Two-Dimensional Emotion Evaluation with Multiple Physiological Signals

Jyun-Rong Zhuang[1(✉)], Ya-Jing Guan[1], Hayato Nagayoshi[1], Louis Yuge[2], Hee-Hyol Lee[1], and Eiichiro Tanaka[1]

[1] Graduate School of Information, Production and Systems, Waseda University, 2-7 Hibikino, Wakamatsu-ku, Kita-Kyushu, Fukuoka 808-0135, Japan
gary_zhuang@akane.waseda.jp, {guanyajing1995, h_nagayoshi}@fuji.waseda.jp, {hlee, tanakae}@waseda.jp

[2] Graduate School of Biomedical and Health Sciences, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8551, Japan
ryuge@hiroshima-u.ac.jp

Abstract. Extended roles of robots for activities of daily living (ADL) lead to researchers' increasing attention to human-robot interaction. Emotional recognition has been regarded as an important issue from the human mental aspect. We are developing an assistive walking device which considers the correlation between physical assistance and mental conditions for the user. To connect the assistive device and user mental conditions, it is necessary to evaluate emotion in real-time. This study aims to develop a new method of two-dimensional valence-arousal model emotion evaluation with multiple physiological signals. We elicit users' emotion change based on normative affective stimuli database, and further extract multiple physiological signals from the subjects. Moreover, we implement various algorithms (k-means, T method of MTS (Mahalanobis Taguchi System) and DNN (deep neural network)) for determining the emotional state from physiological data. Finally, the findings indicate that deep neural network method can precisely recognize the human emotional state.

Keywords: Emotion evaluation · Physiological data · Rehabilitation
Promotion of exercise

1 Introduction

With the advance of science and technology, the issue of the Human-Robotic Interaction (HRI) has received substantial attention. Many researchers have dedicated to improving the people's quality of life by employing the HRI technique. In human factors engineering, emotion is a significant feature to communicate with people; besides, people behavior (e.g. attitudes) may be dominated by emotion. Hence, studying emotions is crucial to the understanding of human behavior [1]. With an increase in the aging population worldwide, over the past few years, many studies have devoted to designing the assistive device for the patients, disabled person and elderly, and led them returning their normal life; thus, assistive device played a key role in the human rehabilitation field. However, most of the user intend assistive device complied

S. Fukuda (Ed.): AHFE 2018, AISC 774, pp. 158–168, 2019.
https://doi.org/10.1007/978-3-319-94944-4_18

with their mind since the assistive device is very unfamiliar to the user while wearing it. For improving the user feeling, the assistive device may enable understanding user's mental, and then follow moderately to human emotion change. It is thus necessary to investigate the emotion evaluation for human.

Several devices have been developed to identify emotional states utilizing "outer" signals, for instance, human voice, facial expressions and gestures. Nevertheless, there are limitations to these methods since it can easily control or fake these physiological signals to confuse the device occasionally. Accordingly, there are potential and actual physiological signals might be hidden or misjudged. Human emotional changes accompanied by many variations in physiological features which are difficult to be pretended. These signals thus are called "inner" signals due to their special characteristic. Therefore, analyzing the "inner" signals are great approach for researching human emotion state. There are several inner" signals, for example, through measuring human heartbeat, the instrument could compute the LF/HF ratio which was employed to assess the emotional arousal [2]. LF is an abbreviation for low frequency power which has a in connection with sympathetic component; HF is an abbreviation for high frequency power which reflect cardiac parasympathetic nerve's activity. In addition, electromyography (EMG) is an electrodiagnostic method for assessing and recording the muscle's electrical signals. Using EMG, it is able to measure the people facial muscles' activities when they located in different emotional states [3]. Furthermore, electroencephalogram (EEG) is an approach applied to record brain's electrical activity. Researching brain wave is valid for comprehending and recognizing various physiological signals. Especially, brain wave is proper signals to detect human emotion state, for instance, pleasure and arousal [1]. As indicated in previous studies, utilizing the physiological signals, it can efficiently evaluate human emotion. For evaluating the emotion, in the 1980s, Russell [4] proposed a circumplex model of affect to distinguish the human emotion via the two-dimensional plane. In this model, the horizontal dimension showed pleasure evaluation that utilized to describe the level of happy and unhappy; On the other hand, the vertical dimension exhibited arousal evaluation that illustrated the level of excited and sleepy. In our laboratory, we have developed assistive walking apparatuses which focused on assisting the ankle joint. This assistive apparatus can raise the equipped foot automatically via stretch reflex mechanism. However, currently, this assistive device only assists users using physical method [5]. In general, training and rehabilitation are not an easy and willing thing for the elderly or patients. Hence, the users' mental aspect needs to be considered when they equipped the device for training or rehabilitation [6].

In this study, we prepared the emotional stimuli experiments to inspire users' feelings and collect multiple physiological signals. The purpose of this paper is to analyze multiple physiological signals and employ several methods for precise evaluating the emotional state, and further mapping on the arousal-pleasure dimensional model.

2 Valence-Arousal Dimensional Model

Emotional models enabled to be characterized via two major features which were valence and arousal [4]. Figure 1 reveals the valence-arousal dimension model for recognizing the emotional state [4]. On the valence axis, we can judge the degree of attraction of a specific event from the human. Additionally, on arousal axis, we can respond the psychological state of the human (e.g., heart rate and blood pressure). In this study, we recognized the emotion state based on this indicated two-dimensional model.

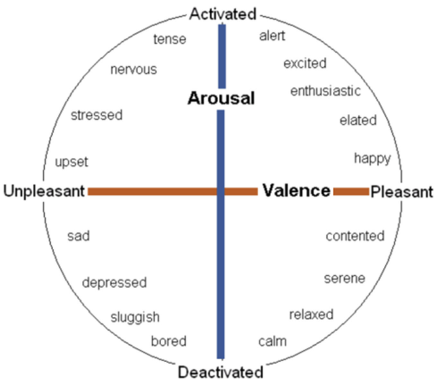


Fig. 1. Valence-arousal dimensional model

3 Experimental Setup

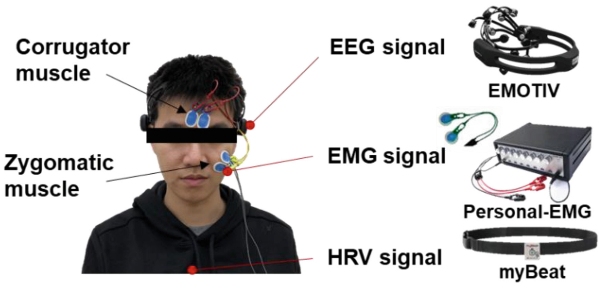
3.1 Normative Affective Stimuli

To elicit the real emotion, the physiological signals can effectively express emotion as the reference. Thus, we attempted to explore the correlation between physiological signals and emotion via the normative affective stimuli. Samson et al. [7] proposed a film library, a collection of stimuli, which available research the emotion variation. These films were presented to 411 subjects for a large online study to verify film clips which reliably elicited the subject emotion [7]. The rating of valence, arousal described each film. The selected film clips for the experiment were varied (with different grade of valence and arousal) to elicit various emotions from the subject. We attempted to achieve all the emotion assumptions. Therefore, we prepared 30 stimuli film through this library to elicit different emotions.

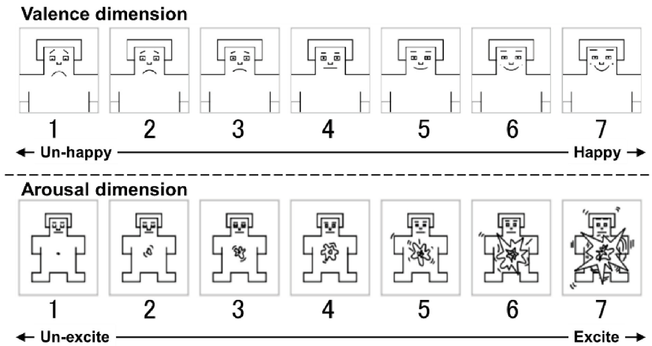
3.2 Experimental Device

In this study, we employed three physiological signals as human emotional judgments, there are EEG (Electroencephalogram), EMG (Electromyography) and HRV (Heart

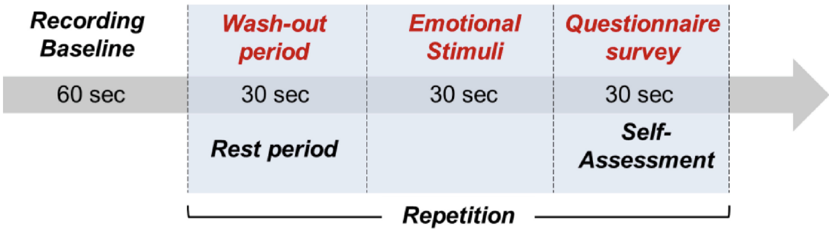
Rate Variability). We used the Emotiv, Personal EMG, and MyBeat to detect the EEG, EMG, and HRV, respectively, as shown Fig. 2(a).



(a) Three physiological signals detector are used. The Emotiv, Personal EMG, and MyBeat to detect the EEG, EMG, and HRV, respectively.



(b) Self-assessment manikins applied in this emotional elicited experiment.



(c) The emotion experimental sequence which is as follows: 1) baseline recording; 2) rest period; 3) emotional stimuli; 4) questionnaire survey.

Fig. 2. (a) Three physiological signals detector are used; (b) Self-assessment manikins (c) Emotion experimental sequence

3.3 Experimental Protocol

We conducted experiments with 20 healthy participants (the age between 21 to 27 years old) wearing three physiological signals detector while watching the film clips. Before starting the experiment, we would record the MVC (maximum voluntary contraction) of each subject. Next, the experimental protocol was as follows: (1) Let subjects clam and rest for 1 min and record these physiological signals as the baseline that specifies the reference level. (2) Subjects watched the selected film clips (30 s for each film clip). After watching each film clips, subjects had 30 s to finish corresponding questionnaire survey (self-assessment as shown in Fig. 2(b)) and have 30 s take a rest as the wash-out period (to eliminate the effect of the previous film clip). Figure 2(c) shows the whole emotion experimental sequence.

3.3.1 Self-assessment

20 Participants would grade (from 1 to 7) their own feelings on valence and arousal dimension, respectively, after watching each film clip. Figure 2(b) demonstrates the SAM (self-assessment manikins [8]) questionnaire. SAM is visualization questionnaire that has vivid graphs, which let subjects grade the emotion on two dimensions easily.

4 Methodology

4.1 Emotional Level Identification

In this study, we processed the emotional scores (1–7) which was divided 3 levels on each valence and arousal dimension. Figure 3 illustrates the nine-emotional state. On valence axis, the 3 levels were following: 1–3, 4 and 5–7 were mapped to “un-pleasure”, “neutral” and “pleasure”, respectively. On arousal axis, the 3 levels were considered as: 1–3 was mapped to “de-excite”, 4 was mapped to “neutral” and 5–7 was mapped to “excite”.

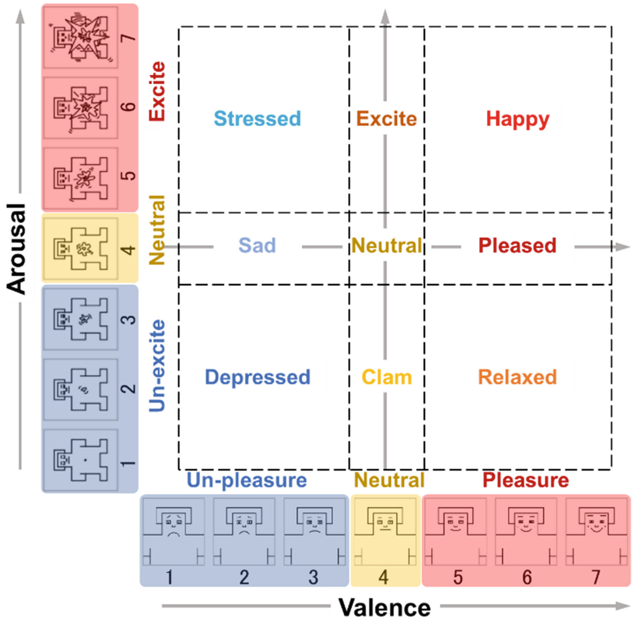


Fig. 3. Nine emotional state

4.2 Physiological Features Selection

We obtained the many raw physiological data from three physiological detectors (EEG, EMG, and HRV). Furthermore, after data processing, we extracted the 17 physiological signals as our physiological features, as shown in Table 1. To eliminate the individual difference, all selected features divided by baseline physiological value.

4.3 Classification Algorithm

4.3.1 K-Means Clustering

During our initial emotion evaluation research, we implemented the unsupervised clustering method of k-means algorithm for analyzing the physiological features and dividing the data set into several natural clusters [6]. By using k-means, it is unnecessary to give the signals with the class labels. The input signals would iteratively be calculated their centroid point (mean value of each group) until several groups could be divided evidently. Each group's feature can be thus captured [6]. We proved the emotion can be mapped on two-dimensional model into four separated areas (three groups on valence axis and two groups on arousal axis) as shown in Fig. 4. In our previous results [6], although k-means algorithm can divide the emotion to four areas, the number of subject and emotional stimuli method would significantly affect the results. To achieve more precisely results, we further expand the data set and use new emotional stimuli method (film library [7]).

Table 1. 17 extracted physiological features

	Features
1	(Max. EMG)/(MVC of corrugator muscle)
2	Max. EMG of corrugator muscle
3	(Max. EMG)/(MVC of zygomatic muscle)
4	Max. EMG of zygomatic muscle
5	Max. LF/HF
6	Max. variance of LF/HF
7	Max. heart rate
8	Min. variance of LF/HF
9	Mean LF/HF
10	Mean variance of LF/HF
11	Standardized LF/HF
12	Standardized heart rate
13	Mean Theta wave's power spectrum
14	Mean Alpha wave's power spectrum
15	Mean low Beta wave's power spectrum
16	Mean high Beta wave's power spectrum
17	Mean Gamma wave's power spectrum

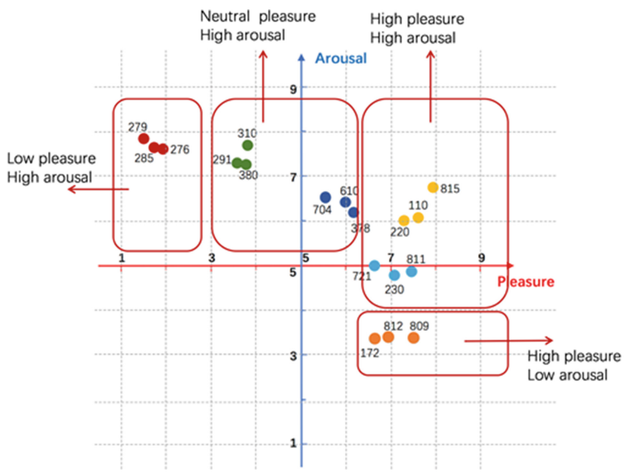


Fig. 4. Four separated groups on two-dimensional model by using k-means clustering

4.3.2 Mahalanobis Taguchi

The Mahalanobis Taguchi System (MTS) is the binary classification algorithms for diagnosis and predict multivariate data. This method utilizes the Mahalanobis distance to measure correlations between the multivariate variables, and use Taguchi method to assess accuracy of estimations via the constructed scale. To continue the emotion evaluation, we increased the number of subjects and number of physiological features, and further utilizing the T method (one of the approaches in MTS) to conduct the emotional classification. The advantages of the T method are that it does not use the Mahalanobis distance, and it can resolve mathematical constraints, and it is possible to judge the direction on emotional dimension using multivariate physiological data. We can attain overall predicted output through the weighted integration with the corresponding signal to noise (S/N ratio). In T method, we need to construct the measurement scale. Following the measured results, we can construct the measurement scale by selected physiological features together with questionnaire results, as shown in Table 2. To confirm threshold for classification, this work used a half of the difference value of two distributions' mean values.

Table 2. T method measurement scale.

	Feature 1	Feature 2	⋯	Feature k	Pleasure axis	Excite axis
Sample 1	⋯	⋯	⋯	⋯	1	2
Sample 2	⋯	⋯	⋯	⋯	3	5
:	:	:	:	:	:	:
Sample n	⋯	⋯	⋯	⋯	7	3

4.3.3 Deep Neural Networks

Deep neural network (DNN) is special form of Artificial Neural Network (ANN). DNN's structure comprise of the multiple hidden layers between the input and output layers. DNN used stacked autoencoders, which can process highly non-linear problems. By using this method, it can train and learn from the all data through multiple layers to calculate repeatedly, and finally obtain the great predictions. In this study, we faced the multivariate variables situations, which is better to apply the DNN to evaluate the emotional state.

We implement the emotion recognition classification of the valence and arousal state, respectively, by using DNN algorithm, and further combined two calculated results as final emotional states judgments. In our DNN system, two classifiers separately employed a stack of three autoencoder with softmax layers as shown in Fig. 5. Additionally, hidden layer 1 has 60 hidden nodes, hidden layer 2 has 30 hidden nodes, and hidden layer 3 has 20 hidden nodes. To reduce the calculated time and obtain high accuracy, it is necessary to screen the important physiological features. Thus, all physiological features would be calculated via attribute selection algorithm [9]. Finally, nine selected features used in valence classifier, and 11 selected features used in arousal classifier. The output features of the hidden layer 3 are utilized as input features for the softmax layer that can be trained as parameters at the same time. After the network completed learning of the weight and bias parameters in softmax classifier, the algorithm must fine-tune all the weights and bias parameters in the whole network simultaneously. The fine-tuning process enabled to improve all weights for all layers in the network. We employed the self-assessed emotional states (valence and arousal) as the basic fact.

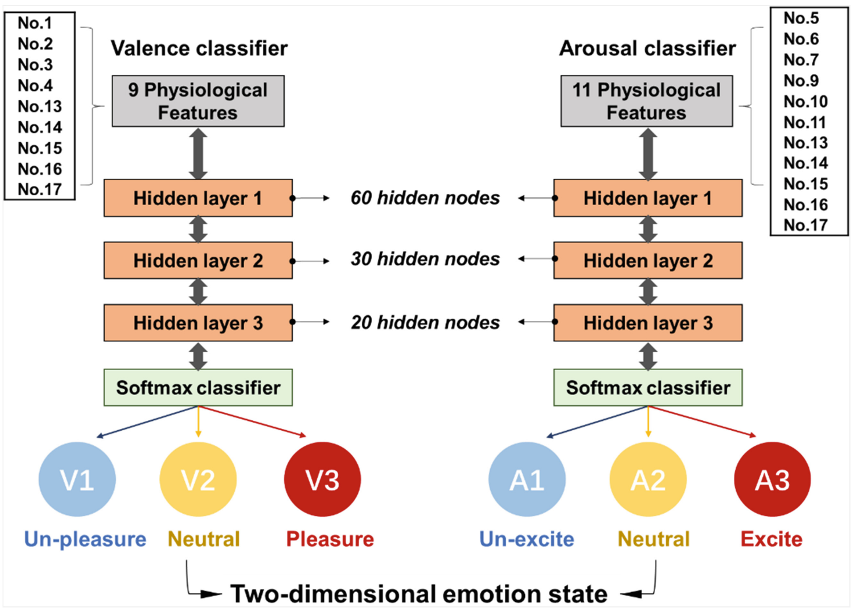
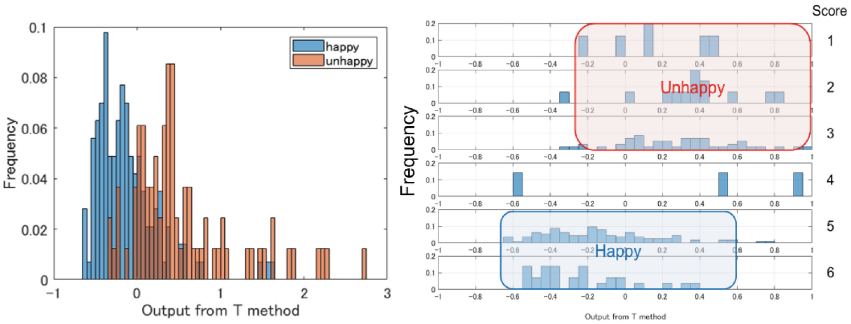


Fig. 5. Two-dimensional emotional DNN classifier structure

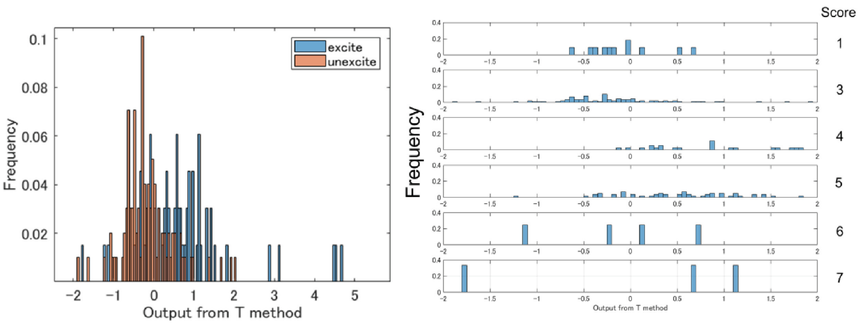
5 Results and Discussion

5.1 The Calculated Results of MTS

In T method, the unit space selection would significantly affect the calculated results. We respectively applied the MTS method to the valence and arousal state. On valence scale, we specified the unit space was score 7. The findings reveal that the valence state can effectively be divided into two groups (happy and unhappy); besides, we also found the distribution of each score that exhibits the single directional tendency from negative (Score 5 and 6) to positive direction (Score 1, 2 and 3), as shown in Fig. 6(a). However, we specified the unit space was score 2 on arousal scale, and the classification results demonstrate the slight lack of precision as shown in Fig. 6(b). To classify results of the arousal scale effectively, we think it may need appropriate unit space to obtain vivid two distributions (excite and un-excite). The classification accuracy of T method was 77% for valence state, and 47% for arousal state.



(a) The distribution of valence state that unit space is score 7



(b) The distribution of arousal state that unit space is score 2

Fig. 6. Calculated results of T method. (a) The distribution of valence state that unit space is score 7; (b) The distribution of arousal state that unit space is score 2

5.2 The Calculated Results of DNN

Deep neural network (DNN) can effectively process the multivariate variables problems for emotion recognition. In our DNN procedure, we utilized the three hidden layers to learn the relevant features from the input physiological signals and further predict the two emotional state. The results reveal that the DNN can achieve the classification accuracy of 69.8% and 77.4% for dividing three valence states and three arousal states by using 17 physiological features. Moreover, by using attribute selection algorithm, we attained the important features (9 features for valence classifier and 11 features for arousal classifier) which can be as input signals of DNN on two classifiers. Thus, the findings show the DNN with selected features can increase the classification accuracy to 79.2% and 81.1% for classifying three valence states and three arousal states. Figure 7 shows the comparison classification accuracy through using the different algorithm.

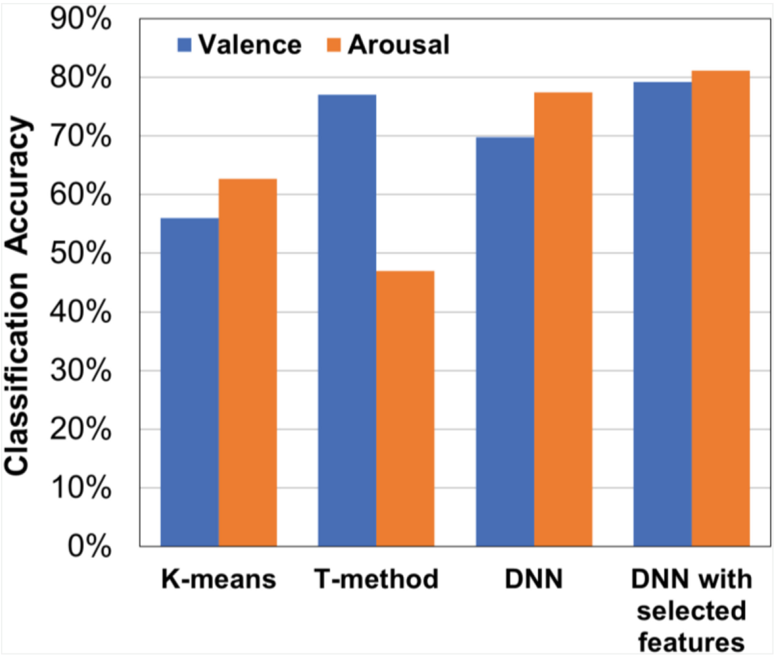


Fig. 7. Four methods' classification accuracy comparison on two emotional states

6 Conclusions

This study aims to develop a new emotion evaluation method for the two-dimensional valence-arousal model by using multiple physiological signals. We applied the k-means, T method of MTS and DNN for confirming the emotional state from physiological data. In previous results, we can employ k-means to divide three groups on valence axis and two groups on arousal axis by using less subjects' measured results. We further enlarged the number of subjects, and the findings showed it can successful divide three groups on each axis; however, the classification accuracy only can reach to 56% and 62.7% on valence and arousal state. Furthermore, we employed the T method of MTS, and we found this method can possible distinguish the valence state (the classification accuracy is 77%). However, on arousal state, this method was less significant (the classification accuracy is only 47%). Finally, we proposed our DNN procedure for processing the emotion evaluation. By using DNN with selected features, we can clearly classify three groups for each dimension; besides, the classification accuracy can achieve to 79.2% and 81.1% for valence and arousal state, respectively. To conclude, the findings indicate that DNN (deep neural network) can precisely recognize and effectively divide the human emotional state. In the future, we will progress the real-time emotional recognition for combining the assistive walking device to improve the feeling of rehabilitation and assistance.

References

1. Chanel, G., Kronegg, J., Grandjean, D., Thierry, P.: Emotion assessment: arousal evaluation using EEG's and peripheral physiological signals. In: International Workshop on Multimedia Content Representation, Classification and Security, pp. 530–537. Springer, Heidelberg (2006)
2. Tanaka, E., Muramatsu, K., Osawa, Y., Saegusa, S., Yuge, L., Watanuki, K.: A walking promotion method using the tuning of a beat sound based on a two-dimensional emotion map. In: Proceedings of the AHFE 2016 International Conference on Affective and Pleasurable Design, pp. 519–525. Walt Disney World, Florida, USA, 27–31 July 2016 (2016)
3. Turner, J.R.: For distinguished early career contribution to psychophysiology: award address 1988. Psychophysiology **26**(5), 497–505 (1989)
4. Russell, J.A.: A circumplex model of affect. J. Pers. Soc. Psychol. **39**(6), 1161–1178 (1980)
5. Tanaka, E., Muramatsu, K., Watanuki, K., Saegusa, S., Yuge, L.: Development of a walking assistance apparatus for gait training and promotion of exercise. In: 2016 IEEE International Conference on Robotics and Automation (ICRA 2016), Stockholm, Sweden, pp. 3711–3716, 16–21 May 2016 (2016)
6. Zhang, Z.Q., Tanaka, E.: Affective computing using clustering method for mapping human's emotion. In: 2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM2017), Munich, Germany, pp. 235–240, 3–7 July 2017 (2017)
7. Samson, A.C., Kreibig, S.D., Soderstrom, B., Wade, A.A., Gross, J.J.: Eliciting positive, negative and mixed emotional states: a film library for affective scientists. Cogn. Emot. **30**(5), 827–856 (2015)
8. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry **25**(1), 49–59 (1994)
9. Karegowda, A.G., Manjunath, A.S., Jayaram, M.A.: Comparative study of attribute selection using ratio and correlation based feature selection. Int. J. Inf. Technol. Knowl. Manag. **2**(2), 271–277 (2010)