Advances in Engineering Research, volume 164
8th International Conference on Manufacturing Science and Engineering (ICMSE 2018)

Personalized modeling of pressure recognition based on Heart Rate Variability

ZHENG Gang[1,a], ZHANG Feng[1,b], SUN Ying[1,2]

[1]Tianjin Key Laboratory "Intelligent Computing and Software New Technology", Tianjin university of technology, TianJin 300384,China

[2]China Civil Aviation Information Technology Research Base, China Civil Aviation University, TianJin 300300,China

[a]zhenggang@tjut.edu.cn, [b]zhangf_tut@163.com

Keywords: stress; heart rate variability; time series; model fusion.

Abstract. Heart Rate Variability is the description of instant duration change between two linked heartbeat in a series of continuing ECG signal. It was found that it can be used to recognize people's psychological stress. Characteristic parameters selecting and recognition model have becoming the kernel of study. This article introduced chaotic parameters of time series to be features used in stress recognition, and by the help of a proposed multi-model fusion pressure identification model, stress could be recognized more accuracy from the experiment results. And the recognition accuracy rate improved by 10%, and reached 84.5% and 89.6% on different data set respectively.

Introduction

The rapid development of artificial intelligence and wearable devices has attracted people's attention in fields such as medicine and psychology. Various researches based on physiological signals have emerged one after another, such as stress recognition[1], sentiment analysis[2], disease diagnosis[3], etc. . Analyzing and modeling the physiological signals of people in different states, and quickly and efficiently identifying different stress, affective states and physiological diseases in a short period of time have an important role in human life.

The physiological signals commonly used in stress recognition include heart rate, brain electricity, blood pressure, respiration, and galvanic skin reactions. Heart rate variability (HRV)[4]reflects chan ges in the periodic heartbeat cycle and has a strong correlation with autonomic and sympathetic chan ges in the human body. Psychology studies[5] indicate that it is an important indicator of stress moni toring. The researchers[6] compared the pressure monitoring methods of various biological signals a nd found that the HRV based pressure detection method works best. Therefore, pressure identificatio n using HRV has higher accuracy and reliability. At present, pressure recognition based on HRV mai nly extracts relevant pressure features from the time domain and frequency domain for model analysi s. Related research work includes: Based on the improved Support Vector Machine (SVM) psycholo gical stress assessment algorithm[7], the clustering information is assigned to the loss function of the support vector machine by clustering the samples first; Strew color word experiments obtain HRV d ata and extract short-term(32s)HRV time and frequency domain features and higher-order statistics of these features. Using PNN and K nearest neighbors(KNN) classifiers, select different smoothing parameters and K values for pressure recognition[8] Using the time domain and frequency domain fe atures extracted by the HRV, the pressure and non-stress states are identified by Linear Discriminant Analysis (LDA) and Logistic regression(LR), and the classifier and HRV features with optimal detec tion pressure are sought[9]; a fuzzy model interpretation is proposed. The method of autonomic nerv ous system was used to evaluate stress[10]. The experimental data was modeled by fuzzy clustering method. The concept of psychological stress was quantified and the functional relationship between psychological stress and autonomic nervous system was established; the subjects were recorded by r emote camera. The facial blood volume pulse, from which the heart rate and frequency domain of th e HRV are acquired, the relaxation and stress states are identified by Naive Bayesian and SVM algor ithms[11]. However, the literature[7-11] has the following deficiencies: no better HRV features are excavated, and the time domain and frequency domain information do not express the pressure state well; the classification model is too single and can not recognize various pressure states well .

Concerning the deficiencies in the literature[7-11]. While extracting the effective HRV time domain and frequency domain information, the HRV nonlinear parameter analysis method was introduced to find pressure-related characteristic parameters. A stacking modeling method is proposed, which adopts a hierarchical structure to overcome the problems of poor generalization capability of pressure recognition models, low stability of the model, and low recognition rate.

Overall plan design

The overall design of the experiment mainly includes five parts: stress induction, data acquisition, ECG data pre-processing, HRV feature extraction and model construction. the stress induction level, ECG data and HRV characteristics are within the same time period. The change of presentation synchronicity, that is, the manner in which the pressure is regulated, not only can cause changes in the degree of pressure of the subject, but also can obtain synchronous changes in the characteristics of the HRV from the ECG signal. In addition, this experiment adopts a reasonable inducing pressure scheme to generate and record objective data in real time during the stress induction process. Through the comparative analysis of this data, it provides a relatively objective assessment of the stress level for the synchronously collected HRV data.

ECG signal is a kind of weak electrical signal with low SNR, and it will be interfered by many kinds of noise. Therefore, the analysis of HRV needs to remove the noise generated during the ECG signal acquisition process, to obtain an effective ECG signal can detect and locate the R wave peak position of the ECG signal, and then extract the HRV features. This paper uses wavelet decomposition and selects Coif4 as the wavelet basis function [12] to decompose the ECG signal by 7 layers and perform threshold denoising on each layer. Threshold function is a different treatment method that is set when the wavelet coefficient exceeds or falls below the threshold. Commonly used methods include hard threshold and soft threshold. This paper combines the above two methods in wavelet denoising of ECG signals, and uses a compromised threshold function to denoise.

In this paper, short-term ECG signal(2 min) is used to identify different pressure levels. Because the heart cycle is relatively small, the incorrectly identified R point is more likely to affect the correctness of the recognition. Therefore, the wavelet transform with higher accuracy is used to detect R waves in this paper. , As shown in Figure 1.

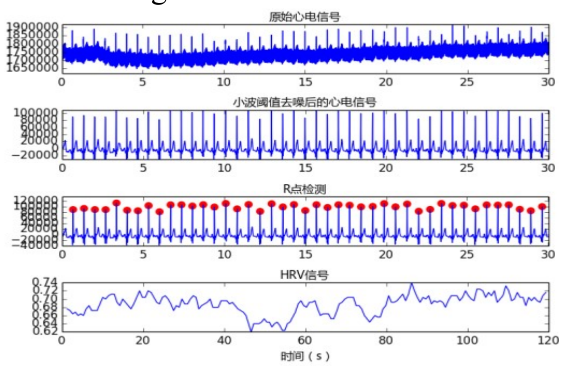

Fig.1 Data preprocessing

Feature extraction

The current analysis of HRV mainly focuses on the time domain and frequency domain. Commonly used characteristic parameters are shown in Table 1. The time domain analysis focuses on the statistics of RR interval sequence indicators, which reflect the degree of HRV transient change, but it loses the timing information contained in the HRV, and the available HRV information is limited. The frequency domain analysis focuses on the separation of physiological factors that cause HRV changes. The autonomic nervous system can be measured quantitatively, but the non-stationary

HRV is considered as a stationary random signal and cannot reflect the real process of the HRV; therefore, this paper introduces the HRV nonlinear analysis, extracts some nonlinear features, and explores the time domain. Frequency domain analysis of hard-to-find HRV timing information.

Table 1 HRV time domain and frequency domain features

	feature	Feature description
Time Domain	SDNN	Standard deviation of normal sinus beat interval (NN)
	RMSSD	Root mean square of consecutive adjacent NN intervals
	SDSD	The standard deviation of the difference between all adjacent NN intervals
	NN50	The difference between adjacent NN intervals in all NN intervals is greater than 50ms
	PNN50	The value of NN50 as a percentage of the total number of NN intervals
Frequency domain	VLF	HRV power density spectrum extremely low frequency energy
	LF	HRV power density spectrum low frequency energy
	HF	HRV power density spectrum high frequency energy
	LF/HF	HRV power density spectrum low-frequency energy and high-frequency energy ratio

In HRV-based stress recognition research, researchers mostly consider the pressure information p arameters from the time domain and frequency domain of HRV. However, HRV is not tapped as the temporal information of time series. This article through the nonlinearity of HRV Analysis, the follo wing non-linear parameters can be applied to pressure recognition.

(1) LZ complexity: It shows the degree of similarity between a finite sequence and a random sequence[13]. The greater the complexity of a sequence, the closer it becomes to a random sequence. The LZ complexity of the HRV signal can reflect the amount of information of the ECG signal and reveal the relevant laws of ECG activity.

(2) Lyapunov exponent It can indicate the average convergence or divergence measure around the HRV phase space[14].

(3) Approximate Entropy (ApEn): It reflects the complexity of the system. The larger the value, the closer to a random signal, the greater the complexity and the richer frequency components. On the contrary, it indicates that the signal is closer to the periodic state, and its complexity is reduced, which means that the randomness of the system is reduced and the regularity is increased[15]. One of the important quantitative indicators that can characterize the chaotic motion of HRV signals in phase space.

(4) Kolmogorov entropy: Describes the rate of evolution of chaotic orbits over time[16], and can therefore be used to measure the rate of change of HRV signal states over time.

Fusion Model Construction

Through the study of pressure identification model, it is found that single model pressure ide ntification often has many defects, for example, it is sensitive to abnormal data, it is easy to produce over-fitting phenomenon, and the system is too low in robustness, compared to other signals, becaus e it is weak. Signals, so ECG signals are more prone to the above situation, so this paper proposes a stacked fusion model solution to solve the above problems.

In this paper, by comparing the effects of different single models, the base model chooses K-nearest neighbor (KNN) and support vector machine (SVM) based on distance, logistic regression (L R) based on probability model, Gradient based on Boosting idea of tree model[17]. The five models of Boosting (GBDT) and Bagging thought Random Forest (RF). The distance-based model works we ll on nonlinear data and has high accuracy. However, it is sensitive to anomalous data and has poor r esults in sample unbalance. However, the tree model is insensitive to abnormal data and has high cla ssification accuracy, but the model can be arbitrarily formulated. Combining; probability-based mod el calculation is low complexity and easy to implement, but the model is easy to under-fit. On this ba sis, the stacked fusion model can synthesize the advantages of the above single model and solve its s hortcomings, making the stability of the pressure identification system improve. The overall framew ork is shown in Figure 2.

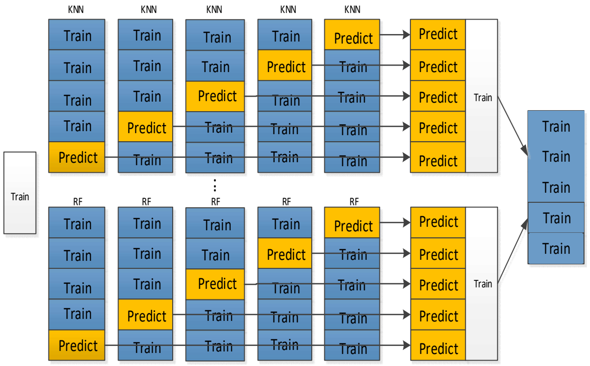

Fig.2-1 train data

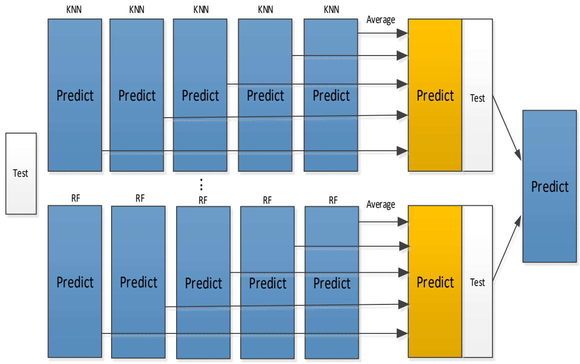

Fig.2-2 test data

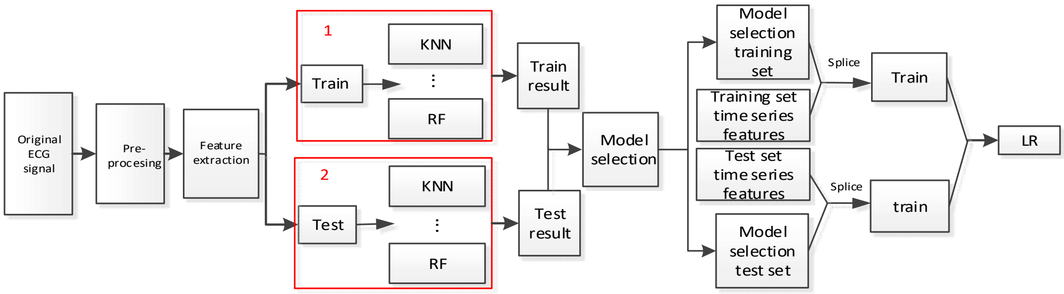

Fig.2 Fusion model structure

The model construction process: 1)Perform cross-validation of different proportions on the original data set to obtain the optimal feature set as the data set of the fusion model. 2) Divide the data set into a training set and a test set in a ratio of 3:1. On the first floor, for the training set, perform a 5-fold cross validation on each base model to obtain the M*5 prediction matrix. M is the training. Number of samples. For the test set, the prediction is performed at each iteration of each base model and the average value is taken as the prediction data of the current model. Finally, a prediction matrix of N*5 is obtained. N is the number of test samples. 3) Sort the model similarity by calculating the maximum information coefficient (MIC) of different base models[18], and select the prediction data of the three models with the smallest similarity as the input of the second layer. 4) Splice the first layer's prediction data and the HRV's time series features to obtain the second layer's training set and test set. 5) Use the LR model in the second tier to predict the final pressure level.

Experiment summary and analysis

In order to verify the validity of the verification of nonlinear parameters in pressure recognition and the advantages of the fusion model compared to the general model, this paper compares the experimental results to verify the different feature sets under calm and low pressure, calm and medium pressure, calm and high pressure conditions. Different model recognition effects.

This article uses the mobile phone rhythm master to induce individual psychological pressure, collect the ECG data of the subjects under different game difficulty, and record the game parameters and facial expressions of the subjects. In this process, changes in the rhythm and complexity of the game not only induce changes in the subject's stress level, but also cause simultaneous changes in the HRV signal. Because the rhythm and complexity of the game can be objectively evaluated by recording the game player's operating parameters and statistically and comparatively, the use of video game-induced psychological stress can provide relatively objective stress label assessment for HRV data. The data description is shown in Table 2.

Table 2 Describe experimental data set

Data description	parameter
The amount of data	1,000
Stimulus source	Rhythm master
Data duration	2min
Pressure type	Calm, low pressure, medium pressure, high pressure

In the fusion model construction process, the base model needs to be selected, and from the five base models, the three similar models with the least similarity are selected through a certain evaluation index for structural integration. This paper chooses to use the maximum information coefficient to measure the similarity between the base models. The similarity between the base models is shown in Figure 3.

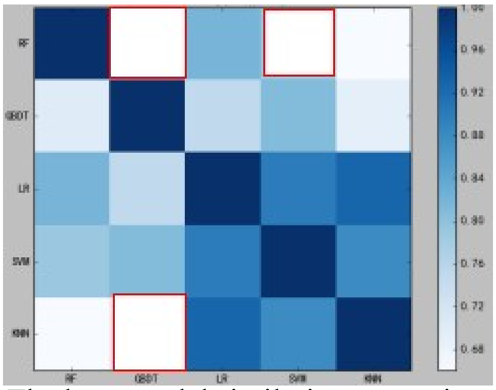

Fig.3 The base model similarity comparison chart

Figure 3 shows that in the first layer structure of the fusion model, the three models RF, SVM, an d GBDT have the lowest similarity, and the fusion effect is better. Through the analysis of the accura cy of the corresponding pressure recognition of the three models, the results are analyzed. The coeffi cients of the models are 0.5, 0.35, and 0.15, respectively.

In comparison experiments, different models are identified by comparing different pressure conditions(calm, low pressure, medium pressure, and high pressure)under different feature sets (time-frequency domain feature sets, time-frequency domain feature sets, and nonlinear feature sets). The rate is shown in Table 3 below.

Table 3 different feature set various models under different stress state recognition accuracy

		SVM(%)	KNN(%)	LR(%)	GBDT(%)	RF(%)	Fusion model(%)
Calm/low pressure	Feature set 1	56. 2	54. 6	55. 1	59. 5	60. 2	64. 5
	Feature set 2	58. 3	57. 9	59. 8	62. 3	63. 8	68. 2
Calm/medium pressure	Feature set 1	64. 3	65. 6	63. 8	66. 5	65. 7	72. 5
	Feature set 2	69. 9	72. 1	70. 8	74. 2	75. 1	80. 3
Calm/high pressure	Feature set 1	71. 2	74. 3	75. 8	79. 5	80. 2	84. 5
	Feature set 2	79. 5	80. 9	80. 4	82. 6	81. 7	89. 6

Note: Feature set 1 in Table 3 represents the HRV time domain frequency domain feature set, and feature set 2 represents the HRV time domain, frequency domain, and nonlinear feature set.

Table 3 shows that compared with the HRV time domain frequency domain feature, the nonlinear feature parameters are rich in more pressure information, so after adding the nonlinear features, the model recognition rate is improved significantly. This shows that the HRV based on chaotic time series is not The linear parameter is pressure identification.

Effect characteristic parameters. In the model comparison test, the fusion model proposed in this paper is compared with the commonly used models in the five pressure identification studies in different stress state recognition comparisons. The fusion model is compared with other model knowledge.The rate of others is significantly increased. Therefore, the modeling method proposed in this paper is effective.

Due to the imbalance of positive and negative samples in the above classification model, we use the ROC curves of different models to verify that the modeling method proposed in this paper has a relatively large improvement compared to the traditional method.

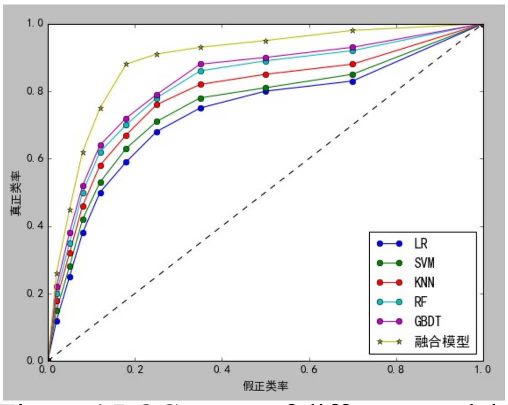

Figure.4 ROC curve of different models under calm and low pressure conditions

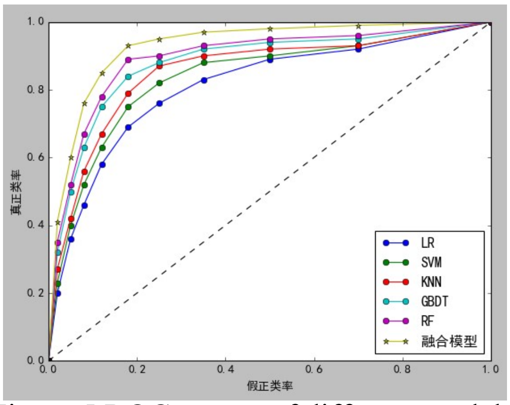

Figure.5 ROC curves of different models under calm and medium pressure conditions

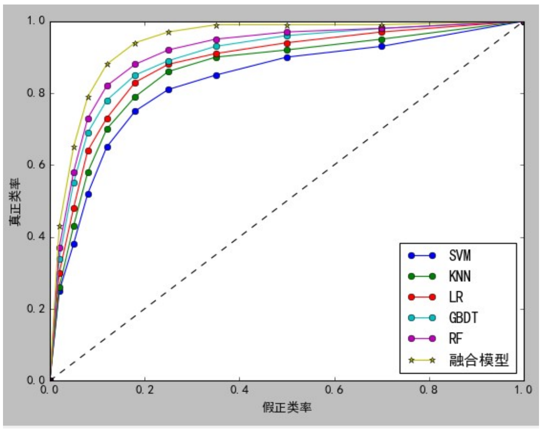

Figure.6 ROC curve of different models under calm and high pressure conditions

From the simulation results shown in Figure 4-6, under different pressure conditions, the AUC value (area between the ROC curve and the coordinate axis) corresponding to the modeling method proposed in this paper is larger than other models, and we can see that the modeling proposed in this paper The method can improve the recognition rate of different pressure states.

Conclusion

Based on the research of HRV pressure identification, this paper introduces the HRV nonlinear analysis and digs out the stress-related characteristic parameters. Experiments prove that the nonlinear features are effective. The stacked pressure identification modeling method proposed in this paper solves the problem of low accuracy of general model recognition and poor robustness of the system through structured modeling techniques, and significantly improves the system recognition rate. Compared with the traditional methods, the proposed method can increase the recognition rate of the same pressure state by up to 10%.

References

[1] Liao W,Zhang W,Zhu Z, et al.A real-time human stress monitoring system using dynamic bayes ian network[C].Computer Vision and Pattern Recognition-Workshops,2005.CVPR Workshops. IEEE Computer Society Conference on IEEE,2005:70~70.

[2] Nardelli M, Valenza G, Greco A, et al. Recognizing Emotions Induced by Affective Sounds through Heart Rate Variability[J]. IEEE Transactions on Affective Computing, 2015, 6(4):385-394.

[3] Wang X, Gu Y, Xiong Z, et al. Electronic Skin: Silk-Molded Flexible, Ultrasensitive, andHighly Stable Electronic Skin for MonitoringHuman Physiological Signals (Adv. Mater. 9/2014)[J]. Advanced Materials, 2014, 26(9):1309-1309.

[4] Marek M P D. Heart Rate Variability[J]. Annals of Noninvasive Electrocardiology, 2006,1(1):151-181.

[5] Dishman R K, Nakamura Y, Garcia M E, etal. Heart rate variability, trait anxiety, and perceived stress among physically fit men andwomen[J]. International Journal of Psychophysiology, 2000, 37(2): 121~133.

[6] Sharma N, Gedeon T. Objective measures, sensors and computational techniques for stress recognition and classification: a survey[J]. Computer Methods & Programs in Biomedicine, 2012, 108(3): 1287~1301.

[7] Li Miao, Zhang Yunpeng, Li Honghong, et al. Evaluation of psychological pressure for individual differences[J]. Chinese Journal of Biomedical Engineering, 2014, 33(1):45-50.

[8] Choi J, Gutierrez~Osuna R. Using heart rate monitors to detect mental stress[C]. Wearable and Implantable Body Sensor Networks, 2009. BSN 2009. Sixth International Workshop on. IEEE, 2009: 219~223.

[9] Kaur B, Durek J J, O'Kane B L, et al. Heart rate variability (HRV): an indicator of stress[C]. Proceedings of SPIE - The International Society for Optical Engineering, 2014: 91180V~91180V~8.

[10] Healey J A, Picard R W. Detecting stress during real-world driving tasks using physiological sensors[J]. Intelligent Transportation Systems IEEE Transactions on, 2005, 6(2): 156~166.

[11] Sandulescu V, Andrews S, Ellis D, et al. Stress Detection Using Wearable Physiological Sensors[M]. Artificial Computation in Biology and Medicine. Springer International Publishing, 2015: 526-532.

[12] Kumar B, Satyanarayana P S. Compression and Denoising - Comparative Analysis from still Images using Wavelet Techniques[J]. Itsocietyindia Org.

[13] Mérai L, Niederreiter H, Winterhof A. Expansion complexity and linear complexity of sequences over finite fields[J]. Cryptography & Communications, 2016:1-9.

[14] Johansson M, Rantzer A. Computation of Piecewise Quadratic Lyapunov Functions for Hybrid Systems[J]. IEEE Transactions on Automatic Control, 1998, 43(4):555-559.

[15] Pincus S. Approximate entropy (ApEn) as a complexity measure[J]. Chaos, 1995, 5(1):110.

[16] Grassberger P, Procaccia I. Estimation of the Kolmogorov entropy from a chaotic signal[J]. Physical Review A, 1983, 28(4):2591-2593.

[17] Alexey N, Alois K. Gradient boosting machines, a tutorial[J]. Frontiers in Neurorobotics,2013, 7(7):21.

[18] Reshef D N, Reshef Y A, Finucane H K, etal. Detecting Novel Associations in Large Data Sets[J]. Science, 2011, 334(6062):151