EEG-Based Emotion Recognition with Manifold Regularized Extreme Learning Machine

Yong Peng, Jia-Yi Zhu, Wei-Long Zheng and Bao-Liang Lu* *Senior Member, IEEE*

***Abstract*— EEG signals, which can record the electrical activity along the scalp, provide researchers a reliable channel for investigating human emotional states. In this paper, a new algorithm, manifold regularized extreme learning machine (MRELM), is proposed for recognizing human emotional states (positive, neutral and negative) from EEG data, which were previously evoked by watching different types of movie clips. The MRELM can simultaneously consider the geometrical structure and discriminative information in EEG data. Using differential entropy features across whole five frequency bands, the average accuracy of MRELM is 81.01%, which is better than those obtained by GELM (80.25%) and SVM (76.62%). The accuracies obtained from high frequency band features (β, γ) are obviously superior to those of low frequency band features, which shows β and γ bands are more relevant to emotional states transition. Moreover, experiments are conducted to further evaluate the efficacy of MRELM, where the training and test sets are from different sessions. The results demonstrate that the proposed MRELM is a competitive model for EEG-based emotion recognition.**

I. INTRODUCTION

Emotion is an overall performance of human's consciousness, which can significantly affect human's action towards peripheral environment and thus it always plays an important role in our daily lives especially in human-human interaction. It is relatively easier for people to recognize others' emotional states. However, how to detect and model users' emotional states with advanced artificial intelligence techniques is a challenging topic within the human-machine interaction community [1]. Recently, many researchers from neuroscience, psychology, neural engineering and computer science fields have been focusing on emotion recognition based on machine learning techniques [2], [4].

Both physiological and non-physiological signals can be used for emotion recognition. The former signals including electroencephalography (EEG), electromyogram (EMG), electrocardiogram (ECG) and respiration signals are considered to be more effective and reliable because human cannot control them intentionally. Among them, EEG-based emotion recognition has received increasing attention recently. Li and Lu showed that γ band EEG signals are more suitable to classify happiness and sadness with high time resolution [3]. Wang *et al.* used ISOMAP to estimate the emotional states and the obtained trajectory is consistent with the transition of emotional states [4]. Deep Belief Networks (DBN), one of the most popular deep learning models, was employed to classify two types of emotional states (positive and negative) from EEG [5].

In this paper, we propose a new classifier, termed Manifold Regularized Extreme Learning Machine (MRELM), to perform emotion recognition based on the differential entropy (DE) feature, which has been shown effective for depicting the emotional information in EEG [6]. MRELM can take both data geometrical structure and discriminative information into consideration. Extensive experimental results show that MRELM outperforms GELM (an ELM variant previously proposed in [7]) and SVM in EEG-based emotion recognition.

II. EXPERIMENTS

A. Stimuli

In order to evoke emotions of subjects, we chose several 4-minute movie clips as stimuli. These movie clips could be divided into three categories: positive, neutral and negative. There were 15 clips in each session and 5 clips for each emotional state. All the movies were in Chinese and easy to understand. The stimuli were only made up of popular movies which are *After Shock*, *Lost in Thailand*, *Just Another Pandora's Box*, *World Heritage in China*, *Untitled Remembering 1942 Project*, and *Flirting Scholar*. We supposed that the 4-minute movie clip could contain an vivid and relatively complete story, so that subjects were able to maintain certain typical emotional state during the 4 minutes.

B. Subjects

Three men and three women aged between 20 to 27 participated in the experiment for three times each, with the interval of about one week. They were all right-handed and had no history of mental illness. The experiments were performed in the day time and subjects were asked to have adequate sleep in the day before experiment. Before the experiments, subjects did sign informed consent forms, and that the experiment followed the Helsinki ethical guidelines.

C. Procedure

A 62-channel electrode cap according to the extended international 10-20 and ESI NeuroScan System were used to record EEG data with sampling rate 1000Hz. 15 movie clips were played with a 10s rest and a 15s hint between

Y. Peng, J.-Y. Zhu, W.-L. Zheng and B.-L. Lu are with Center for Brain-Like Computing and Machine Intelligence, Department of Computer Science and Engineering, Shanghai Jiao Tong University and Key Lab. of Shanghai Education Commission for Intelligent Interaction and Cognitive Engineering, 800 Dongchuan Road, Shanghai 200240 China.

*Corresponding author: Bao-Liang Lu (bllu@sjtu.edu.cn). This work was partially supported by the 973 Program of China (No. 2013CB329401), the NSFC (No. 61272248), and the Sci. & Tech. Commission of Shanghai Municipality (No. 13511500200). The first author was supported by China Scholarship Council and this work was mainly done when he was visiting the EECS Dept., University of Michigan, Ann Arbor.

two clips. During the rest time, subjects were asked to fill a form as feedback to show whether the emotional states were successfully evoked. Figure 1 is the experimental procedure.

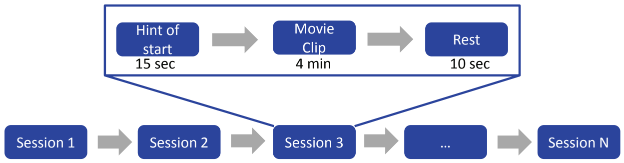

Fig. 1. Procedure of stimuli playing.

III. METHODS

A. Feature Extraction and Preprocessing

Due to the effectiveness of DE in modeling emotion information from EEG signals [6], we chose DE instead of energy spectrum (ES) [8] as feature. The definition of DE is as follows

$$h(X) = -\int_X f(x)\log(f(x))dx,$$

where time series $X \sim \mathcal{N}(\mu, \sigma^2)$, and the length of EEG sequence is fixed. Thus, DE $h(X)$ can be calculated by

$$\int_{-\infty}^{+\infty} -\frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}} \log \frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi\sigma^2}} dx.$$

It can be easily found that the DE definition is equivalent to the logarithm of ES and thus we have $h(X) = \frac{\log(2\pi e\sigma^2)}{2}$. We extracted features on the five common frequency bands of EEG. They are δ(1-3Hz), θ(4-7Hz), α(8-13Hz), β(14-30Hz) and γ(31-50Hz). Short-time Fourier transform (STFT) with 1s non-overlapping Hanning window was used to calculate the average DE features of each channel on these bands. Each frequency band signal has 62 channels and thus 310 dimensional features were obtained for each example. Since the effective experimental time lasted for 57 minutes, we finally got 3400 examples for each session.

Linear Dynamic System (LDS) [9] was used to remove the rapid changes of EEG features and get more reliable examples.

B. Classification

In this section we introduce the MRELM model formulation. ELM was proposed for training single layer feed forward neural networks (SLFNs) [10], which generates the input weights and hidden layer bias randomly. Given training samples $\{(\mathbf{x}_i, \mathbf{t}_i)\}_{i=1}^N$, where $\mathbf{x}_i = (x_{i1}, \cdots, x_{id})^T$ is the input and $\mathbf{t}_i = (t_{i1}, \cdots, t_{im})^T$ is the target. Assuming L is the number of hidden units, the ELM output function is

$$f_L(\mathbf{x}) = \sum\nolimits_{i=1}^{L} \beta_i h_i(\mathbf{x}) = \mathbf{h}(\mathbf{x})\boldsymbol{\beta}, \tag{1}$$

where $\boldsymbol{\beta} = [\beta_1, \beta_2, \cdots, \beta_L]^T$ is the output weight matrix between the hidden layer and the output layer, and $\mathbf{h}(\mathbf{x}) = [h_1(\mathbf{x}), \cdots, h_L(\mathbf{x})]$ is the output (row) vector of the hidden layer with respect to the input $\mathbf{x}$. $\mathbf{h}(\mathbf{x})$ actually maps the data from the d-dimensional input space to the L-dimensional hidden layer feature space (*ELM feature space*) H.

ELM aims at minimizing the training error as $\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|^2$, where $\mathbf{H}$ is the hidden layer representation

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}(\mathbf{x}_1) \\ \mathbf{h}(\mathbf{x}_2) \\ \vdots \\ \mathbf{h}(\mathbf{x}_N) \end{bmatrix} = \begin{bmatrix} h_1(\mathbf{x}_1) & h_2(\mathbf{x}_1) & \cdots & h_L(\mathbf{x}_1) \\ h_1(\mathbf{x}_2) & h_2(\mathbf{x}_2) & \cdots & h_L(\mathbf{x}_2) \\ \vdots & \vdots & \vdots & \vdots \\ h_1(\mathbf{x}_N) & h_2(\mathbf{x}_N) & \cdots & h_L(\mathbf{x}_N) \end{bmatrix}.$$

The output weight matrix $\boldsymbol{\beta}$ can be estimated analytically as

$$\hat{\boldsymbol{\beta}} = \arg\min_{\beta} ||\mathbf{H}\boldsymbol{\beta} - \mathbf{T}||_2^2 = \mathbf{H}^{\dagger}\mathbf{T}. \tag{2}$$

GELM [7], a variant of ELM, which has the following objective

$$\min_{\boldsymbol{\beta}} \ \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|_2^2 + \lambda_1 \mathrm{Tr}((\mathbf{H}\boldsymbol{\beta})^T \mathbf{L}(\mathbf{H}\boldsymbol{\beta})) + \lambda_2 \|\boldsymbol{\beta}\|_2^2, \tag{3}$$

only considers the label consistency property of data and enforces the output of samples in the same class to be similar. However, many researches showed that learning performance can be significantly enhanced if the geometrical structure of data is exploited [11] and thus both geometrical structure and discriminative information were proven effective in discriminative tasks.

Therefore, MRELM emphasizes both above mentioned aspects by defining the within-class graph G^w and between-class graph G^b. The corresponding affinity matrices are

$$W_{ij}^w = \begin{cases} 1, & \text{if } \mathbf{x}_i \in \mathcal{N}_{k1}(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in \mathcal{N}_{k1}(\mathbf{x}_i) \\ & \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ are from the same class} \\ 0, & \text{otherwise.} \end{cases}$$

and

$$W_{ij}^b = \begin{cases} 1, & \text{if } \mathbf{x}_i \in \mathcal{N}_{k2}(\mathbf{x}_j) \text{ or } \mathbf{x}_j \in \mathcal{N}_{k2}(\mathbf{x}_i) \\ & \mathbf{x}_i \text{ and } \mathbf{x}_j \text{ are from different classes} \\ 0, & \text{otherwise.} \end{cases}$$

The two graph Laplacian matrices $\mathbf{L}_w$ and $\mathbf{L}_b$ are defined as $\mathbf{L}_w = \mathbf{D}^w - \mathbf{W}^w$ ($D_{ii}^w = \sum_i W_{ij}^w$) and $\mathbf{L}_b = \mathbf{D}^b - \mathbf{W}^b$ ($D_{ii}^b = \sum_i W_{ij}^b$) respectively. Based on the definitions of $\mathbf{L}_w$ and $\mathbf{L}_b$, we need to minimize $\mathrm{Tr}((\mathbf{H}\boldsymbol{\beta})^T \mathbf{L}_w (\mathbf{H}\boldsymbol{\beta}))$ (to retain the data geometric structure) and maximize $\mathrm{Tr}((\mathbf{H}\boldsymbol{\beta})^T \mathbf{L}_b (\mathbf{H}\boldsymbol{\beta}))$ (to enforce the discriminative information) simultaneously. Therefore, the new graph Laplacian $\mathbf{L}_{new}$ is designed as $(\mathbf{L}_b^{-1/2})^T \mathbf{L}_w \mathbf{L}_b^{-1/2}$ to cover both aspects together.

As a result, we formulate the objective of MRELM as

$$\min_{\boldsymbol{\beta}} \ \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\|_2^2 + \lambda_1 \mathrm{Tr}((\mathbf{H}\boldsymbol{\beta})^T \mathbf{L}_{new} (\mathbf{H}\boldsymbol{\beta})) + \lambda_2 \|\boldsymbol{\beta}\|_2^2, \tag{4}$$

and its solution as

$$\boldsymbol{\beta}^* = (\mathbf{H}^T\mathbf{H} + \lambda_1 \mathbf{H}^T \mathbf{L}_{new} \mathbf{H} + \lambda_2 \mathbf{I})^{-1} \mathbf{H}^T \mathbf{T}.$$

IV. RESULTS AND DISCUSSIONS

We got about 3400 samples for each session and then choose about 2000 samples as training set, the rest in the same session as test set. In order to investigate the stability of MRELM model, we also choose data from one session as training set and the data from another session as test set. Since we have collected three sessions data for each subject, training sets and test sets could be from the same subject's different sessions, but also from different subjects.

A. Experimental Paradigm 1

For each subject, we use training sets and test sets from the same session. Tables I and II show the results of linear-SVM, GELM and MRELM classifiers using DE features on δ, θ, α, β and γ frequency bands as input. Obviously, the performance of MRELM is consistently better than those of GELM and SVM in most cases. If using all frequency band features, the average accuracy across all subjects of MRELM is 81.01%, which is nearly 1% improvement w.r.t GELM (80.25%) and 4.5% w.r.t. SVM (76.62%). Similar results can be found on each frequency band. If comparing the results of different bands, we can find that the accuracies based on β and γ band signals are much higher than remaining bands. This reflects that the transition of emotional states may have closer connection to these two frequency bands.

Table III is the average confusion matrix of three classifiers based on 310 DE features. From this table, we can see that the positive and neutral states are much easier to be recognized while the negative state is difficult to estimate. The MRELM can obtain a 5%-7% accuracy improvement when estimating the negative state w.r.t. GELM and SVM.

B. Experimental Paradigm 2

In this experiment, we use training sets and test sets from different sessions. As shown in Tables IV and V, MRELM still achieves the best average results. The accuracies whose training and test sets are from different sessions is obvious lower than those whose training and test sets are from the same session. The average accuracy for all subjects in experiment 1 is 81.01% while this value is 72.76% for experiment 2. Though there is much loss in accuracy, 72.76% is still a relatively good result for three-class emotion recognition problem which implies that the transition of EEG patterns among different sessions are stable for the same subject.

For evaluating the models based on subject-independent features and investigating the stability of common patterns across subjects, we use D1 as training set and test the data from different sessions of other subjects. The best performance of test sets from subjects A, B, C, E and F are A1(75.02%), B1(69.46%), C2(69.33%), E3(64.73%) and F3(69.86%), respectively. It suggests that the transition of EEG patterns still have strong underlying regularity. And this regularity makes that learning models can be generally used for EEG-based emotion recognition among different subjects.

Figure 2 shows the more obvious visualization of average performance obtained by three classifiers over different setting of training and test sets.

V. Conclusions

In this paper, we worked on recognizing the emotional states (positive, neutral and negative) using EEG data. A new classifier, manifold regularized extreme learning machine (MRELM), was proposed to classify the differential entropy features. Experimental results demonstrated that MRELM is an excellent classifier for EEG-based emotion recognition. Morevoer, we had several observations: 1) the β and γ band features are more related to the transition of emotional states;

TABLE I

Emotion recognition accuracies(%) of SVM(M1), GELM(M2) and MRLEM(M3) for six subjects (A,B,C,D,E,F).

A	Session 1			Session 2			Session 3		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
δ	49.93	54.84	57.88	37.57	46.10	46.03	46.75	50.14	54.48
θ	60.26	61.20	62.21	49.35	49.78	55.85	58.31	54.26	55.28
α	65.17	70.01	71.89	54.41	55.35	57.51	48.13	54.26	55.28
β	**84.10**	85.19	85.19	65.46	66.18	68.71	57.15	**66.26**	**68.93**
γ	81.50	**86.64**	**88.01**	67.27	**75.07**	**75.87**	59.54	61.92	65.32
Total	82.59	84.39	85.26	**75.65**	70.09	72.40	**59.90**	63.95	65.39

B	Session 1			Session 2			Session 3		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
δ	53.47	58.09	62.14	38.73	51.30	54.41	52.02	54.55	59.25
θ	57.59	63.44	65.25	55.92	58.89	60.33	52.38	58.82	57.80
α	72.83	82.73	83.74	65.75	65.10	65.82	65.10	71.32	71.82
β	**90.17**	88.08	89.96	69.44	**69.65**	71.89	**78.97**	**82.30**	**84.39**
γ	89.52	**90.90**	91.19	**70.66**	69.22	69.73	77.24	77.75	79.55
Total	88.15	89.45	**92.63**	65.82	69.15	**72.47**	71.28	79.48	79.33

C	Session 1			Session 2			Session 3		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
δ	50.79	58.45	60.33	35.77	40.97	40.03	44.73	46.97	49.78
θ	69.44	67.05	69.29	49.57	50.58	51.88	43.93	40.75	42.41
α	61.13	61.34	66.11	50.43	52.89	54.55	49.21	45.07	46.60
β	**77.24**	79.19	78.25	90.03	90.75	92.34	58.60	54.62	59.61
γ	76.37	80.92	77.82	89.45	89.96	90.46	59.18	58.45	60.26
Total	76.52	**82.37**	**83.53**	**91.11**	**92.99**	**93.14**	**61.20**	**67.85**	**60.48**

D	Session 1			Session 2			Session 3		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
δ	75.87	78.61	79.62	60.33	58.96	61.92	58.09	62.72	64.81
θ	73.92	84.90	79.48	56.00	61.34	58.89	55.78	60.91	63.66
α	70.16	88.01	87.79	80.56	85.33	86.06	80.27	90.10	90.25
β	92.99	**96.89**	**97.18**	88.09	95.30	95.74	97.18	**96.82**	96.82
γ	90.68	96.60	96.89	**91.98**	**96.89**	**97.25**	96.32	95.74	96.32
Total	**96.68**	96.68	97.11	91.04	**96.89**	96.82	**97.25**	96.53	**97.54**

E	Session 1			Session 2			Session 3		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
δ	58.89	56.50	57.88	55.85	58.45	60.41	48.70	53.18	54.19
θ	66.47	65.17	63.87	40.25	43.61	47.62	40.10	43.71	45.59
α	46.89	58.02	61.06	34.39	51.30	51.95	60.69	63.58	67.56
β	67.12	74.64	76.37	53.90	**74.35**	75.07	63.08	73.77	**75.65**
γ	**76.89**	**80.35**	**81.36**	**70.66**	73.92	**76.66**	63.29	66.98	68.14
Total	70.01	73.19	75.94	60.19	73.19	75.43	**73.99**	**74.57**	71.10

F	Session 1			Session 2			Session 3		
	M1	M2	M3	M1	M2	M3	M1	M2	M3
δ	69.65	72.25	72.04	45.16	40.25	41.33	55.85	57.73	58.17
θ	58.24	59.47	60.79	46.82	49.21	50.29	63.44	64.38	64.31
α	60.48	64.74	67.05	53.11	57.08	57.87	66.84	69.15	69.94
β	**73.19**	80.20	81.72	**59.25**	57.88	59.32	88.29	91.18	93.43
γ	69.80	**85.98**	**87.14**	58.22	57.08	59.39	**93.86**	**94.29**	**94.08**
Total	**73.19**	84.32	85.55	56.50	**59.25**	**62.28**	87.50	90.10	91.76

"Total" means concatenating features from all five frequency bands.

TABLE II

Average performances of SVM, GELM and MRELM in experimental paradigm 1 (mean±std%).

Freq. Band	Mean±Std(%)		
	SVM	GELM	MRELM
δ	52.12±10.46	55.56± 9.60	**57.48± 9.71**
θ	55.43± 9.46	57.93±10.49	**58.93± 9.01**
α	60.31±12.10	65.85±13.34	**67.38±12.87**
β	75.24±14.00	79.07±12.94	**80.59±12.17**
γ	76.84±12.76	79.93±13.24	**80.82±12.66**
total	76.62±13.12	80.25±11.92	**81.01±12.24**

TABLE III

CONFUSION MATRICES OF SVM, GELM AND MRELM IN EXPERIMENTAL PARADIGM 1 (MEAN±STD%).

SVM	Negative	Neutral	Positive
Negative	**58.73±31.53**	24.16±22.28	17.11±15.91
Neutral	10.97±11.68	**79.18±16.24**	9.85±9.52
Positive	6.58±7.75	2.80±4.08	**90.62±10.15**
GELM	Negative	Neutral	Positive
Negative	**60.63±29.23**	24.01±20.02	15.36±14.87
Neutral	7.66±12.08	**85.63±15.83**	6.71±7.95
Positive	3.93±6.26	3.00±3.88	**93.07±7.54**
MRELM	Negative	Neutral	Positive
Negative	**65.98±24.18**	20.46±19.02	13.55±13.36
Neutral	8.79±11.83	**83.85±15.68**	7.35±8.19
Positive	4.30±7.86	3.61±3.51	**92.08±8.74**

TABLE IV

EMOTION RECOGNITION ACCURACIES(%) OF SVM(M1), GELM(M2) AND MRELM(M3) IN EXPERIMENTAL PARADIGM 2.

ALGs.		A1*	A2	A3		B1	B2	B3
M1	A1	82.59	53.83	47.25	B1	88.15	37.21	61.05
M2		84.39	**63.85**	55.49		89.45	61.56	**77.07**
M3		**85.26**	62.64	**62.85**		**92.63**	**63.82**	61.61
M1	A2	**67.27**	**75.65**	48.34	B2	51.95	65.82	64.52
M2		66.91	70.09	50.79		72.90	69.15	73.70
M3		65.76	72.40	**60.93**		**75.22**	**72.47**	**76.81**
M1	A3	37.57	53.61	59.90	B3	68.42	52.24	71.82
M2		**74.06**	**65.10**	63.95		69.95	**65.46**	**79.48**
M3		68.03	62.55	**65.39**		**75.10**	65.09	79.33
M1	D1	66.33	52.46	41.98	D1	49.71	58.02	57.66
M2		**75.87**	60.40	51.01		67.27	49.21	**59.47**
M3		75.02	**62.11**	**64.35**		**69.46**	**61.83**	59.16
ALGs.		C1	C2	C3		D1	D2	D3
M1	C1	76.52	82.88	67.62	D1	96.68	**91.11**	88.22
M2		82.37	**89.02**	67.05		96.68	89.60	**88.58**
M3		**83.53**	81.73	**77.31**		**97.11**	90.87	87.27
M1	C2	55.92	91.11	61.71	D2	**90.17**	91.04	**96.89**
M2		67.77	92.99	74.71		88.08	**96.89**	95.23
M3		**76.87**	**93.14**	**79.32**		90.07	96.82	95.37
M1	C3	76.52	75.29	61.20	D3	76.95	92.49	97.25
M2		75.29	80.42	67.85		80.49	**95.95**	96.53
M3		**79.38**	**85.45**	**70.48**		**82.20**	92.96	**97.54**
M1	D1	58.89	57.15	53.76	D1	96.68	**91.11**	88.22
M2		**65.82**	**71.75**	**63.76**		96.68	89.60	**88.58**
M3		65.72	69.33	62.11		**97.11**	90.87	87.27
ALGs.		E1	E2	E3		F1	F2	F3
M1	E1	70.01	58.31	**57.15**	F1	73.19	44.65	51.45
M2		73.19	66.69	53.11		84.32	42.34	59.54
M3		**75.94**	**73.95**	57.13		**85.55**	**58.30**	**59.90**
M1	E2	54.99	60.19	45.09	F2	59.03	56.50	44.51
M2		68.28	73.19	51.08		77.60	59.25	64.38
M3		**75.58**	**75.43**	**60.84**		**79.08**	**62.28**	**78.34**
M1	E3	47.69	47.69	73.99	F3	60.69	58.89	87.50
M2		**62.43**	**58.16**	**74.57**		74.35	59.47	90.10
M3		59.66	55.36	71.10		**76.38**	**65.43**	**91.76**
M1	D1	58.67	40.03	46.03	D1	48.19	48.98	70.89
M2		61.20	52.96	60.84		**52.82**	**59.25**	**71.75**
M3		**63.07**	**55.63**	**64.73**		51.80	55.83	69.86*

"A1" means the first session data of subject A, and so on. For example, the value 69.86 in bottom right corner is obtained by feeding D1 as training data and F3 as test data into MRELM.

TABLE V

AVERAGE PERFORMANCES OF SVM, GELM AND MRELM IN EXPERIMENTAL PARADIGM 2 (MEAN±STD%).

ALGs.		1	2	3
SVM	1	81.89±10.01	61.33±21.34	62.13±14.65
GELM		85.07±7.79	68.86±18.01	66.91±13.86
MRELM		**86.67±7.38**	**71.89±12.62**	**67.67±11.90**
SVM	2	63.22±14.20	73.39±15.15	60.18±19.91
GELM		73.59±8.16	76.93±13.86	68.32±16.81
MRELM		**77.10±7.83**	**78.76±13.38**	**75.27±13.01**
SVM	3	61.31±15.97	63.37±17.18	75.28±14.70
GELM		72.71±6.11	70.76±14.66	78.75±12.66
MRELM		**73.46±8.28**	**71.14±14.65**	**79.27±12.86**
SVM	D1	56.36±7.45	51.32±7.30	54.06±11.26
GELM		64.60±8.46	58.71±8.61	63.30±9.29
MRELM		**65.01±8.64**	**60.95±5.63**	**64.04±3.94**

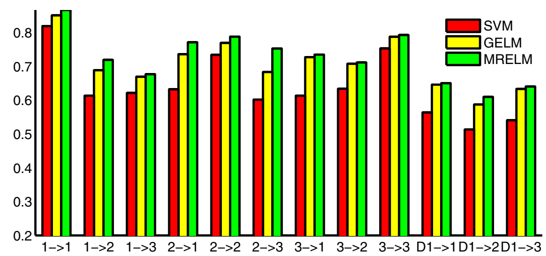

Fig. 2. Average performances of 3 classifiers in experimental paradigm 2.

2) positive state are easiest to be estimated than the other two states; 3) the connection between emotional states and EEG is stable among different sessions and different subjects.

REFERENCES

[1] R. Cowie, E. Douglas-Cowie, N. Tsapatsoulis, G. Votsis, S. Kollias, W. Fellenz, and J.-G. Taylor, "Emotion recognition in human-computer interaction", *IEEE Signal Proc. Mag.*, vol.18, no.1, pp.32-80, 2001.

[2] R.A Calvo, and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications", *IEEE Trans. Affect. Comput.*, vol.1, no.1, pp.18-37, 2010.

[3] M. Li, and B.-L. Lu, "Emotion classification based on gamma-band EEG", in *Proc. 31st Ann. Int. Conf. IEEE EMBS*, pp.2037-2040, 2009.

[4] X.-W. Wang, D. Nie and B.-L. Lu, "Emotional state classification from EEG data using machine learning approach", *Neurocomputing*, vol.129, pp.94-106, 2014.

[5] W.-L. Zheng, J.-Y. Zhu, R.-N. Duan, Y. Peng and B.-L. Lu, "EEG-based emotion classification using deep belief networks", To appear in *Proc. Int. Conf. Multim. & Expo*, 2014. (`http://bcmi.sjtu.edu.cn/~pengyong/`)

[6] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification", in *Proc. 6th Int. IEEE/EMBS Conf. Neural Eng.*, pp.81-84, 2013.

[7] Y. Peng, S. Wang, X. Long, and B.-L. Lu, "Discriminative graph regularized extreme learning machine for face recognition", To appear in *Neurocomputing* (`http://bcmi.sjtu.edu.cn/~pengyong/`).

[8] L.-C. Shi, Y.-Y. Jiao, and B.-L. Lu, Differential entropy feature for EEG-based vigilance estimation, in *Proc. 35th Ann. Int. Conf. IEEE EMBS*, pp. 6627-6630, 2013.

[9] L.-C. Shi and B.-L. Lu, "Off-line and on-line vigilance estimation based on linear dynamical system and manifold learning", in *Proc. 32nd Ann. Int. Conf. IEEE EMBS*, pp.6587–6590, 2010.

[10] G.-B. Huang, Q.-Y. Zhou, and C.-K. Siew, "Extreme learning machine: theory and applications", *Neurocomputing*, vol.70, no.1, pp.489-501, 2006.

[11] N. Guan, D. Tao, Z. Luo and B. Yuan, "Manifold regularized discriminative nonnegative matrix factorization with fast gradient descent", *IEEE Trans. Image Process.*, vol.20, no.7, pp.2030-2048, 2011.