Research on audio test of Chinese national musical instruments based on subjective perception and objective audio characteristics

Yang Qing

College of Arts, Hunan University of Arts and Science, Changde, Hunan, 415000, China

Keywords: Subjective perception; objective audio characteristics; Chinese national musical instruments; audio testing

Abstract: Music and voice interaction become the necessary interface of all kinds of personal and home devices, and users' pursuit of high sound quality becomes the driving force of product quality. Higher quality sound requires higher precision measurement, but the industrial trend is ruthlessly reducing production costs. So better and faster become the requirements of sound test system. From the perspective of subjective perception score and objective audio characteristics, this paper studies the audio test of 37 common Chinese national musical instruments. According to the subjective perception scores of musical instruments, a visual subjective audio description method is proposed, and the similarity matrix of subjective timbre is obtained by calculating the differences between audio perception vectors. Then the objective audio features representing timbre are extracted, and the differences between features are calculated to obtain the similarity matrix of objective timbre features.

1. Introduction

Audio is defined as the acoustic perception that distinguishes different sounds under the same pitch and intensity. As far as musical instruments are concerned, audio is the most important basis for distinguishing different musical instruments. The most intuitive way to describe audio perception is to use discrete subjective descriptors to describe audio [1]. In the literature of Western orchestras and Chinese orchestras, many descriptive words describing the sound characteristics of musical instruments are recorded, such as "bright", "sharp", "soft", "dark", "thick" and "hoarse". Researchers often screen audio descriptors through subjective or objective selection methods. For example, subjective selection methods include word selection method, mandatory selection evaluation method, etc. [2]. Objective selection methods include correlation analysis, cluster analysis, multi-dimensional scale analysis, etc. The advantage of using single descriptors to describe musical instrument audio is intuitive and in line with people's perception. The disadvantage is that there is no degree of descriptors, and the non independent relationship between descriptors is ignored [3]. Since 1870s, scholars have been looking for robust objective features to explain timbre perception, that is, using objective audio features to represent timbre. Early studies on timbre perception generally believed that spectral centroid is one of the important dimensions affecting timbre perception [4]. In addition, spectral deviation, spectral density, attack time, attack center time, sustain time, amplitude spectrum envelope and spectral flux are common characteristics of timbre [5]. In order to improve the efficiency of audio signal transmission and storage, codecs are widely used in digital audio systems, and the introduction of codec algorithms makes the traditional objective measurement indicators no longer correspond to the perceived sound quality [6]. At the current level of technology, subjective evaluation is the most reliable method for testing perceptible audio quality, but subjective evaluation is time-consuming and laborious, requires very strict test design and test environment, and cannot meet the needs of on-site testing. Audio test is a typical example of large quantity and wide area, easy to get started and accurate and difficult. As far as the tested objects are concerned, from large speakers to small mobile phone headsets, they all bring information and pleasure through audio [7].

  DOI: 10.25236/icallh.2021.051

2. Audio test

The spectrum centroid of audio is used to describe the center of gravity of spectrum power spectrum, which indicates that the range of pre dominant frequency is the coefficient of power spectrum. For fast calculation, coefficients below 62.5hz have been combined. Spread spectrum is the mean square value of the logarithm power spectrum of the center of gravity for each frame, which is used to describe the shape of the power spectrum. The standard deviation of the average of all frames is used to describe the spread spectrum of audio signal. For most people, the output of audio devices is good. It is loud and undistorted. Ring, that is, sensitivity, that is, the input level is converted into the output level or the measurement of the digital sound size. The sound level is measured by sound pressure level in engineering. Sound pressure level is the ratio between the pressure of sound wave and the pressure of a very small sound. This ratio has a wide range, so it is usually expressed by logarithm. Therefore, the unit of sound pressure level is dB. There are some differences in the audio frequency of different kinds of musical instruments. For example, plucked instruments have obvious bulges in the concord dimension. Most wind instruments have obvious depressions in the crisp dimension. In the same category of musical instruments, the audio of different musical instruments is not the same. For example, in bowed stringed instruments, Zhonghu has a softer voice, and there is little difference in the average score of each descriptor, and the corresponding timbre spider diagram is rounded and approximately circular. Jinghu has a higher average score on crisp and bright description words, and a lower average score on dark and thick description words. Therefore, the audio has obvious fluctuations. This method overcomes the shortcomings of using a single descriptive word to describe national musical instruments in the past, and the description of timbre is vivid, intuitive and more in line with people's perception. Due to the characteristics of electro-acoustic equipment, the output amplitude response of each frequency audio frequency is different after passing through the equipment, so there is a frequency response test. Frequency response test verifies the ratio of output signal to input signal at each frequency. In fact, the test response is to test the frequency response of multiple frequency points in the nominal working frequency range. For horn acoustic test is the ratio of sound pressure level to input voltage, for power amplifier is the frequency curve of the ratio of output voltage to input voltage. Figure 1 shows the typical principle of audio test, in which figure (a) is the test of power amplifier and other electronic components, and figure (b) is the test of loudspeaker and other sound components.

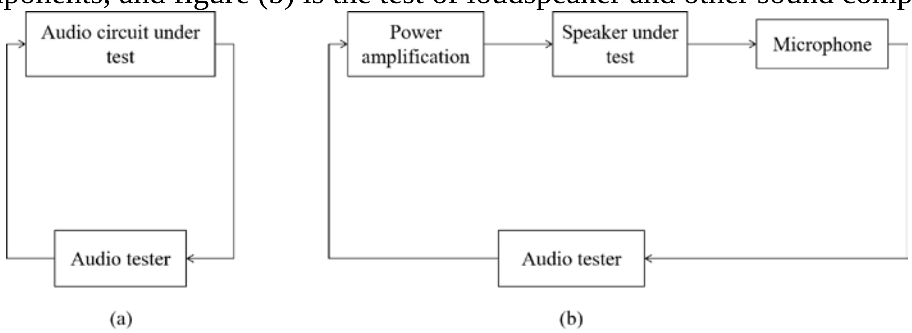

Figure 1 Audio test principle

3. Research on audio test of Chinese national musical instruments based on subjective perception and objective audio characteristics

This experiment uses the subjective evaluation data set of national musical instrument audio. The data set contains sample data of 37 ethnic musical instruments, and each instrument contains a piece of performance of 3~4s. The sampling rate is 44lkHz, and the sampling bit depth is 16bit. According to the categories, musical instruments can be divided into bowed instruments, wind instruments, plucked instruments and percussion instruments. Through a series of word selection correlation analysis, cluster analysis, etc., 16 musical instrument audio descriptors are obtained: slender, bright, dim, sharp, thick, thin, thick, crisp, shriveled, hoarse, rough, pure, harmonious, full, Soft and cloudy. The 16 audio descriptors can represent the whole audio perception more completely. Table 1 lists the Chinese national instruments and their classifications. The subjective evaluation process was that 34 participants perceived the music segments of each instrument for 3-4s, and scored 16 auditory attributes, with a score range of 1-9.

Table 1 List of Chinese National Musical Instruments and Their Classification

Bowed instrument	Wind instrument		Plucked stringed instrument	Percussion
Gao Hu	Bangdi	Bass Suona	Liuqin	Cloud gong
Erhu	Qudi	Executives	Soprano Ruan	Chimes
Zhonghu	New flute	Middle tube	Pipa	
Gehu	Big flute	Low tube	Dulcimer	
Beigehu	Treble sheng	Contrabass	Zhongruan	
Jinghu	Alto Sheng	Bau	Da Ruan	
Banhu	Bass sheng	Collapse	Guzheng	
	Treble suona	Xiao	Guqin	
	Alto Suona		Sanxian	
	Tenor Suona		Konghou	

Based on the description method of audio perception score, the ith instrument can be expressed as vector V_{i16} according to the average score result. By calculating the Euclidean distance between 37 instruments in the data, the difference matrix D_{sub} based on subjective perception score can be obtained. The dimension is 37×37, and the subscript "sub" is a subjective abbreviation, indicating that the matrix is calculated based on subjective perception score. In order to express the similarity, this paper adopts the method of normalization and difference from 1, and the element value of the similarity matrix $S_{sub}.S_{sub}$ can be approximated as 0~1. The closer to 1, the more similar the audio scores of the two instruments.

In the similarity matrix of Chinese national musical instrument audio based on objective audio features, the feature vector similarity of similar musical instrument audio is more obvious. With similar spectrum and audio quality, these instruments are easily misclassified by each other. For example, erhu, Zhonghu and Gaohu, which belong to the same playing instrument, are wrongly divided into Gaohu and Zhonghu before adding the new eigenvalue. After adding the new eigenvalue, there are 9 wrongly divided into Gaohu and 31 wrongly divided into Zhonghu, indicating that the addition of the new eigenvalue does not play a significant role in the differentiation of the family instruments, However, it is noticed that the number of misclassification of these three musical instruments to other musical instruments decreases, which indicates that the addition of new eigenvalues can help to distinguish musical instruments of different families. At the same time, it is also noted that after adding new feature values to guzheng, pipa and Ruan, which are both playing instruments, the number of instances correctly classified by them increases, while the number of instances wrongly classified to other family instruments decreases, indicating that the addition of new feature values has a good effect on distinguishing the instruments played. Then, if you want to measure the audio characteristics accurately and quickly, you can use discrete frequency sweep. Simply put, it is to traverse all test frequencies with one excitation, and obtain the response of all frequencies at one time. Although from the signal point of view, each frequency is successively experienced, but only one occurrence-acquisition-analysis is performed, which makes the test process significantly faster than a single frequency cycle.

Specifically in the audio test, the test methods that can be implemented include frequency traversal, multi audio test, discrete sweep and continuous sweep. Table 2 summarizes the characteristics of each method.

Table 2 Comparison of audio measurement methods

	Speed	Precision	Harmonic analysis	Machinery defect inspection
Frequency point traversal	Slow	High	Stand by	Difficult
Multi-audio	Fastest	Low	Difficult	Difficult
Discrete frequency sweep	Faster	High	Stand by	Difficult
Continuous frequency sweep	Fast	High	Stand by	Stand by

4. Conclusions

In this paper, the subjective auditory characteristics of 37 Chinese national musical instruments are studied, and a visual audio test method of national musical instruments is put forward through the subjective perception scores of musical instruments, and the similarity matrix of subjective timbre is obtained by calculating the differences between timbre perception vectors. In a word, under the background of advanced manufacturing, more and more test measurements have changed from simple readings to engineering improvement and scientific research innovation. Subjective perception is easier to operate than objective audio features, but the interference of individual subjective differences cannot be ruled out. It has a certain degree of subjectivity and is difficult to fully reproduce. The objective audio features use a large number of complex signal processing and scientific algorithms, and the conclusions are objective, which can be reproduced anytime and anywhere, and effectively save manpower and material resources. In addition, we can further study the causes of cross group similarity to test whether the playing order has an impact on the sound similarity of these instruments. We can also conduct subjective experiments in low, medium, high range and limit range respectively to calculate the subjective perception and objective characteristics similarity of different sound areas of the same instrument, So we can get the subjective characteristics and corresponding objective characteristics of each instrument in different sound areas. This will provide more detailed parameters for the music information system of Chinese national musical instruments.

Acknowledgements

Hunan Philosophy and Social Science Fund Project 2020,A study on the discourse pattern of & quot; telling Chinese story well & quot; in Contemporary National instrumental music, Serial number: (20YBA195)

References

[1] Li Zijin, Jiang Yiliang, Liang Xiaojing. Research on the timbre similarity of Chinese national musical instruments based on subjective perception scores and objective audio characteristics. Journal of Fudan University (Natural Science Edition), vol. 59, no. 3, pp. 90- 97, 2020.

[2] Pan Fei. A Preliminary Study on the Sound Power Test of the Yangqin, a Traditional Chinese Musical Instrument. New Movies and Dramas, vol. 177, no. 1, pp. 187-190, 2018.

[3] Hu Yaowen, Long Hua, Sun Jun, et al. Research on the Classification of Musical Instruments Based on Audio Features. Software Guide, vol. 17, no. 6, pp. 17-21, 2018.

[4] Cheng Jun, Jixiang, Ma Yunfeng, et al. Expressive recognition of video advertisements fused with video and audio features. Journal of China Academy of Electronics, vol. 14, no. 1, pp. 102-104, 2019.

[5] Zhou Jinao, Long Hua. Multilingual classification algorithm based on audio feature parameters. Communication Technology, vol. 51, no. 10, pp. 86-91, 2018.

[6] Zhao Xiuwen, Liu Wuying, Li Fuyu, et al. Efficient music retrieval method based on audio fingerprint features. Journal of Armed Police Engineering University, no. 4, pp. 27-32, 2018.

[7] Qiu Kaibin, Li Jianliang. Audio feature extraction method for UAV recognition. Noise and vibration control, vol. 38, no. 2, pp. 188-192, 2018.