A Design of Memory-based Learning Classifier usign Genteic Strategy for Emotion Classification

Momory-based Learing Classifier for Emotion Classification

Byoung-Jun Park, Eun-Hye Jang,
Sang-Hyeob Kim, Chul Huh
IT Convergence Technology Research Laboratory
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
{bj_park, cleta4u, shk1028, chuh}@etri.re.kr

Myung-Ae Chung
IT Convergence Services Future Technology Research
Electronics and Telecommunications Research Institute
Daejeon, Republic of Korea
machung@etri.re.kr

Abstract*—In this study, we discuss emotion classification for seven kinds of emotion (happiness, sadness, anger, fear, disgust, surprise, stress) in the psycho-physiological research. Seven emotions are evoked by stimulus formed on audio-visual film clips, and then physiological signals of autonomic nervous system responses are measured for the reaction of stimulation. Additionally, seven different emotions will be classified by the proposed classification methodology using physiological signals. We introduce a classification methodology on memory-based learning that dwells upon the usage of genetic strategy (Genetic Algorithms). Genetic algorithms (GAs) take selection problems of instances and features of memory into two level optimization processes. In the first level, GAs chooses *P* % of instances as a set of memory comes from instances with *c* classes. In the second level of the optimization process, GAs is instrumental in the formation of a core set of features that is a collection of the most meaningful and discriminative components of the original feature space. In classification problems, it becomes important to carefully select instances and establish a subset of features in order to achieve a sound performance of a classifier. The study offers a complete algorithmic framework and demonstrates the effectiveness of the approach for the classification of seven emotions. Numerical experiments show that a suitable selection of instances and a substantial reduction of the feature space could be accomplished and the classifier formed in this manner is characterized by high classification accuracy for the seven emotions based on physiological signals.

Keywords-memory-based learning; emotion classification; pysiological signals; genetic algorithms

I. INTRODUCTION

Recently, the most popular research in the field of emotion recognition on human-computer interaction is to recognize human's feeling using various physiological signals. In the psycho-physiological research, it is known that there is strong correlation between human emotion state and physiological reaction. Emotion plays an important role in contextual understanding of messages from others in speech or visual forms. For affective communication between user and computer, it has to consider how emotions can be recognized and expressed during human-computer interaction and emotion recognition is one of the key steps towards emotional intelligence in advanced human-machine interaction. Psychologists and engineers have tried to analyze facial expressions, vocal emotions, gestures, and physiological signals in an attempt to understand and categorize emotions. Many previous studies on emotion have reported that there is correlation between basic emotions such happiness, sadness, anger, fear, etc. and physiological responses [1]-[3]. Recently, emotion recognition using physiological signals has been performed by various machine learning algorithms (e.g., Fisher's Linear Discriminant, k-Nearest Neighbor algorithm, Neural Networks, and Support Vector Machine , etc.[4]-[8]).

The objective of this study is to achieve dataset on multi-physiological signals for seven emotions induced by an emotional stimulus and to develop memory-based learning classifier driven by genetic algorithms. Firstly, in order to get physiological signals for emotions, we use 10 different emotional stimuli set to induce seven emotions, i.e., happiness, sadness, anger, fear, disgust, surprise and stress under the same conditions. We identify emotion-specific physiological responses induced by these emotional stimuli. To induce each emotion, ten emotional stimuli sets which have been tested their suitability and effectiveness, are used in experiment. Physiological signals, namely, Skin temperature (SKT), photoplethysmography (PPG), electrodermal activity (EDA) and electrocardiogram (ECG) are acquired by MP100 Biopac system Inc. (USA) and analyzed to extract features for emotional pattern dataset.

For emotion classification, we use one of the techniques of evolutionary optimization, namely, genetic algorithms (GAs) [9]-[10]. In order to improve classification speed and accuracy of a classifier, suitable formations of a set of instance and features are required. GAs embraces two optimization processes, that is, choosing *P* % of instances as a set of memory comes from patterns with seven classes (emotions) and forming a core set of features that is a collection of the most meaningful and discriminative components of the original feature space. Numerical experiments are carried out and it is shown that a suitable selection of prototypes and a substantial reduction of the feature space could be accomplished that is also accompanied with a higher classification accuracy.

 ISBN: 978-1-61208-340-7

The study is organized into five sections. We outline the measurements of physiological signal induced by seven emotion stimuli in Section II. Section III covers all necessary development issues by the optimization scheme. Numeric experimental studies are presented in Section IV while Section V offers some concluding comments.

II. STIMULI AND PHYSIOLOGICAL SIGNALS FOR SEVEN EMOTIONS

In this study six males (20.8 years±1.26) and six females (21.2 years±2.70) students participated. None of the subjects reported any history of medical illness or psychotropic medication and any medication that would affect the cardiovascular, respiratory, or central nervous system. A written consent was obtained before the beginning

A. *Emotional stimuli*

As shown in Fig. 1, seventy emotional stimuli (7 emotions x 10 set) are used to successfully induce emotions. Emotional stimuli are selected the 2-4 min long audio-visual film clips which are captured originally from movies and TV. Audio-visual film clips have widely used because these have the desirable properties of being readily standardized, involving no deception, and being dynamic rather than static [11]-[14].

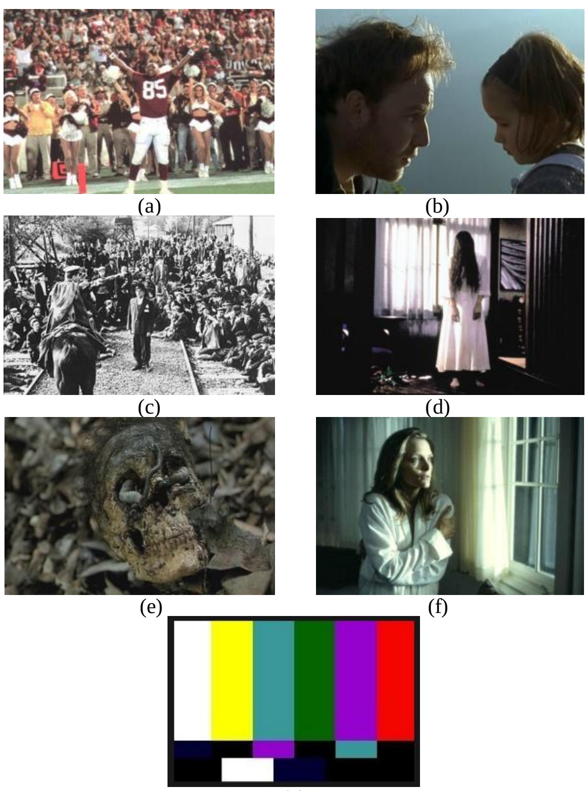


Figure 1. Examples of emotional stimuli; (a) Happiness: victory, wedding, laughing, etc., (b) Sadness: death of parents/lover, separation, longing for mother, etc., (c) Anger : massacre, beating, attack, etc., (d) Fear : ghost, haunted house, scare, etc., (e) Disgust: body in pieces, vomiting, etc., (f) Surprise: sudden or unexpected scream etc., and (g) Stress : audio/visual noise on screen, etc.

TABLE I. SUITABILITY AND EFFECTIVENESS OF EMOTIONAL STIMULI

Emotion / set	Happiness	Sadness	Anger	Fear	Disgust	Surprise	Stress
1	100% (8.4)	92% (9.5)	75% (9.7)	75% (10)	75% (10.2)	75% (9.3)	92% (9.3)
2	100% (8.9)	100% (9.1)	75% (9.9)	100% (9.9)	92% (10.8)	92% (9.7)	100% (9.1)
3	100% (8.8)	100% (8.7)	75% (9.7)	83% (9.8)	92% (9.9)	100% (9.7)	100% (8.8)
4	100% (9.6)	100% (9.7)	75% (9.5)	92% (9.6)	100% (10.4)	100% (9.9)	100% (8.9)
5	100% (9.6)	100 (9.3)	92% (9.8)	92% (9.7)	92% (9.7)	83% (9.6)	100% (9.3)
6	100% (9.3)	100% (9.3)	92% (9.4)	92% (9.7)	100% (10.3)	83% (9.6)	100% (8.8)
7	100% (9.3)	75% (8.9)	92% (8.9)	83% (9.6)	100% (9.3)	100% (9.5)	92% (9.3)
8	92% (8.0)	100% (9.0)	83% (9.2)	100% (9.3)	83% (10.2)	83% (9.4)	100% (9.3)
9	100% (9.7)	100% (9.2)	92% (9.5)	100% (9.3)	100% (10.1)	83% (8.6)	100% (9.1)
10	92% (8.8)	100% (9.3)	92% (9.7)	75% (8.7)	100% (10.1)	75% (10.3)	100% (9.3)

The used audio-visual film clips are examined their suitability and effectiveness by preliminary study. After being presented each film clip, twenty-two college students rate the category and intensity of their experienced emotion on emotional assessment scale. The result shows that emotional stimuli have the suitability of 93% and the effectiveness of 9.5 point on average. The suitability of each stimulus is ranged from 75 to 100% and from 8.4 to 10.4 point in the effectiveness as shown in Table I. The suitability of emotional stimuli means the consistency between the target emotions designed to induce each emotion and the categories of participants' experienced emotion. The effectiveness is determined by the intensity of emotions reported and rated by the subjects on a 1 to 11 point Likert-type scale (e.g., 1 being "least happy" or "not happy" and 11 being "most happy").

B. *Physiological signals*

The dataset of physiological signals such as SKT, EDA, PPG, and ECG are collected by MP100 Biopac system. SKT is an important and effective indicator of emotion states and reflects autonomic nervous system activity. Variations in the SKT mainly come from localized changes in blood flow caused by vascular resistance or arterial blood pressure. EDA is a method of measuring the electrical conductance of the skin, which varies depending on the moisture of the skin, caused by sweat. PPG is a signal derived from light absorption changes in pulse oximeters in contact with the skin. The change in the light absorption reflects a change in the blood flow rate. Thus, the PPG signal is subordinate to pulse pressure, i.e., the difference between systolic and diastolic pressure in the arteries. Lastly, ECG is a transthoracic (across the thorax or chest) interpretation of the electrical activity of the heart over a period of time, as detected by electrodes attached to the surface of the skin and recorded by a device external to the body.

 ISBN: 978-1-61208-340-7

TABLE II. TABLE TYPE STYLES

Physiological signals			Features
EDA			SCL, NSCR, meanSCR
SKT			meanSKT, maxSKT
PPG			meanPPG
ECG	Time domain	Statistical parameter	meanRRI, stdRRI, meanHR, RMSSD, NN50, pNN50
		Geometric parameter	SD1, SD2, CSI, CVI, RRtri, TINN
	Frequency domain	FFT	FFTapLF, FFTapHF, FFTnLF, FFTnHF, FFTLF/HFratio
		AR	ARapLF, ARapHF, ARnLF, ARnHF, ARLF/HFratio

SKT electrodes are attached on the first joint of non-dominant ring finger and on the first joint of the non-dominant thumb for PPG. EDA is measured with the use of 8 mm AgCl electrodes placed on the volar surface of the distal phalanges of the index and middle fingers of the non-dominant hand. Electrodes are filled with a 0.05 molar isotonic NaCl paste to provide a continuous connection between the electrodes and the skin. ECG electrodes are placed on both wrists and one left ankle with two kinds of electrodes, sputtered and AgCl ones. The left-ankle electrode is used as a reference.

The signals are acquired for 1 minute long baseline state prior to presentation of emotional stimuli and 2-4 minutes long emotional states during presentation of the stimuli. The obtained signals are analyzed for 30 seconds from the baseline and the emotional state. The emotional states are determined by the result of participant's self-report.

C. Feature extraction

For seven emotions, 28 features extracted from the physiological signals and used to analysis are shown in Table II. Skin conductance level (SCL), average of skin conductance response (meanSCR) and number of skin conductance response are obtained from EDA. The mean (meanSKT) and maximum skin temperature (maxSKT) and the mean amplitude of blood volume changes (meanPPG) are gotten from SKT and PPG, respectively. ECG is analyzed in the view point of time domain and frequency domain. Analysis in time domain is divided into the statistical and the geometric approaches and in frequency domain is dealt with FFT and AR.

III. DESIGN OF MEMORY-BASED LEARINIG CLASSIFIER

For the classification of seven emotions, the proposed classifier is a type of memory-based learning that uses only specific instances to solve classification problem. Namely, the classifier is a method for classifying objects based on the closest training samples called memorized instances in the feature space. This classifier embraces two selection problems to classify a new pattern to a class. One is the selection of instance to be memorized and another is feature selection.

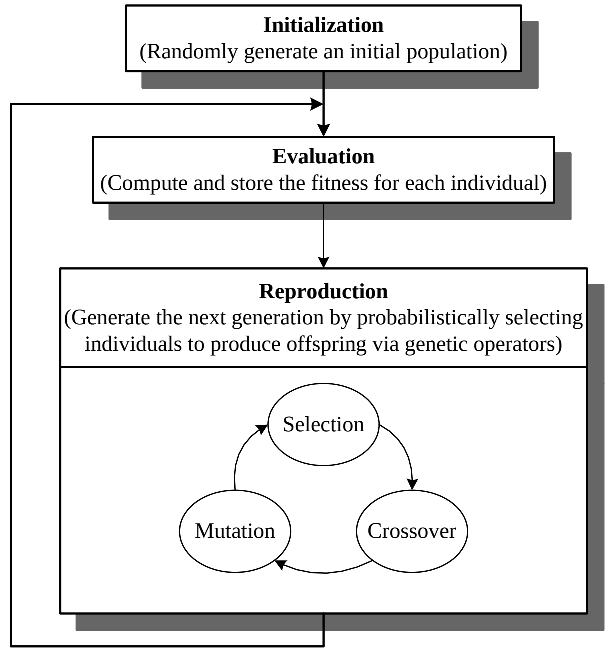


Figure 2. General Flowchart of Genetic Algorithms

To format instance to be memorized, we start with choose *P* % of patterns using GAs. The classifier generates classification predictions using only P % of patterns. The classifier does not use any model to fit and only is based on distance between a pattern and memorized instance. Given a set of N instance, the classifier finds the one instance closest in feature space to an unknown pattern, and then assigns the unknown pattern to the class label of its nearest instance. The underlying distance between a pattern and an instance is measured by weighted Euclidean one,

$$\|\mathbf{x}-\mathbf{y}\|^2 = \sum_{i=1}^{n} \frac{(x_i - y_i)^2}{\sigma_i^2} \qquad (1)$$

where **x** and **y** are the two patterns in the *n*-dimensional space and σ_i is the standard deviation of the *i*-th feature whose value is computed using the instance set.

Secondly, once the instances to be memorized have been formed, we reduce feature space by choosing a core set of features encountered in the problem. Those features are regarded as the most essential ensemble of features that, organized together, exhibit the highest discriminatory capabilities. We use GAs to choose *d* % of features which minimizes the classification error.

GAs have proven to be useful in optimization of such problems because of their ability to efficiently use historical information to obtain new solutions with enhanced performance and a global nature of search supported there [9][10]. GAs are also theoretically and empirically proven to support robust search in complex search spaces. The search of the solution space is completed with the aid of several genetic operators. There are three basic genetic operators used in any GAs - supported search that is reproduction,

 ISBN: 978-1-61208-340-7

crossover, and mutation. Reproduction is a process in which the mating pool for the next generation is chosen. Individual strings are copied into the mating pool according to their fitness function values. Crossover usually proceeds in two steps. First, members from the mating pool are mated at random. Second, each pair of strings undergoes crossover as follows: a position *l* along the string is selected uniformly at random from the interval [1, *l*-1], where *l* is the length of the string. Two new strings are created by swapping all characters between the positions *k* and *l*. Mutation is a random alteration of the value of a string position. In a binary coding, mutation means changing a zero to a one or vice versa. Mutation occurs with small probability. Those three operators, combined with the proper definition of the fitness function, constitute the main body of the genetic computing. A general flowchart is visualized in Fig. 2.

As a generic search strategy, the GAs has to be adjusted to solve a given optimization problem. There are two fundamental components that deserve our attention: a fitness function, and a representational form of the search space. Given the wrapper mode of instance formation and feature selection, we consider the minimization of the classification error as a suitable fitness measure. There could be different choices of the search space considering that an optimal collection of features could be represented in several different ways. Here, we adopt the representation scheme of the search space in the form of the *n*-dimensional unit hypercube (*N* is the number of patterns while n denotes the number of features of emotional dataset). The content of a chromosome is ranked viz. each value in this vector is associated with an index the given value assumes in the ordered sequence of all values encountered in the vector (The elements of a chromosome for prototypes and features are ranked separately.). Considering that we are concerned with *d* % of all features, we pick up the first $d \times n$ ($0<d<1$) entries of the vector of the search space. This produces a collection of features forming the reduced feature space. This mechanism of the formation of the feature space is portrayed in Fig. 3. For the entire patterns, the prototype formation is carried out in the same manner as we encountered in the feature selection.

Given the instance formation and the feature selection, we consider the minimization of the classification error to be a suitable fitness measure of GAs.

$$Fitness = 1 / \left(1 + \frac{\text{Number of misclassified patterns}}{\text{Number of patterns}} \right) \qquad (2)$$

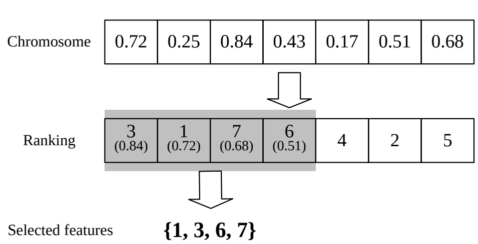


Figure 3. GAs formation of the reduced feature space; here, n=7 while the assumed reduction of the space results in 4 features viz. {1, 3, 6, 7}.

There are two stopping conditions for seven emotion classification: (a) the algorithm terminates if the objective function does not improve during the last 100 generations, otherwise (b) it terminates after 500 iterations. The size of the population is related to the dimensionality of the search space.

IV. Results of Emotion Classification

The numerical studies presented here provide some experimental evidence behind the effectiveness of the GAs approach. The detailed setup of an extensive suite of experiments is reflective of the methodology we outlined in the previous sections. The two essential parameters that we use in the assessment of the performance of prototype and feature selection are the percentage of features (denoted by "*d*") forming the core of the reduced feature space and the percentage of the data forming the instance set (*P*) optimized by the GAs. The results are reported for the testing data sets for various values of "*P*" and "*d*".

In this paper, for the optimization of the classification of seven emotions, GAs uses the serial method of real type, roulette-wheel in the selection operator, one-point crossover in the crossover operator, and uniform in the mutation operator. More specifically, we use the following values of the parameters: maximum number of generations is 500; the number of populations is 100, crossover rate is 0.75, and mutation probability is set to 0.1.

For the seven emotions dataset given as multi-physiological signals, the relationship between the percentage of features used in the GAs optimization, values of "*P*" and the resulting classification accuracy is presented in Table III. Here, "No. of F" is the number of selected features for *d* % of entire features, "AVG" and "STD" indicate average and standard deviation, respectively. The classification accuracy was computed over 10-fold realization of the experiments, namely, for each combination of the values of the parameters (*d* and *P*), the experiments was repeated 10 times by running GAs. Fig. 4 shows the result of fitness function of GAs for *d*=30 and *P*=70. There is an evident effect of the "optimal" subsets of the instance and the feature space: clearly the subsets leads to better classification results when compared with the outcomes of the classifier operating on the entire feature space.

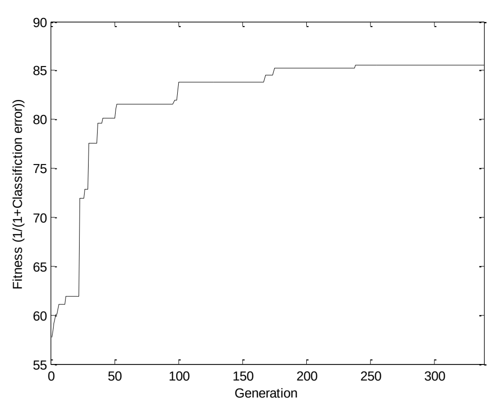
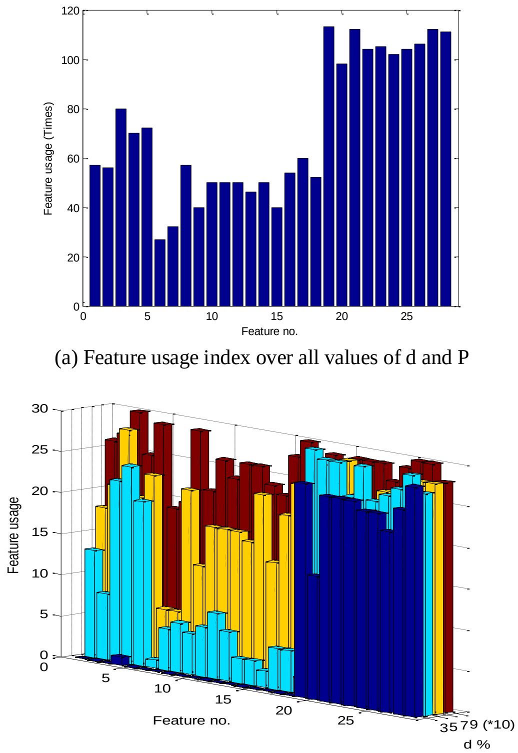


Figure 4. Fitness function of GAs for d=30 and P=70



(a) Feature usage index over all values of d and P

(b) Feature usage index over all P for each d

Figure 5. Cumulative number of occurrence of individual features

TABLE III. CLASSIFICATION ACCURACY FOR SEVEN EMOTIONS

d % (No. of F)	*P* % 30	*P* % 50	*P* % 70	AVG±STD over *P*
30(8)	44.4 ± 1.96	64.0 ± 0.92	83.0 ± 2.26	63.8 ± 16.15
50(14)	27.7 ± 2.05	36.9 ± 5.56	38.6 ± 3.22	34.4 ± 6.17
70(20)	24.3 ± 0.83	27.6 ± 0.69	31.9 ± 1.64	27.9 ± 3.32
90 (28)	24.5 ± 0.84	26.6 ± 0.95	30.2 ± 1.27	27.1 ± 2.60
100(28)	24.6 ± 0.85	26.4 ± 0.95	31.1 ± 1.31	27.4 ± 2.98

TABLE IV. COMPARISON OF THE CLASSIFICATION ACCURACY OF THE PROPOSED METHOD AND OTHER METHODS

Method		Accuracy (%)	Features
CART		21.7 ± 3.6	28
C4.5		16.1 ± 1.2	28
kNN		45.3 ± 2.3	28
FLD		20.3 ± 1.8	28
NN		18.0 ± 1.0	28
PNN		16.3 ± 1.9	28
RBFs		17.4 ± 1.2	28
SOM	Supervised	16.3 ± 1.9	28
	Unsupervised	15.4 ± 3.2	28
SVM		17.8 ± 3.1	28
Proposed Methodology		83.0 ± 2.26	8

With the increasing values of "*d*", the classification accuracy of seven emotions decreases substantially; in the case of *P*=30% it drops from 44.4 to 24.3 when increasing the number of features from 10% to 70%. Conclusively, the use of all features dropped accuracy of classification for seven emotions. Changes in the values of "*P*" have far less effect on the classification rate, however, the distinguished result was occurred in *d*=30 and *P*=70. From these results, we observe that the number of suitable features is 8 and 70 % of dataset are required as the prototype for the seven emotion recognition using multi-physiological signals. We report the number of occurrences of the features in Fig. 5. This indicator becomes more illustrative and offers an interesting view at the suitability of the features when forming various reduced feature spaces and using different prototype set sizes. In case of *d*=30 % for the classification of seven emotions, namely, the number of features is 8, we have gotten that classification accuracy is 83 % for P=70.

For the classification of seven emotions, Table 4 contrasts the classification accuracy (%) of the proposed method with other well-known methods studied in the literatures [5]-[8]. These machine learning algorithms are standard algorithms, which have been applied to various and lots of fields. For more information, refer to [5]-[8].

As abovementioned, the experiments are reported for the 10 times using a split of data into 70% training and 30% testing subsets, namely, 70% of the whole patterns are selected randomly for training of all methods and the remaining patterns are used for testing purposes. The results are averaged over 10 times for testing dataset. The experimental results reveal that the proposed approach and the resulting model outperform the existing methods both in terms of the simpler structure and better prediction (generalization) capabilities on feature space reduced 70% of entire feature space.

While the experimental results provide sound evidence behind the selection process showing that the reduced feature spaces led to the better classification results than those obtained in some previous studies, they are also quite revealing in showing that the reduction of the feature space could exhibit different effectiveness. In some cases, the reduction of the dimensionality of the feature space could be high but there could be cases where the elimination of subsets of features could not be strongly justifiable

V. CONCLUSION

In this study, we have discussed the acquisition of multi-physiology signals using emotion stimuli and the design of a classification methodology for seven emotions. The emotion stimuli used to induce a participant's emotion were evaluated for their suitability and effectiveness. The result showed that emotional stimuli have the suitability of 93% and the effectiveness of 9.5 point on average. Twenty eight features have been extracted by means of the statistical and the geometric approaches in time and frequency domain from physiological signals such as EDA, SKT, PPG and ECG. These signals have been induced by emotional stimuli. In order to improve the classification accuracy using physiological signals for seven emotions, we have

 ISBN: 978-1-61208-340-7

introduced a memory-based learning classifier using the evolutionally mechanism for the seven emotions expressed by physiological signals. The optimization process of forming the instance and the feature space is reflective of the conjecture on the importance of forming a set of instance and a core set of features whose discriminatory capabilities emerge through their co-occurrence in these set. The methodology of feature selection becomes legitimate considering that we immediately see the result of the reduction of the feature space being translated into the corresponding classification rate. The use of the instance is also justifiable considering that this classification scheme is the simplest that could be envisioned in pattern classification. The proposed classifier will lead to better chance to recognize human emotions by using physiological signals in the emotional interaction between man and machine.

ACKNOWLEDGMENT

This research was supported by the Converging Research Center Program through the Ministry of Science, ICT and Future Planning, Korea (2013K000329).

REFERENCES

[1] R. W. Picard, E. Vyzas, and J. Healey, "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State," IEEE Transactions Pattern Analysis and Machine Intelligence, vol. 23, no.10, 2001, pp.1175-1191.

[2] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion Recognition Using Bio-Sensors: First Step Towards an Automatic System," Affective Dialogue Systems, Tutorial and Research Workshop, 2004.

[3] F. Nasoz, K. Alvarez, C. L. Lisetti, and N. Finkelstein, "Emotion Recognition from Physiological Signals for Presence Technologies," International Journal of Cognition, Technology and Work, Special Issue on Presence, vol. 6, no. 1, 2003, pp. 4-14.

[4] M. Murugappan, N. Ramachandran, and Y. Sazali, "Classification of human emotion from EEG using discrete wavelet transform," Journal of Biomedical Science and Engineering, vol. 3, 2010, pp. 390-396.

[5] J. R. Quinlan, C4.5 Programs for Machine Learning, San Mateo, CA: Morgan Kaufmann, 1992.

[6] R. O. Duda, P. E. Hart, and D. G. Stork, Pattern classification, 2nd ed. Wiley-Interscience, New York, 2000.

[7] P. D. Wasserman, Advanced Methods in Neural Computing, New York, Van Nostrand Reinhold, 1993, pp. 35–55.

[8] S. Parsa and S. A. Naree, "A New Semantic Kernel Function for Online Anomaly Detection of Software," ETRI Journal, vol.3 , no. 2, 2012, pp. 288-291.

[9] D. E. Goldberg, Genetic Algorithms in search, Optimization & Machine Learning, Addison-Wesley, 1989.

[10] S. A. Taghanaki, M. R. Ansari, B. Z. Dehkordi, and S. A. Mousav, "Nonlinear Feature Transformation and Genetic Feature Selection: Improving System Security Decreasing Computational Cost," ETRI Journal, vol.34, no.6, 2012, pp 847-857.

[11] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," Cognition and Emotion, vol. 9, pp. 87-108, 1995.

[12] R. S. Lazarus, J. C. Speisman, A. M. Mordkoff, and L. A. Davidson, "A Laboratory study of psychological stress produced by an emotion picture film," Psychological Monographs, vol. 76, 1962, pp. 553.

[13] M. H. Davis, J. G. Hull, R. D. Young, and G. G. Warren, "Emotional reactions to dramatic film stimuli: the influence of cognitive and emotional empathy," Journal of personality and social psychology, vol. 52, 1987, pp. 126-133.

[14] D. Palomba, M. Sarlo, A. Angrilli, A. Mini, and L. Stegagno, "Cardiac responses associated with affective processing of unpleasant film stimuli," International Journal of Psychophysiology, vol. 36, 2000, pp. 45-57.