Advanced Science and Technology Letters
Vol.122 (ASP 2016), pp.247-252
http://dx.doi.org/10.14257/astl.2016.122.47

Improvement of Spectral Fingerprint for Audio Content Recognition

Jeongseok Jo[1], Sungjun Han[2] and Jongweon Kim[3]

[1] Dept. of Copyright Protection, Sangmyung University, Seoul, Korea
[2]Dept. of Contents and Copyright, Sangmyung University, Seoul, Korea.
[3]Dept. of Contents and Copyright, Sangmyung University, Seoul, Korea, Corresponding Author
{jeongs,sungjun}@cclabs.kr, {jwkim}@smu.ac.kr

Abstract. In this paper, we proposed an improved fingerprinting algorithm for audio content recognition. There are many feature points as fingerprints in the spectral components and this is a burden causes about the feature database size and retrieval time. The main fingerprints are spectral peaks and distances between the peaks and this is famous open source algorithm. As mentioned above, there are too many components and we studied to reduce the spectral features without loss of the matching rate. Human auditory system cannot distinguish some peaks in the same bark band and we applied 24 bark bands to spectral components. We chose representative peaks in a bark band and then made fingerprint using these peaks. Proposed algorithm improved the search speed was doubled and the DB size reduced by half.

Ke**ywords:** audio content recognition, fingerprint, bark band

1 Introduction

Many people listen to the music using the device such as MP3 player, tablet and the mobile phone. The music is distributed in the right holder. The right holder produces the music. Then, they have the rights for the music and it is protected by the copyright law.

The music is produced to the digital music from music producers by using the digital production devices and the software. The digital music is easily copied and transmitted from the illegal consumer to others by using BitTorrent, P2P and the shared web server [1].

The illegal consumers obtain profit instead of the original owners like music producers. Therefore, we use the audio content recognition to protect the copyright infringement for the music using fingerprint algorithm. If the unique codes such as watermark are inserted into the music and the right holder has the unique feature generated fingerprint algorithm for this music in the database, the right holder can

ISSN: 2287-1233 ASTL

track the music distributed by the illegal consumer by using the audio content recognition like fingerprint algorithm.

The previous fingerprint method is that WILLDREVO-dejavu has the good identification ability regarding the music recognition [2]. Nevertheless, the recognition speed of this method is slow to track the music because this method generates many fingerprints. Thus, it should reduce the speed of the recognition and the storage. Therefore, we improve the performance using bark band for these problem.

This paper is divided into 4 sections. The section 1 explains the related work. The section 2 represents the proposed recognition method. The section 3 shows the experiment and the result. The section 4 is the conclusion.

2 The related work

2.1 The WILLDREVO-dejavu

The WILLDREVO-dejavu [3] called as the dejavu use the hash as the feature of fingerprint. This method process is shown as fig 1.

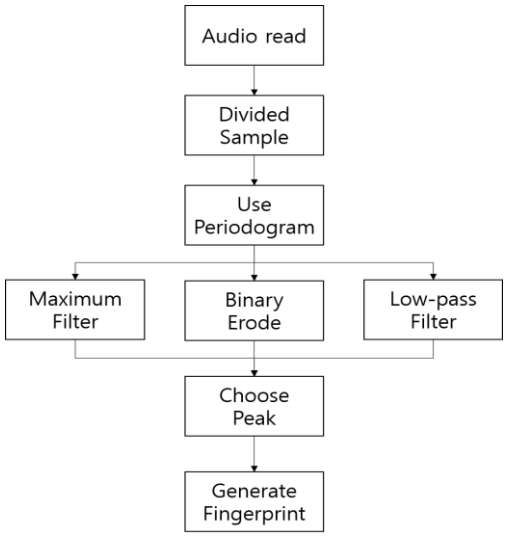

Fig. 1. The process of the previous method

In this method, the audio samples are divided into 4096 samples and move as much as the window size after the half of samples are overlapped. Then, Samples are transformed by periodogram. The data transformed by periodogram have the location of the time and frequency and the amplitude values. For this data, this method performs the maximum filter, the binary erosion and the low-pass filter. These extract peaks and reinforce the extracted peaks in order to be tolerated for the audio signal attack. The extracted peaks are used to generate to the hash values by SHA-1 hash algorithm. These converted hash become fingerprints.

3 The proposed method

In the proposed method, we improve the dejavu algorithm using bark band [4]. The dejavu method generates many peaks. Therefore, the recognition speed is decreased and many storages are occupied in the database. However, it has the good performance for the music recognition. We improve the recognition speed and the volume of storages is decreased by reducing peaks. The proposed method process is shown as fig 2.

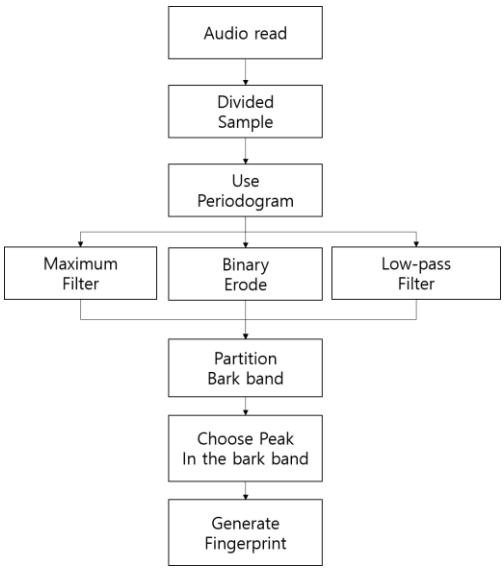

Fig. 2. The process of the proposed method

In the proposed method, we use bark band. The bark band is defined as a psychoacoustical scale proposed by Eberhard Zwicker in 1961. He measures the subjective loudness and divides the frequency according to perceptually equal distances [5]. He discovers that the bark scale is linear below 500 Hz. After these process, he separates the bark ranges included the number of 24 scale values [6,7].

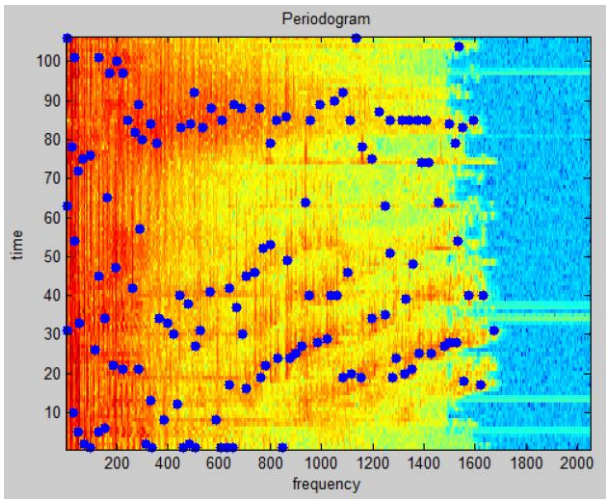

Fig. 3. The peaks of the previous method for 5secs music

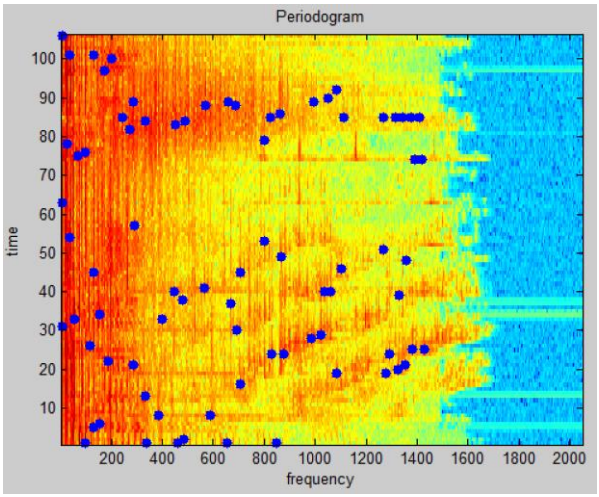

Fig. 4. The peaks of the proposed method for 5secs music

We use the characteristic of the bark band. We divide the frequency of the peak as shown as fig 3 using the bark bank. After using the bark band, we select the representative peaks located higher the center frequency among the bark scale as shown in fig 4. We generate fingerprint using SHA-1 hash algorithm through the generated peaks.

4 Experiment result

In the experiment, we use the number of 30 audio contents. We compose the query audio composed of 1, 3, 5 seconds from the original audio contents. To compare between the methods, we rebuild the open source code using the MATLAB [8]. We compare the creation speed for fingerprints, the recognition speed and the storage of the database of previous method with these for the proposed. The experiment result is displayed in the table 1.

Table 1. The experiment result for the compassion of two algorithm

List		The previous method	The proposed method
The creation speed for fingerprints		2324 sec	1638 sec
The occupied storage in the database		22.2MB	10.4MB
The recognition speed	1sec	About 47 sec	About 20 sec
	3sec	About 47 sec	About 20 sec
	5sec	About 47 sec	About 20 sec

In the experiment result, the recognition rate for the previous method and the proposed method is 100%. The proposed method uses a small amount of peaks than the previous method, the creation speed for fingerprints, the occupied storage in the database and the recognition speed is higher than the previous method about 2 times.

5 Conclusion

We improve the previous method using bark band to reduce the number of peaks. We compare the performance of the previous method with that of the proposed method. We decrease the creation speed for fingerprints, the occupied storage in the database and the recognition speed. The performance of the proposed method is improved about 2 times.

Although we improve the previous method, we do not find the time index. However, the previous method can find the time index. Thus, we will study to expand reducing the number of peaks and finding the time index.

Acknowledgments. This research project was supported by Ministry of Culture, Sports and Tourism (MCST) and from Korea Copyright Commission in 2015.

References

1. Jin, X., Kim, J.: Digital Music Piece Identification using Local Maxima. In: Advanced Science and Technology Letters, vol.52, pp.65--69, AICT, (2014).
2. Jo, J., Kim, J.: Improvement of a music identification algorithm for time indexing .: Advanced Science and Technology Letters, vol.120, pp.765--769, (2015)
3. Wang, A.: An Industrial Strength Audio Search Algorithm. In: ISMIR. p. 7—13, (2003).
4. Haitsma, J., Kalker, T.: A Highly Robust Audio Fingerprinting System. In: ISMIR. p. 107--115, (2002).
5. Zwicker, E.: Subdivision of the audible frequency range into critical bands. In: The Journal of the acoustical Society of America, vol.33, pp. 248, (1961).
6. Traunmüller, H.: Analytical expressions for the tonotopic sensory scale. In: The Journal of the Acoustical Society of America, pp.97—100, (1990).
7. Sound Perception: The Science of Sound Design, http://home.ieis.tue.nl
8. Lartillot O., Toiviainen, P.: A Matlab toolbox for musical feature extraction from audio. In: Proc. the 10th Int. Conference on Digital Audio Effects, pp. 10--15, (2007).

