, K.S.: fNIRS-based brain-computer interfaces: a review. *Front. Hum. Neurosci.* **28**, 1–15 (2015)
8. Herff, C., Heger, D., Fortmann, O., Hennrich, J., Putze, F., Schultz, T.: Mental workload during n-back task-quantified in the prefrontal cortex using fNIRS. *Front. Hum. Neurosci.* **7**, 1–9 (2013). Article no 935
9. Chakraborty, A., Konar, A., Halder, A., Kim, E.: Emotion control by audio-visual stimulus using fuzzy automata. In: *IEEE International Conference on Fuzzy Systems (FUZZ)*, pp. 1–8. IEEE, July 2010
10. Lahiri, R., Rakshit, P., Konar, A., Nagar, A.K.: Evolutionary approach for selection of optimal EEG electrode positions and features for classification of cognitive tasks. In: *IEEE Congress on Evolutionary Computation (CEC)*, pp. 4846–4853, July 2016
11. De, A., Konar, A., Samanta, A., Biswas, A., Ralescu, A.L., Nagar, A.K.: Cognitive load classification tasks from hemodynamic responses using type-2 fuzzy sets. In: *Fuzz IEEE* (2017)