

Analysis of Tibetan Folk Music Style Based on Audio Signal Processing

Ma Ying, Li Kaiyong^{*}, Hou Jiayu, Ga Zangjia

National Experimental Teaching Demonstration Center of Communication Engineering, Qinghai University for Nationalities, Xining, China

Email address:

1037125248@qq.com (Ma Ying), 492302219@qq.com (Li Kaiyong), 1493564749@qq.com (Hou Jiayu), 345592115@qq.com (Ga Zangjia)

^{*}Corresponding author

To cite this article:

Ma Ying, Li Kaiyong, Hou Jiayu, Ga Zangjia. Analysis of Tibetan Folk Music Style Based on Audio Signal Processing. *Journal of Electrical and Electronic Engineering*. Vol. 7, No. 6, 2019, pp. 151-154. doi: 10.11648/j.jeeec.20190706.13

Received: November 11, 2019; **Accepted:** November 16, 2019; **Published:** December 3, 2019

Abstract: National folk music has different styles, has extremely strong regional and national characteristics, and has a high cultural and artistic value. It carries the profound connotation of national culture. Music has non-semantic symbolicity and strong ambiguity, which makes the related research topics of music signals more challenging than speech signals. With the rapid increase of the number of digital music, due to the complexity of music itself, the ambiguity of the definition of the category of music and the limitation of the understanding of the characteristics of human auditory perception, Therefore, the analysis of the characteristics of folk music is a prerequisite for realizing the rapid and effective retrieval of folk music resources, and plays an important role in audio signal processing. However, there are few studies on the classification and information extraction of folk music. The article is based on the St-EN and St-ZCR feature extraction of the three styles of music in Aze, Le, and playing and singing in Amdo Tibetan folk music. Three kinds of musical styles have adopted a time-domain analysis is briefly analyzed Amdo Tibetan folk music signal, by extracting signal features music, We can find short-term energy than the short-time average zero-crossing rate of all types of music more clearly reflect the unique characteristics of the signal.

Keywords: National Folk Music, Extraction of Music Features, Classification of Music, Short-term Energy, Short-term Zero-crossing Rate

1. Introduction

National folk music is a cultural product of a nation, produced under years of living labor and different environments. Tibetan folk music is diverse and rich in content. [1] It expresses the production, labor, and character traits of Tibetan people and shows the unique social culture of Tibetan areas in a colorful and unique tone. Due to the unique natural ecology and human history of the Tibetan community, the Tibetan traditional music reflects diversity and integrity, regardless of its artistic features or social functions. [2]

Therefore, Tibetan traditional music can be divided into three types: folk music, religious music, and court music. Folk music is the most important part of Tibetan traditional music, and it is also the most widely distributed, the most diverse and the most abundant. [3]

2. Introduction of Tibetan Folk Music Style

2.1. Classification of Tibetan Folk Music

The classification of Tibetan folk music mainly includes: 1. Folk songs are called "folk songs", one of the most rich musical genres in Tibetan traditional music. 2. Singing and dancing music refers to singing accompanied by dancing, or dancing accompanied by singing, which is an art form of singing and dancing. 3. Instrumental music is primarily folk songs or song and dance music tunes performed by Musical Instruments. 4. The drama music is mainly the Tibetan traditional drama Ajilam music, including four types: Character vocal music, drumstick lining music, the chanting rhyme with white tone and the interspersed singing and dancing music. 5. Rap music is a special art form which is based on literature.

2.2. Characteristics of Amdo Tibetan Folk Songs

Amdo Tibetan folk music genre is rich, there are le, aze, three types of playing and singing. In the literature in ref. 1 and 2 combined with the feature of amduo Tibetan folk culture and folk songs, appreciates the amdo Tibetan music feature of rich folk music. Literature appreciation, 2 in the Tibetan language is the meaning of "song", also calls the toasting song in amdo Tibetan areas, and is widespread and the most current mining a kind of folk song, also is the most popular a kind of folk songs [2]; A ze is the meaning of a small folk song and dance, mainly popular in the agricultural areas of gansu and qinghai or half rural or half pastoral areas, mostly in wedding ceremony and other festive festivals, singing accompanied by simple movements, slow rhythm steps, melodious melody [1]; As a typical representative of amdo Tibetan folk songs, playing and singing is rich in various forms, with fresh and lively music, with a little dance characteristics and more feather modes [2]. In this paper, music feature extraction method will be adopted to appreciate and analyze the style of Tibetan folk music in amdo, which will lay a better foundation for the classification and retrieval of folk music.

3. Analysis of Tibetan Folk Music Style Based on MATLAB

3.1. Analysis in Time-domain

3.1.1. Short-term Energy

The short-term energy of signals can show the variation trend of signal amplitude. Besides being applied to endpoint detection, the main purpose is to distinguish unvoiced and voiced segments. The short-term energy of the n th frame signal can be expressed as

$$E_n = \sum_{m=0}^{N-1} x_n^2(m) \quad (1)$$

This is a function for measuring the amplitude change of signals. It is very sensitive to high levels. [4, 5]

3.1.2. Short-term Average Zero-crossing Ratio

The short-term zero-crossing rate represents the number of times that a frame signal waveform passes through the horizontal axis. For continuous signals, zero-crossing means that the time-domain waveform passes through the time axis; while for discrete signals, if the adjacent sample values change symbols, it is also called zero-crossing. [6, 7]

The short-term zero-crossing ratio of the defined signal is

$$Z_n = \frac{1}{2} \sum_{M=0}^{N-1} |\text{sgn}[x_n(m)] - \text{sgn}[x_n(m-1)]| \quad (2)$$

The short-term zero-crossing rate is the most widely used to judge high and low sounds in music signal analysis. [8, 9]

3.1.3. Short-term Autocorrelation

The short-term autocorrelation of the defined signal is

$$R_n(k) = \sum_{M=0}^{N-1-K} x_n(m)x_n(m+k) \quad (0 < k \leq K) \quad (3)$$

Here K is the maximum delay point. [10, 11]

In signal processing, the autocorrelation function is periodic. The pitch period can be estimated by the position of the first peak in the autocorrelation function. [12, 13]

3.2. The Characteristics of Tibetan Folk Songs

3.2.1. Analysis of the Characteristics of "Aze" Music Signal

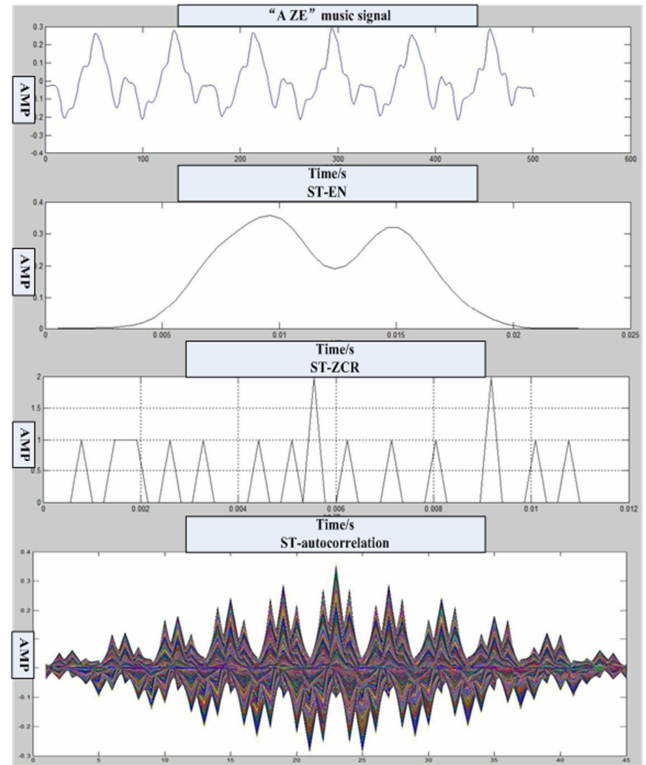


Figure 1. Feature analysis of Aze music signal.

It can be seen from figure 1 that one frame of music signal of "Aze" has obvious high and low rhythm change, which can be seen that the treble duration is slightly higher than the bass, the tone is high, the rhythm is slower; the short-term energy appears the two highest The amplitude point, one frame of music signal has obvious two energy increase and decrease regular; the short-term average zero-crossing rate is relatively stable at first, then it changed to bass after a high pitch appeared in the middle, and then a high pitch appeared again after a period of stability, which was somewhat similar to the short-term energy change. The short-term autocorrelation was most obvious in the periodicity of the three time domain analyses.

3.2.2. Analysis of the Characteristics of "Le" Music Signal

It can be seen from figure 2 that the music signal of "Le" is the same as "Aze", and there is also a significant change in the high and low sounds. The difference is that the treble lasts longer than the bass duration, the tonality is melodious, and has the character of long cavity. The short-term energy is different from "Aze". Only one highest amplitude point appears in one frame, that is to say, the obvious increase and decrease of energy in the singing process.; the short-term

average zero-crossing rate , it starts off smooth but relatively dispersed , "Le" has the meaning of laziness. In the middle, there was a high pitch followed by a stable change. Compared with the hasty singing style in the early stage, "le" was free to stretch, and the short-term autocorrelation function showed a changing rule. The melody was euphemistic and passionate.

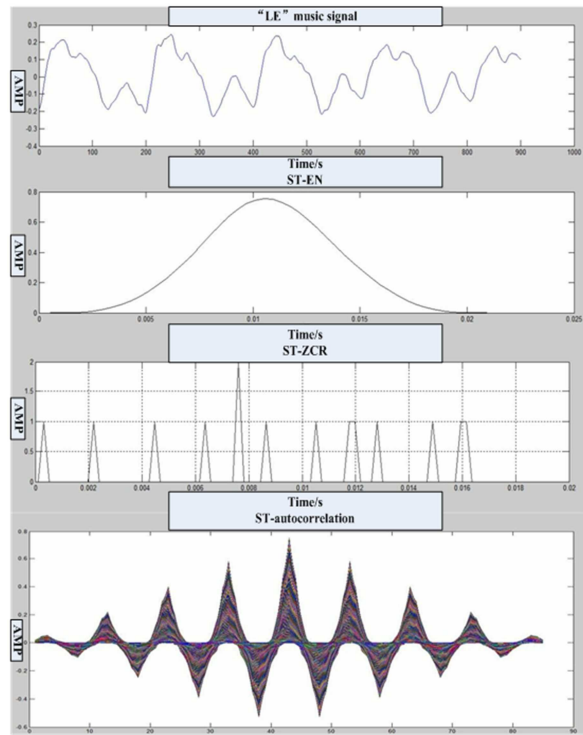


Figure 2. Characteristic analysis of Le music signal.

3.2.3. An Analysis of the Characteristics of the Music Signal in "Playing and Singing "

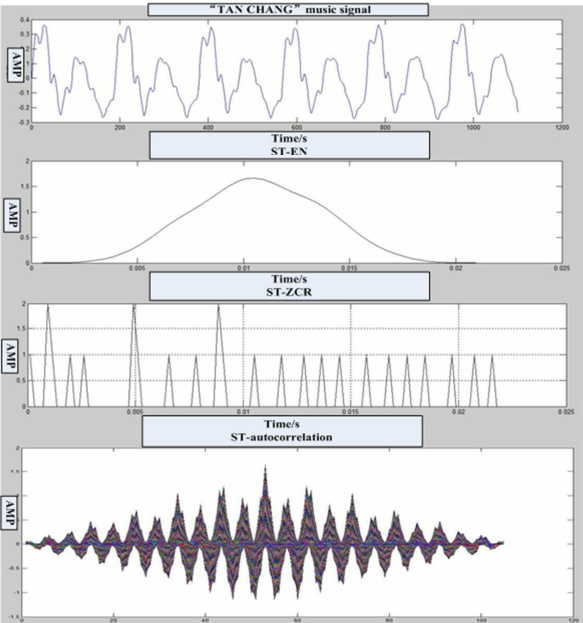


Figure 3. Time Domain Analysis of Music signal of playing and singing.

It can be seen from figure 3 that the music signal of "playing and singing " has a certain change law but different from "Aze", the treble duration is balanced with the bass duration, and it can be clearly seen that the high and low pitch are faster and the rhythm is more cheerful, short, crisp; The short-term energy is the same as "Le", there is a maximum amplitude point , and there is a clear law of increase and decrease; short-term average zero-crossing rate is basically consistent with "Le", the previous period changes smoothly but relatively scattered in the late stage, there is also a high-pitched change in the middle; the short-time autocorrelation function is basically close to the sinusoidal curve, and the periodicity is strong.

4. Analysis and Comparison of Simulation Results

Table 1. Time-domain eigenvalues extracted from three types of music signal.

The time domain eigenvalue	Music type		
	Aze	Playing and singing	le
Analysis method			
Short-time energy	0.3583	1.6630	0.753
Short-time average zero crossing rate	2	2	2

5. Conclusion

All three music styles adopt the time-domain analysis method to intercept a frame of music signal and add Windows for analysis. According to the simulation results, the largest short-term energy is "playing and singing". "Playing and singing " is to sing a tonic or a backbone sound first, and then to freely extend or to sing a extremely short, clear and coherent throat trill on the basis of the relatively stable position of the larynx in the upper and lower levels. The "Le" melody is euphemistic, the tone is passionate, and the lyrics are mostly expressed as ambitious goals and grand ambitions and blessing. It has more solo forms. The "Aze" tunes are very rich and very characteristic. There are songs such as "playing and singing " and "le", such as free-stretching "introduction", long cavities and vibrato, and the light rhythm of Tibetan songs and dances. It has become a unique style in the folk songs of Amdo Tibetan Region. By extracting the characteristic values of music signal, we can find that the short-term energy can reflect the unique characteristics of various music signals more clearly than the short-term average zero-crossing rate.

Music has non-semantic symbolicity and strong ambiguity, which makes the research on music signal more challenging than voice signal. The music analysis of ethnic minorities is still in the research stage, and its potential social and economic value is self-evident. From the current research status and application requirements, in the foreseeable future, content-based automatic music classification and retrieval will continue to be the main research direction in the field of music analysis and recognition. [14, 15]

Fund Project

Qinghai Province Key R&D and Transformation Project 2019-GX-170.

References

- [1] Zhuo Macuo. Research on the Arrangement of Tibetan Folk Music [J]. Drama House, 2016 (11): 59-60.
- [2] Luo Hongyu, Li Wei. Categories and characteristics of the qinghai anduo Tibetan folk songs [J]. Journal of Qinghai normal university (philosophy and social sciences), 2012, 34 (06): 104-107.
- [3] Jue. The Classification and Distribution of Tibetan Traditional Music and Related Natural Information and Humanities [J]. Musical Art (Journal of Shanghai Conservatory of Music), 2015 (01): 123-136.
- [4] Zhang Chunlei, Zeng Xiangyang, Wang Shuguang. Endpoint detection based on critical band power spectrum variance [J]. Acoustic technology 2012, 31 (02): 204-208.
- [5] Zhao Li. Speech Signal Processing (Second Edition) [M]. Mechanical Industry Press, 2010. 37-44.
- [6] Sun ke. Research on feature extraction and classification of Chinese national music [D]. Donghua university, 2017.
- [7] Feng Shang, Xu Zhongliang, Ma Lin, Li Haifeng. Study on the sparse analysis method of music signals [J]. Journal of Fudan University (Natural Science), 2018, 57 (03): 328-334.
- [8] Xu Jincai. Research on audio signal separation based on deep neural network [D]. Beijing University of Posts and Telecommunications, 2018.
- [9] Guan Xin. Research on algorithm for automatic classification of music signals [D]. Tianjin University, 2009.
- [10] Li Min, Zhang Weiwei, Jiang Mingxin, Liu Yan. Design of Music Analysis and Synthesis Based on MATLAB [J]. Journal of Dalian Nationalities University, 2010, 12 (03): 269-271.
- [11] Qin Jing. Research and application of music retrieval technology based on content and semantics [D]. Dalian University of Technology, 2018.
- [12] Deng Jianguang, Pan Xiaoheng, Lin Yuzhi. Review of instrument recognition based on acoustic characteristics [J]. Journal of dongguan institute of technology, 2012 19 (03): 58-64.
- [13] Zhou mingzhong, ji lixin. VAD algorithm based on average amplitude and weighted zero-crossing rate and its FPGA implementation [J]. Journal of university of information engineering, 2010, 11 (06): 713-718.
- [14] Y. Fuji, K. Yamamoto, and S. Nakagawa, Hidden Conditional Neural Fields for Continuous Phome Speech Recognition, IEICE Trans. Inf. & Syst., Vol. 95-D, No. 8, pp. 2094-2104, Aug. 2012.
- [15] T. N. Sainath, D. Nahamoo, B. Ramabhadran, D. Kanevsky, V. Goel, and P. M. Shah, Exemplar-based sparse representation phone identification features," Proc. ICASSP, pp. 4492-4495, 2011.