Le-kai ZHANG, Shou-qian SUN, Bai-xi XING, Rui-ming LUO, Ke-jun ZHANG, 2019. Using psychophysiological measures to recognize personal music emotional experience. *Frontiers of Information Technology & Electronic Engineering,* 20(7):964-974. https://doi.org/10.1631/FITEE.1800101

Using psychophysiological measures to recognize personal music emotional experience

Key words: Music; Emotion recognition; Physiological signals; Wavelet transform

Corresponding author: Bai-xi XING
E-mail: sisyxing@gmail.com

Motivation

1. We aim to use psychophysiological measures for emotion-based music recognition, since human emotional states can influence physiological changes, and, in turn, such physiological features can be used to reflect human emotions when listening to music.
2. The relationship between the physical state of the individual and the music emotional experience may make the computer be aware of personal music preference for effective music recommendation.
3. We will collect audio data and physiological data and perform music emotion recognition using different machine learning methods to verify the results.

Main idea

1. Physiological variations and musical stimuli could be linked by emotional properties.

2. We compared different algorithms to reveal the emotion patterns in different data fusions, which thus helped achieve the best performance by NuSVR. In the meantime, we found that the data set with all features had the best performance in modeling.

3. The results provided a promising way to make the computer be aware of personal music preference for effective music recommendation.

Method

1. Music features including the melfrequency cepstral coefficient (MFCC), centroid, flux, and roll-off, and physiological features including electrodermal activity (EDA), photoplethysmography (PPG), skin temperature (SKT), respiration (RSP), and pupil diameter (PD) variation information were all collected to form a database for emotion recognition modeling.

2. Based on the existing multimodal feature data sets, we compared LR, RR, SVR (linear kernel), SVR (RBF kernel), SVR (poly kernel), decision trees, KNN, MLP, and NuSVR with the most relevant feature data set to find the optimal model, separately.

Major results

1. Emotion recognition based on different methods

Table 5 Comparison of different algorithms for emotion recognition results

Algorithm	Arousal		Valence	
	MSE	CC	MSE	CC
Decision tree	0.047 78	0.2395	0.037 27	0.2335
SVR (linear kernel)	0.046 49	0.4371	0.027 29	0.5762
Linear regression	0.043 59	0.4465	0.026 75	0.5812
Ridge regression	0.043 34	0.4474	0.026 74	0.5813
MLP	0.055 73	0.5471	0.035 51	0.6103
SVR (poly kernel)	0.030 99	0.6426	0.025 93	0.6574
k-nearest neighbors	0.027 22	0.6774	0.019 10	0.7179
SVR (RBF kernel)	0.026 76	0.7031	0.017 67	0.7605
NuSVR	**0.023 23**	**0.7347**	**0.014 85**	**0.7902**

Major results

2. Comparison results of single-modality emotion recognition

Table 6 Comparison results of single-modality emotion recognition

Feature	Arousal		Valence	
	MSE	CC	MSE	CC
RSP	0.0509	0.202	0.0407	0.205
EDA	0.0468	0.275	0.0368	0.271
PD	0.0421	0.413	0.0367	0.303
Music	0.0437	0.425	0.0373	0.420
PPG	0.0479	0.333	0.0337	0.425
SKT	**0.0452**	**0.491**	**0.0342**	**0.539**

Major results

3. Comparison results of multi-modality emotion recognition

Table 7 Comparison results of multi-modality emotion recognition

Modality fusion	Feature	Arousal		Valence	
		MSE	CC	MSE	CC
All physiological signals	EDA, PPG, R, SKT, PD	0.0363	0.540	0.0256	0.611
Best three physiological signals	PPG, SKT, PD	0.0351	0.553	0.0246	0.616
Best three physiological signals & musical features	PPG, SKT, PD, Music	0.0312	0.623	0.0201	0.721
All physiological signals & musical features	EDA, PPG, R, SKT, PD, Music	**0.0293**	**0.6499**	**0.0157**	**0.7735**

Conclusions

1. A physiological feature database and a music emotion feature database were built.

2. Based on these two databases, we compared LR, RR, SVR (linear kernel), SVR (RBF kernel), SVR (poly kernel), decision trees, k-nearest neighbors (KNN), MLP, and NuSVR to reveal the emotion patterns in different data fusions, which thus helped achieve the best performance by NuSVR.

3. In all the data fusion comparisons, the data set with all the features (music features and all physiological features) had the best performance in modeling.