Identification of Optimal Emotion Classifier with Feature Selections Using Physiological Signals

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

Jin-Hun Sohn
Department of Psychology & Brain Research Institute
Chungnam National University
Daejeon, South Korea
jhsohn@cnu.ac.kr

Abstract—The purpose of this study is to identify optimal algorithm for emotion classification which classify seven different emotional states (happiness, sadness, anger, fear, disgust, surprise, and stress) using physiological features. Skin temperature, photoplethysmography, electrodermal activity and electrocardiogram are recorded and analyzed as physiological signals. For classification problems of the seven emotions, the design involves two main phases. At the first phase, Particle Swarm Optimization selects P % of patterns to be treated as prototypes of seven emotional categories. At the second phase, the PSO is instrumental in the formation of a core set of features that constitute a collection of the most meaningful and highly discriminative elements of the original feature space. The study offers a complete algorithmic framework and demonstrates the effectiveness of the approach for a collection of selected data sets.

Keywords—emotion classification, physiological signals, prototypes, feature selections, particle swarm optimization

I. INTRODUCTION

Recently, the most popular research in the field of emotion recognitions is to recognize human’s feeling using various physiological signals. In the physio-psychological research, it is known that there is strong correlation between human emotion state and physiological reaction. Psychologists and engineers have tried to analyze facial expressions, vocal emotions, gestures, and physiological signals in an attempt to understand and categorize emotions [1][2]. Most of all, physiological signals have been used to recognize human's emotions and feelings because the signals acquisition by non-invasive sensors is relatively simple, physiological responses induced by emotion are less sensitive in social and cultural difference, and there is a strong relationship between physiological reactions and emotional and affective states of humans [3].

Many previous studies on emotion have reported that there is correlation between basic emotions such as happiness, sadness, anger, fear, etc. and physiological responses [3][4]. Recently, emotion recognition using physiological signals has been performed by various machine learning algorithms, that is, Fisher’s Linear Discriminant (FLD), k-Nearest Neighbor algorithm (kNN), Support Vector Machine (SVM) and so on [5-8]. One of most widely applied machine learning method is instance-based learning (IBL [9]) which was shown to perform well in a number of challenging learning tasks.

In this paper, we introduce identification of optimal emotion classifier with feature selections using physiology signals induced by emotional stimuli for seven emotion classifications (happiness, sadness, anger, fear, disgust, surprise, and stress). To induce each emotion, ten emotional stimuli sets which have been tested for 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. To complete their efficient development of a classifier for the seven emotions, we use one of the techniques of evolutionary optimization, namely, particle swarm optimization (PSO). In order to improve classification speed and classifier accuracy for the seven emotions, suitable formations of a set of prototypes and a core set of features are required. PSO embraces two level optimization processes. In the first level, PSO choose P % of patterns as a set of prototypes comes from patterns with seven emotional categories. In the second level of the optimization process, PSO 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. Numerical experiments were 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 using physiological signals for the seven emotions.

II. MEASUREMENTS OF PHYSIOLOGICAL SIGNALS AND EMOTIONAL STIMULI

In this section, we will deal with experiments for the induction of the seven emotions and acquisition of physiological signals on an emotion induced by an emotional stimulus. 6 males (20.8 years±1.26) and 6 females (21.2 years±2.70) college students participated in this study. 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 of the experiment.

A. Experimental Settings and Procedures

The laboratory is a room of 5m x 2.5m size having a sound-proof (lower than 35dB) of the noise level where any outside noise or artifact are completely blocked. A comfortable chair is placed in the middle of the laboratory and 38 inch TV monitor set for presentation of film clips is placed in front of the chair. An intercommunication device is placed to the right side of chair for subjects to communicate with an experimenter. A CCTV is installed on the top of the monitor set to observe participant’s behaviours and their behaviours were storage through the monitor and a video cassette recorder outside the laboratory as shown in Fig.1.

Prior to the experiment, subjects are introduced to detail experiment procedures and have an adaptation time to feel comfortable in the laboratory setting. Then they are attached electrodes on their wrist, finger, and ankle for measurement of physiological signals. Physiological signals are measured for 60 sec prior to the film clip presentation (baseline) and for 2 to 4 min during the presentation of the film clips (emotional state), then for 60 sec after presentation of the film clips as recovery term as shown in Fig. 2. Subjects rate the emotion that they experienced during presentation of the film clip on the emotion assessment scale. This procedure is conducted on each of the seven emotions for 10 times. 730 physiological signal data except for severe artefact effect by movements, noises, etc. are used for analysis.

B. Emotional Stimuli and Physiological Signals

To successfully induce the seven emotions (happiness, sadness, anger, fear, disgust, surprise and stress), seventy emotional stimuli, which consist of 10 sets for the seven emotions, are used in the experiments. Emotional stimuli are constituted 2~4 min long audio-visual film clips which are captured originally from movies, documentary and TV shows such as victory, wedding, laughing, etc. for happiness, death of parents/lover, separation, longing for mother, etc. for sadness, massacre, beating, attack, etc. for anger, ghost, haunted house, scare, etc. for fear, body in pieces, vomiting, etc. for disgust, sudden or unexpected scream etc. for surprise, and audio/visual noise on screen, etc. for stress. 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. They also have a relatively high degree of ecological validity, in so far as emotions are often evoked by dynamic visual and auditory stimuli that are external to the individual [10-12].

The suitability and effectiveness of emotional stimuli are examined in preliminary study prior to an experiment. 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”). Twenty-two college students, that are different group from participants in the experiment, categorize their experienced emotion into seven emotion and estimate intensity of their categorized emotion on emotional assessment scale after being presented each film clip. The Table 1 shows the results of the suitability and effectiveness of emotional stimuli gotten from in preliminary study. The emotional stimuli have the suitability of 93% and the effectiveness of 9.5 point on average of 10 sets as shown in the results. The suitability of each stimulus is ranged from 75 to 100% and the effectiveness comes out from 8.4 to 10.4 point as shown in results.

The dataset of physiological signals such as skin temperature (SKT), electrodermal activity (EDA), photoplethysmography (PPG), and electrocardiogram (ECG) are collected by MP100 Biopac system Inc. (USA). 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.
TABLE I. SUITABILITY AND EFFECTUALNESS OF EVOKING EMOTIONS

<table>
<thead>
<tr>
<th>Set</th>
<th>Emotions</th>
<th>Happiness</th>
<th>Sadness</th>
<th>Anger</th>
<th>Fear</th>
<th>Disgust</th>
<th>Surprise</th>
<th>Stress</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td>100% (8.4)</td>
<td>92% (9.5)</td>
<td>75% (9.7)</td>
<td>75% (10)</td>
<td>75% (10.2)</td>
<td>75% (9.3)</td>
<td>92% (9.3)</td>
<td>83% (9.5)</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>100% (8.9)</td>
<td>100% (9.9)</td>
<td>100% (9.9)</td>
<td>100% (9.9)</td>
<td>100% (9.9)</td>
<td>100% (9.9)</td>
<td>100% (9.9)</td>
<td>100% (9.9)</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>100% (8.8)</td>
<td>100% (8.7)</td>
<td>75% (9.7)</td>
<td>83% (9.8)</td>
<td>92% (9.9)</td>
<td>100% (9.7)</td>
<td>100% (8.8)</td>
<td>93% (9.3)</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>100% (9.6)</td>
<td>100% (9.7)</td>
<td>75% (9.5)</td>
<td>92% (9.6)</td>
<td>100% (10.4)</td>
<td>100% (9.9)</td>
<td>100% (8.9)</td>
<td>95% (9.7)</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>100% (9.6)</td>
<td>100% (9.3)</td>
<td>92% (9.8)</td>
<td>92% (9.7)</td>
<td>92% (9.7)</td>
<td>83% (9.6)</td>
<td>100% (9.3)</td>
<td>94% (9.6)</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>100% (9.3)</td>
<td>75% (8.9)</td>
<td>92% (8.9)</td>
<td>83% (9.6)</td>
<td>100% (9.3)</td>
<td>100% (9.5)</td>
<td>92% (9.3)</td>
<td>92% (9.3)</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>100% (9.7)</td>
<td>100% (9.2)</td>
<td>92% (9.5)</td>
<td>100% (9.3)</td>
<td>100% (10.1)</td>
<td>83% (8.6)</td>
<td>100% (9.1)</td>
<td>96% (9.4)</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>92% (8.8)</td>
<td>100% (9.3)</td>
<td>92% (9.7)</td>
<td>75% (8.7)</td>
<td>100% (10.3)</td>
<td>75% (10.3)</td>
<td>100% (9.5)</td>
<td>91% (9.5)</td>
</tr>
</tbody>
</table>

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 by AcqKnowledge (Ver. 3.8.1) software (USA). The emotional states are determined by the result of participant’s self-report (scene that emotion is most strongly expressed during presentation of each stimulus). 28 features extracted from the physiological signals and used to analysis are as follows: SCL, NSCR, meanSCR, meanSKT, maxSKT, meanPPG, meanRRI [ms], stdRRI [ms], meanHR [1/min], RMSSD [ms], NN50 [count], pNN50 [%], SD1 [ms], SD2 [ms], CSI, CVI, RRI, TINN [ms], FFTapLF, FFTapHF, FFTnLF, FFTnHF, FFTLF/HFRatio, ARapLF, ARapHF, ARnLF, ARnHF, and ARLF/HFRatio.

III. IDENTIFICATION OF OPTIMAL EMOTION CLASSIFIER

A. Two level processes for a classifier with feature selection

For the classification of the seven emotions, the proposed classifier is a type of instance 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 patterns called prototypes in the feature space. This classifier embraces two selection problems to classify a new pattern to a class. One is the selection of prototype patterns and another one is feature selection.

In light of these observations, we adopt two level optimization processes for the formation of the prototypes and the feature space. To format prototype, we start with choose P % of patterns using particle swarm optimization (PSO). 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 prototypes. Given a set of N prototypes, the classifier finds the one prototype closest in feature space to an unknown pattern, and then assigns the unknown pattern to the class label of its nearest prototype. The underlying distance between a pattern and a prototype is measured by weighted Euclidean one, that is

$$d(x, y) = \frac{1}{\sigma_i} \sum_{i=1}^{n} \left(\frac{x_i - y_i}{\sigma_i}\right)^2. \quad (1)$$

Where \(x\) and \(y\) are the two patterns in the n-dimensional space and \(\sigma_i\) is the standard deviation of the i-th feature whose value is computed using the prototype set.

Secondly, once the prototypes 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. Often their number could be quite limited in comparison with the dimensionality of the overall feature space. One can consider d% of the total number of features, say 10%, 20%, etc. The features forming the core have to be considered altogether. We use PSO to choose d% of features which minimizes the classification error.

Overall, the algorithm can be outlined as the following sequence of steps

Step 1: Randomly generate “N” particles, \(p_i\), and their velocities \(v_i\). Each particle in the initial swarm (population) is evaluated using the objective function. For each particle, set \(p_{best} = p_i\) and search the best particle of \(p_{best}\). Set the best particle associated with the global best, \(g_{best}\).

Step 2: Adjust the inertia weight, \(w\). Typically, its values decrease linearly over the time of search. We start with \(w_{max}=0.9\) at the beginning of the search and move down to \(w_{max}=0.4\) at the end of the iterative process,
where $\text{iter}_{\text{max}}$ denotes the maximum number of iterations of the search and “$t$” stands for the current index of the iteration.

Step 3: Given the current values of $\text{gbest}$ and $\text{pbest}$, the velocity of the $i$-th particle is adjusted following (3). If required, we clip the values making sure that they are positioned within the required region.

$$v_i = wv_i + c_1(r_1(pbest_i - p_i) + c_2(r_2(gbest_i - p_i))$$

Step 4: Based on the updated velocities, each particle changes its position using the expression

$$p_i = p_i + v_i$$

Step 5: Move the particles in the search space and evaluate their fitness both in terms of $\text{pbest}$, and $\text{gbest}$.

Step 6: Repeat from Step 2 to Step 5 until the termination criterion has not been met. Otherwise return $\text{gbest}$ as the solution found.

B. Prototypes and features versus PSO

The underlying principle of PSO[13] involves a population-based search in which individuals representing possible solutions carry out a collective search by exchanging their individual findings while taking into consideration their own experience and evaluating their own performance. PSO involves two competing search strategy aspects [14][15]. One deals with a social facet of the search; according to this, individuals ignore their own experience and adjust their behavior according to the successful beliefs of individuals occurring in their neighborhood. The cognition aspect of the search underlines the importance of the individual experience where the element of population is focused on its own history of performance and makes adjustments accordingly.

PSO is conceptually simple, easy to implement, and computationally efficient. Unlike many other heuristic techniques, PSO has a flexible and well-balanced mechanism to enhance the global and local exploration capabilities [15]. The basic elements of PSO technique are performance index (fitness), particles, best particles, and velocity. The particle is moving in the search space with some velocity which plays a pivotal role in the search process.

As PSO is an iterative search strategy, we proceed until there is no substantial improvement of the fitness or we have exhausted the number of iterations allowed in this search. As a generic search strategy, the PSO 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 prototype 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. 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. Considering that we are concerned with d % of all features, we pick up the first dxn (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.

IV. NUMERICAL EXPERIMENTS

The numerical studies presented here provide some experimental evidence behind the effectiveness of the PSO 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

| TABLE II. | CLASSIFICATION ACCURACY (AVERAGE AVG AND STANDARD DEVIATION STD %) REGARDED AS A FUNCTION OF "D" AND "P" FOR THE SEVEN EMOTIONAL DATASET |
|------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| d %        | P %             | AVG±STD over P  |
| (No. of F) | 30              | 50              | 70              | 30              | 50              | 70              |
| 30 (8)     | 58.3 ± 3.89     | 79.5 ± 10.93    | 90.3 ± 6.36     | 76 ± 15.42      |
| 50 (14)    | 45.5 ± 4.15     | 54.1 ± 5.04     | 72.4 ± 4.39     | 57.3 ± 12.24    |
| 70 (20)    | 35.5 ± 1.92     | 39.9 ± 3.41     | 47.3 ± 3.74     | 40.9 ± 5.81     |
| 100 (28)   | 35.5 ± 1.86     | 39.4 ± 1.9      | 49.9 ± 2.15     | 41.6 ± 6.46     |
“d”) forming the core of the reduced feature space and the percentage of the data forming the prototype set (P) optimized by the PSO. The results are reported for the testing data sets for various values of “P” and “d”.

The numeric values of the parameters of the PSO were either predetermined (considering some existing guidelines available in the literature, cf. [16][17]) or it becomes selected experimentally. More specifically, we used the following values of the parameters: maximum number of generations is 500; swarm size is 150; maximal velocity, \( v_{\text{max}} \), is 20% of the range of the corresponding variables; \( w_{\text{max}} = 0.4; w_{\text{max}} = 0.9 \); and acceleration constants \( c_1 \) and \( c_2 \) are set to 2.0. The inertia weight factor “w” has been regarded as a linearly decreasing function of iterations (optimization time).

The seven emotional classes consist of happiness, sadness, anger, fear, disgust, surprise and stress. The proposed prototype-based classifier is evaluated by 10 times repeated random sub-sampling validation for seven emotion classification. The 70% of the whole emotional patterns are selected randomly for training and the remaining patterns are used for testing purposes. The results are reported by presenting the average and standard deviation of the classification accuracy obtained over 10 repetitions of the experiment for the test dataset. When reporting results, we concentrate on the determination of relationships between the collections of features and obtained classification rates. We also look at the optimal subsets of features constructed with the use of the method.

For the seven emotions dataset given as multi-physiological signals, the relationship between the percentage of features used in the PSO optimization, values of “P” and the resulting classification accuracy is presented in Table II. 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 PSO.

With the increasing values of “d”, the classification accuracy of the seven emotions decreases substantially; in the case of P=30% it drops from 58.3 to 35.5 when increasing the number of features from 30% to 70%. The similar downward tendency occurs when dealing with any P % and considering the same increase in the percentage of features. Conclusively, the use of all features dropped accuracy of classification for the seven emotions. Changes in the values of “P” have far less effect on the classification rate, however, the distinguished result was obtained 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. 4 (a). The number of occurrences of a given feature is computed across all values of “P” and “d”. Interestingly, there are several dominant features such as FFTnHF (feature24), FFTLF/HFratio (feature 25), ARnLF (feature 26), ARnHF (feature 27), and ARLF/HFratio (feature 28), meanPPG (feature 6), meanHR (feature 9) and CSI (feature 15) are of lowest relevance. Prototypes were picked up by PSO as a supervisor of the classifier for each class and given patterns were assigned into a class through those and selected features. In case of d=30 for the classification of the seven emotions, namely, the number of features is 8, we have gotten that classification accuracy is 90.3 % for P=70. PSO Fig. 4 (b) shows the fitness function of PSO. Therefore, we can consider than the number of core

![Figure 4](image_url)

*Figure 4. Results of feature selections on PSO: (a) feature usage index over all values of d and P, and (b) a fitness function of PSO for d=30 and P=70*

<p>| TABLE III. COMPARISON OF THE CLASSIFICATION ACCURACY OF THE PROPOSED METHOD AND OTHER METHODS (RESULTS FOR TESTING DATA) |
|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Method</th>
<th>Accuracy (%)</th>
<th>Number of Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>CART</td>
<td>21.7 ± 3.6</td>
<td>28</td>
</tr>
<tr>
<td>C4.5</td>
<td>16.1 ± 1.2</td>
<td>28</td>
</tr>
<tr>
<td>kNN</td>
<td>45.3 ± 2.3</td>
<td>28</td>
</tr>
<tr>
<td>FLD</td>
<td>20.3 ± 1.8</td>
<td>28</td>
</tr>
<tr>
<td>NN</td>
<td>18.0 ± 1.0</td>
<td>28</td>
</tr>
<tr>
<td>PNN</td>
<td>16.3 ± 1.9</td>
<td>28</td>
</tr>
<tr>
<td>RBFs</td>
<td>17.4 ± 1.2</td>
<td>28</td>
</tr>
<tr>
<td>SOM Supervised</td>
<td>16.3 ± 1.9</td>
<td>28</td>
</tr>
<tr>
<td>SOM Unsupervised</td>
<td>15.4 ± 3.2</td>
<td>28</td>
</tr>
<tr>
<td>Proposed Methodology</td>
<td>90.3 ± 2.26</td>
<td>8</td>
</tr>
</tbody>
</table>
features is eight and feature 19, 21, 23, 24, 25, 26, 27, and 28 are selected on Fig. 4 (a).

For the classification of the seven emotions, Table III contrasts the classification accuracy (%) of the proposed method with other well-known methods studied in the literatures [5-8]. As abovementioned, the experiments are reported for the 10 times using a split of data into 70%-30% training and 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.

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 the 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. In addition that, we have introduced an instance-based learning classifier with feature selection learned by particle swarm optimization (PSO) mechanism for the seven emotions expressed by multi-physiological signals. The optimization process of forming the prototypes and the feature space is reflective of the conjecture on the importance of forming a set of prototypes 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 prototype is also justifiable considering that this classification scheme is the simplest that could be envisioned in pattern classification.

ACKNOWLEDGMENT

This research was supported by the Converging Research Center Program funded by the Ministry of Education, Science and Technology (No. 2011K000655 and 2011K000658).

REFERENCES