Classification of Emotional States of Interest and Neutral Using Features from Pulse Wave Signal

Sukanya Phongsuphap*, and Akara Sopharak**

* Department of Computer Science, Faculty of Science, Mahidol University, Bangkok 10400, Thailand
(Tel : +66-2-354-4333; E-mail: ccsps@mahidol.ac.th)
** Program in Computer Information System, Burapha University Chantaburi Campus, Chantaburi 22170, Thailanad
(Tel : +66-3-931-0000; E-mail: akara@buu.ac.th)

Abstract: This paper investigated a method for classifying emotional states by using pulse wave signal. It focused on finding effective features for emotional state classification. The emotional states considered here consisted of interest and neutral. Classification experiments utilized 65 and 60 samples of interest and neutral states respectively. We have investigated 19 features derived from pulse wave signals by using both time domain and frequency domain analysis methods with 2 classifiers of minimum distance (normalized Euclidean distance) and *k*-Nearest Neighbour. The leave-one-out cross validation was used as an evaluation method. Based on experimental results, the most efficient features were a combination of 4 features consisting of (i) the mean of the first differences of the smoothed pulse rate time series signal, (ii) the mean of the absolute values of the second differences of the normalized interbeat intervals, (iii) the root mean square successive difference, and (iv) the power in high frequency range in normalized unit, which provided 80.8% average accuracy with *k*-Nearest Neighbour classifier.

Keywords: Emotional state classification, Pattern recognition, Signal processing.

1. INTRODUCTION

Emotional recognition is one of research fields for emotional human-computer interaction or affective computing [1]. Since computers currently have more impact on human work, a computer development with a combination of emotional capability helps increase the efficiency of computers and the human's performance as well. In this research, we aim to give a computer an ability to classify two emotional states of interest and neutral. These emotional states can occur when the subjects are working on a computer, which can be applied to benefit computer users. For instance, in the development of new software, in a stage of software testing with users, if we know their emotional states while using the software such as having an interest or not, it will enable software developers to develop better software. Another example can be seen in distant learning in which teacher cannot directly communicate with their students. This may hinder teachers from knowing students' interest in a subject. Assisting teacher to know students' emotional state enables them to modify or adjust subject content so as to attract students' attention and interest in the subjects, making teaching and learning more effective.

Early research on emotional recognition focused on facial or voice expressions [2-5]. However, the limitation is that humans can control their facial or voice expression. Subsequent research utilized physiological data such as skin conductivity, electromyogram, respiration and heartbeat rate [6-7]. The previous research on physiological data considered a large amount of combined physiological data, and so a large number of tools for data collection were used to communicate with humans, resulting in inconvenience in data collection. This may generate erroneous data, which give rise to errors in further stages. Therefore, in this research, we focus on an analysis using fewer signals. The pulse wave signal is one of promising signals, so it is selected to study here.

2. BACKGROUND

The pulse is a simple straightforward way to determine the contraction rate of the heart. The pulse can be determined manually by palpation, the regular changes in blood pressure that accompany an individual contraction of the heart. The blood volume pulse is often expressed as a wave, and the amplitude of this wave is a reflection of blood flow and arteriole tone. The photoplethysmograph (PPG) is used to measure the blood volume pulse that is found throughout the human body. This pulse wave will result in a change in the volume of arterial blood and capillaries blood with each pulse beat. This change in blood volume can be detected in peripheral parts of the body such as the fingertip by the PPG. A use of PPG requires only one piece of tool, so it provides convenience making the subject feel more comfortable during data collection. Therefore, this paper used pulse wave signals derived from the PPG to analyze emotional states.

3. METHODOLOGY

Our methodology consists of 3 major stages beginning with a stage of signal acquisition, which is a collection of data on pulse wave from a finger sensor, resulting in interbeat intervals of pulse beats. Then, these data are passed through a stage of features extraction, yielding features from both time domain and frequency domain. The final stage is a classification.

3.1 Data Collection

The pulse wave signals were collected from students. There were 19 subjects, with 9 being males and 10 being females. The subjects were invited to the lab setting room and sat in front of the computers with their feet and right hands placed comfortably apart, with the left index-fingers on the sensor. After that, we would create an event to elicit the subjects' emotion corresponding to the experimental setting.

For interested emotion, we wanted the subjects to become concentrated on a particular circumstance in a given period of time. We had the subjects count the number of rounds of continuous pulse wave so that they became concentrated on counting the number of wave rounds appearing on the screen. For neutral emotion, we wanted the subjects not to feel concentrated on or absorbed with a particular circumstance during an experiment. We let the subjects feel relaxed and not anxious or think of a particular issue by having a dialogue with them. Topics in the dialogue concerned general daily life activities and were changed continuously so that the subjects will not concentrate on an issue. If we could observe them to concentrate on any issue for too long, we would distract their attention from that issue. The subjects were allowed to be relaxed and not anxious about any issue. However, we emphasized on emotions occurring from the subjects' true feeling. Consequently, after collecting data on each emotion, the subjects were asked again about their feelings to examine whether or not they were the desired ones. All of skeptical data were eliminated. Finally, we got 65 and 60 samples of interest and neutral states respectively.

A short-time recording approach was used for data collection. Although we analyzed the signal collected in 5 minutes, in the stage of actual data collection, we spent 6 minutes. Since at the beginning and at the end of data collection, the data derived might have an error stemming from the subjects' bodily movements. In order to curtail such errors, we removed signal collected in a 6-minute period for half a minute in the header part and another half a minute in the end part, yielding a 5-minute period signal for further processes.

3.2 Feature Extraction

We had considered a total of 19 features for differentiating two emotional states of interest and neutral. All of these features were calculated from the pulse rate time series signal, or the interbeat intervals obtained from the PPG. The considered features were as follows:

(1) *sdNN*: Standard deviation of all interbeat intervals.
(2) *rmssd*: The square root of the mean of the sum of the squares of differences between adjacent interbeat intervals.
(3) *sdsd*: Standard deviation of differences between adjacent interbeat intervals.
(4) *NN50*: Number of pairs of adjacent interbeat intervals differing by more than 50 ms in the entire recording.
(5) *pNN50*: NN50 divided by the total number of all interbeat intervals.
(6) *PSD*: Total power, frequency range approximately ≤ 0.4 Hz.
(7) *VLF*: Power in very low frequency range, frequency ≤ 0.04 Hz.
(8) *LF*: Power in low frequency range, frequency range: 0.04 - 0.15 Hz.
(9) *HF*: Power in high frequency range, frequency range: 0.15 - 0.4 Hz.
(10) LF_{norm} : LF in normalized unit, $LF \times 100/(LF+HF)$.
(11) HF_{norm} : HF in normalized unit, $HF \times 100/(LF+HF)$.
(12) LFpHF : Ratio *LF/HF*
(13) μ : mean of all interbeat intervals.
(14) δ : mean of the absolute values of the first differences of all interbeat intervals.
(15) δ_{norm} : mean of the absolute values of the first differences of the normalized interbeat intervals.
(16) γ : mean of the absolute values of the second differences of all interbeat intervals.
(17) γ_{norm} : mean of the absolute values of the second differences of the normalized interbeat intervals.
(18) f_1 : mean of the pulse rates smoothed by Hanning window.
(19) f_2 : mean of the first difference of the smoothed pulse rate signal.

Features (1) – (5) are the ones in time domain, which are recommended by the Task Force of the European Society of Cardiology and the North American of Pacing and Electrophysiology [8]. They are statistical time-domain features which can be divided into two classes: first, the ones derived from direct measurements of the interbeat intervals, and second, the ones derived from the differences between interbeat intervals.

Features (6) – (12) are obtained by frequency domain analysis method. These features are also recommended by the Task Force [8]. They provide the basic information of how power or variance distributes as a function of frequency. The spectral components are calculated on the basis of short-term recording analysis.

Features (13) – (19) are the ones proposed by Picard et al. [6]. They can be computed directly from time domain data.

3.3. Classification

This stage is a use of features extracted from the pulse rate time series signal or interbeat interval data as explained in Section 3.2 to classify emotional states. Two classifiers, normalized Euclidean distance and *k*-Nearest Neighbour, were used and compared.

3.3.1. Normalized Euclidean Distance

The normalized Euclidean distance classifier [9] was used to calculate the distance between the features of the unknown signal to the mean features of the classes, and assigned the unknown signal to the class that was a closest distance. The normalized Euclidean distance was computed by the following equation.

$$d_k^2 = (\boldsymbol{X} - \boldsymbol{\mu}_k)^T \boldsymbol{\sigma}_k^{-1} (\boldsymbol{X} - \boldsymbol{\mu}_k) \quad (1)$$

where $\boldsymbol{X}$ is a feature vector of signal data,
$\boldsymbol{\mu}_k$ is a mean feature vector of the kth class,
$\boldsymbol{\sigma}_k^{-1}$ is the inverse matrix of the diagonal matrix of the kth class with diagonal elements given by variances.

3.3.2. *k*-Nearest Neighbour

The *k*-Nearest Neighbour classifier (*k*-NN) [10] was used to classify unknown signal data into emotional classes by considering the nearest distance of neighbouring signals. The *k*-NN algorithm worked by assigning a feature vector of the unknown signal to the class most frequently represented by the *k* nearest samples of known classification. Here, we calculated the distance by using the normalized Euclidean distance with k = 3, 5, 7, 9, and 11.

4. EXPERIMENTAL RESULTS

We performed experiments to find the most efficient features from all of 19 features (see Section 3.2). At first, a relationship of each pair of features was considered in order to group features to reduce the number of combinations. To find out relationship value, correlation coefficient values (r) of all pairs of features were calculated from all of collected signal data. Then, we grouped features with $| r | >= 0.7$ within the same group, obtaining 7 groups of features as shown in Table 1. Then, related features of each group were considered again to find out a representative of each group by performing classification experiments by using individual features. We selected the feature that gave the highest accuracy in classification as a representative feature for each group. Table 2 showed the representative features.

Table 1 Seven groups of features.

Group no.	Features
1	μ, f_1
2	f_2
3	δ_{norm}, γ_{norm}
4	*sdNN*, δ, γ, *rmssd*, *sdsd*, *NN50*, *pNN50*
5	*PSD*, *LF*, *HF*
6	*VLF*
7	LF_{norm}, HF_{norm}, *LFpHF*

Table 2 Representative feature in each group.

Group no.	Representative feature
1	f_1
2	f_2
3	γ_{norm}
4	*rmssd*
5	*PSD*
6	*VLF*
7	HF_{norm}

The seven representative features can create totally 127 possible combinations. We had performed experiments for classifying interest and neutral emotional states by using each of the 127 possible combinations.

For normalized Euclidean distance classifier, the highest correct classification rates by using single features and combinations of two features to seven features were shown in Table 3. Experimental results showed that a combination of 4 features consisting of f_2, γ_{norm}, *rmssd,* and *PSD* yielded the highest correct classification rate of 75.2 %.

Table 3 The highest correct classification rates using normalized Euclidean distance.

No. of Features	Best feature combination	Accuracy rate (%)
1	(γ_{norm})	66.4
2	(γ_{norm}, *rmssd*)	72.0
3	(f_2, γ_{norm}, *rmssd*)	74.4
4	(f_2, γ_{norm}, *rmssd, PSD*)	75.2
5	(f_2, γ_{norm}, *rmssd, VLF,* HF_{norm})	74.4
6	(f_1, f_2, γ_{norm}, *rmssd* , *PSD, VLF*) or (f_1 , f_2 , γ_{norm} , *rmssd* , *PSD,* HF_{norm}) or (f_2 , γ_{norm} , *rmssd, PSD, VLF,* HF_{norm})	73.6
7	(f_1, f_2, γ_{norm}, *rmssd* , *PSD, VLF* , HF_{norm})	73.6

For *k*-NN classifier, the classification results were shown in Table 4. We got three best sets of features as follows:

(1) A combination of f_2 , γ_{norm} , *rmssd, VLF,* and HF_{norm} performed well for *k*=3.

(2) A combination of f_2 , γ_{norm} , *rmssd,* and HF_{norm} performed well for *k*=5,7 and 9.

(3) A combination of f_1, f_2 , γ_{norm} , *rmssd,* and HF_{norm} performed well for *k*=11.

However, the highest accuracy rate of 82.4% was obtained by using one of the following two combinations:

(1) a combination of four features consisting of f_2 , γ_{norm} , *rmssd*, and HF_{norm} , or

(2) a combination of five features consisting of f_1 , f_2 , γ_{norm} , *rmssd,* and HF_{norm} .

Table 4 The highest correct classification rate for each *k* value of *k*-NN classifier.

K	Features	Accuracy rate (%)
3	(f_2 , γ_{norm} , *rmssd, VLF,* HF_{norm})	81.6
5	(f_2 , γ_{norm} , *rmssd,* HF_{norm})	81.6
7	(f_2 , γ_{norm} , *rmssd,* HF_{norm})	81.6
9	(f_2 , γ_{norm} , *rmssd,* HF_{norm})	82.4
11	(f_1 , f_2 , γ_{norm} , *rmssd,* HF_{norm})	82.4

When considering average accuracy rates by varying *k*=3, 5, 7, 9, and 11. The combination of 4 features consisting of f_2 , γ_{norm} , *rmssd*, and HF_{norm} gave the highest average accuracy rate of 80.8%, followed by a combination of f_1 , f_2 , γ_{norm} , *rmssd,* and HF_{norm}, and a combination of f_2 , γ_{norm} , *rmssd, VLF,* and HF_{norm} giving 80.2% and 79.2% average accuracy rates respectively. The accuracy rates of all of three sets of features were summarized in Table 5. We can notice that the common features of the three sets of features are the ones giving the highest average accuracy rate. They are f_2 , γ_{norm} , *rmssd*, and HF_{norm} .

Table 5 Summary of the correct classification rates using *k*-NN classifier and average accuracy rate for each feature set.

Features	Accuracy rate (%)					Average (%)
	k=3	*k*=5	*k*=7	*k*=9	*k*=11	
f_2 , γ_{norm}, *rmssd, VLF,* HF_{norm}	81.6	80.0	78.4	79.2	76.8	79.2
f_2 , γ_{norm}, *rmssd,* HF_{norm}	77.6	81.6	81.6	82.4	80.8	80.8
f_1 , f_2 , γ_{norm}, *rmssd,* HF_{norm}	79.2	79.2	80.0	80.0	82.4	80.2

When comparing the two classifiers, the *k*-Nearest Neighbour classifier yielded the significantly higher correct classification rate than the normalized Euclidean distance classifier did, 80.8% vs. 75.2%.

5. CONCLUSIONS

This paper focused on a classification of two emotional states of interest and neutral. A total of 19 features from pulse wave signal were investigated. From an initial experiment, we obtained seven features as representatives of the 19 features. Then, 127 possible combinations of the representatives were used to classify the two emotional classes by using two classifiers of normalized Euclidean distance and *k*-Nearest Neighbour, and used the leave-one-out test method for evaluation. Experimental results revealed that pulse wave signal had potential to be used for classifying interest and neutral emotional states. The suitable features were a combination of 4 features consisting of (i) the mean of the first differences of the smoothed pulse rate time series signal (f_2), (ii) the mean of the absolute values of the second differences of the normalized signal (γ_{norm}), (iii) the root mean square successive difference (*rmssd*), and (iv) the power in high frequency range in normalized unit (HF_{norm}). It provided 80.8% average accuracy by using *k*-Nearest Neighbour classifier.

ACKNOWLEDGMENTS

The authors would like to thank all of subjects for their good collaboration.

REFERENCES

[1] R. W. Picard, *Affective Computing*, MIT press, Cambridge, MA, USA, 1997.

[2] L. Yuanzhong, H. Kobatake, “Extraction of facial sketch images and expression Transformation based on FACS,” *Proc. of the IEEE International Conference on Image Processing (ICIP1995)*, Vol. 3, pp. 520-523, 23-26 Oct 1995.

[3] T. Otsuka, J. Ohya, “Recognizing multiple persons’ facial expression using HMM based on automatic extraction of significant frame from image sequences,” *Proc. of the IEEE International Conference on Image Processing (ICIP1997)*, Vol. 2, pp. 546-549, 26-29 Oct 1997.

[4] Z. Guojun, J. H. L. Hansen, J F Kaiser, “Classification of speech under stress based on features derived from the Nonlinear teager energy operator,” *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP1998)*, Vol. 1, pp. 549-552, 12-15 May 1998.

[5] Y. Tian, T. Kanade, J. F. Cohn, “Recognizing action units for facial expression analysis,” *IEEE Trans on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 2, pp. 97-115, 2001.

[6] J. Healey, R. W. Picard, “Digital processing of affective signals,” *Proc. of the IEEE international Conference on Acoustics, Speech and Signal Processing (ICASSP1998)*, Vol. 6, pp. 3749-3752, 12-15 May 1998.

[7] R. W. Picard, E. Vyzas, J. Healy, “Toward machine emotional intelligence: Analysis of affective physiological state,” *IEEE Trans on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 10,pp. 1175-1191, 2001.

[8] Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology, “Heart Rate Variability: Standards of measurement, physiological interpretation, and clinical use,” *European Heart Journal*, Vol. 17, No. 3, pp. 354-381, 1996.

[9] D. R. Wilson, T. R. Martinez, “Improved Heterogeneous Distance Functions,” *Journal of Artificial Intelligence Research*, Vol. 6, pp. 1-34, 1997.

[10] J. Schurmann, *Pattern Classification: A Unified View of Statistical and Neural Approaches*, John Wiley & Sons, New York, USA, 1996.