EMOTION RECOGNITION FROM PERIPHERAL PHYSIOLOGICAL SIGNALS ENHANCED BY EEG

Shiyu Chen *Zhen Gao* *Shangfei Wang*⋆

University of Science and Technology of China
School of Computer Science and Technology
230027, Hefei, Anhui

ABSTRACT

Current multi-modal emotion recognition from physiological signals requires electroencephalogram(EEG) signals and peripheral physiological signals during both training and test. Compared with the peripheral physiological signals, it is more difficult to obtain EEG signals in our daily life. Therefore, we propose a novel approach to recognize emotions from peripheral signals by using EEG features as privileged information, which is only available during training. During training, first, peripheral physiological features and EEG features are extracted. Then, we construct a new peripheral physiological feature space using canonical correlation analysis with the help of EEG features. Finally we train a support vector machine(SVM) to map the new peripheral physiological features to the emotion labels. During test, only peripheral physiological features are used to recognize emotions from the constructed peripheral physiological feature space with the trained SVM model. The experimental results on two benchmark databases show that our proposed approach using EEG features as privileged information outperforms the method which recognizes emotions merely from the peripheral physiological signals.

***Index Terms*—** Emotion Recognition, EEG, Peripheral Physiological Signals, Privileged Information, Canonical Correlation Analysis

1. INTRODUCTION

Emotion recognition from physiological signals has attracted increasing attention due to its great potential during human-computer interaction. Current research recognizes emotions from either one physiological modality [1][2] or the combination of several physiological signals (i.e. both peripheral signals and EEG signals) [3][4][5]. The EEG signals reflect emotion changes on the central nervous system, while the peripheral signals reflect the emotion influence on the autonomic nervous system. These two systems have intrinsic relations. Therefore, emotion recognition by EEG and peripheral signal fusion is promising. However, compared with peripheral signals, it is much difficult and expensive to obtain the EEG signals which are collected using professional equipments. Therefore, current multi-modal emotion recognition from physiological signals, which requires electroencephalogram(EEG) signals and peripheral physiological signals during both training and test, is not practical.

Based on the above considerations, we propose a new approach to classify emotions from peripheral signals with the help of EEG signals. During training, both EEG signals and peripheral physiological signals are required. A peripheral physiological feature space is constructed with the help of EEG features using canonical correlation analysis (CCA). During test, only the constructed peripheral physiological feature space is required. Thus, the EEG signals are used as privileged information [6] [7], which is only available during training to help peripheral physiological signals to construct a better feature space. SVM is adopted to recognize users' emotions from the constructed peripheral physiological features. Experimental results on two benchmark databases demonstrate that our approach can improve the recognition performance on both valence and arousal spaces. To the best of knowledge, this is the first work to recognize emotion from peripheral physiological signals by using EEG as privileged information.

2. THE FRAMEWORK OF OUR METHOD

The framework of our method is shown as Fig. 1. We first extract features from peripheral signals and EEG signals. In the training phase, we create a new feature space for peripheral features using CCA with the help of EEG features and use the new feature to train a SVM classifier. In the test phase, we map the extracted peripheral features to the new feature space. Then, we predict the emotion state for each sample based on the new features. The EEG features are only available during training to help the peripheral features to construct a better

⋆ This is the corresponding author. This work has been supported by the National Science Foundation of China (Grant No. 61175037, 61228304, 56561473270), and the project from Anhui Science and Tech-nology Agency (1106c0805008).

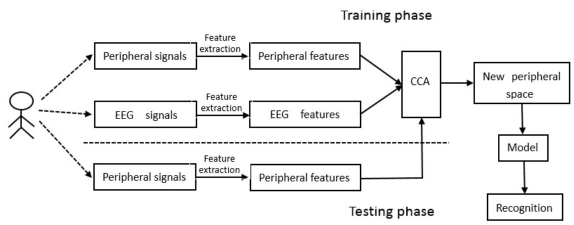

Fig. 1. The framework of our method

feature space as privileged information.

2.1. Feature extraction

2.1.1. Peripheral features

Peripheral signals include electrooculogram(EOG) signals, electromyography(EMG) signals, electrocardiograph(ECG) signals, galvanic skin response(GSR) signals, respiration(RSP) signals, skin temptation(TEMP) signals and plethysmograph(PLET) signals. Before extracting features, these physiological signals should be preprocessed using band-pass filters. For different physiological signals, different features are extracted. The details of extracted features are summarized in Table 1.

2.1.2. EEG features

To restrain the noise, we preprocessed EEG signals adopting a band-pass filter with a lower cutoff frequency of 0.3Hz and a higher cutoff frequency of 45Hz. Then the spectral power from theta (4 Hz $< f <$ 8 Hz), slow alpha (8 Hz $< f <$ 10 Hz), beta (12 Hz $< f <$ 30 Hz), and gamma (30Hz $< f$) bands are extracted for EEG signals as features.

2.2. Constructing a new peripheral feature space

In this paper, we use CCA to construct a new peripheral feature space with the help of EEG features. The original peripheral features and EEG features are denoted as P and E respectively. We use CCA to find the linear projection matrix a and b, so that the transformed peripheral features F and the transformed EEG features G have maximum correlation with each other as shown in Eq. 1.

$$\rho = \frac{F \cdot G}{\sqrt{F' \cdot F}\sqrt{G' \cdot G}} \ . \tag{1}$$

where F and G are defined as follows.

$$F = Pa \ . \tag{2}$$

$$G = Eb \ . \tag{3}$$

F will be used as the new peripheral features for training and test. The relationships between P, E, a, b are follows:

$$(P^T P)^{-1} P^T E (E^T E)^{-1} E^T P a = m^2 a \ . \tag{4}$$

Table 1. All features extracted from peripheral signals

Signal	Extracted features	filters
EOG	Energy, mean and variance	0.4Hz
EMG	Energy, mean and variance	1Hz
ECG	HRV, root mean square of the mean squared difference of successive beats, standard deviation of beat interval change per respiratory cycle, 14 spectral power in the bands from [0, 1.5]Hz, low frequency [0.01, 0.08]Hz, medium frequency [0.08, 0.15]Hz and high frequency [0.15, 0.5]Hz components of HRV power spectrum, Poincare analysis features(2 features)[1]	1Hz
GSR	Mean, mean of the derivative, mean of the positive derivatives, proportion of negatives in the derivative, number of local minima, 10 spectral powers within 0-2.4Hz	3Hz
RSP	Band energy ratio, average respiration signal, mean of the derivative, standard derivation, range of greatest breath, 10 spectral powers within 0-2.4Hz, average and median peak to peak time	0.45Hz
TEMP	Mean, mean of the derivative, spectral powers in 0-0.1 Hz and 0.1-0.2 Hz	3Hz
PLET	Average and standard derivation of HRV and inter-beat intervals, energy ratio between 0.04-0.15 Hz and 0.15-0.5 Hz, spectral power in 0.1-0.2 Hz, 0.2-0.3 Hz, 0.3-0.4 Hz, 0.01-0.08 Hz, 0.08-0.15 Hz and 0.15-0.5 Hz components of HRV	0.45Hz

$$(E^T E)^{-1} E^T P (P^T P)^{-1} P^T E b = m^2 b \ . \tag{5}$$

Where m = $a^T P^T E b$. m is defined as the maximum eigenvalue of $(P^T P)^{-1} P^T E (E^T E)^{-1} E^T P$ and a is the corresponding eigenvector. Meanwhile, b is the eigenvector of $(E^T E)^{-1} E^T P (P^T P)^{-1} P^T E$. For details, please refer to [7] [8]. The projection matrix a will reflect the relation between peripheral features P and EEG features E.

In the training phase, we learn the projection matrix a and we use the F to train the classifier for emotion recognition. In the test phase, we first use a to project the peripheral features to new feature space and use the new feature space to predict the emotion states.

2.3. Classifier and emotion recognition

We adopt the SVM to recognize the users' emotions from the constructed peripheral features. The EEG features as privileged information are only used in the training phase to construct a better feature space for the peripheral features. In the test phase, we predict the users' emotions only with the constructed peripheral features. In the process of classification, radial basis function is used as kernel function.

3. EXPERIMENTS

3.1. Experimental conditions

To validate the performance of our method, two benchmark databases are adopted: the DEAP database[9] and the MAHNOB-HCI database[10].

The DEAP database records seven kinds of physiological signals, EEG, EOG, EMG, ECG, GSR, RSP, TEMP and PLET, from 32 participants during their watching music videos. The MAHNOB-HCI database includes five kinds of physiological signals, EEG, ECG, GSR, RSP and TEMP, from 27 participants during their watching 20 videos. The emotional self-assessment of both databases are in nine-scale evaluations from 1 to 9 for valence and arousal. In our work, we translate the ratings as positive or high if they are larger than 5, otherwise, we translate them as negative or low. Thus, we get 533 and 1216 EEG segments from the MAHNOB-HCI database and the DEAP database respectively. Specifically, for valence, 289 positive and 244 negative EEG segments, for arousal, 268 high and 265 low EEG segments are from the MAHNOB-HCI database. For valence, 672 positive and 544 negative EEG segments, for arousal, 726 high and 490 low EEG segments are from the DEAP database.

To validate the effectiveness of our proposed method, we conduct two group experiments, i.e. emotion recognition from a peripheral signal and emotion recognition from peripheral signal combination. For each group experiment, we compare three methods: emotion recognition from peripheral signals only, emotion recognition from peripheral signals after principle component analysis (PCA) and our method. For the first method, peripheral features described in Section 2.1 are used. For the second method, peripheral features after feature selection using PCA are used. A 10-fold cross-validation is applied to the experiments. Two metrics, accuracy and averaged F1-score, are adopted to evaluate the performance of emotion recognition.

3.2. Experimental results and analyses

3.2.1. Emotion recognition from a peripheral signal

The results of emotional recognition from each peripheral signal on the DEAP database and the MAHNOB-HCI database are shown in Table 2 and Table 3 respectively.

From Table 2 and Table 3, we find that compared with the method recognizing emotion merely from the peripheral signal, our method using EEG features as privileged information has better performance on valence and arousal spaces recognition on the two databases, since both accuracies and F1-scores of our method are higher than those of recognizing emotion from peripheral signal only in most cases.

Table 2. Emotional recognition results on DEAP

	valence		arousal	
	accuracy	F1-score	accuracy	F1-score
EOG	0.5551	0.5105	0.5691	0.5055
EOG+PCA	0.5518	**0.5309**	0.5403	0.4920
Ours	0.5428	**0.5332**	0.5567	**0.5346**
EMG	0.5444	0.5208	0.5526	0.4967
EMG+PCA	0.5304	0.5063	0.5461	0.4928
Ours	**0.5543**	**0.5441**	**0.5650**	**0.5374**
RSP	0.5411	0.5271	0.6053	0.5718
RSP+PCA	0.5304	0.5172	0.5707	0.5452
Ours	**0.5748**	**0.5626**	0.6044	**0.5808**
GSR	0.5337	0.5241	0.5526	0.5176
GSR+PCA	0.5099	0.5027	**0.5584**	**0.5349**
Ours	**0.5567**	**0.5458**	**0.5863**	**0.5582**
PLET	0.5543	0.5258	0.5806	0.5213
PLET+PCA	0.5304	0.5047	0.5600	0.5027
Ours	0.5419	0.5199	**0.6209**	**0.5761**
TEMP	0.5518	0.5262	0.5740	0.5157
TEMP+PCA	0.5411	**0.5328**	0.5403	**0.5174**
Ours	**0.5543**	**0.5406**	**0.5970**	**0.5395**
ALL	0.5510	0.4915	0.6151	0.5106
ALL+PCA	**0.5526**	**0.5253**	0.5757	**0.5519**
Ours	**0.5814**	**0.5752**	**0.6271**	**0.6018**

For both valence and arousal recognition, using PCA has negative impact on the performance of emotion recognition. From Table 2 and Table 3, we can find that when using P-CA, accuracies are decreased on the DEAP database and the MAHNOB-HCI database in most cases. The goal of PCA is to find a set of values of linearly uncorrelated variables called principal components using an orthogonal transformation. P-CA find an orthogonal peripheral physiological space with no extern information, while CCA find an orthogonal peripheral physiological space enhanced by EEG. CCA can take advantage of the information from EEG in the training phase. Furthermore, using PCA might cause information loss since part of the faithful features have been removed during feature selection. Therefore, our method using CCA can construct a better feature space.

The performance of arousal recognition is better than valence recognition on both databases, since both accuracies and F1-scores of arousal recognition are higher than those of valence recognition for all the cases. It may indicate that it is easier to recognize arousal than valence.

To further evaluate the effectiveness of the proposed method, we conduct the above 10-fold cross-validation experiment 10 times, and apply t-test to check whether the improvement using the privileged information is significant or not based on the accuracy on valence and arousal space.

Table 3. Emotional recognition results on MAHNOB-HCI

	valence		arousal	
	accuracy	F1-score	accuracy	F1-score
ECG	0.5197	0.5033	0.5816	0.5793
ECG+PCA	0.5122	0.5000	0.5704	0.5702
Ours	**0.5460**	**0.5368**	**0.6360**	**0.6360**
RSP	0.5235	0.5167	0.5779	0.5779
RSP+PCA	**0.5328**	0.4961	0.5253	0.5230
Ours	**0.5572**	**0.5478**	**0.6079**	**0.6197**
GSR	0.5422	0.5373	0.6098	0.5996
GSR+PCA	0.5216	0.5142	0.5779	0.5649
Ours	**0.5760**	**0.5561**	0.6060	**0.6059**
TEMP	0.5122	0.5081	0.5328	0.5328
TEMP+PCA	**0.5328**	0.4903	0.5253	0.5252
Ours	**0.5516**	**0.5476**	**0.5403**	**0.5401**
ALL	0.5066	0.4774	0.6004	0.6002
ALL+PCA	**0.5159**	0.3863	0.5647	0.5641
Ours	**0.5741**	**0.5514**	**0.6135**	**0.6125**

As shown in Table 4, the p-values on the DEAP database and HCI database are less than 0.05 in most cases, demonstrating that the improvement of emotion recognition caused by using EEG as privileged information is significant.

3.2.2. Emotion recognition from peripheral signal combination

Since there are six and four kinds of peripheral signals on the DEAP database and the MAHNOB-HCI database respectively, the number of peripheral signal combinations is very larger. Due to the page limit, we only conduct emotion recognition from one kind of peripheral signal combination, i.e. all the peripheral signals provided by the database. The results of emotional recognition from peripheral signal combination are shown in the last line of Table 2 and Table 3.

From the two tables, similar observations as those of emotion recognition from a peripheral signal can be obtained. First, the emotional recognition results using EEG as privileged information are superior to the results merely using peripheral features. Second, the emotional recognition results using PCA are not always better. Third, the performance of arousal recognition is better than valence recognition.

To further evaluate the effectiveness of our model, we conduct the above 10-fold cross-validation experiment 10 times, and apply t-test to check whether the improvement using the privileged information is significant or not based on the accuracy. From the last line of Table 4, we can find that the performance of our method using EEG as privileged information increase significantly compared with that of merely using EEG features on both valence and arousal spaces, since all the p-values are less than 0.05.

Table 4. Results of t-test

		valence	arousal
DEAP	EOG	0.8137	**0.0011**
	EMG	**0.0085**	**0.0429**
	RSP	**0.0334**	0.0533
	GSR	**3.3842E-05**	**6.5233E-10**
	PLET	**2.5358E-04**	**9.5327E-08**
	TEMP	0.4065	**2.6405E-05**
	ALL	**1.0683E-05**	**2.0358E-04**
MAHNOB-HCI	ECG	**7.7740E-04**	**6.9241E-10**
	RSP	**8.9676E-05**	**1.1788E-04**
	GSR	0.2818	**0.0066**
	TEMP	**1.7356E-04**	**0.0276**
	ALL	**6.5797E-09**	**0.0061**

Table 5. Comparison results on DEAP

		Ours	peripheral signals [9]
valence	accuracy	0.581	0.627
	F1-score	0.575	0.608
arousal	accuracy	0.627	0.570
	F1-score	0.582	0.533

3.3. Comparison with related works

Koelstra et.al [9] proposed a method for emotion recognition using peripheral physiological signal combination with the same experimental conditions. We compare with their works on the DEAP database. The details are shown in Table 5. On valence space, the results are slightly worse than their works. On arousal space, compared with their work only using peripheral signals, accuracy and F1-score are increased by 0.057 and 0.049. Totally, our method is comparable with their works and even better than theirs. On the MAHNOB-HCI database, most works divided emotions into three categories. Hence, it is unable to compare with them directly.

4. CONCLUSIONS

In this paper, we propose a new emotion recognition method from peripheral signals with the help of EEG signals. Specifically, EEG signals, which have intrinsic relations in physiology with peripheral signals, are used as the privileged information, which is only required during training. The CCA is adopted to model the relationships between peripheral signals and EEG signals, and thus construct a better peripheral feature space enhanced by EEG signals during training. Then, SVM is trained to map the new peripheral physiological features to the emotional labels. Experimental results on two benchmark databases show that our method can improve the recognition performance for each individual peripheral signals and their combination on both valence and arousal spaces, verifying the effectiveness of our proposed method.

5. REFERENCES

[1] Jonghwa Kim and Elisabeth André, "Emotion recognition based on physiological changes in music listening," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 12, pp. 2067–2083, 2008.

[2] Panagiotis C Petrantonakis and Leontios J Hadjileontiadis, "Emotion recognition from eeg using higher order crossings," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 14, no. 2, pp. 186–197, 2010.

[3] Zahra Khalili and Mohammad Hasan Moradi, "Emotion recognition system using brain and peripheral signals: using correlation dimension to improve the results of eeg," in *Neural Networks, 2009. IJCNN 2009. International Joint Conference on*. IEEE, 2009, pp. 1571–1575.

[4] Kevin C Tseng, Bor-Shyh Lin, Chang-Mu Han, and Psi-Shi Wang, "Emotion recognition of eeg underlying favourite music by support vector machine," in *Orange Technologies (ICOT), 2013 International Conference on*. IEEE, 2013, pp. 155–158.

[5] Raul Fernandez, "Stochastic modeling of physiological signals with hidden markov models: A step toward frustration detection in human-computer interfaces," M.S. thesis, Citeseer, 1998.

[6] Vladimir Vapnik and Akshay Vashist, "A new learning paradigm: Learning using privileged information," *Neural Networks*, vol. 22, no. 5, pp. 544–557, 2009.

[7] Yachen Zhu, Shangfei Wang, and Qiang Ji, "Emotion recognition from users' eeg signals with the help of stimulus videos," in *Multimedia and Expo (ICME), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1–6.

[8] Bruce Thompson, "Canonical correlation analysis," *Encyclopedia of statistics in behavioral science*, 2005.

[9] Sander Koelstra, Christian Muhl, Mohammad Soleymani, Jong-Seok Lee, Ashkan Yazdani, Touradj Ebrahimi, Thierry Pun, Anton Nijholt, and Ioannis Patras, "Deap: A database for emotion analysis; using physiological signals," *Affective Computing, IEEE Transactions on*, vol. 3, no. 1, pp. 18–31, 2012.

[10] Mohammad Soleymani, Jeroen Lichtenauer, Thierry Pun, and Maja Pantic, "A multimodal database for affect recognition and implicit tagging," *Affective Computing, IEEE Transactions on*, vol. 3, no. 1, pp. 42–55, 2012.