Original Article

www.jmss.mui.ac.ir

The Emotion Recognition System Based on Autoregressive Model and Sequential Forward Feature Selection of Electroencephalogram Signals

Sepideh Hatamikia, Keivan Maghooli, Ali Motie Nasrabadi[1]

Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, [1]Faculty of Engineering, Shahed University, Tehran, Iran

Submission: 15-11-2013 Accepted: 14-04-2014

ABSTRACT

Electroencephalogram (EEG) is one of the useful biological signals to distinguish different brain diseases and mental states. In recent years, detecting different emotional states from biological signals has been merged more attention by researchers and several feature extraction methods and classifiers are suggested to recognize emotions from EEG signals. In this research, we introduce an emotion recognition system using autoregressive (AR) model, sequential forward feature selection (SFS) and K-nearest neighbor (KNN) classifier using EEG signals during emotional audio-visual inductions. The main purpose of this paper is to investigate the performance of AR features in the classification of emotional states. To achieve this goal, a distinguished AR method (Burg's method) based on Levinson-Durbin's recursive algorithm is used and AR coefficients are extracted as feature vectors. In the next step, two different feature selection methods based on SFS algorithm and Davies–Bouldin index are used in order to decrease the complexity of computing and redundancy of features; then, three different classifiers include KNN, quadratic discriminant analysis and linear discriminant analysis are used to discriminate two and three different classes of valence and arousal levels. The proposed method is evaluated with EEG signals of available database for emotion analysis using physiological signals, which are recorded from 32 participants during 40 1 min audio visual inductions. According to the results, AR features are efficient to recognize emotional states from EEG signals, and KNN performs better than two other classifiers in discriminating of both two and three valence/arousal classes. The results also show that SFS method improves accuracies by almost 10-15% as compared to Davies–Bouldin based feature selection. The best accuracies are %72.33 and %74.20 for two classes of valence and arousal and %61.10 and %65.16 for three classes, respectively.

Key words: *Autoregressive model, classification, Davies–Bouldin index, electroencephalography, emotional models, sequential forward feature selection*

INTRODUCTION

Considering the importance of emotions in managing daily life of an individual, the need for designing brain computer interface (BCI) systems which can explore brain signals and detect user's emotional states for people with disabilities is growing. In such systems, revealing some of emotions like fear and stress can play a vital role for these patients in dangerous situations. Emotions affect cognition, perception, memory, attention, and learning process. People need to express their emotions to communicate with others; also daily activities, are entirely influenced by the individual's feelings.[1,2] In order to evaluate electroencephalogram (EEG) signals correlated to emotions, different categorizations of emotional states have been proposed that two of them are more common: Discrete model, consists of six basic emotions suggested by Ekman *et al*. and two-dimensional continues model suggested by Russell.[3,4] In Ekman model, six discrete emotions acceptable for all nationalities and cultures, are include: Happiness, sadness, fear, anger, surprise, and disgust. In this model, emotions do not have any coherence to each other and belong to distinct emotional classes. In Russell continues model, two-dimensional model is defined for continuous representation of two axes: Valence and arousal. Valence axis ranges from pleasant to unpleasant, while arousal axis ranges from exciting to calm and all emotions are distributed in two-dimensional spaces corresponding to their degree of valence and arousal.

In neurophysiologic researches, biological signals has shown a relationship between emotions and physiological activities.[5,6] Several studies have been accomplished by

Address for correspondence:
Sepideh Hatamikia, Department of Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran.
E-mail: shatamikia11@gmail.com

approach of emotion recognition using different physiological signals such as heart rate, skin conductance, respiration rate, electrocardiogram, and electromyogram.[7-10] Furthermore, many systems have been proposed for emotion recognition using facial expressions and speech signals.[11,12] However, in recent years, most authors are focused on EEG signals for designing the emotion recognition systems. Murugappan *et al*.[13] proposed a method using statistic features from EEG frequency bands and wavelet transform to classify five discrete emotions (fear, disgust, and neutral, happy, and surprise), and they achieved average accuracy of %79.17. In another work of this author,[14] different discrete emotions were classified by energy-based wavelet features and the average accuracy of %83.26 was obtained. Chung and Yoon.[15] suggested emotion recognition system based on power spectral density (PSD) features; they classified two and three emotional classes of valence and arousal classes and achieved the average accuracy of 66.6% and 53.4% for two and three classes of valence levels, respectively. Soleymani *et al*.[16] suggested a system using EEG, pupillary response and gaze distance for classification of three classes of valence and arousal dimension based on PSD features. They obtained the best accuracies of 76.4 and 68.5 for two and three classes of valence and arousal classes.

Bastos Filho *et al*.[17] proposed an emotion recognition system based on three feature extraction methods from EEG signals to discriminate three emotional states: Stress, calm and normal state. Extracted features were PSD, statistical and high order crossing (HOC), and the best accuracy of %70.1 was achieved using PSD features. Furthermore, other feature extraction methods based on statistic, nonlinear and energy logarithmic features are introduced in different publications; the summary of some results and applied feature extraction methods of recently works are shown in Table 1.

As it has been demonstrated, several feature extraction methods are suggested by different researchers in order to classification emotional classes, but there have not been any studies completed using feature extraction based on AR model. The goal of this study is evaluating the performance of AR features in the classification of two and three emotional states over valence/arousal model. To perform this approach, different AR orders of Burg's method based on Levinson-Durbin's recursive algorithm were extracted and AR coefficients were used as features vectors; afterward, two feature selection methods based on sequential forward feature selection (SFS) algorithm and Davies–Bouldin index were used in order to decrease redundancy of features and computation time. Then, selected features were given to K-nearest neighbor (KNN), quadratic discriminant analysis (QDA) and linear discriminant analysis (LDA) classifier to classify different emotional classes. The procedure of suggested emotion recognition system is as shown in Figure 1.

The organization and structure of the study is defined as follows: Materials and methods section is consisting of the research methodology by describing data acquirement, preprocessing, AR feature extraction, SFS method and classification process; in the next section, the research results are presented, and conclusion of the study is represented in discussions and conclusions.

MATERIALS AND METHODS

Data Input

In this research, EEG signals of a publicly available dataset for emotion analysis using physiological signals (DEAP) is used which is collected by Koelstra *et al*.[28] This database includes EEG and peripheral signals from 32 subjects: 16 women and 16 men with the average age of 26.9. EEG signals were recorded according to 10-20 standard system from 32 position include: Fp1, AF3, F3, F7, FC5, FC1, C3, T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, C4, T8, CP6, CP2, P4, P8, PO4, and O2. During the experiments, EEG signals were down sampled (128 Hz) and filtered (between 4.0 and 45 Hz), also eye artifacts were eliminated by blind source separation technique. Koelstra *et al*.[28] provided the preprocessed data and the raw data for each subject. In this study, the preprocessed data were applied to evaluate the proposed method. In the experiments, 41 min music video inductions were represented for each participants and degree of valence and arousal was ranged by using self-assessment manikins (SAM) questionnaire.[29] SAM is a distinguished questionnaire that visualizes the degree of valence and arousal dimensions by manikins and participants should choice one number from 1 to 9 written below the manikins as shown in Figure 2.

In this study, emotional states are divided into two and three classes of valence and arousal dimensions according to the participant's SAM ratings. Valence and arousal level subdivisions are as shown in Table 2.[15]

Feature Extraction

Autoregressive model

The AR model has high ability in representing and modeling the characteristics and information inside a signal. AR model is frequently used in different approaches toward processing EEG signals such as: BCI designs,[30-32] classification of schizophrenic patients,[33] estimation of hypnosis levels,[34] determination of sleep stages,[35] analysis of anesthesia[36] and classification of epilepsy diagnosis.[37] In AR model, each sample is obtained from the summation of previous weighted samples according to Eq. 1. The model order is determined by the number of weights, which are called AR coefficients.

$$x(t) = -\sum_{i=1}^{p} a_i x(t-i) \tag{1}$$

Table 1: Some classification results of publications related to emotion recognition

Author	Reference	Feature extraction method	Number of emotional classes	Average accuracy %	Year
Khosrowabadi and Rahman	[18]	Fractal dimension include: Higuchi, Minkowski Bouligand, and fractional Brownian motion	4	81	2010
Nie *et al*.	[19]	Log band energy of FFT	2	89.22	2011
Takahashi and Tsukaguchi	[20]	Power at each frequency band and mean of raw signals	2	62.3	2003
Petrantonakis and Hadjileontiadis	[21]	HOC	6	62.58	2011
Schaaff and Schultz	[22]	Statistical features	3	47.11	2009
Mampusti *et al*.	[23]	Statistical features	4	54.09	2011
Khosrowabadi *et al*.	[24]	Mutual information and magnitude squared coherence	3	79	2010
Park *et al*.	[25]	Power of frequency bands, cerebral asymmetry, and coherences	3	66.3	2013
Kwon *et al*.	[26]	Power difference between left and right hemispheres in alpha and gamma band	2	64.78	2013
Hosseini and Khalilzadeh	[27]	Wavelet method, fractal dimension by Higuchi's algorithm and correlation dimension	2	79.20	2010
Murugappan *et al*.	[13]	Wavelet method	5	79.17	2009
Chung and Yoon	[15]	PSD	3	52.2	2012

FFT – Fast fourier transform; HOC – High order coefficients; PSD – Power spectral density

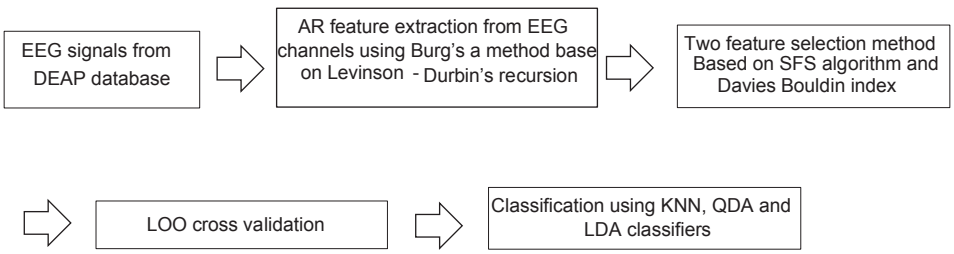

Figure 1: Block diagram of the experimental process

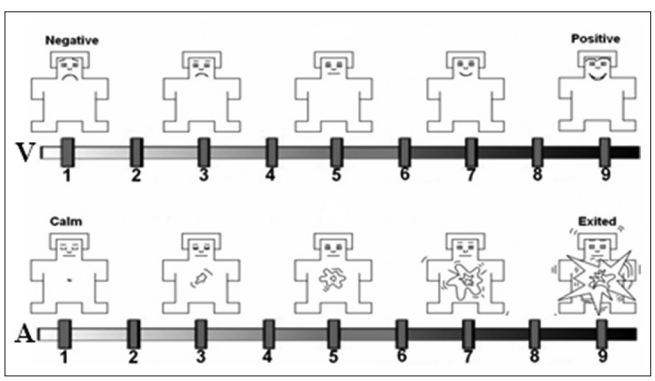

Figure 2: Self-assessment manikins scales of valence (above) and arousal (bellow)

Where, P is the model order and AR coefficients are denoted as $a_i\,(i=1,\ldots,p)$. In this paper, AR coefficients are obtained by applying Burg method.[38] In Burg method, AR reflection coefficients are estimated by minimizing the sum of forward and backward forecasted errors.

k_p is The pth reflection coefficient which is a criterion of the correlation between $x(t)$ and $x(t-p)$. By applying the Levinson–Durbin recursion algorithm, these reflection coefficients k_i can be converted, into AR parameters according to Eq. 2:

$$a_{p,i} = \begin{cases} a_{p-1,i} + k_p a^*_{p-1,p-i} & i = 1,\ldots,p-1 \\ k_p & i = p \end{cases} \tag{2}$$

In the pth level of Burg method, after estimating previous reflection coefficients $k_1, \ldots k_{p-1}$ through a recursive process, k_p reflection coefficient is determined.

At each level, reflection coefficient is calculated as bellow:

$$k_p = \frac{-2\sum_{t=p+1}^{N} e_{f,p-1}(t)\, e^*_{b,p-1}(t-1)}{\sum_{t=p+1}^{N}\left[\left|e_{f,p-1}(t)\right|^2 + \left|e^*_{b,p-1}(t-1)\right|^2\right]} \tag{3}$$

where, $e_{f,p-1}$ and $e_{b,p-1}$ are forward and backward forecasted errors for $(p-1)^{th}$ order of the model.[39] In the present work, AR coefficients from different orders of Burg's method based on Levinson–Durbin recursion algorithm were extracted as feature vectors, and the results of classification accuracies were compared.

Table 2: SAM ratings for affective class

Two classes		Three classes		
High	Low	High	Medium	Low
$S_r \geq 5$	$S_r < 5$	$S_r \geq 6$	$6 > S_r \geq 4$	$S_r < 4$

S_r is the rating of SAM questionnaire; SAM – Self-assessment manikins

Feature selection

In order to decrease complexity of computing and redundancy of features, different feature selection methods have been proposed. In this study, we used two approaches of feature subset selection: Scalar feature selection based on Davies–Bouldin index and vector feature selection based on SFS.

Feature selection based on the Davies–Bouldin index

In this case, feature selection is performed based on the values of Davies–Bouldin index.[40] The principle of this measure is based on two basis parts of data clustering: Minimizing inter-class distance (the distance among all data in a class), and maximizing intra-class distance (the distance between classes).

Mathematically, the Davies–Bouldin index is given as follow:

$$\mathrm{DBI} = \frac{1}{M}\sum_{\substack{i=1 \\ i\neq j}}^{M} \max\left[\frac{\mathrm{diam}(C_i)+\mathrm{dim}(C_j)}{d(i,j)}\right] \quad (4)$$

Where, $\mathrm{diam}(C_i)$ is the maximum distance between all pairs of samples in class i, d (i, j) is the distance between the center of class i and class j and *M* is the number of classes. Lower values of DBI index indicate less cluster overlap and thus higher class separation, while higher values show lower class discrimination.

In the experiment performed here, at first, Davies–Bouldin index for each feature was computed; then, features were ranked in descending order of criterion values. Finally, the features with the lowest ranking were selected.

Feature selection based on SFS method

SFS algorithm is one of the simplest feature subset selection methods. To achieve the best feature set, this algorithm is subsequently added to the first set of features which is initially empty. According to Figure 3, at first feature set A is considered empty and does not include any feature. Then, this algorithm seeks one of the features has the most influence in improving the fitness and adds the feature with the highest fitness x*; next, it seeks for the second feature that combination of it with the first selected feature results in the best. This procedure continues until adding a new feature does not increase the performance. Finally, A is considered the best feature set. Here, classification accuracy is considered as the fitness of a feature set.[41]

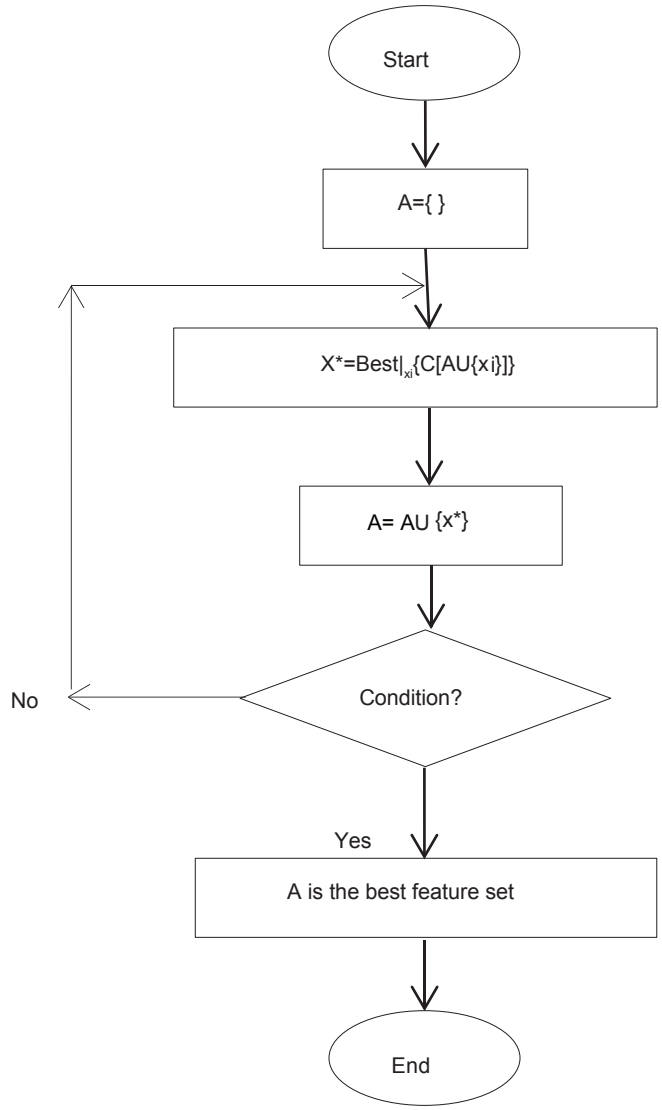

Figure 3: Procedure of feature selection using sequential forward feature selection method

Classification

K-nearest neighbor

K-nearest neighbor is a simple classifier that has been utilized in many pattern recognition applications. In this classifier, the class label of a new test sample is determined with respect to the labels of the nearest training samples. *k* closest training samples to a new test sample are determined and the label of a test sample is specified according the most repeatedly labels of these k closest samples. The number of the nearest neighbors (K) is required to be determined for the classification process. In this study, different K values were inspected and the *k* value with the best classification accuracy was selected.

Linear discriminant analysis

Linear discriminant analysis is one of the most distinguished classifiers in statistic, machine learning and pattern recognition. This classifier discovers a linear combination of

features that separates or determines two or more classes of events or objects. This classifier finds a one-dimensional subspace in which the classes are commonly well separated by a linear separating hyperplane. The discriminant function is defined as follow:

$$d_k(x) = 2\mu_k^T \sum_k^{-1} \mu_k - \mu_k^T \sum_k^{-1} \mu_k - 2\log\pi_k \quad (5)$$

Where, x is the set of measurements, k is the class of data, π_k is the prior probability and Σ_k is the covariance matrix.[42,43]

Quadratic discriminant analysis

Quadratic discriminant analysis classifier is a generalized version of LDA classifier. In this classifier, unlike LDA, covariance of each class is not considered identical. Further, the surface that separates the subspaces will be conical. The discriminant function is given by:[43]

$$d_k(x) = (X - \mu_k)^T \sum_k^{-1} (X - \mu_k) + \log|\Sigma| - 2\log\pi_k \quad (6)$$

And the discriminant rule is as follow:

$$d_k(x) = \min_{1 \le k \le K} d_k(x)$$

EXPRIMENTAL RESULTS

In this paper, 32 channels of EEG signals from 32 participants during watching emotional inductions were used to evaluate the proposed methodology. To ensure the assumption of stationary, EEG signal of each channel was divided into1 s windows and AR coefficients were extracted for each window. The classification results based on different orders of AR model, different feature selection methods and different classifiers were compared. Two different feature subset selection methods were used for decreasing the redundancy of features. In the first method, Davies–Bouldin index was computed for each feature and features with the smaller values were selected. Then, the performance of these selected features was evaluated using classification accuracies of different classifiers. In the second method, SFS algorithm was applied. In this method, the best subset of features with the best classification accuracy was selected. In this study, three classifiers include KNN, QDA and LDA were used to classify two and three classes of valence and arousal levels. All the data were divided into test and training set and the leave-one-out cross validation was used to validate the performance of classification results. In this cross validation method, feature vectors of one participant were used as the test data and the feature vectors of others were used for training the model. This process is repeated until all participants are used as the test data; finally, the average of all participants' classification accuracies was considered.

The classification results of two and three classes of valence and arousal using Davies–Bouldin and SFS feature selection methods through different orders of AR model and different classifiers are shown in Tables 3-6.

Comparison Between Classifiers

As we said earlier, three classifiers were used for evaluating the performance of proposed method. According to Tables 3-6, nearly similar results have been obtained using different AR model orders; but, the best classification accuracies

Table 3: Classification accuracy of two classes of valence and arousal using different orders of Burg's method and Davies-Bouldin based feature selection method

Classifier	KNN (%)		QDA (%)		LDA (%)	
Model order (p)	Valence	Arousal	Valence	Arousal	Valence	Arousal
6	56.33	57.97	55.64	56.95	51.64	53.98
7	56.61	58.34	54.65	56.25	48.18	53.75
8	57.28	58.67	55.96	56.01	52.45	54.74
9	58.66	58.23	58.26	58.91	55.23	54.47
10	56.33	59.22	57.78	57.42	50.68	55.98
11	57.12	55.64	52.29	54.37	51.19	55.48

KNN – K-nearest neighbor; QDA – Quadratic discriminant analysis; LDA – Linear discriminant analysis

Table 4: Classification accuracy of two classes of valence and arousal using different orders of Burg's method and SFS feature selection method

Classifier	KNN (%)		QDA (%)		LDA (%)	
Model order (p)	Valence	Arousal	Valence	Arousal	Valence	Arousal
6	71.17	68.61	68.39	64.37	60.27	63.81
7	70.13	71.72	67.64	65.26	61.82	62.31
8	67.97	70.86	63.12	65.78	59.47	63.12
9	67.66	69.53	67.04	67.83	60.38	59.64
10	72.33	74.20	70.35	69.26	63.22	65.54
11	69.02	72.03	65.54	65.32	62.19	59.03

KNN – K-nearest neighbor; QDA – Quadratic discriminant analysis; LDA – linear discriminant analysis; SFS – Sequential forward feature selection

Table 5: Classification accuracy of three classes of valence and arousal using different orders of Burg's method and Davies-Bouldin based feature selection method

Classifier	KNN (%)		QDA (%)		LDA (%)	
Model order (p)	Valence	Arousal	Valence	Arousal	Valence	Arousal
6	45.64	49.38	43.19	46.29	40.98	46.03
7	44.32	48.82	46.78	49.38	37.74	45.39
8	51.17	51.56	51.87	52.14	48.28	49.63
9	46.29	48.65	41.19	47.59	39.40	43.38
10	50.39	53.67	50.97	49.63	47.18	49.78
11	48.71	52.67	51.39	50.39	46.29	49.03

KNN – K-nearest neighbor; QDA – Quadratic discriminant analysis; LDA – Linear discriminant analysis

Table 6: Classification accuracy of three classes of valence and arusal using different orders of Burg's method and SFS feature selection method

Classifier	KNN (%)		QDA (%)		LDA(%)	
Model order (*p*)	Valence	Arousal	Valence	Arousal	Valence	Arousal
6	55.22	58.20	50.25	53.42	49.78	52.18
7	54.32	57.18	53.26	55.76	50.15	49.12
8	61.10	65.16	57.42	57.18	51.20	52.36
9	59.64	60.85	56.23	56.62	49.31	51.90
10	56.94	57.78	52.18	53.42	49.86	50.63
11	57.35	55.16	55.28	51.27	51.20	48.76

KNN – K-nearest neighbor; QDA – Quadratic discriminant analysis; LDA – Linear discriminant analysis; SFS – Sequential forward feature selection

for two valence and arousal classes using Davies–Bouldin feature selection were %58.66 and %59.22 for KNN classifier, %58.26 and %57.42 for QDA classifier and %55.23 and %55.98 for LDA classifier by model order $P = 9,10$, respectively. Furthermore, the best classification accuracies for two valence and arousal classes using SFS method were %72.33 and %74.20 for KNN classifier, %70.35 and 69.26 for QDA classifier and %63.22 and %65.54 for LDA classifier by model order $P = 10$. According to the results of Tables 5 and 6, the best classification accuracies of three classes of valence and arousal using Davies–Bouldin feature selection method were %51.17 and %53.67 for KNN classifier,%51.87 and %52.14 for QDA classifier and %48.28 and %49.63 for LDA classifier by model order $P = 8,10$. Furthermore, the best classification accuracies for three classes of valence and arousal using SFS method were %61.10 and %65.16 for KNN classifier, %57.42 and %57.18 for QDA classifier and %51.20 and %52.36 for LDA classifier by using model order $P = 8$. The results show that the best classification accuracies for both feature selection methods are obtained using KNN classifier; while, the lowest classification accuracies are belonged to LDA classifier.

Comparison between feature selection methods

In this research, we used two feature selection methods; scalar feature selection based on Davies–Bouldin index and vector feature selection based on sequential forward selection (SFS) algorithm. According to the results of Tables 3-6, the best accuracies for two classes of valence and arousal using Davies–Bouldin feature selection are %58.66 and %59.22, and for three classes are %51.17 and %53.67; while, the best accuracies using SFS method are %72.33 and %74.20, for two classes and %61.10 and %65.16 for three classes, respectively.

The results show that feature selection based on SFS method has improved the classification accuracies by almost 10% to 15% as compared to Davies–Bouldin based feature selection. Scalar feature selection based on Davies–Bouldin index has lower complexity and computing time than SFS method; but, the obtained results are not significant. It seems the reason is ignoring features correlations in procedure of feature subset selection; because scalar feature selection methods treat features individually and ignore the feature associations, while SFS method considers correlations between features in selecting the best feature subset.

DISCUSSIONS AND CONCLUSIONS

As we mentioned earlier, there have not been any studies in previous works completed on emotion classification using AR features. Because of comprehensive ability of AR model to discover the characteristics of the signals, we decided to evaluate the performance of this kind of features in recognizing of emotions. In this study, we examined two different feature selection methods based on Davies–Bouldin index and SFS algorithm, and the classification results of three different classifiers (KNN, QDA and LDA) were compared through both feature selection methods. In order to estimate AR parameters, several methods such as Yull–Walker, Burg, and Covariance methods are proposed. In this study, the Burg's method based on Levinson–Durbin recursion was used because of its higher ability of minimizing both forward and backward forecasted errors compared to other methods. Selecting the order of the model is an important issue to model the signal; hence, different model orders were examined and the classification results were compared. We used SFS and Davies–Bouldin based feature selection methods because of their low complexity and simple procedure to remove the redundant features. comparison of classification accuracies of Tables 3-6 show that SFS method perform better than feature selection based on Davies–Bouldin index; also KNN classification results are better than other classifiers.

According to our proposed system, AR coefficients are efficient in discrimination of two and three valence/arousal classes of emotions, and in case of two classes, the proposed technique shows the better classification accuracy than three classes. We evaluated our method with EEG signals of available database for emotion analysis (DEAP). Until now, limited articles have been published using DEAP database; Koelstra *et al*.[28] proposed a system based on power spectral features from EEG signals, Fisher criterion for feature selection and naïve Bayse classifier; They achieved the average accuracies of %57.6 and %62 for two classes of valence and arousal using this database. In another study using DEAP database, Yoon and Chung[44] designed an emotion recognition system based on Fast Fourier transform feature extraction, Pearson correlation coefficient for feature selection and Bayes classifier. They obtained the average accuracies of %70.9 and %70.1 for two classes of valence and arousal, and %55.4 and %55.2 for three classes. Bastos Filho *et al*.[17] proposed an emotion classification method to classify three emotional states: Stress, calm and normal using DEAP database. They used PSD, statistical and HOC features, and the best accuracy of %70.1 was achieved using PSD features. Chung and Yoon[15] proposed

an emotion recognition method using Bayes classifier based on a weighed-log-posterior probability function and power spectral features using this database and the best accuracies of 66.6% and 53.4% were obtained for two and three classes of valence dimension, respectively.

Compared with previous studies using DEAP database, our new proposed method has shown higher classification accuracy. Almost in all these studies, feature extraction method is based on power spectral features. Our research showed that AR features are efficient and have similar classification accuracies to power spectral features in distinguishing affective emotional states. our suggested method based on AR features, SFS method and KNN classifier has improved the classification accuracy rate in the classification of valence/arousal classes by almost %2–%4 as compared to the best reported classification accuracies using DEAP database; until now, the highest achieved accuracies using DEAP database for two classes of valence/arousal space are %70.9 and %70.1[44] whereas, the classification accuracies of our proposed method are %74.20 and %72.33. Furthermore, in comparison with other new studies with other databases,[16,21] our new proposed method has demonstrated higher classification accuracy with lower computational complexity. However, for the real situation, the classification accuracy must be improved higher. In the future, it is purposed to develop a system with higher classification accuracy and investigate another feature extraction, feature selection and classification methods to improve the performance and classification accuracy rate.

REFERENCES

1. Rosalind P. Affective computing: From laughter to IEEE. IEEE Trans Affect Comput 2010;1:11-7.
2. Picard R, Vyzas E, Healey J. Toward machine emotional intelligence: Analysis of affective physiological state. IEEE Trans Pattern Anal Mach Intell 2001;23:1175-91.
3. Ekman P, Friesen WV, O'Sullivan M, Chan A, Diacoyanni-Tarlatzis I, Heider K, *et al*. Universals and cultural differences in the judgments of facial expressions of emotion. J Pers Soc Psychol 1987;53:712-7.
4. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980;39:1161-78.
5. Wagner J, Kim J, Andre E. Implementing and comparing selected methods for feature extraction and classification. IEEE ICME International Conference, Amsterdam; 2005 July. p. 940-3.
6. Nasoz F, Lisetti CL, Alvarez K, Finkelstein N. Emotion recognition from physiological signals for user modeling of affect. 9th International Conference on User Model, Pittsburg, USA; 2003 June, 22-26.
7. Cacioppo JT, Tassinary LG. Inferring psychological significance from physiological signals. Am Psychol 1990;45:16-28.
8. Ekman P, Levenson RW, Freison WV. Autonomic nervous system activity distinguishes among emotions. J Exp Soc Psychol 1983;20:195-216.
9. Egon L, Broek V, Schutt M, Westerink J, Herk J, Tuinenbreijer K. Computing emotion awareness through facial electromyography. Human Computer Interaction (HCI/ECCV). Human-Computer Interaction – ECCV. Workshop on HCI, May 13, Graz, Austria. 2006. p. 52-63.
10. Kim K, Bang S, Kim S. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society EMBS/BMES Conference. Proceedings of the Second Joint. Vol. 1. 2002. p. 50-1.
11. Bung H, Furui S. Automatic recognition and understanding of spoken languages: A first step toward natural human machine communication. Proc IEEE 2000;88:1142-65.
12. Cowie R, Douglas Cowie E, Tsapatsoulis N, Votsis G, Kollias S, Fellenz W, *et al*. Emotion recognition in human-computer interaction. IEEE Signal Process Mag 2001;18:32-80.
13. Murugappan M, Nagarajan R, Yaacob S. Comparison of different wavelet features from EEG signals for classifying human emotions. IEEE Symposium on Industrial Electronics and Applications (ISIEA), Kuala Lumpur, Malaysia; 2009 Oct 4-6.
14. Murugappan M, Ramachandran N, Sazali Y. Classification of human emotion from EEG using discrete wavelet transform. J Biomed Sci Eng 2010;3:390-6.
15. Chung SY, Yoon HJ. Affective classification using Bayesian classifier and supervised learning. 12th International Conference on Control, Automation and System (ICCAS), Island, 2012 Oct 17-21, p. 1768-1771.
16. Soleymani M, Pantic M, Pun T. Multimodal emotion recognition in response to videos. IEEE Trans Affect Comput 2012;3:211-23.
17. Bastos Filho T, Ferreira A, Atencio A, Arjunan S, Kumar D. Evaluation of feature extraction techniques in emotional state recognition. IEEE Proceedings of 4th International Conference on Intelligent Human Computer Interaction, Kharagpur, India; 2012 December, 27-29.
18. Khosrowabadi R, Rahman AW. Classification of EEG correlates on emotion using features from Gaussian Mixtures of EEG Spectrogram. International Conference on Information and Communication Technology for the Muslim World (ICT4M), Jakarta; 2010 Dec, 13-14.
19. Nie D, Wang XW, Shi LC, Lu BL. EEG-based emotion recognition during watching movies. Proceedings of the 5th International IEEE EMBS Conference on Neural Engineering Cancun, Mexico, 2011 April-May. p. 667-70.
20. Takahashi K, Tsukaguchi A. Remarks on Emotion Recognition from Multi-Modal Bio-Potential Signals. IEEE Conferance on Industrial Technology; 2003 Oct 5-8. p. 1654-59.
21. Petrantonakis PC, Hadjileontiadis LJ. A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition. IEEE Trans Inf Technol Biomed 2011;15:737-46.
22. Schaaff K, Schultz T. Towards an EEG-based emotion recognizer for humanoid robots. The 18th IEEE International Symposium on Robot and Human Interactive Communication. Toyama, Japan, 2009 Sept-Oct. p. 792-6.
23. Mampusti ET, Ng JJ, Quinto JI, Teng GL, Suarez MT, Trogo RS, Measuring academic affective states of students via brainwave signals. Third International Conference on Knowledge and Systems Engineering (KSE); 2011 Oct. p. 226-31.
24. Khosrowabadi R, Heijnen M, Wahab A, Quek HC. The dynamic emotion recognition system based on functional connectivity of brain regions. IEEE Intelligent Vehicles Symposium University of California, San Diego, CA, USA; 2010 June 21-24.
25. Park MS, Oh HS, Jeong H, Sohn JH. EEG-based emotion recognition during emotionally evocative films. International Winter Workshop on Brain-Computer Interface (BCI), Gangwo; 2013 Feb, 18-20.
26. Kwon M, Kang JS, Lee M. Emotion classification in movie clips based on 3D fuzzy GIST and EEG signal analysis. International Winter Workshop on Brain-Computer Interface (BCI), Gangwo; 2013 Feb, 18-20.
27. Hosseini SA, Khalilzadeh MA. The IEEE International Conference on Biomedical Engineering and Computer Science (ICBECS), Wuhan, China; 2010 April, 1-6.
28. Koelstra S, Muhl C, Soleymani M, Lee JS, Yazdani A, Ebrahimi T, *et al*. IEEE transactions on affective computing. DEAP: A database for emotion analysis. Using Physiol Signals 2012;3:18-31.

29. Bradley MM, Lang PJ. Measuring emotion: The self-assessment manikin and the semantic differential. J Behav Ther Exp Psychiatry 1994;25:49-59.
30. Faradji F, Ward RK, Birch GE. A brain-computer interface based on mental tasks with a zero false activation rate. Proceedings of the 4th International IEEE EMBS Conference on Neural Engineering Antalya, Turkey; 2009 April-May. p. 355-88.
31. Huan NJ, Palaniappan R. Classification of mental tasks using fixed and adaptive autoregressive models of EEG signals, Proceedings of the 2 International IEEE EMBS Conference on Neural Engineering Arlington, Virginia; 2005 March, 16-19.
32. Huan NJ, Palaniappan R. Neural network classification of autoregressive features from electroencephalogram signals for brain-computer interface design. J Neural Eng 2004;1:142-50.
33. Sabeti M, Katebi SD, Boostani R, Price GW. A new approach for EEG signal classification of schizophrenic and control participants. Expert Syst Appl 2011;38:2063-71.
34. Elahi Z, Boostani R, Motie Nasrabadi A. Estimation of hypnosis susceptibility based on electroencephalogram signal features. Sci Iran 2013;20:730-7.
35. Ubeyli ED, Cvetkovic D, Holland G, Cosic I. Analysis of sleep EEG activity during hypopnoea episodes by least squares support vector machine employing AR coefficients. Expert Syst Appl 2010;37:4463-7.
36. Ni Z, Wang L, Meng J, Qiu F, Huang J. EEG signal processing in anesthesia feature extraction of time and frequency parameters. Procedia Environ Sci 2011;8:215-20.
37. Han M, Sun L. EEG signal classification for epilepsy diagnosis based on AR model and RVM. International Conference on Intelligent Control and Information Processing, Dalian, China; 2010 August, 13-15.
38. Burg JP. A new analysis technique for time series data. NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater Acoustics. Netherlands, 1968 August 12-23.
39. Stoica P, Moses RL. Introduction to Spectral Analysis. USA: Prentice Hall; 1997.
40. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell 1979;1:224-7.
41. Theodoridis S, Koutroumbas K. Pattern Recognition. Amsterdam, Elsevier; 2003.
42. Herman P, Prasad G, McGinnity TM, Coyle D. Comparative analysis of spectral approaches to feature extraction for EEG-based motor imagery classification. IEEE Trans Neural Syst Rehabil Eng 2008;16:317-26.
43. Bhattacharyya S, Khasnobish A, Chatterjee S, Konar A, Tibarewala DN. Performance analysis of LDA, QDA and KNN algorithms in left-right limb movement classification from EEG data. Proceedings of International Conference on Systems in Medicine and Biology, Kharagpur, India; 2010 December, 16-18.
44. Yoon HJ, Chung SY. EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm. Comput Biol Med 2013;43:2230-7.

How to cite this article: Hatamikia S, Maghooli K, Nasrabadi AM. The Emotion Recognition System Based on Autoregressive Model and Sequential Forward Feature Selection of Electroencephalogram Signals. J Med Sign Sence 2014;4:194-201.

Source of Support: Nil, **Conflict of Interest:** None declared

BIOGRAPHIES

Sepideh Hatamikia was born in 1989. She received the B.Sc. degree in biomedical engineering from Islamic Azad University, Dezful branch and the M.Sc. degree in Biomedical Engineering, Science and Research Branch, Islamic Azad University, Tehran, Iran in 2011 and 2013, respectively. Her research interests include biomedical signal processing, pattern recognition, analysis of brain signals and EEG-based brain-computer interface.

E-mail: shatamikia11@gmail.com

Keivan Maghooli has received his B.Sc. in electronic engineering from the Shahid Beheshti University, Tehran, Iran, M.Sc. in biomedical engineering from the Tarbiat Modaress University, Tehran, Iran, and Ph.D. in biomedical engineering from the Research and Science branch, Azad University, Tehran, Iran, majoring in Signal Processing and Artificial Intelligence. He has been with the Biomedical Faculty at Research and Science branch, Azad University, Tehran, Iran, since 2000, where he is currently an Assistant Professor and Head of Bioelectric Department.

E-mail: k_maghooli@srbiau.ac.ir

Ali Motie Nasrabadi received a BS degree in Electronic Engineering in 1994 and his MS and PhD degrees in Biomedical Engineering in 1999 and 2004 respectively, from Amirkabir University of Technology, Tehran, Iran. Since 2005, he has joined to Shahed university and now he is an associate professor in the Biomedical Engineering Department at Shahed University, in Tehran, Iran. His current research interests are in the field of Biomedical Signal Processing, Nonlinear Time Series Analysis and Evolutionary Algorithms. Particular applications include: EEG Signal Processing in Mental Task Activities, Hypnosis, BCI and Epileptic Seizure Prediction.

E-mail: nasrabadi@shahed.ac.ir