KURAM VE UYGULAMADA EĞİTİM BİLİMLERİ EDUCATIONAL SCIENCES: THEORY & PRACTICE

Received: January 21, 2018
Revision received: June 22, 2018
Accepted: June 25, 2018

www.estp.com.tr
DOI 10.12738/estp.2018.6.182 ⬩ December 2018 ⬩ 18(6) ⬩ 2822-2832

Research Article

Application of Emotion Cognitive Model in Interactive National Music Education

Mingming Yang
Renmin University of China

Abstract

Musical emotion cognition can analyze the connotation of music on the semantic level, so it plays a significant role in human-computer interaction, multimedia music teaching and computer music research. This paper proposes the emotion cognitive model based on Hervner emotion-ring theory, and establishes a musical feature space with 10 description indexes by using the emotion recognition evaluation method. In order to verify the proposed model, we design a music emotion cognition experiment. Through the simulation calculation data and experimental data, it is found that the vector distance is less than 0.035, which is in line with the error requirements. Finally, we apply the cognitive model to the interactive national music education platform, realizing the accurate search of personalized music.

Keywords

Emotion Cognition • Modeling • Interactive • National Music Education

[1]**Correspondence to:** Mingming Yang (MA), School of Arts, Renmin University of China, Beijing 100001, China. Email: 15313973999@qq.com

Citation: Yang, M. M., Application of Emotion Cognitive Model in Interactive National Music Education. *Educational Sciences: Theory & Practice, 18(6),* 2822-2832. http://dx.doi.org/10.12738/estp.2018.6.182

With the development of multimedia technology, big data and artificial intelligence technology, music education faces both major opportunities and challenges (Cano & Sanchez-Iborra, 2015; Mills, 1999; Smith, Smaill & Wiggins, 2000). Different from the traditional face-to-face music education, modern music education mostly adopts online, digital entertainment and human-computer interaction technologies (Faiola & Matei, 2010; Culén, Mainsah & Finken, 2014, Talsik, 2015). The computer emotion cognition technology has brought a huge application platform. Among them, musical emotion cognition has become the core part of music media content understanding and human-computer interaction. Currently, it has been widely used in entertainment robots, VR game industry and music education, which makes emotion recognition technology for music is being carried out in quantity. We have known that the emotional connotation belongs to a higher level of musical features with no specific standard. At present, there are mainly two major modelling methods of emotion cognition: knowledge-based model and data-based model. For the former, there are two main sources of music knowledge, namely, the mathematical equations between the key variables established in the study of music psychology and other measurable variables, as well as a knowledge model of expert system form of people's practical experience, qualitative and quantitative analysis of the cognition process of music psychology. Among them, Katayose put forward a simulated emotion cognition analysis model based on heuristic rules for Japanese pop-music (Katayose, Imai & Inokuchi, 1988; Katayose, 1990). T. Liu used the expert system to conduct emotion cognition research on chime music in ancient Chu regions (Liu, Sun & Pan, 2004). The method based on data model uses statistics analysis or machine learning theory to establish statistical or learning models of key variables and other measurable variables, featured in uniform model structure, simple modelling method, convenient operation and maintenance, and high precision. And Liu put forward a hierarchy-based classification system (Liu, Zhang & Zhu, 2003). Zhuang used BP neural network to emotionally classify Chinese popular music (Feng, Zhuang & Pan, 2001). Peng Qiong compared the neural network emotion cognitive model with the statistical emotion cognitive model (Peng, 2007).

In terms of the musical emotion cognitive model, Hevner put forward the famous emotion ring model which divides emotion into eight basic categories according to discrete semantics (Hevner, 1936). The model has been widely used. Later, Thayler established a two-dimensional emotion model that describes the emotional response in a stress-energy manner (Thayer, 1989). Tellegen and Waston expanded Thayler's model to describe the emotion by using happiness and engagement (Yun & Chang, 2009). However, Hevner's model still possesses strong vitality because of its simple and discrete structure.

To sum up, the paper starts with the study of the musical features affecting of musical emotion cognition, finds out the musical features that have a great effect on the music emotion classification, and establishes the corresponding knowledge system. This paper proposes the emotion cognitive model based on Hervner emotion-ring theory, and establishes a musical feature space with 10 description indexes by using the emotion recognition evaluation method. In order to verify the proposed model, a musical emotion cognition experiment is designed. The cognitive model is applied to the interactive national music education platform to realize the accurate search of personalized music.

Modelling of emotion cognition

Emotion recognition

Emotion-ring theory was put forward by Hevner in 1935. The theory divides the main musical emotions into eight categories with a total of 67 words for distance description. Through the emotion ring shown in Fig.1, emotional recognition of musical works is achieved through this series of discrete words.

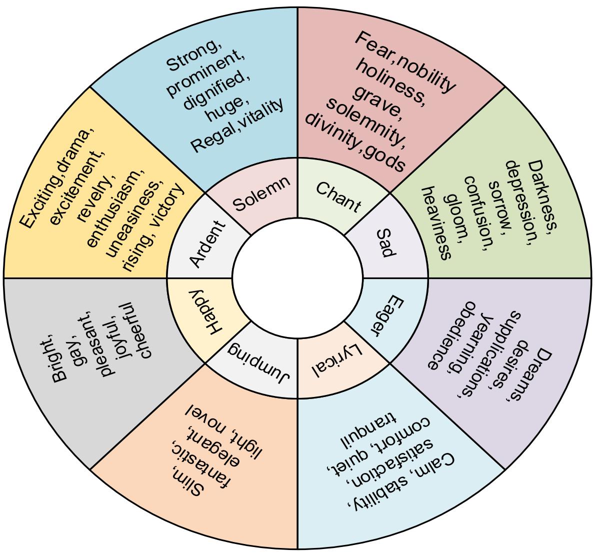

Figure 1. Hevner emotion-ring model.

Mathematically, we treat emotion recognition as a mathematical mapping process that maps Hevner's musical feature space that doesn't indicate emotional states into the 8-dimensional emotional space. In this emotional space, each of defined components represents the similarity degree between the corresponding emotional subclasses in music and emotion ring, which is actually a fuzzy paradigm. Therefore, the mapping has the fuzzy expression of musical emotion. The mapping relation between musical emotion cognition and reasoning system is shown in Fig.2.

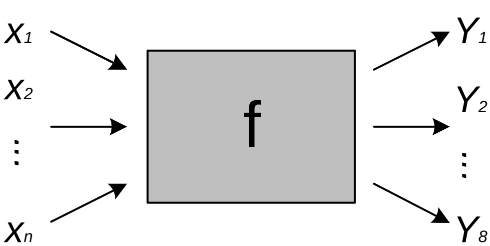

Figure 2. Mapping relation.

In the theory of musical emotion expression, it is pointed out that the main task of emotion recognition is not to classify emotions, but to judge the similarity of emotions. The mapping rules of emotion recognition are the cognitive discriminating formulas and discriminating rules established by the system through finding the regularity of musical emotion recognition through some learning strategies based on some samples obtained by emotion labeling experiment. When meeting new music, the emotion recognition determines the musical emotion vector according to the discriminating formulas and discriminating rules.

Evaluation method

Essentially, emotion recognition is a mapping process. Thus, the emotion recognition system can be evaluated through the mapping accuracy and the mapping speed. The mapping speed depends on the complexity of mapping rules; and the reference to evaluate the mapping accuracy is music emotion cognition experiment. That is: the closer the recognition result of machine labeling on musical emotion and artificial recognition is, the higher the accuracy of cognitive reasoning is. The above practice implies the accuracy of the evaluation system. The accuracy rate refers to the ratio of music period of all computer recognition results matching artificial recognition results to the music period to be recognized. The judgment formula of accuracy can be expressed as:

$$P = \frac{N_{identify}}{N_{total}} \times \% \quad (1)$$

where, P is judgment of accuracy rate; $N_{identify}$ is the number of correct recognitions; N_{total} is the total number of music periods. When the distance between artificial recognition emotion E_h and the automatic recognition emotion E_r is less than a certain threshold, we think that the correct recognition is reached. It meets the following criteria:

$$D(E_h, E_r) < D_{treshhold} \quad (2)$$

where, $D_{treshold}$ is threshold of distance. In this paper, the threshold is limited to 0.05.

Musical feature space

In this work, statistical processing is conducted on these perceptual attributes of music based on time shaft. Then, ten musical features and corresponding formalized definitions are extracted by using related theories of music combined with the actual features of different types of music. According to the variation regularities of basic elements such as pitch interval, pitch and sound intensity, these features use the weighted average based on time distribution to represent the directional index, and the standard deviation to represent the stability index. They are:

The range index *Range*, as shown in (3), is used to describe the average pitch.

$$Range = \frac{1}{n}\sum_{i=1}^{n} Pirch_i \tag{3}$$

The strength index *Intensity*, as shown in formula (4), used to describe the average intensity.

$$Intensity = \frac{1}{n}\sum_{i=1}^{n} Intensity_i \tag{4}$$

The speed index *Speed*, to describe the music speed, is directly extracted from the passage.

The intensity stability index Sta_{Int}, as shown in formula (5), obtained from the standard intensity difference.

$$Sta_{Int} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Intensity_i - Intensity_{ave})} \tag{5}$$

Where, $Intenstity_{ave}$ represents the average intensity.

Intensity directional index D_{Int}, describing the change of intensity, obtained from the distribution of intensity defference, given by equation (6),

$$D_{Int} = \frac{\frac{1}{n}\sum_{i=1}^{n}[(Intensity_{i+1} - Intensity_i)Dur_i]}{Dur_{mus} - Dur_n} \tag{6}$$

The melodic directional index D_{pit} is obtained by the time value distribution of pitch difference, such as formula (7),

$$D_{pit} = \frac{\frac{1}{n}\sum_{i=1}^{n}(Inteval_i \cdot Dur_i)}{Dur_{mus} - Dur_n} \tag{7}$$

Pitch stability index Sta_{pit}, obtained by the standard deviation of pitch, such as (8),

$$Sta_{pit} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Pitch_i - Range)^2} \tag{8}$$

Interval stability index $Sta_{Interval}$, obtained by the standard deviation of absolute value of interval, as shown in formula (9),

$$Sta_{Interval} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(|Interval_i| - |Interval_{av}|)^2} \tag{9}$$

Interval span index $Span_{Interval}$, obtained by the mean value of interval absolute value, such as formula (10),

$$Span_{Interval} = \frac{1}{n-1}\sum_{i=1}^{n-1}|Interval_i| \tag{10}$$

Note density index $Note_{Dens}$ is the number of notes contained in each bar, as shown in formula (11),

$$Note_{Dens} = \frac{Note_N}{Bar_N} \tag{11}$$

According to the above music characteristic indexs, we can establish the corresponding music characteristic space.

Experimental Research

Experimental design

To verify the musical emotion cognitive model, a musical emotion cognition experiment has been designed so as to investigate the emotional attributes contained in the musical works and to measure the listeners' understanding of the emotional similarities of the musical compositions so that the fuzzy distribution vector of musical emotion cognition with typical style and emotionally representative music is obtained. In this experiment, according to the psychological feelings for musical emotion, each experimental object can elect several words suitable for expressing the emotional connotation of the music in the basic language value set, that is, each experimental object selects subset contained in the basic language value set. As a result, the choice of all objects can form a random set defined on basic language value set. Thus, for each basic language value, we can calculate the number of times that it is selected, namely, the number of the set containing the language value, and then the frequency of its selection is obtained. This frequency can be used as a measure of the corresponding dimension in fuzzy distribution vector of musical emotion cognition.

According to the cognitive behavior of musical emotion and psychological measurement requirements, we follow the following principles when designing the musical emotion cognition experiment: (1) Inviting music teachers to participate in the collection process of music experimental samples to ensure a higher content validity; (2) Experimental material is only music sound, and doesn't include any knowledgeable factor, that is, it doesn't list the title of the relevant period, the author and other information; (3) Music period style covers as wide as possible to reduce the effect of the object's preference on the experimental results; (4) The content of period has audibility and is suitable for different test populations; (5) the test items are randomly distributed, and there is no fixed order in difficulty and familiarity; (6) The test time is strictly controlled in 20 minutes, otherwise it will make the object produce psychological fatigue, thus affecting the accuracy and reliability of the test.

Schoen found in the research that the relationship between the musical works and the types of emotional reactions is basically not affected by age, training, experience, listening time and occasion. Therefore, to facilitate the experiment, we set the main experimental objects as student groups who are divided into two groups with professional education and non-professional education to conduct comparative study. The music sample is established by the following two stages:

First of all, we invite professional musician college students who have been educated for many years as music collectors to collect 430 different types of MIDI multi-track songs under the guidance of a full-time teacher through major music download sites. Then, we examine the emotional expression of music and production effects, and select out 200 songs. The distribution of selected songs is shown in Table 1, of which most are classical music and Chinese folk music.

Table 1
Distribution of Experimental Music

Range	Number	Ratio	Example
National	22	8%	Two springs, why not return
Religion	16	9.3%	Christmas songs, Buddhist songs
Game	34	10.7%	The seven dragon balls, Blue The Danube
Classic	81	34%	Carmen
Movie	77	38%	Lion King, Romance of the Three Kingdoms

Then, music pre-processing is performed. The subsection of music is conducted artificially in accordance with musical form analysis theory to get 679 periods with independent emotional semantics. 360 periods of 20-50 seconds are roughly divided into eight emotional categories, ad are randomly divided into 10 groups. Each group has 36 periods with average length of 35 seconds, and then the experimental music periods sample is obtained.

Based on the applications of Web, our users of experimental system log in through the program interface, fill in some basic information, and can enter the network survey. After playing the music, the objects can select an emotional subclass with reference to the basic language value set. The experiment was conducted from September to October 2016 and by the way of network survey. The participants involved in the experiment were divided into professional and non-professional categories. Professional objects came from the music colleges and music teachers and students of ordinary universities. Non-professional objects were students and music lovers from ordinary universities. The ratio between the two types of objects was 1: 2. The total number of participants was 798 with a male female ratio about 1: 1.

Data analysis method

In the experiment, the method of calculating fuzzy distribution is used to get the judgment data of music periods which will form a music emotional subclass. The formula of fuzzy distribution is as follows:

$$v^l = \frac{N^l}{N} \tag{12}$$

where, v^l is the lth fuzzy distribution of emotional subclass in the emotion vector.

After obtaining the emotional fuzzy distribution data of all the samples by the formula (11), we need to further increase the similarity constraint of Hevner's ring of musical emotion, that is, we will process the data based on the multi-source emotion fusion method according to the calculation of the musical emotion language values to get emotional vector of emotional connotation of sample periods.

Experimental results and analysis

Emotion-labelled data of professional and non-professional objects are obtained through emotional similarity constraint processing. And the emotional vector distance of professional and non-professional objects of each period sample can be obtained by using the formula of musical emotion vector distance. The distribution is shown in Fig.3.

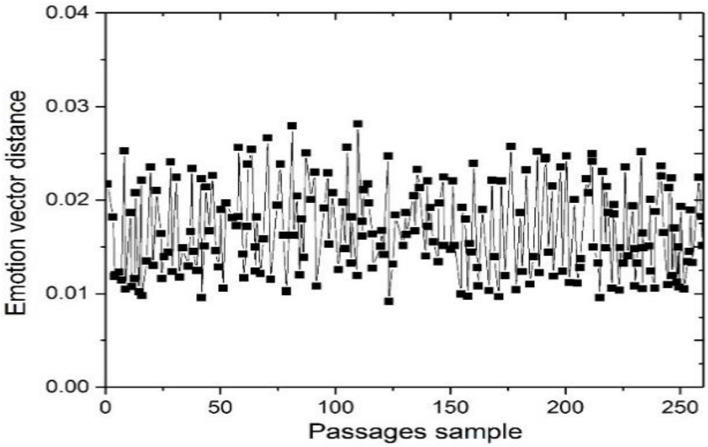

Figure 3. Emotional vector distance.

Table 2
Music Contribution Value and Variance

No.	Characteristic value	Variance contribution rate%	Cumulative contribution rate%
1	2.8045	28.045	28.045
2	1.8243	18.243	46.288
3	1.0856	10.856	57.144
4	0.9479	9.479	66.623
5	0.8687	8.687	75.310
6	0.7353	7.353	82.663
7	0.6181	6.181	88.844
8	0.5384	5.384	94.228
9	0.3451	3.451	97.679
10	0.2321	2.321	100

In the judgments of the professional and non-professional objects about the musical emotional connotation of the first experimental period, the emotional vector distance meets the normal distribution. At the same time, the correlation coefficients of the two sets of data are calculated to get distribution as shown in Table 1, of which the judgment about the emotional element of “sacredness, sadness, happiness” in music is highly correlated, which shows that the consistency is good. However, the average correlation coefficients reach and the judgment is basically the same, indicating that the professional background has little effect on the judgment of musical emotional connotation.

Application in interactive national music education

In recent years, a large number of researches have focused on the retrieval system which enables the users to search the music while only remembering the song melody “based on the humming melody”. However, music melody retrieval is only performed on the physical characteristics of music so that retrieval cannot be achieved when users don't remember the melody. In the high-level semantic features of music, emotional semantics are higher-level features exceeding melodic semantics, so it is necessary to fully consider the emotional needs of users in the retrieval. Therefore, at the application level, we apply the emotion recognition model to interactive national music education. Aiming at the accurate recognition of musical emotion, we design a semantic retrieval scheme of musical emotion for interactive music education.

As shown in Fig.4, it is the whole structure of music retrieval with input condition of emotional vocabulary or the users' own hobby and emotional experience. Emotional search for music databases is a new way of music retrieval with a higher level of semantics and intelligence than systems that use melodic similarity retrieval. Hence, the emotion recognition model proposed in this paper and the method of music retrieval are not only conceptual innovations, but also breakthroughs in methods.

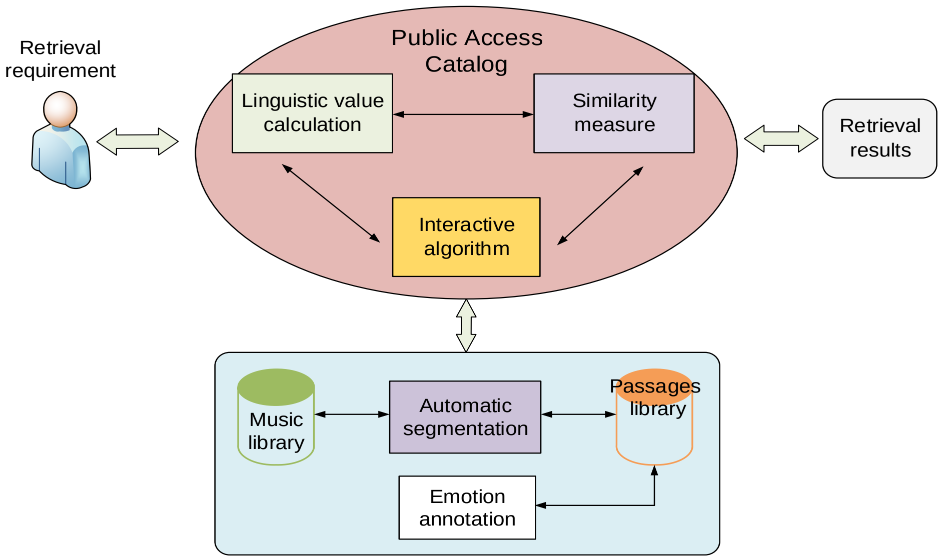

Figure 4. Application of emotional semantic retrieval in interactive music education.

Through the statistical analysis of psychological experiment, the musical emotion vector represented by the basic emotional language value set is established, and all the music periods in the music library are mapped to the emotional space through emotional cognition experiment. The statistical analysis of the experimental results itself contains the process of averaging, so it can be seen as an analysis of public emotion. The retrieval carried out in this space forms public emotion retrieval. We use an interactive algorithm to achieve personalized emotional retrieval to integrate the users' emotional needs into the interactive evolutionary process.

Conclusion

In this paper, we establish an emotion cognitive model based on Hervner's Emotion Ring theory. The model possesses 10-dimensional musical feature space, which can accurately grasp the music's physical characteristics of periods. Through the discriminating method, we find that the model's emotion judgment is relatively accurate. According to the emotion cognition distance of several music samples, the deviation is less than 0.035. In addition, variance calculation results of the descriptors in the feature space also show that the prediction of the cognition model accords with the actual choice of the experimental objects.

The musical feature space is established through the analysis of musical features to construct the musical emotion cognitive model. This is an automatic recognition process of emotional content of musical works, which is different from the traditional method of emotion clustering. The model in this paper can be interpreted as the semantic similarity between the musical emotion and the basic language value set, which conforms to the actual situation of people's cognition of musical emotion cognition.

Finally, we use the emotion cognitive model to design a semantic-driven music retrieval method. And then we talk about its application in the interactive online music education system, which has achieved good results.

References

Culén, A. L., Mainsah, H. N., & Finken, S. (2014). Design practice in human computer interaction design education. *International Conference on Achi*, 300-306.

Cano, M. D., & Sanchez-Iborra, R. (2015). On the use of a multimedia platform for music education with handicapped children: A case study. *Computers & Education*, *87*(C), 254-276. http://dx.doi. org/10.1016/j.compedu.2015.07.010

Faiola, A., & Matei, S. A. (2010). Enhancing human-computer interaction design education: Teaching affordance design for emerging mobile devices. *International Journal of Technology & Design Education*, *20*(3), 239-254. http://dx.doi. org/10.1007/s10798-008-9082-4

Feng, Y., Zhuang, T., & Pan, H. (2001). Query similar music by correlation degree. *Multimedia Information Pressing*, *2195*, 885-890. http://dx.doi. org/doi.org/10.1007/3-540-45453-5116

Hevner, K. (1936). Experimental studies of the elements of expression in music. *The American Journal of Psychology*, *48*(2), 246-268. http://dx.doi. org/10.2307/1415746

Katayose, H., Imai, M., & Inokuchi, S. (1988). Sentiment extraction in music. *International Conference on Pattern Recognition*, *2*, 1083-1087. http://dx.doi. org/10.1109/ICPR.1988.28447

Katayose, H. (1990). Expression extraction in virtuoso music performances. *International Conference on Pattern Recognition*, *1*, 780-784. http://dx.doi. org/10.1109/ICPR.1990.118216

Liu, T., Sun, S., & Pan, Y. (2004). Emotional recognition for chime bell music. *IEEE International Conference on Systems*, *1*, 568-573. http://dx.doi. org/10.1109/ICSMC.2004.1398360

Liu, D., Zhang, N., & Zhu, H. (2003). CAD system of music animation based on form and mood recognition. *Pattern Recognition & Artificial Intelligence*, *16*(3), 283-288.

Mills, J. (1999). Using a Large Data Set for Research in Music Education: Music in Years 10 and 11. *Bulletin of the Council for Research in Music Education*, *17*(141), 103-107.

Peng, Q. (2007). Research on computer analysis and automatic recognition technology of musical emotion. *Ph.D thesis, Shanghai Jiao Tong University,* Smith, M., Smaill, A., & Wiggins, G. A. (2000). Music education: An artificial intelligence approach. *Workshops in Computing*, 221-239. http://dx.doi. org/10.1007/978-1-4471-3571-5

Talsik, E. (2015). The investigation of readiness for e-learning of pre-service music teachers in Turkey. *Anthropologist, 21*(12), 263-270.

Thayer, R. (1989). The biopsychology of mood and arousal. *Quarterly Review of Biology, 1*, 195-222. http://dx.doi. org/10.2307/1415746

Yun, S., & Chang, D. Y. (2009). Speech emotion recognition via a max-margin framework incorporating a loss function based on the Watson and Tellegen's emotion model. *IEEE International Conference on Acoustics*, 4169-4172. http://dx.doi. org/10.1109/ICASSP.2009.4960547