Emotion Recognition Based on Multi-Variant Correlation of Physiological Signals

Wanhui Wen, Guangyuan Liu, Nanpu Cheng, Jie Wei, Pengchao Shangguan, and Wenjin Huang

Abstract—Emotion recognition based on affective physiological changes is a pattern recognition problem, and selecting specific physiological signals is necessary and helpful to recognize the emotions. Fingertip blood oxygen saturation (OXY), galvanic skin response (GSR) and heart rate (HR) are acquired while amusement, anger, grief and fear of 101 subjects are individually elicited by films. The affective physiological changes in multi-subject GSR, the first derivative of GSR (FD_GSR) and HR are detected by the multi-variant correlation method. The correlation analysis reveals that multi-subject HR, GSR and FD_GSR fluctuations respectively have common intra-class affective patterns. In addition to the conventional features of HR and GSR, the affective HR, GSR and FD_GSR fluctuations are quantified by the local scaling dimension and applied as the affective features. The multi-subject affective database containing 477 cases is classified by a Random Forests classifier. An overall correct rate of 74 percent for quinary classification of amusement, anger, grief, fear and the baseline state are obtained.

Index Terms—Affective pattern recognition, local scaling dimension, multi-variant correlation, physiological signal

1 Introduction

Due to the important role of emotion in decision making, problem solving, negotiation and many other social activities, more and more researches on computer-based human emotion recognition have been done and applied in multimedia digital entertainment [1], safe driving [2], [3], health care [4], social security [5] and other fields.

Although the fabric of emotion itself is elusive, the scientists gradually reach an agreement on the main process of emotion. They think that emotion is a series of reactions when individuals face a situation related to their main target. Lang [6] believes that emotions are multidimensional, including the changes in cognition, experience, central and peripheral physiology and behavior response system, and he proposes a two dimensional (arousal/valence) space to describe emotion. Ekman et al. [7] proposes a discrete model of emotion, and he thinks the basic discrete emotions are joy, surprise, disgust, grief, anger and fear. Although there are many other models of emotion, the models proposed by Lang and Ekman are most widely accepted.

Picard et al. [8] believes that the user's emotional changes are embodied in speech [9], facial expressions [10], body posture [11], central nervous system and autonomic nerve physiological activities [12], etc. Voice and facial expression can be deliberately hided by the user at some social occasions, but autonomic nerve physiological activities can not be controlled by the user's subjective intention. Since autonomic nerve physiological activities are honest emotion indicators, many physiological measures are used to identify emotions in literature. The intensity of emotion can affect the skin conductance (SC), heart rate (HR), electrocardiography (ECG) and respiration, and the valence of emotion is reflected by the activities of different facial muscles [12]. Although most researches tend to use a variety of emotion measures to identify emotions, Agrafioti et al. [13] has said that too many emotion measurements would cause interference to the subjects and not be conducive to practical applications. Therefore, he only adopts the emotion features of the ECG signal [13].

1.1 Related Research Work

Researchers in the field of affective physiological pattern recognition usually use some emotion elicitation materials to elicit certain emotions of the subject(s), record their physiological measures, extract affective features from these physiological measures and classify the affective physiological data through certain kinds of classifiers. Table 1 lists part of the related work in affective physiological pattern recognition.

Katsis et al. [4], Gross and Levenson [14], Rani et al. [15], Kim and Andre [16], etc. use IAPS images, music, cognitive tasks, complicated mathematical problems, film clips, questionnaires and other methods to induce emotions. The benefit of IAPS images to induce emotion is that they have experimentally determined degree of arousal and valence. However, they may not be able to evoke some strong emotions. Kim et al. [17] thinks that static images as visual stimuli are insufficient to effectively induce emotion, and he has designed a combination of blue light, background sad music and sad story narration to induce sadness. Kim and Andre [16] has applied music to evoke positive/negative and high/low arousal emotions and found that electrocardiogram

- *W. Wen, G. Liu, J. Wei, P. Shangguan, and W. Huang are with the School of Electronic and Information Engineering, Southwest China University, Chongqing 400715, China. E-mail: {cwenwanh, liugy, hybyyhcx, talent}@ swu.edu.cn, shangguanpengchao@163.com*
- *N. Cheng is with the School of Materials Science and Engineering, Southwest China University, Chongqing 400715, China. E-mail: cheng_np@swu.edu.cn.*

Manuscript received 24 Aug. 2013; revised 22 May 2014; accepted 25 May 2014. Date of publication 3 June 2014; date of current version 23 July 2014. Recommended for acceptance by J.-M. Fellous.
For information on obtaining reprints of this article, please send e-mail to: reprints@ieee.org, and reference the Digital Object Identifier below.
Digital Object Identifier no. 10.1109/TAFFC.2014.2327617

TABLE 1
Related Work in Affective Physiological Pattern Recognition

Author	Emotion elicitation method	Emotions elicited	# subjects	Measures	Data analysis technique	Correct rates
Katsis [4]	IAPS	Relaxed, neutral, startled, apprehensive, very apprehensive	2 men, 3 women	Blood volume pressure (BVP), heart rate (HR), Galvanic skin response (GSR), respiration (RSP)	Artificial Neural networks (NN), Support vector machine (SVM), random forest (RF), neuro-fuzzy system (NFS)	84.3 percent, leave-one-out cross-validation (LOOCV)
Picard [8]	Personal imagery	No emotion, anger, hate, grief, platonic love, romantic love, joy, reverence	1 actress	BVP, HR, RSP, Facial electromyogram (EMG), Skin conductance (SC)	Sequential floating forward search (SFFS), Fisher Projection (FP)	81 percent, LOOCV
Agrafioti [13]	video game, IAPS	Active arousal, passive arousal & valence	42/32	Electrocardiogram (ECG)	Empirical mode decomposition, Hilbert-Huang transform	78.43 percent for active arousal, LOOCV
Rani [15]	Computer based cognitive tasks	Frustration, engagement, anxiety, boredom, anger	8 women, 7 men	photoplethysmogram (PPG), ECG, EMG, electrodermal activity (EDA), temperature	K-nearest neighbor (KNN), regression tree, Bayesian networks and SVM	82.81 percent-88.86 percent, LOOCV
Kim [16]	music	Positive, negative, high/low arousal	3 male	EMG, ECG, SC, RSP	Linear discriminant analysis, emotion specific multilevel dichotomous classification	95 percent: subject dependent; 70 percent: subject independent, LOOCV
Kim [17]	Audio, visual and cognitive approach	Sadness, anger, stress, surprise	125 children/ 50 children	ECG, skin temperature (ST), EDA	SVM	78.4 percent: 3 classes; 61.8 percent: 4 classes, five-fold cross-validation (CV)
Bailenson [19]	videos	Amusement, sadness, neutral	41/151	Face tracking, ECG, SC, somatic activity	WEKA software package	mono-subject: 94 percent: amusement; 98 percent: sadness, two-fold CV
Lisetti [20]	Movie clips, difficult math questions	Sadness, anger, fear, surprise, frustration, amusement	3 women, 26 men	GSR, HR, temperature (TEM)	KNN, discriminant function analysis, Marquardt backpropagation	73.9 percent-91.7 percent, LOOCV for kNN
Koelstra [21]	videos	Arousal, valence, liking	32	Electroencephalogram(EEG), electrooculogram (EOG), EMG, GSR, BVP, TEM, RSP	Single-trial voting	61.6 percent-64.7 percent
Hagg [22]	IAPS	High, medium, low arousal, positive, negative	1	EMG, EDA, ST, BVP, ECG, RSP	NN	96.58 percent: arousal; 89.93 percent: valence, 7:1.5:1.5 for training, test and validation
Wagner [23]	music	Anger, surprise, sadness, pleasure	1	EMG, ECG, SC, RSP	KNN, linear discriminant function, multilayer perceptron	92.05 percent, LOOCV
Fleureau [24]	Audio-video clips	Event detection, valence detection	2 women, 8 men	PPG, GSR, EMG	Gaussian process classifier	87.41 percent, 85.71 percent: mono-user; 84.04 percent, 85.09 percent: multi-user, two-fold CV

(ECG), respiration, skin conductance and electromyogram (EMG) can identify these emotions.

After comparing a variety of emotion eliciting methods, Jonathan et al. [18] has found that film as an emotion elicitation method requires more cognitive participation and can induce the strong emotional feeling. As the duration of film clips usually lasts much longer than the feeling of emotion, many researchers have adopted additional data selection to pick up the affective data from the whole experiment data recording. Bailenson et al. [19] codes the facial expressions of the subjects second by second with a laboratory software to rate the amount of amusement and sadness, and the average ratings are used as the criterion for affective data selection. The problem of this method is that some emotions do not have obvious facial expressions. Lisetti and Nasoz [20] determines the most likely time slots of intended emotion by film plots. In order to avoid affective data selection, Koelstra et al. [21] has selected short-length videos for emotion elicitation. In that work, the duration of the videos is only 1 minute, and additional data selection is not necessary. However, the relatively short duration of video may not be effective for the elicitation of some emotions [18].

After gathering the affective data, many data analysis techniques have been applied to affective feature selection and affective data classification in literature. Since these techniques are applied to deal with different datasets which contain different types of bio-signals and involve dissimilar emotion classes, it is hard to determine the best technique for affective data analysis. According to Breiman [25], the random forests (RM) method used by Katsis et al. [4] has three precious advantages. First, it does not over-fit in prediction because of the Law of Large Numbers. Second, it is an accurate classifier because of the randomness of input data selection and feature selection during its construction process. Furthermore, it can measure the importance of affective features by permutation.

The correct classification rate is an important index to evaluate the performance of the emotion recognition system. However, the recognition results in many related works may not be compared directly, as different definitions of correct recognition rates have been used in them. Despite this, among the emotion recognition systems in literature, the correct rates of mono-subject data are usually better than those of multi-subjects data, given the certain types of emotions. Although the correct recognition rates of user-independent systems are not as good as those of user-dependent systems, they have stronger generalization performance.

1.2 The Present Work

In this work, films are applied to elicit strong feeling of amusement, anger, grief and fear. As emotions cause body changes such as the perspiration of palms and the extraction of muscles [12], we expect that the subjects may have similar body changes when they watch the same film and experience similar emotional feelings. Therefore, before feature extraction, a multi-variant correlation method derived from random matrix theory (RMT) is used to detect the common affective physiological patterns in multi-subject HR, GSR, the first derivative of GSR and fingertip oxygen saturation (OXY). The aim of the correlation analysis is to explore the similarity of the physiological data which are acquired from a group of subjects when they individually watch the same emotion eliciting video materials. We compare the similarity of multi-subject physiological data during emotion elicitation with that of finite-length uncorrelated random data and with that of the baseline state. The comparison can reveal the influence of the emotion on the similarity of multi-subjects' certain kind of physiological signal and locate the affective physiological changes. After correlation analysis, a method named local scaling dimension (LSD) is applied to characterize the detailed fluctuation structure of the affective physiological changes. And then, the conventional affective features and the LSD affective physiological features in various time scales are extracted and applied to a random forests to distinguish amusement, anger, grief, fear and the baseline state.

The following sections are organized as: Section 2 is about the experiment settings of gathering affective data. In Section 3, we apply the multi-variant correlation analysis to the acquired physiological data and find out the influence of each emotion on the fluctuations of the physiological signals. In Section 4, affective physiological changes are located according to the findings in Section 3, and affective features are extracted using the method of local scaling dimension. Section 4 also reports the emotion classification results, and the random forests classifier gives the importance of each affective feature. Section 5 gives the discussion and conclusions.

2 Acquisition of Affective Data

According to the online audiences' comments and ratings, we select four videos from the popular film website to elicit amusement, grief, anger and fear, respectively. And then, we choose the healthy first-year university students as the subjects of emotion elicitation and affective data acquisition.

2.1 The Subjects

The subjects have been recruited from the freshmen. Before data acquisition, the subjects have been required to provide the details of the personal information, including their name, sex, age, etc. Each subject has also signed an informed consent and filled out the Eysenck personality questionnaire which can reflect the personality traits of the subjects. If the physiological signals of the subjects have pathological changes, they are refused to take part in the affective data acquisition. The subjects' information and all other experimental data are managed by our emotion data acquisition management system (EDAMS) installed on a data storage server.

2.2 The Emotion Elicitation Materials

Four films are selected from the popular film websites. The films are easy to understand, have lots of playing records and are respectively remarked as amusement, grief, anger and fear by the audience. twenty one members of the research group have pre-assessed the above four films and chosen some clips which can respectively evoke strong feelings of the above four emotions. The first clip is a flash animation combined with the comic talk voice, the second clip is from a movie named Tokyo trial, the third clip is a micro film about the loss of family love, and the last clip which is

TABLE 2
The Arrangement of the Video Materials

Time	Video #1 (for grief elicitation)	Video #2 (for amusement elicitation)	Video #3 (for fear elicitation)	Video #4 (for anger elicitation)
0 s-10 s	Black screen	Black screen	Black screen	Black screen
11 s-40 s	Welcome to take part in our experiment. You will see several scenery pictures. Please relax yourself.	Welcome to take part in our experiment. You will see several scenery pictures. Please relax yourself.	Welcome to take part in our experiment. You will see several scenery pictures. Please relax yourself.	Welcome to take part in our experiment. You will see several scenery pictures. Please relax yourself.
41 s-70 s	The scenery picture is going to play. Please wait a moment.	The scenery picture is going to play. Please wait a moment.	The scenery picture is going to play. Please wait a moment.	The scenery picture is going to play. Please wait a moment.
71 s-190 s	12 scenery pictures with background soft music	12 scenery pictures with background soft music	12 scenery pictures with background soft music	12 scenery pictures with background soft music
191 s-200 s	Black screen	Black screen	Black screen	Black screen
201 s-229 s	Experimental guidance: "The following is a film clip. Please fully feel the emotion when you watch it."	Experimental guidance: "The following is a film clip. Please fully feel the emotion when you watch it."	Experimental guidance: "The following is a film clip. Please fully feel the emotion when you watch it."	Experimental guidance: "The following is a film clip. Please fully feel the emotion when you watch it."
230 s-239 s	Experimental guidance: "Time takes people away, cherish the family love from now on."	The flash animation starts at 230 s.	The film clip starts at 230 s.	Experimental guidance: "a war criminal whose army slaughtered millions of innocent people refuses to confess his guilt."
240 s-	The micro film starts at 240 s and ends at 831 s.	The flash animation ends at 416 s.	The film clip ends at 478 s.	The film clip starts at 239 s and ends at 511 s.
After the first play of the film clip	Experimental guidance: "Please watch the film again, and press the button when you recall strong feeling of emotion occurred during the first play." (8 seconds)	Experimental guidance: "Please watch the film again, and press the button when you recall strong feeling of emotion occurred during the first play." (9 seconds)	Black screen (10 seconds). Experimental guidance: "Please watch the film again, and press the button when you recall strong feeling of emotion occurred during the first play." (10 seconds)	Experimental guidance: "Please watch the film again, and press the button when you recall strong feeling of emotion occurred during the first play." (8 seconds)
After the guidance	The second play of the micro film	The second play of the flash animation	The second play of the film clip	The second play of the film clip
Final scene	Thank you for watching the video material. (5 seconds)	Thank you for watching the video material. (5 seconds)	Black screen (10 seconds). Thank you for watching the video material. (5 seconds)	Thank you for watching the video material. (5 seconds)

composed of two segments is from a horror film named the Ring. During the selection of the film clips, the film plots which have aroused certain target emotion of the group members are labeled as emotion-eliciting film plots.

The experiment instructions, 12 scenery pictures with soft background music and the emotion elicitation clip have been synthesized into one video material. During the experiments, the subjects follow the experiment instructions in the video and have not been interrupted by the experimenter. Before emotion elicitation, the scenery pictures and the soft music are arranged to make the subjects in a baseline state of low arousal and peacefulness. An emotion elicitation film clip is played twice in one video material. During the first play, the subjects are asked to fully feel their emotion evoked by the clip. During the second play, they are asked to recall the emotional feelings and press a button on the armrest of the subject's seat to report the occurrence of the strong emotional feelings. The arrangement of the video materials is given in Table 2.

2.3 The Data Acquisition Environment

The data acquisition laboratory is divided into the rest room, the control room and the test room. The rooms are separated by walls having the function of sound insulation. The acquisition of personal information of the subjects and the pre-experiment instruction is done in the rest room.

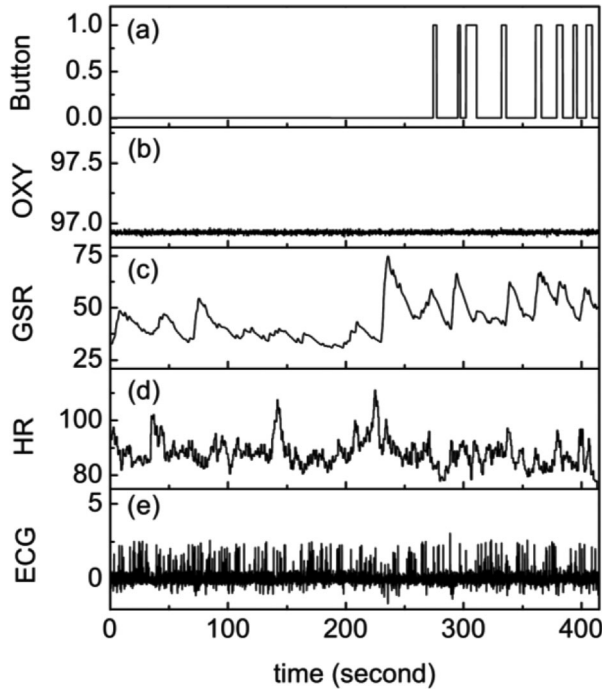


Fig. 1. The button status and the physiological recordings of an amusement elicitation experiment. The button status "1" in (a) reveals the occurrence of strong emotional feeling.

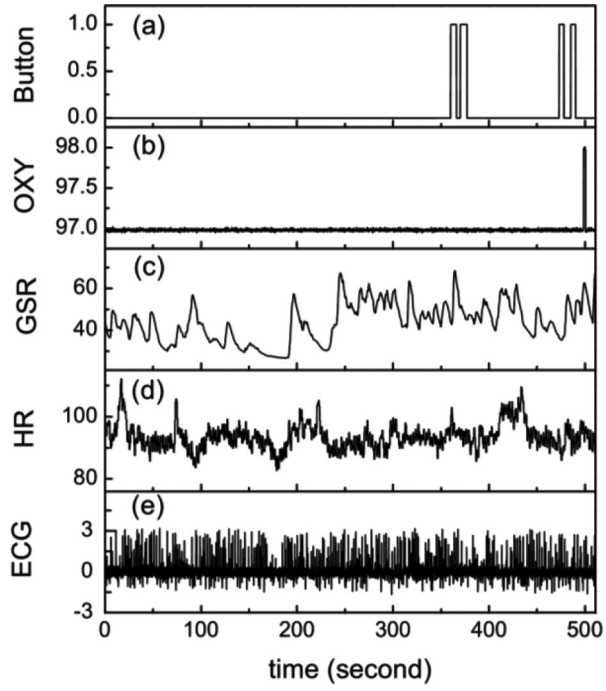


Fig. 2. The button status and the physiological recordings of an anger elicitation experiment. The button status "1" in (a) reveals the occurrence of strong emotional feeling.

There are one data storage server, one video monitor, one video-playing computer and one physiological data acquisition computer in the control room. The video playing computer and the physiological data recording computer are synchronized by a single chip microcomputer system, and the time difference between the start of video playing and the start of physiological data recording is controlled to be less than 1 second. The video playing computer has two display equipments, the first one being a 19-inch display in the control room and the second one being a 51-inch flat-screen TV hanging on the opposite wall of the subject's seat in the test room. A video probe of the video monitor is installed near the upper left of the TV.

The physiological data are acquired by Biopac MP150 multi-channel physiological recorder and the Acknowledge3.8.2 application software. The modules of OXY, ECG and GSR are respectively used to acquire fingertip blood oxygen saturation, electrocardiogram and galvanic skin conductance (GSR) of the fingers. 0.5 Hz high pass filter and 35 Hz low pass filter have been set for the ECG data acquisition, and 10 Hz low pass filter has been set for the GSR data acquisition. Through the built-in function of Acknowledge3.8.2, the R-R heart rate is calculated by ECG peak detection.

2.4 Data Acquisition

During the play of the experiment video material, three kinds of data have been recorded from individual subject. The first data are the physiological signals acquired at a sampling rate of 400 Hz, including fingertip blood oxygen saturation, ECG and GSR. The second data are the video recorded by the video monitor from which the facial expression and possible body movement can be marked by the experimenter. The clock of the video monitor is adjusted to be the same to that of the video playing computer. The third data acquired at a sampling rate of 10 Hz are the status of the button fixed on the armrest of the subject's seat. The button status acquisition starts synchronously at the start of the video play. '0' of the button status represents the default status of the button, and '1' of the button status reveals the occurrence of strong emotional feeling of the subject.

After synchronous data acquisition, the subjects are required to fill in an emotional experience scale which is composed of 27 emotional adjectives respectively describing the feeling of amusement, anger, grief, fear and calmness. The adjectives in the same emotion category are synonyms and are linearly quantified into 5 scores. Score 1 represents no emotion, score 5 describes very strong emotion, and larger score means stronger emotion. Besides the score of the emotional feelings, the subjects are also required to answer three questions: (1) Did you understand the story content? (2) Have you ever seen the film? (3) Did you concentrate when you were watching the video material? If a subject has answered no for (1), yes for (2) or no for (3), the data corresponding to the subject are labeled as non-effective data.

At the end of the experiment, the experimenter asks the subjects several questions according to their score in the emotional experience scale: (1) At which plot did you experience the strongest emotion? (2) Which emotion category did the strong feeling belong to? (3) Was the strong emotional feeling hard to categorize? (4) Did you correctly press the button to report your strong emotional feeling? After the experiment, all data are uploaded to the data storage server and managed by the EDAMS. Figs. 1, 2, 3, and 4 show the button status and the physiological recordings of amusement, anger, grief and fear elicitation experiments, and the data in these figures belong to one subject.

The four emotion elicitation tasks are divided into two steps. The first step is amusement elicitation followed by

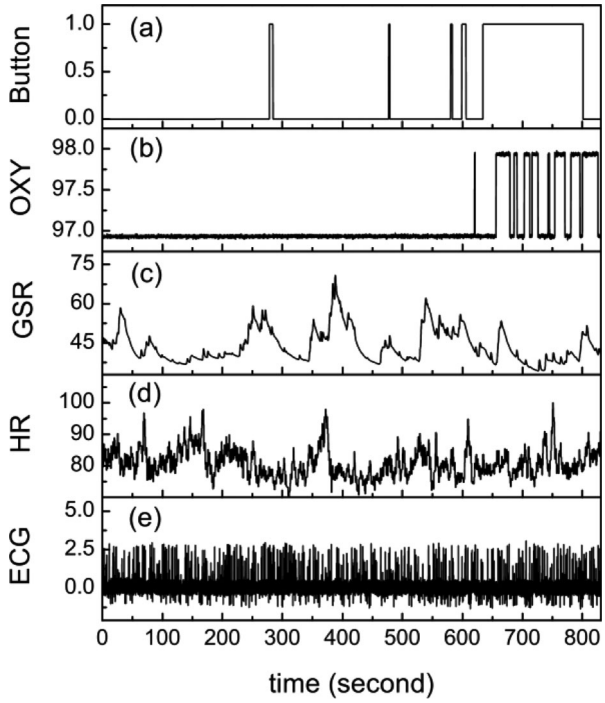


Fig. 3. The button status and the physiological recordings of a grief elicitation experiment. The button status "1" in (a) reveals the occurrence of strong emotional feeling.

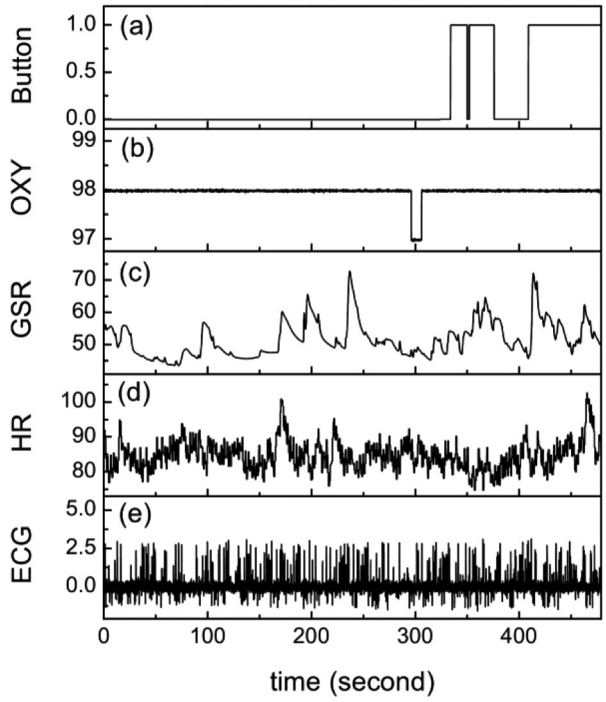


Fig. 4. The button status and the physiological recordings of a fear elicitation experiment. The button status "1" in (a) reveals the occurrence of strong emotional feeling.

grief elicitation, and the second step is anger elicitation followed by fear elicitation. There are one-month interval between the two steps and twenty-minute rest between the two adjacent emotion elicitation tasks. As the elicited emotions include extreme negative emotions, e.g. grief and fear, the subjects are informed the possible emotional feeling and they can choose to quit before the experiment. As a result, after we filter the electrode failure data and the non-effective self-report data, three subjects have been elicited the above four emotions, seven subjects have been elicited three of the four emotions, 16 subjects have been elicited two of the four emotions and 75 subjects only have been elicited one of the four emotions.

3 Similarity of the Affective Physiological Data

3.1 The Multi-Variant Correlation Method

Muller et al. [26] has introduced a linear but multivariate approach for the analysis of interrelations between data channels of a T-dimensional recording. This approach is based on the zero-lag correlation matrix and techniques known from random matrix theory. The Pearson coefficient which quantifies the similarity between two data channels is used as a basic measure in order to construct an interdependence matrix of N data channels. In his work, Muller et al. [26] gives the evidence that although the Pearson coefficient is a bivariate measure, the interrelation between N data channels can be estimated, for any $N \geq 2$.

As to the construction of a user-independent emotion recognition system, the common patterns of many subjects' affective physiological changes are of great interest and importance. Let us suppose that a certain kind of physiological data are recorded while two subjects individually watch the same emotion elicitation video. And then, with the technique of running window, the similarity between the subjects' physiological data at certain time slots can be quantified by the Pearson coefficients which vary between ± 1, the completely correlated or anti-correlated cases. This similarity measure has been applied in [27] to construct a physiological linkage vector of the film consumers in order to realize automatic highlight detection.

If the number of the subjects $N \geq 2$, an equal-time correlation matrix C is constructed by first normalizing the physiological fluctuation time series $M_i, (i = 1, 2, \ldots, N)$ [28]

$$G_i = \frac{M_i - \langle M_i \rangle}{\sigma_i}, \tag{1}$$

where σ_i is the standard deviation, $\langle M_i \rangle$ is the mean, and

$$M_i(j) = \ln X_i(j+1) - \ln X_i(j),\ j = 1, 2, \ldots, T, \tag{2}$$

where X_i is of length $T+1$ and depicts the physiological time series of the ith subject. The measure $M_i(j)$ has several nice properties [29]. First, any multiplicative, time-independent sample bias cancels in the log-ratio. Second, this measure has a natural interpretation in terms of relative signal change since for a small change $M_i(j) \approx [X_i(j+1) - X_i(j)]/X_i(j)$ is simply the relative increment. A large value of $M_i(j)$ reflects a large signal variation while a small value corresponds to an almost constant signal amplitude. In the next step, the correlation matrix C of an $N \times T$ data matrix $G = \{G_i\}$ is calculated by

$$C = \frac{1}{T} G G', \tag{3}$$

where $G'$is the transpose of G.

It is obvious that the correlation matrix C is a real symmetric matrix, and each row or column of C displays the cross correlations of a specific subject's data with all other subjects' data. All the information about the linear two-point correlation structure of two subjects' physiological time series X_i and X_j is contained in the bivariate measures represented by the $N(N-1)/2$ independent matrix elements $C_{ij}, i, j = 1, 2, \ldots, N$. However, it is difficult to follow all matrix element of C in the course of time, and due to the finite value of T, there is a certain number of random correlations of order $\sim 1/\sqrt{T}$ hidden in C_{ij} [26]. For finite T, C_{ij} is a non-zero value even in the case of completely uncorrelated X_i and X_j.

Muller et al. [26] points out that no matter G is Gaussian process or not, the set of eigenvalues and eigenvectors of C represents multivariate quantities which provide information about the characteristics of the joint probability distribution of the underlying process, and the eigen-values measure the degree of similarity between N time series. For stationary, infinitely long time series, if all X_i are completely uncorrelated, all non-diagonal elements of C will be equal to zero, and all N eigenvalues of C will be equal to 1. If all X_i are identical, all entries of C will be equal to one, and there remains only one non-zero eigenvalue which is equal to N. For N highly correlated but not identical time series, a large gap between the largest eigenvalue and all others occurs [26]. Thus, the degree of similarity between N time series depends on how much the largest eigenvalue of C is.

For finitely long random time series, if all G_i are uncorrelated, Edelman [30] gives the eigenvalue distribution of the correlation matrix of normalized $G_i, i = 1, 2, \ldots, N$ in the limit $N \to \infty$ and $T \to \infty$

$$P_{\mathrm{rm}}(\lambda) = \frac{Q}{2\pi} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda}, \tag{4}$$

where $Q \equiv T/N \geq 1$ is fixed and $\lambda \in [\lambda_-, \lambda_+]$ with

$$\lambda_\pm(Q) = 1 + 1/Q \pm 2/\sqrt{Q}. \tag{5}$$

Here, λ_+ predicts the maximum random correlations among N finite-length uncorrelated time series.

Concerning the correlation matrix C of our N subjects' normalized physiological time series, we calculate its maximum eigenvalue $\lambda_{\max}$ and compare it with λ_+. If $\lambda_{\max} > \lambda_+$, there are genuine correlations, i.e., a certain degree of similarity, between N subjects' physiological data. Otherwise, N subjects' physiological data are not similar.

3.2 Physiological Time Series and Their Correlation Matrix

For each of the above four target emotions, 35 physiological recordings have been obtained, and each physiological recording includes the data of OXY, GSR, HR and ECG. There are strong affective physiological responses in these physiological data according to the report of the subjects. Due to the three electrodes measurement of ECG, the P wave and T wave are contaminated by the noise, and we only use ECG to calculate R-R heart rate which is a good emotion indicator in literature. Other commonly used emotion measures are GSR and its phasic part caused by internal or external stimuli. We calculate the first derivative of GSR (FD_GSR) because the high fluctuation of FD_GSR reveals the phasic part of GSR. GSR, FD_GSR, HR and OXY are further analyzed to find out the affective physiological patterns in them.

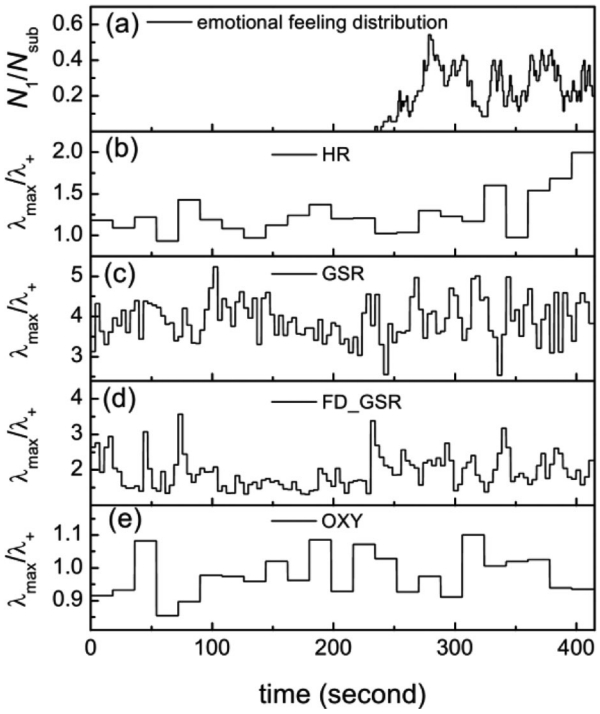


Fig. 5. The emotional feeling distribution and the normalized maximum eigenvalue of HR, GSR, FD_GSR, and OXY in each running window of the amusement data. N_1 is the number of the subjects having strong emotional feelings at that time, and $N\mathrm{sub}$ is the total number of the subjects of amusement.

The 35 physiological recordings in the same emotion category are multi-subject data, and whether the affective physiological changes are similar among multi-subjects will be analyzed by the multi-variant correlation method. Considering the different variation frequency of GSR, FD_GSR, HR and OXY, different time intervals are used when constructing the physiological time series from the original physiological data, 0.1 second for GSR and FD_GSR and 0.5 second for HR and OXY. In each time interval, we calculate the mean of certain original physiological data and obtain 35 equal-time GSR, FD_GSR, HR and OXY time series from 35 subjects' original physiological recordings. The method of running-window is used to divide the whole data recording into many adjacent segments. The parameter Q in Section 3.1 is fixed to 1 ($\lambda_+ = 4$), and this determines the length of the running-window, 3.6 seconds for GSR and FD_GSR and 18 seconds for HR and OXY.

3.3 Similarity of Multi-Subject Affective Physiological Changes

For the data in each running-window, the equal-time correlation matrix C is calculated, and the maximum eigenvalue $\lambda_{\max}$ of C is obtained. Figs. 5, 6, 7, and 8 show the normalized maximum eigenvalue $\lambda_{\max}/\lambda_+$ for the 35 recordings of amusement, anger, grief and fear, respectively. The $\lambda_{\max}/\lambda_+$ in most running windows of HR, GSR and

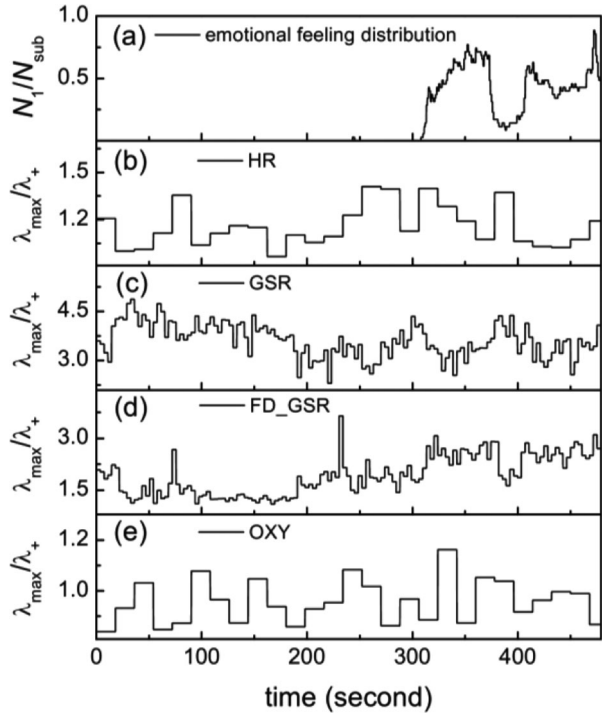


Fig. 6. The emotional feeling distribution and the normalized maximum eigenvalue of HR, GSR, FD_GSR, and OXY in each running window of the fear data. N_1 is the number of the subjects having strong emotional feelings at that time, and $N\mathrm{sub}$ is the total number of the subjects of fear.

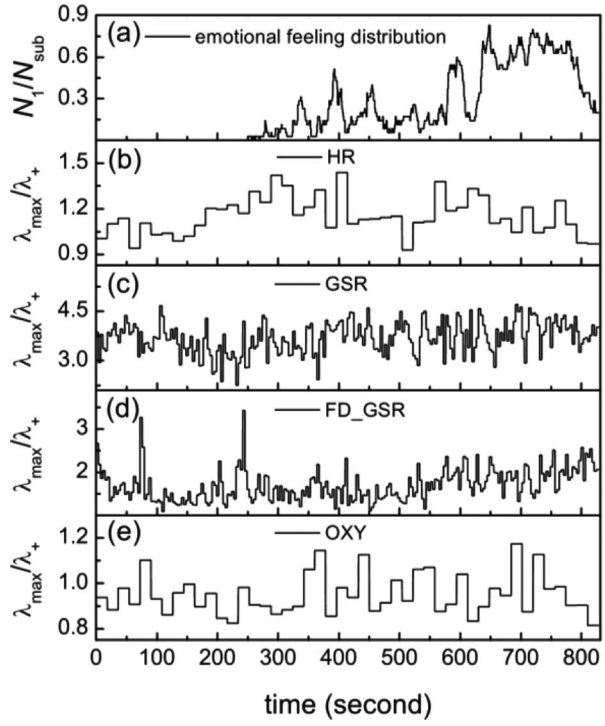


Fig. 8. The emotional feeling distribution and the normalized maximum eigenvalue of HR, GSR, FD_GSR, and OXY in each running window of the grief data. N_1 is the number of the subjects having strong emotional feelings at that time, and N_{sub} is the total number of the subjects of grief.

FD_GSR, are larger than 1, revealing the similarity of these multi-subject physiological data. On the contrary, many $\lambda_{\max}/\lambda_+$ of OXY are smaller than 1, showing the dissimilarity of multi-subject OXY data.

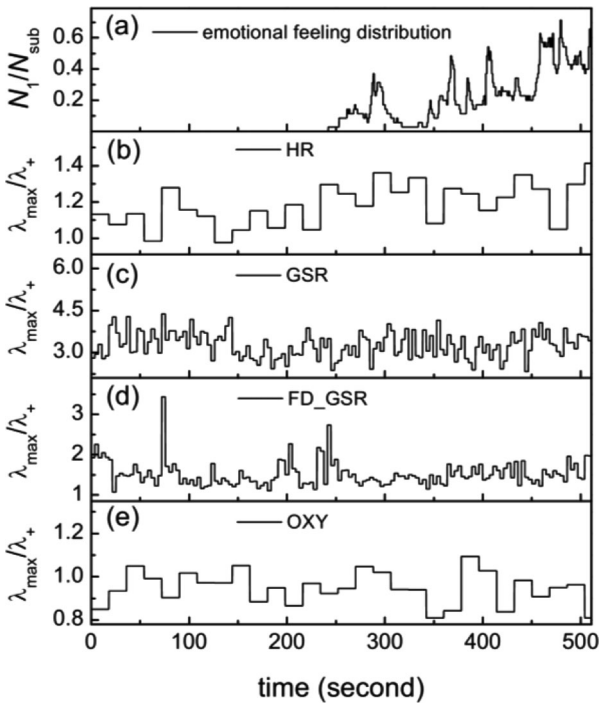


Fig. 7. The emotional feeling distribution and the normalized maximum eigenvalue of HR, GSR, FD_GSR, and OXY in each running window of the anger data. N_1 is the number of the subjects having strong emotional feelings at that time, and $N\mathrm{sub}$ is the total number of the subjects of anger.

In Figs. 5, 6, 7, and 8, the baseline state elicited by the scenery pictures also shows multi-subject similarity in HR, GSR, FD_GSR, and some running windows of OXY. In order to quantify the increased or decreased similarity of physiological data caused by amusement, anger, fear and grief, a

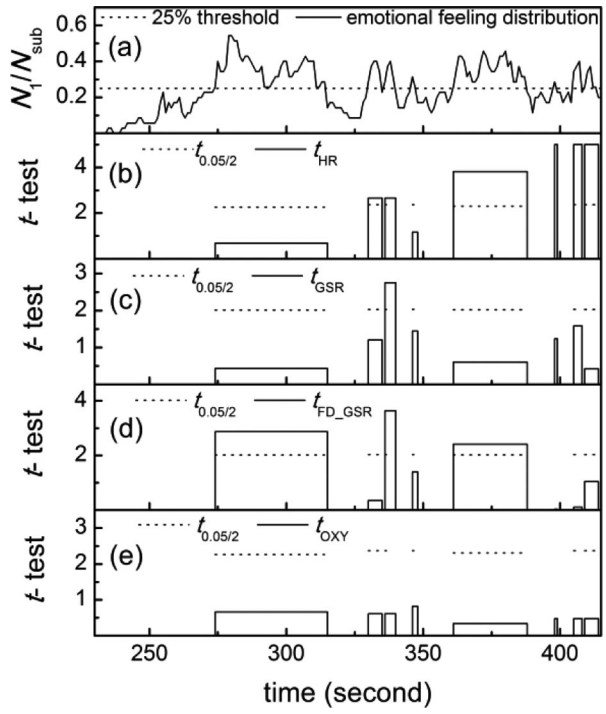


Fig. 9. The t-test of amusement data. $t_{0.05/2}$ is the threshold of the t statistics to reject the null hypothesis.

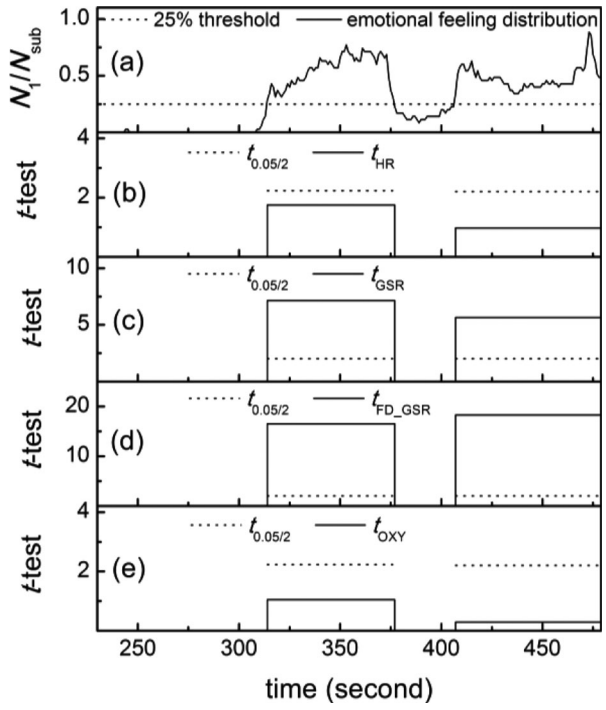


Fig. 10. The t-test of fear data. $t_{0.05/2}$ is the threshold of the t statistics to reject the null hypothesis.

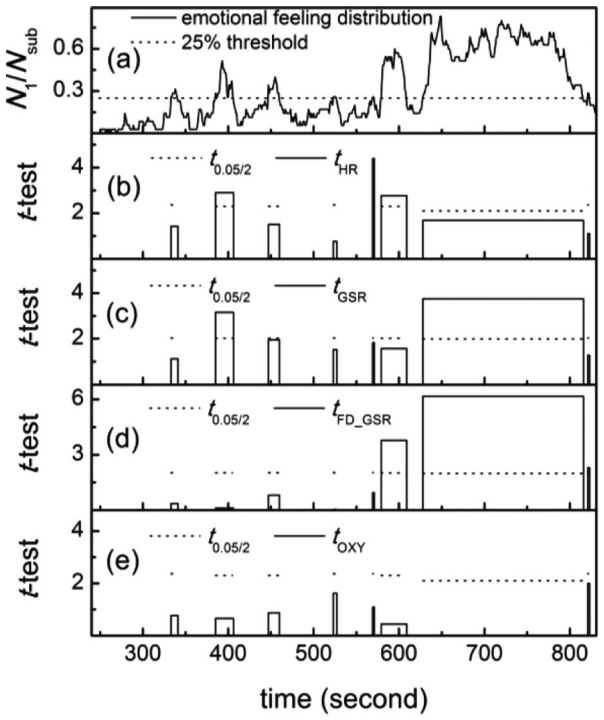


Fig. 12. The t-test of grief data. $t_{0.05/2}$ is the threshold of the t statistics to reject the null hypothesis.

bilateral unpaired two-sample t-test is done by testing the following null hypothesis:

$$\mathrm{H}_0 : \mu_0 = \mu_1, \tag{6}$$

where μ_0 is the mean of $\lambda_{\max}/\lambda_+$ in the scenery pictures duration, and μ_1 is the mean of $\lambda_{\max}/\lambda_+$ in the running windows of being tested data. The t-test is done in the time durations in which more than 25 percent of the subjects have reported strong emotional feelings, and the significance level is 0.05.

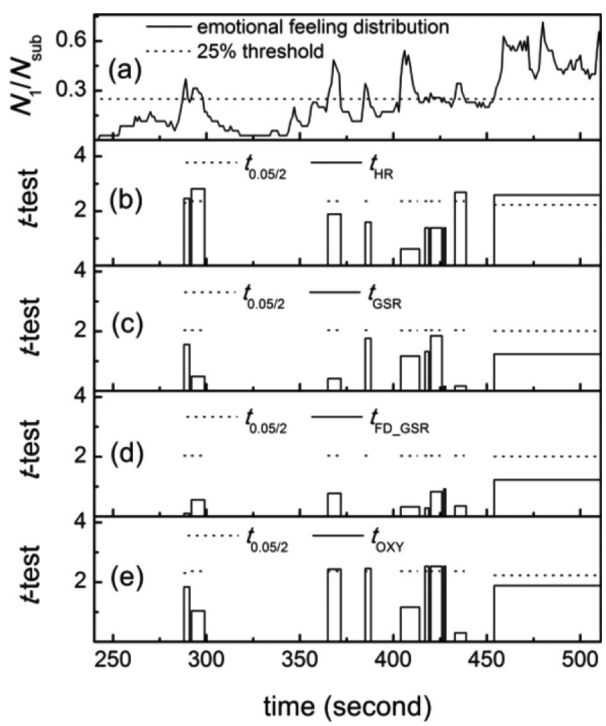


Fig. 11. The t-test of anger data. $t_{0.05/2}$ is the threshold of the t statistics to reject the null hypothesis.

In Figs. 9, 10, 11, and 12, the similarity of multi-subject HR in some of the reported strong amusement, anger and grief durations is significantly different from that of the baseline; the similarity of multi-subject GSR and FD_GSR in some of the reported strong amusement, grief and fear durations is significantly different from that of the baseline; the similarity of multi-subject OXY in all reported strong amusement, grief and fear durations is not significantly different from that of the baseline, and the significantly different similarity of OXY has only shown in several short durations of strong anger feelings, e.g. 365-371 s, 417-418 s, 420-425 s, and 427 s, in which the $\lambda_{\max}/\lambda_+$ are smaller than 1, revealing the uncorrelation of the multi-subject OXY. According to Figs. 5, 6, 7, and 8, except for the mean $\lambda_{\max}/\lambda_+$ of GSR in the fear durations, the mean $\lambda_{\max}/\lambda_+$of HR, GSR and FD_GSR in the significantly different durations are larger than the baseline mean $\lambda_{\max}/\lambda_+$.

3.4 Summary of Intra-Class Affective Physiological Patterns

To sum up, amusement, fear, anger and grief can influence the fluctuations of HR, FD_GSR and GSR at many emotion-eliciting film plots, but not significantly affect the fluctuations of fingertip OXY. Amusement, anger and grief increase the similarity of the multi-subject HR fluctuations, and they respectively cause common intra-class affective HR patterns of the subjects. Amusement, grief and especially fear increase the similarity of the multi-subject FD_GSR fluctuations, and they respectively cause common affective FD_GSR patterns of the subjects. Although

amusement and grief increase the similarity of the multi-subject GSR fluctuations, many non-affective video or audio external stimuli also cause strong correlations between multi-subject GSR. Besides, as is shown in Figs. 5, 6, 7, and 8, the non-affective switch of stimuli type increases the similarity of multi-subject HR and FD_GSR, respectively. The GSR fluctuations of fear are less similar than those of the baseline, indicating larger individual difference of affective changes in them.

4 Emotion Recognition

4.1 Affective Physiological Change Locating

The physiological signal recordings last several hundred seconds, much longer than the duration of the subjects' emotional feeling. Therefore, the following information is applied to locate the affective physiological changes: (1) The subjects have reported the time of strong emotional feeling; (2) According to the results of correlation analysis, there are affective physiological changes in HR, GSR and FD_GSR in certain durations. The similarity of three physiological signals of amusement, grief and fear has significant difference from that of the baseline. However, the time of significantly different similarity of these three physiological signals are not exactly the same. We have chosen the durations in which at least two physiological signals have shown significantly different similarity, i.e. 335-339 s and 360-387 s for amusement, 385-405 s, 578-608 s and 627-815 s for grief, and 313-376 s and 406-479 s for fear. As there is only one kind of physiological signal, HR, has shown increased similarity in the strong anger durations, three time durations have been accordingly chosen to be the affective physiological change time slots, i.e. 291-298 s, 426-427 s and 453-511 s.

For a certain subject's data, we select the data in the overlap between the reported strong target emotional feeling durations and the affective physiological change durations revealed by multi-variant correlation analysis and t-test. Besides, we choose the start of the affective data according to the start of a GSR uprise or the start of a strong HR fluctuation. Although the physiological change of each emotion does not have the same duration, many data analysis methods require the data to be equal time. Considering that the HR change of grief and the FD_GSR change of fear last longer than those of amusement and anger, the length of each affective data segment is set to be 15 seconds in which a complete HR fluctuation of grief and several FD_GSR fluctuations of fear are included. However, the durations of 335-339 s for amusement, 291-298 s and 426-427 s for anger are less than 15 seconds. The 426-427 s has been discarded due to the too short time duration, and the remaining two short time durations are prolonged forward the time to be 15 seconds. For each of the 35 physiological recordings in each of the four emotions, more than one data segment are selected if the subject has experienced strong target emotion in several affective physiological change durations. Besides, the data with much shorter duration of reported strong emotional feeling than 15 seconds are discarded. Finally, we obtain 75 amusement data segments from 35 subjects, 78 anger data segments from 33 subjects, 98 grief data segments from 35 subjects and 92 fear data segments from 31 subjects. One hundred and thirty-four baseline segments of length 15 seconds are also selected from the subjects' data corresponding to the scenery pictures duration.

4.2 Affective Feature Extraction

First, 12 conventional GSR features in [31] are calculated in this work. They are the mean and standard deviation of GSR, the average and root mean square of the first derivative of GSR (AVE_FD_GSR and RMS_FD_GSR), the number, average amplitude, average duration (AVE_D) and maximum amplitude of skin conductance response (SCR), the mean of the absolute values of the first differences of the raw GSR, the mean of the GSR filtered by a Hanning window and the mean of the absolute values of the first and second differences of the normalized GSR (1st_diff_nGSR and 2nd_diff_nGSR). The Hanning window filter is applied to reduce high frequency noisy fluctuations of GSR [8].

Second, the fluctuations of HR, GSR and FD_GSR in a variety of time scales are quantified by local scaling dimension and applied as affective features. LSD calculates the fluctuation strength of the signal in different time scales and can reveal detailed structures of the affective physiological change process. Aiba and Matsuo [32] introduces LSD to describe the fluctuation of the strength function in highly excited nuclei. For a sampled signal

$$S(t) = \sum_i S_i \delta(t - t_i), \tag{7}$$

where $\delta(t)$ is the Dirac sampling function and t_i is the sampling time spot of S_i with

$$\sum_i S_i = 1. \tag{8}$$

The main idea of LSD is: First, the whole segment interval is divided into L bins with length ε, and the binned $\{S_i\}$ are

$$p_n = \sum_{i \in n-\text{th bin}} S_i, n = 1, 2, \ldots, L. \tag{9}$$

Second, the fluctuation of p_n is characterized by moments of order m defined as follows:

$$\chi_m(\varepsilon) \equiv \sum_{n=1}^{L} p_n^m = L\langle p_n^m \rangle, m > 1, \tag{10}$$

where $\langle \cdot \rangle$ is an average of $\chi_m(\varepsilon)$ with respect to the bin boundary. Finally, the LSD is defined at each time scale ε as

$$D_m(\varepsilon) \equiv \frac{1}{m-1} \frac{\partial \log \chi_m(\varepsilon)}{\partial \log \varepsilon}. \tag{11}$$

LSD analyzes how the moments $\chi_m(\varepsilon)$ scale for different time scales. If the signal is uniformly distributed with no fluctuation, i.e., $p_n = 1/L \propto \varepsilon$ for all L and ε, a power law scaling $\chi_m(\varepsilon) \propto \varepsilon^{1 \cdot (m-1)}$ holds for $\chi_m(\varepsilon)$ with respect to the bin width ε, and then $D_m(\varepsilon) = 1$. If the signal is concentrated in a single time spot, i.e., the extreme case of large fluctuation, $\chi_m(\varepsilon) = 1 \propto \varepsilon^{0 \cdot (m-1)}$ for any value of ε since the binned strength is $p_n = 1$ for only one bin and $p_{n'} = 0$ for the others, and then $D_m(\varepsilon) = 0$. If the signal fluctuates between the above extreme cases, $D_m(\varepsilon) \in (0, 1)$ and a smaller $D_m(\varepsilon)$ means larger fluctuations in the signal.

As to our experimental data, the HR fluctuations are more complicated than those of GSR and FD_GSR.

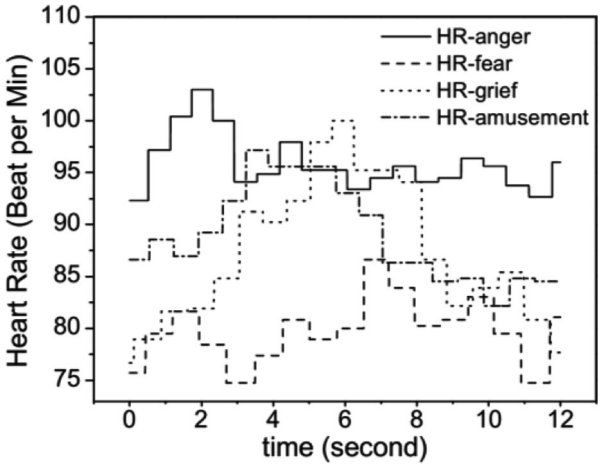


Fig. 13. Four affective HR segments coming from one subject.

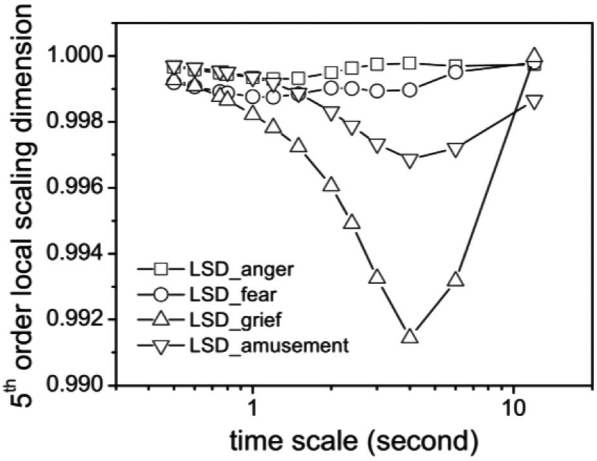


Fig. 14. Fifth order LSD of the HR segments in Fig. 13. For a certain HR segment, the fluctuations vary at different time scales.

Therefore, the second and fifth order LSD of GSR and FD_GSR are calculated in the whole 15 seconds, but the fifth order LSD of HR are calculated in several time durations. First, we select the HR data of 1 second length centered by the maximum variation step of the 15-second HR data; Second, we select 2, 3, ... ,12 seconds data around the maximum variation step of the 15-second HR data; Finally, we select the whole 15 second HR data. Fig. 13 shows four 12-second HR segments respectively belonging to amusement, anger, grief and fear of one subject. Fig. 14 is the fifth order LSD of the HR segments in Fig. 13. Before calculating the LSD, the inversion of each HR segment is appended at the end of the segment in order to quantify the fluctuation in large time scales.

The time scales of LSD are calculated as follows: Let $N_{\mathrm{d}} = L \times \varepsilon \times f$, where f is the number of the data in 1 second with $f = 400$ for HR and $f = 10$ for GSR and FD_GSR. If $f = 400$ and $\{2 \times n_{\mathrm{s1}}\}$ are the numbers which can be divided exactly by N_{d}, the time scales ε of LSD for HR are $\{n_{\mathrm{s1}}/f\}$ with $n_{\mathrm{s1}}/f \geq 0.5$. If $f = 10$ and $\{2 \times n_{\mathrm{s2}}\}$ are the numbers which can be divided exactly by $N_{\mathrm{d}}/2$, the time scales ε of LSD for GSR and FD_GSR are $\{n_{\mathrm{s2}}/f\}$ with $n_{\mathrm{s2}}/f \geq 0.5$.

In addition to $D_5(\varepsilon)$ of HR, normalized GSR, FD_GSR and the absolute FD_GSR (A_FD_GSR), the features in Table 3 are also extracted. Finally, 155 HR features and 43 GSR and FD_GSR features are obtained, and the whole feature set contains 198 features.

4.3 Affective Data Classification

The affective cases described as the vectors of the 198 features construct the original dataset, and the dimensions of the amusement, anger, grief, fear and baseline data sets are respectively 75, 78, 98, 92 and 134. A random forests classifier coming from the RF website of Breiman and Culter [33] has been used to classify the above affective data. RF grows many classification trees using n_{ds} training cases randomly sampled from the N_{DS} original data with replacement and m_{f} input features randomly sampled from M_{OF} original features ($m_{\mathrm{f}} \ll M_{\mathrm{OF}}$). When the training set for the current tree is drawn by sampling with replacement, about one-

TABLE 3
Minimum and Maximum of LSD as Features

Feature	Additional remarks
$\min D_5(\varepsilon), \varepsilon \geq 6$	To quantify the maximum fluctuation of HR in larger time scales than 6 seconds. The data length is also larger than 6 seconds.
$\min D_5(\varepsilon), \varepsilon \geq 10$	To quantify the maximum fluctuation of HR in larger time scales than 10 seconds. The data length is also larger than 10 seconds.
$\min D_5$	The minimum of the above two kinds of features
$\min D_5(\varepsilon)$	To quantify the maximum fluctuation of HR in all time scales
$\min D_5^1(\varepsilon)/\min D_5^i(\varepsilon)$ $i = 2, 3, \ldots, 8$ $\min D_5^1(\varepsilon) \times \min D_5^i(\varepsilon)$ $i = 9, 10, 11, 12, 15$	i is the length of the data. These features are to quantify the relations of the maximum fluctuation of HR in very short time around the maximum HR variation step and the maximum fluctuation of HR in larger time durations.
$\min D_5(\varepsilon)/\min D_2(\varepsilon)$	To quantify the sparseness of GSR uprisings.
$\min D_5(\varepsilon)$& corresponding ε	To quantify the maximum fluctuation of GSR, FD_GSR and A_FD_GSR in all time scales
$\max(D_2(\varepsilon_k) - D_2(\varepsilon_{k-1}))$& $\min(D_2(\varepsilon_k) - D_2(\varepsilon_{k-1}))$ $k = 2, 3, \ldots$	To quantify the maximum fluctuation variation of GSR, FD_GSR and A_FD_GSR.

TABLE 4
The Quinary Classification by RF

	amusement	anger	grief	fear	baseline
Classified as amusement	**42/75 (56.0%)**	9/78 (11.5%)	21/98 (21.4%)	1/92 (1.1%)	1/134 (0.7%)
Classified as anger	14/75 (18.7%)	**42/78 (53.8%)**	6/98 (6.1%)	3/92 (3.3%)	8/134 (6.0%)
Classified as grief	16/75 (21.3%)	9/78 (11.5%)	**67/98 (68.4%)**	10/92 (10.9%)	1/134 (0.7%)
Classified as fear	1/75 (1.3%)	5/78 (6.4%)	1/98 (1.0%)	**78/92 (84.8%)**	0/134 (0.0%)
Classified as baseline	2/75 (2.7%)	13/78 (16.7%)	3/98 (3.1%)	0/92 (0.0%)	**124/134 (92.5%)**

third of the cases called out-of-bag (oob) data are left out of the sample. The classification is done by putting down the oob cases in every tree in the forest and counting the number of votes cast for the correct class. Besides, the importance of each feature for classification can be evaluated by the Gini importance that is often very consistent with the permutation importance measure [33]. The permutation importance measure is to randomly permute the values of feature f in the oob cases, put these cases down the tree, and then subtract the percent age of votes for the correct class in the feature-f-permuted oob data from the percentage of votes for the correct class in the untouched oob data.

The RF has been tested by leave-one-subject-out cross validation, and the emotion recognition rates are defined as follows. Let N_t represent the dimension of the test set of a certain emotion in the quinary classification, N_{rc} be the number of correctly recognized cases of the emotion and $N^i_{\mathrm{re}}, i = 1, 2, 3, 4$ be the number of falsely recognized cases of the emotion. The correct rate R_{c} and error rates R^i_{e} of a certain emotion for quinary classification are calculated:

$$R_{\mathrm{c}} = \frac{N_{\mathrm{rc}}}{N_{\mathrm{t}}} \times 100\%, \tag{12}$$

$$R^i_{\mathrm{e}} = \frac{N^i_{\mathrm{re}}}{N_{\mathrm{t}}} \times 100\%, \quad i = 1, 2, 3, 4\,. \tag{13}$$

The results of the quinary classification are given in Table 4. An overall correct classification rate of 74 percent which is 3.7 times of the random guess correct rate (20 percent) of five emotion states is obtained. In the quinary classification, amusement and grief are the most difficult to distinguish. As the multi-subject physiological data have shown intra-class similarity in the correlation analysis, these classification errors may be caused by the inter-class similarity of multi-subject amusement and grief patterns of GSR, FD_GSR and HR.

Fig. 15 gives the Gini importance of the 198 affective features in the quinary classification. The mean decrease in Gini index of 29 features are larger than 2, three of them being extracted from GSR, 15 of them being extracted from FD_GSR and 11 of them being extracted from HR. Among the above mentioned 29 features, nine of them are conventional features and 20 of them are LSD features. Ten most important features sorted downward are listed in Table 5.

In the real situation, the emotional experience of a specific user may not be available. Therefore, the affective data have been selected only according to the affective physiological change durations found by correlation analysis and t-test, and the corresponding quinary classification result is shown in Table 6. Due to the much disperse distribution of the amusement and anger feeling, the intra-class impurity of their physiological data has respectively increased when the individual subject's emotional feeling report is not used in the affective data selection process, causing the significantly decreased correct rates of these two emotions. Due to the relatively concentrated emotional feeling distribution of grief and fear in the affective physiological change durations, the correct rates of these two emotions are still good when the individual subject's emotional feeling report is absent in affective data selection.

If both the correlation analysis results and the self emotional feeling reports are not referenced during the affective data selection, i.e., the affective data are randomly selected from the physiological recording in the time duration of the emotion elicitation film, the quinary classification result is shown in Table 7, in which the amusement and grief data have been totally missed, and the correct rate of fear is near that of random guess.

5 Discussion and Conclusions

5.1 Discussion

In this paper, we have shown that multi-subject GSR, FD_GSR and HR have specific affective changes for amusement, anger, grief and fear. Moreover, FD_GSR is a better emotion indicator than GSR, as the former has revealed more significantly increased affective intra-class similarity of amusement, grief and fear among multi-subjects. The Gini importance of the features given by the RF also shows that the features extracted from FD_GSR are more important than those directly extracted from GSR. Besides, many LSD features show large Gini importance, indicating that LSD has well characterized the physiological changes of the emotions. Although our physiological data are acquired from 101 subjects, the overall correct ratio of quinary classification is comparable with those of the previous work [17]. The improved classification performance in the current work benefits from the accurate location

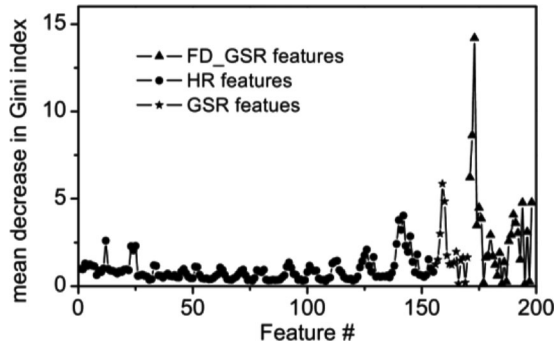


Fig. 15. The Gini importance of the 198 features in the quinary classification.

TABLE 5
The Good Affective Features Found by RF

Feature	Gini importance	Feature	Gini importance
No. of SCR	14.21	A_FD_GSR: $\max(D_2(7.5) - D_2(2.5))$	4.80
RMS_FD_GSR	8.63	A_FD_GSR: $\min\ D_5(\varepsilon)$	4.78
AVE_FD_GSR	6.23	AVE_D_SCR	4.49
1st_diff_nGSR	5.86	A_FD_GSR: $D_5(1.5)$	4.11
2nd_diff_nGSR	4.85	HR_15 s: $D_5(5)$	4.04

TABLE 6
The Quinary Classification of the Affective Data Selected by Correlation Analysis and *t*-Test

	amusement	anger	grief	fear	baseline
Classified as amusement	**23/90 (25.6%)**	8/129 (6.2%)	18/139 (12.9%)	2/151 (1.3%)	4/134 (3.0%)
Classified as anger	17/90 (18.9%)	**47/129 (36.4%)**	7/139 (5.0%)	15/151 (9.9%)	10/134 (7.5%)
Classified as grief	39/90 (43.3%)	12/129 (9.3%)	**104/139 (74.8%)**	13/151 (8.6%)	1/134 (0.7%)
Classified as fear	8/90 (8.9%)	28/129 (21.7%)	9/139 (6.5%)	**121/151 (80.1%)**	1/134 (0.7%)
Classified as baseline	3/90 (3.3%)	34/129 (26.4%)	1/139 (0.7%)	0/151 (0.0%)	**118/134 (88.1%)**

TABLE 7
The Quinary Classification of Randomly Selected Affective Data

	amusement	anger	grief	fear	baseline
Classified as amusement	**0/35 (0%)**	0/33 (0%)	0/35 (0%)	0/31 (0%)	0/134 (0%)
Classified as anger	9/35 (25.7%)	**19/33 (57.6%)**	20/35 (57.1%)	23/31 (74.2%)	2/134 (1.5%)
Classified as grief	0/35 (0%)	0/33 (0%)	**0/35 (0%)**	0/31 (0%)	0/134 (0%)
Classified as fear	24/35 (68.6%)	11/33 (33.3%)	13/35 (37.1%)	**7/31 (22.6%)**	3/134 (2.2%)
Classified as baseline	2/35 (5.7%)	3/33 (9.1%)	2/35 (5.7%)	1/31 (3.2%)	**129/134 (96.3%)**

of affective physiological changes detected by multi-variant correlation analysis.

The differences of our work compared with the previous work are the affective data acquisition setup and the multi-variant correlation analysis of the physiological data. First, the use of popular films ensures that our subjects have experienced strong feeling of the target emotion. Second, the button report of emotion, the emotional experience scale and the interview shortly after the experiment give detailed information of the subjects' emotional feelings, including the duration, the strength and the category of the target emotion. Finally, multi-subject physiological changes caused by emotions have been obtained by the multi-variant correlation analysis. The detailed affective data recording and the multi-variant correlation analysis have helped us to find out where and what the affective physiological changes are, e.g., the HR segments in Fig. 13. Based on the above analysis, we have found that the HR response duration of grief is the longest among those of amusement, grief, anger and fear, and the signal monitoring time should not be less than 15 seconds so that a complete grief HR fluctuation can be observed, as is shown in Fig. 13. Despite this, compared with the signal monitoring time of the previous work, e.g., 50 s in [17] and 160 s in [16], 15-second signal monitoring time in this work is significantly reduced, and a shorter signal monitoring time is better for practical application.

The high accuracy of the multi-subject emotion recognition may decrease in real scenario when: (1) the emotion is aroused by other emotion stimuli; (2) the user is not the same kind of college freshman; (3) no prior knowledge of the emotional activity is available. In the future work, we will acquire data from more general population and apply other emotion elicitation materials other than the materials in this paper, and then these data will be used to train the user-independent emotion recognition system in order to improve its generalization performance in practice.

The facial expression data acquired during data acquisition is not used in emotion recognition in this paper for the reasons: (1) We have found that an emotional facial expression has not always been accompanied by a self-reported strong emotional feeling; (2) The facial expressions of anger and fear elicited by films are hard to distinguish by observation. Despite this, the facial expressions of amusement and grief are good emotion indicators, and we will analyse these data in future.

5.2 Conclusion

The changes of multi-subject affective HR, GSR and FD_GSR have been successfully detected by multi-variant correlation analysis and quantified by local scaling dimension. Affective local scaling dimension features and conventional features of HR, GSR and FD_GSR have been extracted and applied for quinary classification of amusement, anger, grief, fear and baseline state. An overall correct rate of 74 percent for quinary classification have been obtained when applying the HR, GSR and FD_GSR

features to classify a multi-subject database containing 477 affective data cases of 101 subjects. During the emotion classification, the RF classifier has evaluated the importance of the affective features, and the important affective features are applicable to multi-users. The affective HR and GSR data in this work can be acquired without much user inconvenience and are suitable for practical application.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation of China (Grant No. 61103132 and No. 51171156), the key grant Project of Chinese Ministry of Education (Grant No. 311032) and the Fundamental Research Funds for the Central Universities (Grant No. XDJK2013A020).

REFERENCES

[1] R. L. Mandryk, K. M. Inkpen, and T. W. Calvert, "Using psychophysiological techniques to measure user experience with entertainment technologies," *Behav. Inf. Technol.*, vol. 25, no. 2, pp. 141–158, 2006.

[2] J. A. Healey and R. W. Picard, "Detecting stress during real-world driving tasks using physiological sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 6, no. 2, pp. 156–166, Jun. 2005.

[3] C. D. Katsis, N. S. Katertsidis, G. Ganiatsas, and D. I. Fotiadis, "Toward emotion recognition in car-racing drivers: A biosignal processing approach," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 3, pp. 502–512, May 2008.

[4] C. D. Katsis, N. S. Katertsidis, and D. I. Fotiadis, "An integrated system based on physiological signals for the assessment of affective states in patients with anxiety disorders," *Biomed. Signal Process. Control*, vol. 6, pp. 261–268, 2011.

[5] B. Verschuere, G. Crombez, E. H. W. Koster, and K. Uzieblo, "Psychopathy and physiological detection of concealed information: A review," *Psychologica Belgica*, vol. 46, pp. 99–116, 2006.

[6] P. J. Lang, "The emotion Probe: Studies of motivation and attention," *Amer. Psychologist*, vol. 50, pp. 372–385, 1995.

[7] P. Ekman, R. W. Levenson, and W. V. Friesen, "Autonomic nervous system activity distinguishes among emotions," *Science*, vol. 221, pp. 1208–1210, 1983.

[8] R. W. Picard, E. Vyzas, and J. Healey, "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23 no. 10, pp. 1175–1191, Oct. 2001.

[9] E. A. Moataz, S. K. Mohamed, and K. Fakhri, "Survey on speech emotion recognition: Features, classification schemes, and databases," *Pattern Recog.*, vol. 44, no. 3, pp. 572–587, 2011.

[10] Y. V. Venkatesh, A. A. Kassim, J. Yuan, and T. D. Nguyen, "On the simultaneous recognition of identity and expression from BU-3DFE datasets," *Pattern Recog. Lett.*, vol. 33, no. 13, pp. 1785–1793, 2012.

[11] B. Arnrich, C. Setz, R. L. Marca, G. Troster, and U. Ehlert, "What does your chair know about your stress level?" *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 207–214, Mar. 2010.

[12] J. T. Cacioppo, G. G. Berntson, J. T. Larsen, and K. M. Poehlmann, "The psychophysiology of emotion," in *The Handbook of Emotion*, M. Lewis and J. M. Haviland-Jones, Eds., 2nd ed. New York, NY, USA: Guilford Press, 2000, pp. 173–191.

[13] F. Agrafioti, D. Hatzinakos, and A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 102–115, Jan.–Mar. 2012.

[14] J. J. Gross and R. W. Levenson, "Emotion elicitation using films," *Cognition Emotion*, vol. 9, no. 1, pp. 87–108, 1995.

[15] P. Rani, C. Liu, N. Sarkar, and E. Vanman, "An empirical study of machine learning techniques for affect recognition in human-robot interaction," *Pattern Anal. Appl.* vol. 9, pp. 58–69, 2006.

[16] J. Kim and E. Andre, "Emotion recognition based on physiological changes in music listening," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 12, pp. 2067–2083, Dec. 2008.

[17] K. H. Kim, S. W. Bang, and S. R. Kim, "Emotion recognition system using short-term monitoring of physiological signals," *Med. Biol. Eng. Comput.*, vol. 42, no. 3, pp. 419–427, May 2004.

[18] R. Jonathan, D. R. Rebecca, and J. J. Gross, "Emotion elicitation using films," in *Handbook of Emotion Elicitation and Assessment*, R. J. Davidson, P. Ekman, and K. Scherer, Eds. New York, NY, USA: Oxford Univ. Press, 2007, pp. 9–28.

[19] J. N. Bailenson, E. D. Pontikakis, I. B. Mauss, J. J. Gross, M. E. Jabon, C. A. C. Hutcherson, C. Nass, and O. John, "Real-time classification of evoked emotions using facial feature tracking and physiological responses," *Int. J. Human-Comput. Stud.*, vol. 66, pp. 303–317, 2008.

[20] C. L. Lisetti and F. Nasoz, "Using noninvasive wearable computers to recognize human emotions from physiological signals," *EURASIP J. Appl. Signal Process.*, vol. 11, pp. 1672–1687, 2004.

[21] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras,"DEAP: A database for emotion analysis using physiological signals," *IEEE Trans. Affective Comput.*, vol. 3, no. 1, pp. 18–31, Jan.–Mar. 2012.

[22] A. Haag, S. Goronzy, P. Schaich, and J. Williams, "Emotion recognition using bio-sensors: First steps towards an automatic system," in *Proc. Kloster Irsee Tutorial Res. Workshop Affective Dialogue Syst.*, 2004, pp. 36–48.

[23] J. Wagner, J. Kim, and E. Andre, "From physiological signals to emotions: Implementing and comparing selected methods for feature extraction and classification," in *Proc. IEEE Int. Conf. Multimedia Expo.*, 2005, pp. 940–943.

[24] J. Fleureau, P. Guillotel, and H. T. Quan, "Physiological-based affect event detector for entertainment video applications," *IEEE Trans. Affective Comput.*, vol. 3, no. 3, pp. 379–385, Jul.–Sep. 2012.

[25] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[26] M. Muller, G. Baier, C. Rummel, K. Schindler, and U. Stephani, "A multivariate approach to correlation analysis based on random matrix theory," in *Seizure Prediction in Epilepsy: From Basic Mechanisms to Clinical Applications*, B. Schelter, J. Timmer, and A. S. Bonhage, Eds. New York, NY, USA: Wiley, 2008, pp. 209–226.

[27] C. Chenes, G. Chanel, M. Soleymani, and T. Pun, "Highlight detection in movie scenes through inter-users, physiological linkage," in *Social Media Retrieval*, N. Ramzan, R. van Zwol, J.-S. Lee, K. Clüver, and X.-S. Hua, Eds. New York, NY, USA: Springer, 2013, pp. 217–237.

[28] W. H. Wen, Y. H. Qiu, G. Y. Liu, N. P. Cheng, and X. T. Huang, "Construction and cross-correlation analysis of the affective physiological response database," *Sci. China: Inf. Sci.*, vol. 53, no. 9, pp. 1774–1784, 2010.

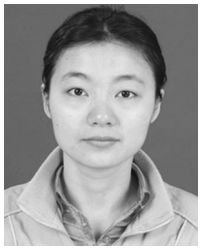
[29] M. Barthelemy, B. Gondran, and E. Guichard, "Large scale cross-correlations in Internet traffic,'' *Phys. Rev. E*, vol. 66, pp. 056110-1–056110–7, 2002.

[30] A. Edelman, "Eigenvalues and condition numbers of random matrices," *SIAM J. Matrix Anal. Appl.* vol. 9, no. 4, pp. 543–560, 1988.

[31] D. Giakoumis, D. Tzovaras, K. Moustakas, and G. Hassapis, "Automatic recognition of boredom in video games using novel biosignal moment-based features," *IEEE Trans. Affective Comput.*, vol. 2, no. 3, pp. 119–133, Jul.–Sep. 2011.

[32] H. Aiba and M. Matsuo, "Scaling analysis of fluctuating strength function," *Phys. Rev. C*, vol. 60, no. 3, pp. 034307–1–034307-15, 1999.

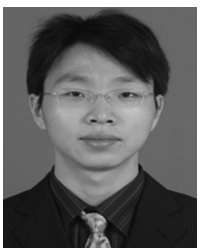
[33] (2013) [Online]. Available: http://www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm



Wanhui Wen received the BE degree in electronic and information engineering, in 2002, the ME degree in computer application, in 2005, and the PhD degree in basic psychology, in 2010, all from Southwest China University. She is currently an associate professor at Southwest China University. Her research interests include affective pattern recognition with interdisciplinary knowledge.

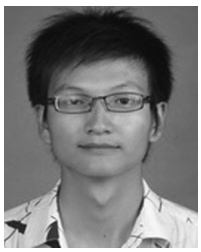
Guangyuan Liu received the PhD degree in circuit and system from the University of Electronic Science and Technology of China, in 1999. He is currently a professor at Southwest China University. His main research interests include computational intelligence and affective computing and the affective state of physiological signals.

Nanpu Cheng received the BS and MS degrees in theoretical physics from Southwest China University, in 1998 and 2001, respectively and the PhD degree in metallic materials from Central-South China University, in 2007. He is currently a professor at Southwest China University, and works at the Computational Materials Laboratory. His research interests include the application of Random Matrix Theory.



Jie Wei is currently working toward the PhD degree at Southwest China University. His current research interests include affective pattern recognition.

Pengchao Shangguan is ccurrently working toward the postgraduate degree at Southwest China University. His current research interests include affective data acquisition and analysis.



Wenjin Huang is currently working toward the postgraduate degree at Southwest China University. His current research interests include affective data acquisition and analysis.