International Journal of Security and Its Applications
Vol. 10, No. 1 (2016), pp.437-448
http://dx.doi.org/10.14257/ijsia.2016.10.1.38

Copyright Management for Mix Music

Jeongseok Jo[1] and Jongweon Kim[2*]

[1]Dept. of Copyright Protection Sangmyung University
[2]Dept. of Contents and Copyright Sangmyung University
[1]jeongs@cclabs.kr, [2]jwkim@smu.ac.kr

Abstract

We propose the copyright management for mix music system that classifies the information of the music used by the producers and records the range for it. The system saves the information of the produced mix music in the database. Mix music means a mixture of musical styles in the field of music. In the course of mix music production, the music producer needs permission of usage of music copyright and pays the whole price of the music. However, the most mix music producers use only part of that music. The payment for the whole of the music is very expensive. Therefore, they use the music illegally, so that they do not have to pay the expensive price. If it becomes possible to pay only that used part, the producers will not use it illegally and also the music market will be invigorated. Many mix music producers experienced inconvenience because they should input much information manually. The program to find the information of used music and input it automatically. In addition, the information about the created mix music is managed automatically in the database for the mix music. This system improves the previous music identification algorithm. We use the continuous combined hash. It reduces the useless matched hash. After identifying music, we select the largest offset that is time index. In the experiment, the performance of music identification is increased and that of the time index for the original music is improved.

Keywords*: Mix music, Music identification, Time index, Copyright management, Audio fingerprint*

1. Introduction

The many people connect their own favor music, and construct a new soundtrack on the internet these days. The Youtube web site offers Mashup API for them to make it easily. Naturally, it is not breach of copyright to create a new music by using it, because it is possible to make use only the contents uploaded on the Youtube web site. However, the mix music used not permitted by the copyright holder is the copyright infringement [1].

Mix music is composed of DJ Mix that special beat makers produce, mash up that makes the mixed music like original one by composing more than two songs of only voice and accompaniment. The medley that sticks several songs by switching only rhythm. Such mix music is regarded to need a system that administers copyright efficiently because many works are too confused in one work to distribute copyright fee. [2-4]

There is a case of violation in the mash up music that is a kind of mix music because of not getting any permission from the writer and copyright holder. We can see it in the case of Bridgeprot v.Dimension in 2005. The court ruled that even one second of utilization of the original music can be illegal in the mash up [3, 5].

The mix music creator should pay for the whole original music even when he uses only a small part. That is why they cannot use it easily if they are not possible to pay for the

* Corresponding Author

ISSN: 1738-9976 IJSIA

music sample. Especially, there are more illegal cases in the latest music because it costs more. At this moment, P2P, BitTorrent are the main web sites to download sound sources and this can occur another infringement of copyright. [6] The case that uses whole original music in mix music is rare. Such a case can be, however, mix music makers do not usually use the whole part. However, music producers pay the whole original music. It is not a burden to popular mix music producer, but it is burdensome to most producers and it motivates them to use it illegally [7].

If it becomes possible to pay only the used part to solve the problems stated above, producers can make wider range of music, and distribute mix music more smoothly. Furthermore, the copyrights of the produced mix music can be protected by registering them. Mix music producers enter music information that they used manually into their mix music. They may not know what music they used when they use a lot of music because they feed input data into the system manually [8].

We can solve such a case by finding the data of used music in the database that the music is saved and input the music data automatically. [9] Mix music producers can make music more conveniently. If the charge for using is administered together, the problem of the charge will be solved by putting the settlement fee of producing music automatically. In this paper, we propose a system that inputs music data automatically when the mix music producer uses it, and informs the used part of that music.

This paper is divided into 4 sections. The Section 1 explains related work. The Section 2 explains copyright management system for mix music. The Section 3 shows the experiment results. The Section 4 is the conclusion.

2. Related Work

2.1. Spectrogram

Spectrogram is a method that visualizes frequencies and the signal is defined as the Fourier transform. It is shown as the two-dimentional image composed time domain and frequency domain. The x axis of spectrogram indicates time and the y axis represents frequency. The colors of that means amplitudes as shown in Figure 1. The amplitude has higher values in the red and the lower values in the blue. Spectrogram can help to easily compare with two audio signal because it displays the amplitude variation freely in the time and frequency domain.[10] The translated form of spectrogram is periodogram. It is commonly used to estimate of the power spectral density. We can analysis the audio signal more in detail because it can estimate the best tempo for music. Therefore, we use periodogram instead of spectrogram. [11]

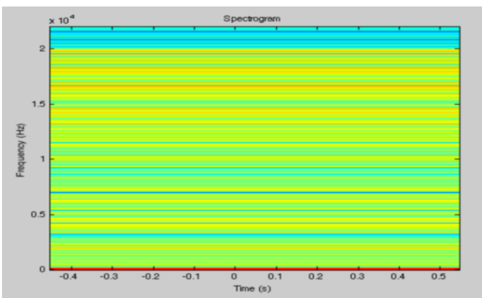

Figure 1. Spectrogram

2.2. Audio Fingerprint

The fingerprint of human was used from 2000 years ago as thumbprints. The fingerprint of human is very specific information for human because each of people cannot have the same fingerprint. [12] Audio fingerprint use this property.

Audio fingerprint means the short feature information extracted by audio signal. [13] It represents the feature of music. It is generated by the the audio signal algorithm to be able to distinguish the audio. It has been commonly used to copyright protection and the music identification. It is tracking the illegal distributed audio in the copyright field and used to identify music in the music identification field.

2.3. WILLDREVO-dejavu Open Source

WILLDREVO-dejavu is an open source that has the music identification and the time indexing function. It has the good performance for music identification and the function to be able to find time index. It uses the hash as fingerprint extracted for music signal to identify music and can find time index through the offset that is the location of the time index. Figure 2 is the WILLDREVO-dejavu open source process. [14]

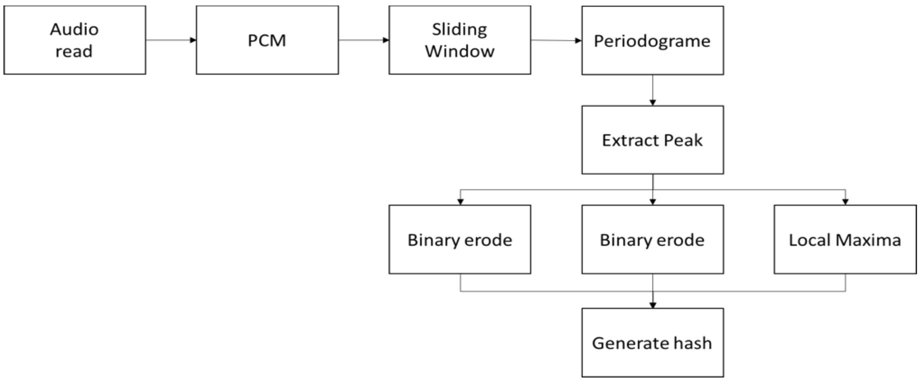

Figure 2. WILLDREVO-dejavu Open Source Process

2.3.1. Peak Finding

In this algorithm, the peak finding is the basis phase to generate hash and offset. The peaks are extracted at the location of the highest amplitude value in the time and frequency domain. These are emphasized by using high pass filter because peaks in the low frequency band are the susceptible to attack. The filtered pecks can be used to the constellation map as shown in Figure 3. [14]

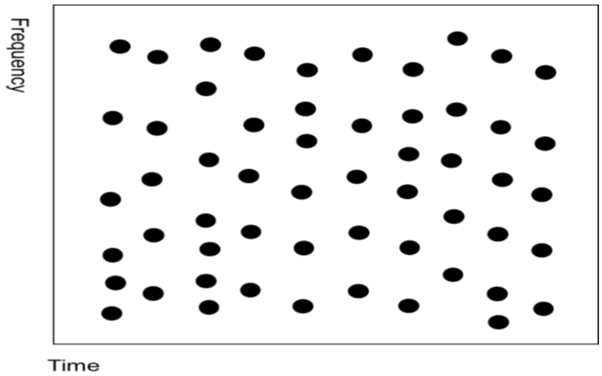

Figure 3. The Constellation Map

2.3.2 Fingerprint Hashing: Despite not like the music, these can have the similar pecks because the music signal is often similar. Therefore, this algorithm use hash function to reduce the overlap of fingerprints because the property of hash is very sensitive. For example, if we change 1 bit for the sentence wanted to convert to hash, the hash is generated to the different value.

The hash value is generated by the distance between the anchor peak and the candidate peaks, the time and the frequency of the anchor peak as shown in Figure 4. [14]

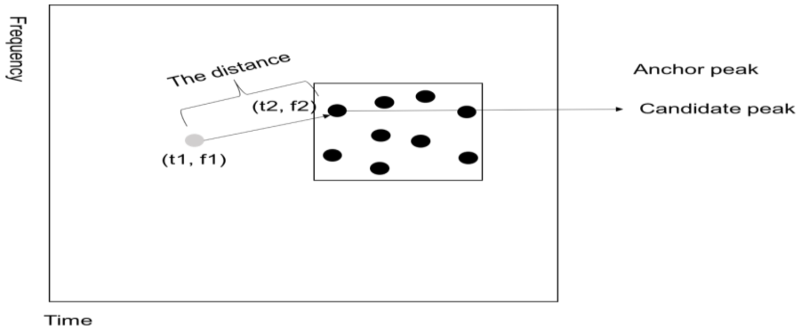

Figure 4. The Distance of Pecks

2.3.3 Robust Constellation: Before generating hash, this algorithm processes the robust constellation to obtain the correct features. The robust constellation is composed by anchor peck and candidate peaks. The anchor peak is selected in the constellation map. In this process, this algorithm fixes the anchor peck to make to the robust peck for noise and selects the candidate peaks that have the higher amplitude than around the anchor peak. Position of the selected candidate peaks is coupled with the front anchor peak in time. The pattern of peaks is composed as the robust constellation map and used to generate hash as shown in Figure 5. [14]

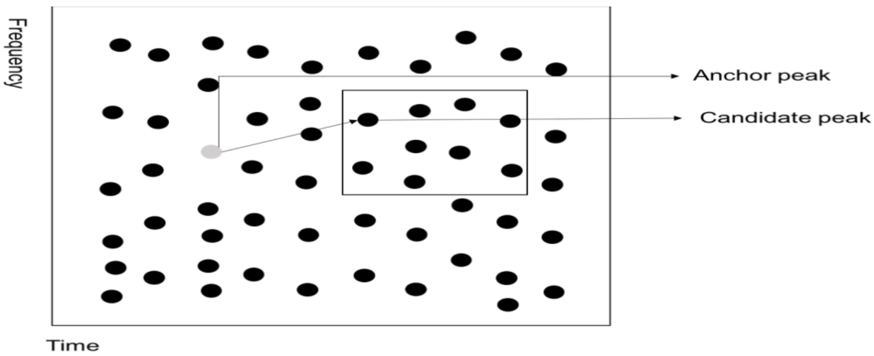

Figure 5. The Robust Constellation

2.3.4 Fast Combinatorial Hashing: In the process of music identification, if we do not use the fast combinatorial hashing, it is very slow because the total of extracted hash are compared with the feature hash of query music. Therefore, this algorithm uses the indexing of constellation map. The constellation map selects the anchor peak in the robust constellation process. The anchor peaks are combined with candidate peak in the target zone. This form is called as the combinatorial hash. The fast combinatorial hashing is process that matches among the combinatorial hash. Through the process, if we identify the music, we identify the anchor peaks and the connected target zone to compare with the combinatorial hash of the original music and that of the query music. [14]

2.4. OpenFP Open Source

OpenFP is the music identification open source as shown in Figure 6. It uses the Bark band and IIRLP as the main algorithms. OpenFP uses libfftw3 library for audio file process. Libfftw3 has many audio processing function as audio processing library.

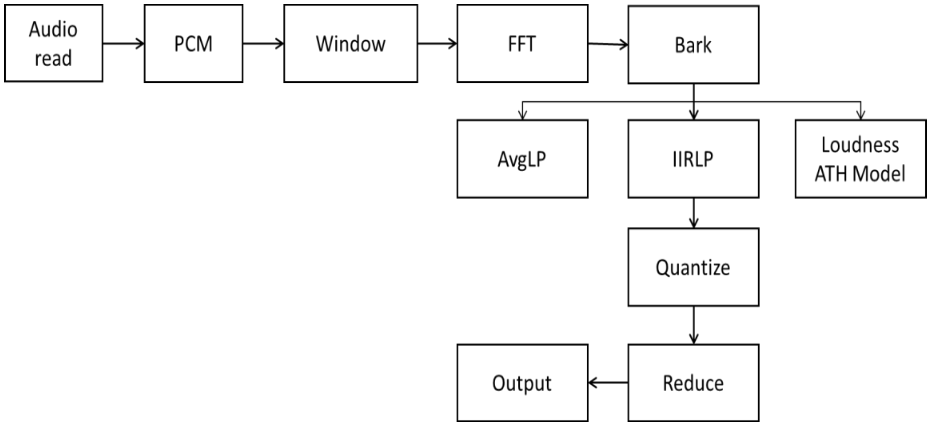

Figure 6. OpenFP Total Process

2.4.1 Fingerprint Extraction: In the openFP, the segmentation is performed twice for the audio sample. The first segmentation is composed as frame. The frame is divided into subframes as shown as Figure 7. The subframes are translated to complex numbers by FFT and calculate using hamming window. The features of subframes are extracted from each of subframes using bark band, lowpass filter and quentization. [15] The bark band reduce power spectrum and lowpass filter decrease the noises. After the bark band and lowpass filter, the subframe is quntized by binary flag. The features of subframes are composed to produce subfingerprints.

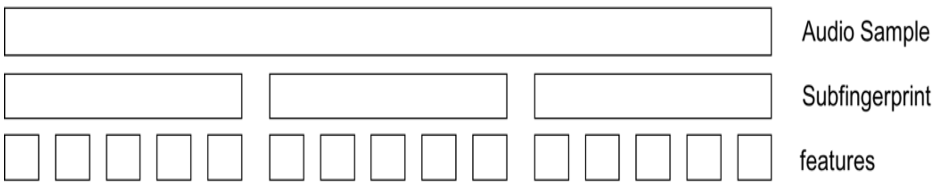

Figure 7. OpenFP Composition

2.4.2. Fingerprint Clustering: The extracted subfingerprints generate the vector using Mel Frequency Censtral Clustring(MFCC). MFCC is a power spectrum divided to frequencies and generated by Mel Frequency Censtral(MFC). The subfingerprints is clustered by MFCC. The generated vector is another feature of the subfingerprints. This algorithm uses the vector of the subfingerprints in the identification process to the first matching features to improve the matching time and then it uses the subfingerprints for the subfingerprint of the matched vector [16].

3. Copyright Management System for Mix Music

The proposed system is the copyright management for the mix music. The mix music producer manually input the music information in the music management system. However, manually inserting the music information is a very hard work for producer from a position because the producer uses and creates many music. Therefore, when the producer makes the mix music, the producer need to automatically input the music information and extract the feature of the created mix music to save in the music management database for the copyright protection of the created work. Figure 8 is the total system process as follow.

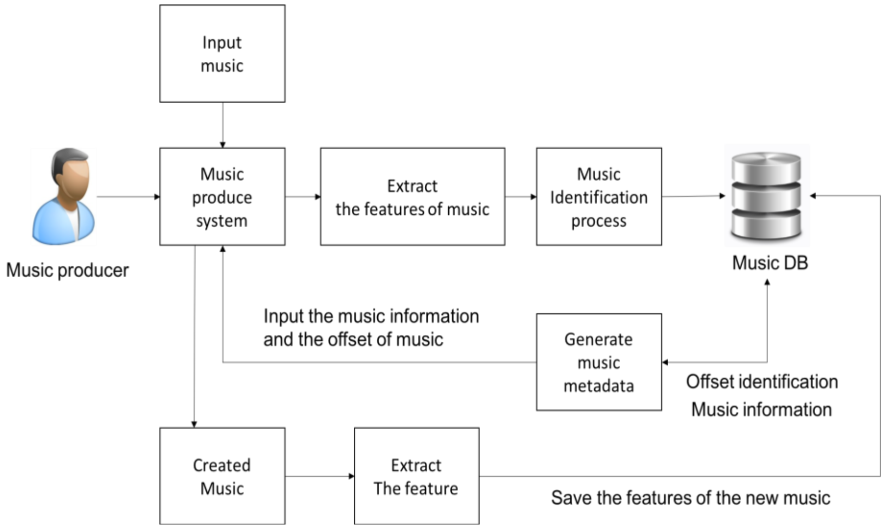

Figure 8. Music System about Music Identification and Time Index

The music information is composed by the music name and the time index of the music and so on. The music name is found by the music identification algorithm. In the cases of time index, finding the time index is identified by the offset. We improve the music identification and the time index function based on the previous method to build the system.

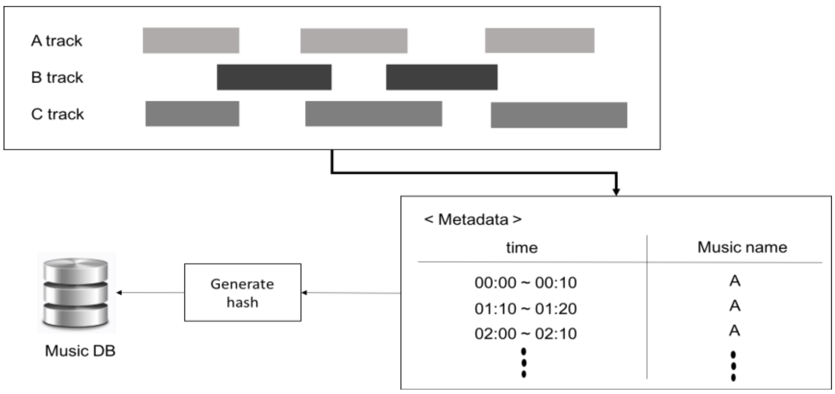

Figure 9. Metadata for Mix Music

3.1. The Music Identification

The previous method identifies all hash for the music that producer uses in the identification process. It can increase miss match rate because the music that producer do not uses has the same hash with the music that producer uses in the case of having the same peaks. Therefore, we need to protect this situation to identify the music exactly. We add the index into the hash. When we identify the music, we find the continuous matched hashes called hash set using index. The size of initial hash set is used the continuous index as the number of five times as shown in Figure 11.

For example, if we cannot identify hash set or the number of the continuous matched hash is less than 20, we should increase the size of the hash set. If we identify many hash set over 500, we should decrease the size of that. We should stop until it is more than 20 less than 500 for the continuous matched hash.

The offset informs us the time index of the original music for the query music. We identify the largest counted offset from the continuous matched hash as shown in Figure 12. The total process of the proposed method is as shown in Figure 10.

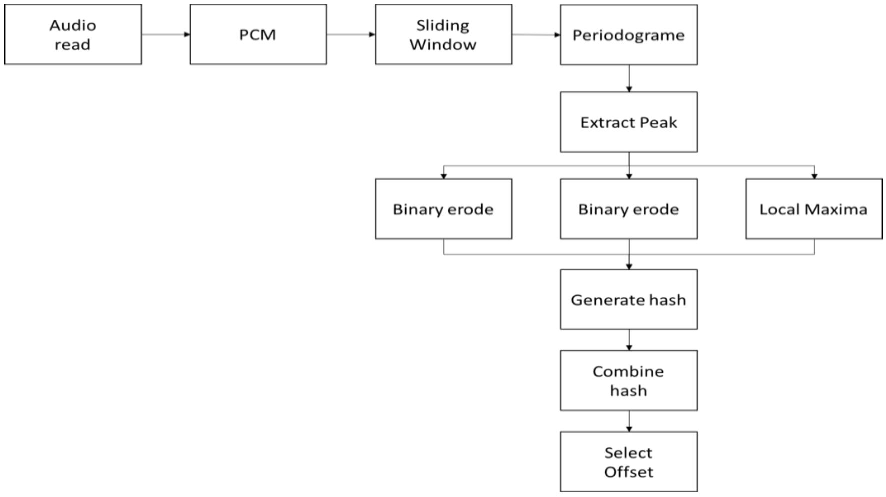

Figure 10. Proposed Algorithm Total Process

The proposed method is not generally different from the previous method. However, we improve the algorithm using the combined hash method and exactly find the time index through the combined hashes and the selected offset.

The combined hash called hash set is listed as a from. The largest count offset is found in the listed hash set. When we define the window size as *NFFT*, window overlap rate as Overlap and audio sample rate as Fs. The found offset is calculated by (1).

$$Second = round(\frac{offset}{Fs} * NFFT * Overlap) \tag{1}$$

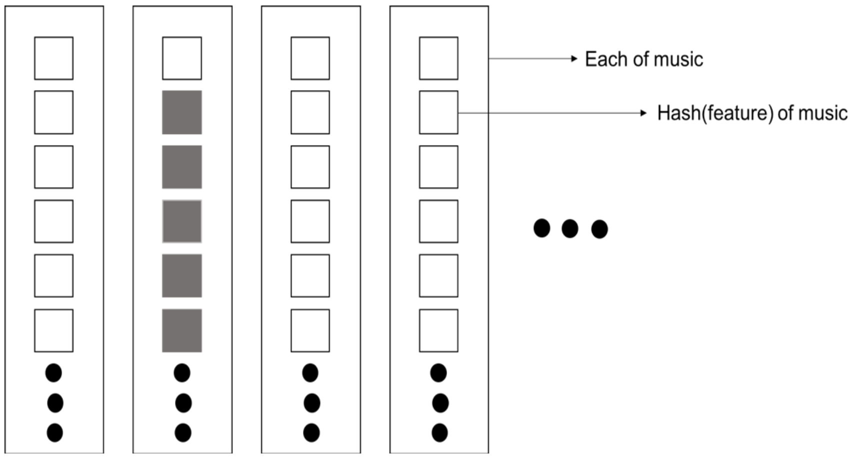

Figure 11. The Combined Hash

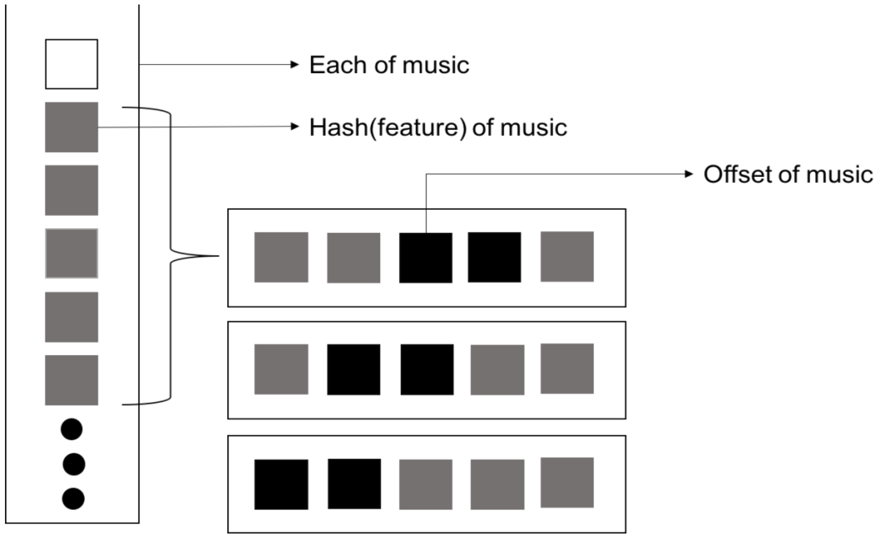

Figure 12. The Searched Time Index

3.2 The Database System

The identified music by hash set and the found location of query music by offset is saved in the database of the music produce system automatically using the database query system as shown Figure 13. [17] The feature of the produced mix music is extracted and saved in the database. These information is shown to the mix music producer.

The previous method is only used to save purpose. Therefore, if the database amount is increased, the identified time is also increased. To prevent this situation, we use the database query identification. We improve that all identification process is executed in the database and the result is only obtained. [18, 19]

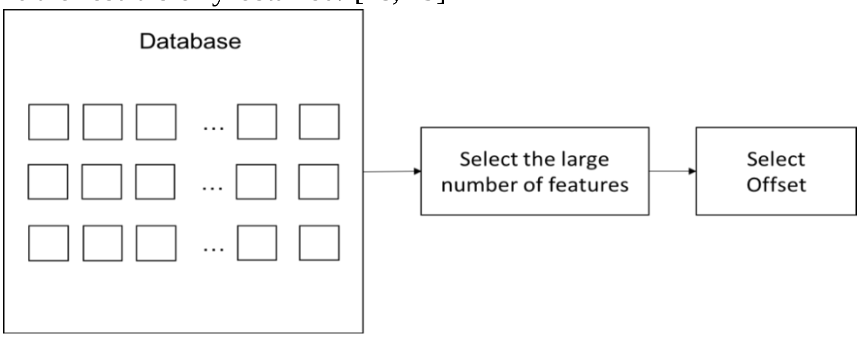

Figure 12. Database Composition

3. Experiment Result

The experiment was aimed at algorithm used in the system and comparison target was the algorithm of WILLDREVO-dejavu open source. It is the previous music retrieval algorithm and this algorithm have time-offset function. To compare smoothly with offered algorithm, we transform "WILLDREVO-dejavu open source" into MATLAB language and we also write the proposed algorithm in MATLAB language. [20]

In the database, there are characteristics of twenty-five music samples. In the nine samples out of 25 samples, the query music is extracted randomly at intervals of two seconds such as 1sec, 3sec and 5sec.

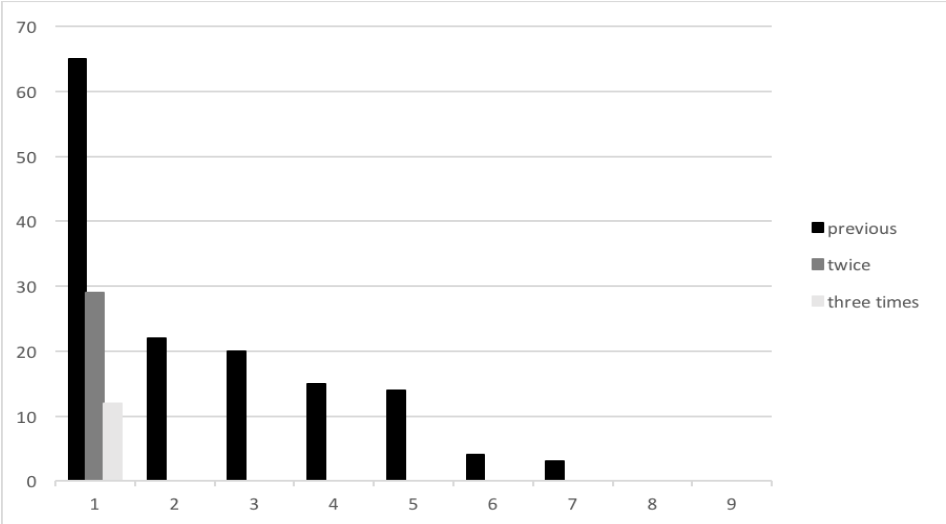

Figure 13. The Effect of Overlap for Hash

The ranking is listed by the number of the matched hash between the hash extracted from the selected 1 second query music and the hash extracted from the original music. The previous method is distributed in the rank two to rank seven. It shows that the hash of others is matched with the hash of the query music except the hash of the original music. However, the proposed method is distributed only rank one. It shows that the propose method is more stable than the previous method.

The previous method compares approximately the number of 1000 offsets. It produces many useless offsets as shown in Figure 15. However, the proposed method compares within the number of 500 offsets. It generates necessary offsets to use the finding time index. As a result, the proposed method is reduced for the comparison times and we also reduce the miss matched rate for offset.

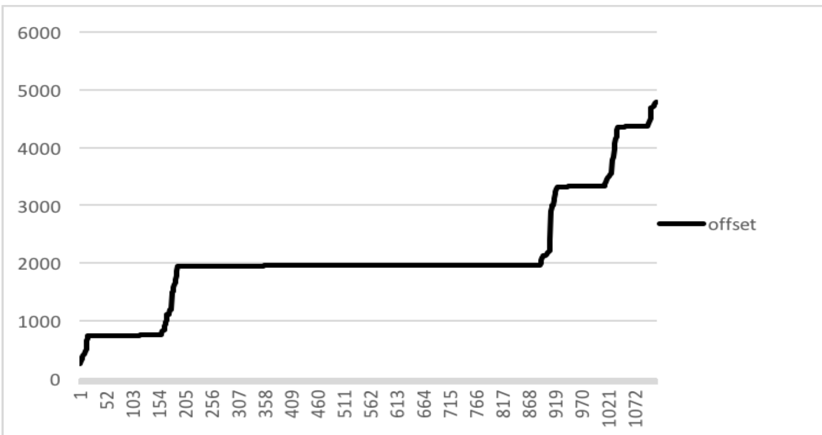

Figure 14. The Offset Distribution for the Previous Method

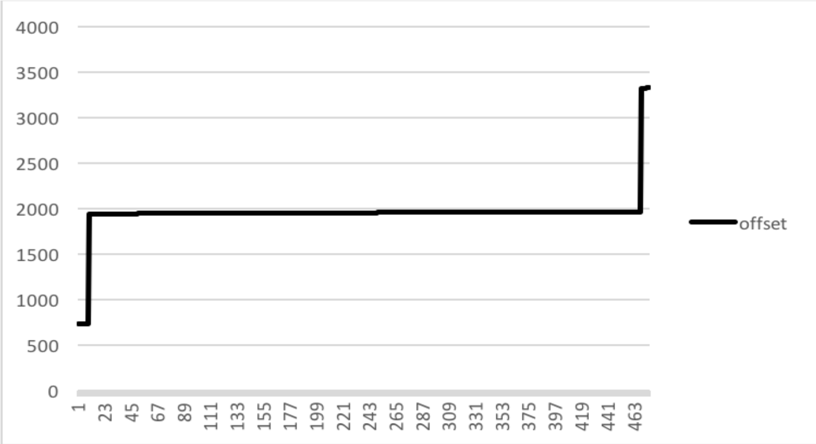

Figure 15. The Offset Distribution for the Proposed Method

Table 1. Experiment Result

List		WILLDREVO-dejavu	OpenFP	The proposed method
Recognition ratio	1sec	80%	0%	80%
	3sec	100%	0%	100%
	5sec	100%	100%	100%
Smallest units		1sec	5sec	1sec
Processing speed		About 70sec	About 1sec	About 3sec

The experiment results show that WILLDREVO-dajavu open source use database system only for storage purpose. Therefore, when it identifies the music, the duration for identifying the music and time index is about 70 seconds. OpenFP has the best search time about 1second. However, openFP cannot be identified for the time index and the query music for 1second and 3seconds. Thus, we cannot use this algorithm in the proposed system. The duration for identifying the music and time index of the proposed method is about 1second. It is better than WILLDREVO-dejavu for processing speed. The reason of that the proposed method is faster than WILLDREVO-dejavu for processing speed is that the proposed method uses the search function of the database using query sentences.

4. Conclusion

To build the mix music produce system for the copyright management, we improve the previous algorithm. However, the identification rate is not different with the previous algorithm.

The function of the previous algorithm has the identifying music and the finding time index. However, in the experiment result, it has a few problems about which the extracted hashes for the query music are matched with the hashes of others except for the hashes of the original music. In addition, it influences the performance for the function of the finding time index. Therefore, we improve the function of finding time index. As a result, we decrease the number of times for the comparison and we obtain more exactly the time index set.

The precious algorithms only use the database as the storage purpose. Thus, the identification time is slow. To improve it, we use the database query system. As a result, we increase the identification time about 20 times. In the study result, the identification rate of the proposed algorithm is insufficient. Therefore, we improve it through the continuous study.

Acknowledgments

This research project was supported by Ministry of Culture, Sports and Tourism (MCST) and from Korea Copyright Commission in 2015. This paper is a revised and expanded version of a paper entitled “Improvement of a music identification algorithm for time indexing” presented at “The 4th International Conference on Digital Contents and Application” held on December 16-19, 2015 at Jeju Island, Korea

References

[1] Pike and H. George , "Google, YouTube, copyright, and privacy." Information Today, vol.24, No.4 **(2007**), pp.15.
[2] Boone, Christine Emily, "Mashups: History, legality, and aesthetics." **(2011).**
[3] D. Mongillo, "Girl Talk Dilemma: Can Copyright Law Accommodate New Forms of Sample-Based Music, The." Pitt. J. Tech. L. & Pol'y 9, **(2009)**.
[4] Simpson-Jones and Katie, "Unlawful Infringement or Just Creative Expression-Why DJ Girl Talk May Inspire Congress to Recast, Transform, or Adapt Copyright." J. Marshall L. Rev.43, **(2009**).
[5] O'Brien, S Damien and B F. Fitzgerald, "Mashups, remixes and copyright law." Internet Law Bulletin, vol.9, **(2006)**, pp. 17-19.
[6] J Xun and J Kim, “Digital Music Piece Identification using Local Maxima”, Advanced Science and Technology Letters, AICT, vol.52 **(2014),** pp.65-69.
[7] McGranahan and Liam, "Bastards and Booties: Production, Copyright, and the Mashup Community.", Trans: Transcultural Music Review= Revista Transcultural de Música, vol.14 **(2010)**.
[8] Ellis, PW Daniel, and E. Graham Poliner, "Identifyingcover songs' with chroma features and dynamic programming beat tracking." Acoustics, Speech and Signal Processing, 2007., ICASSP 2007, IEEE International Conference on, vol. 4, **(2007)**.
[9] T Nao, "Massh!: a web-based collective music mashup system." Proceedings of the 3rd international conference on Digital Interactive Media in Entertainment and Arts. ACM, **(2008)**, pp.526-527.
[10] M J Bastiaans, , "A sampling theorem for the complex spectrogram, and Gabor's expansion of a signal in Gaussian elementary signals." Optical Engineering, vol.20, **(1981).**
[11] Lartillot, Olivier, P.Toiviainen and T. Eerola, "A matlab toolbox for music information retrieval.", Data analysis, machine learning and applications. Springer Berlin Heidelberg., **(2008)**, pp.261-268.
[12] Haitsma, Jaap, and T. Kalker, "A Highly Robust Audio Fingerprinting System." ISMIR. Vol. 2002 **(2002)**, pp. 107-115.
[13] J Seo, J Haitsma, and T Kalker, "Linear speed-change resilient audio fingerprinting." Proc. IEEE Workshop on Model based Processing and Coding of Audio, **(2002)**.
[14] WANG and Avery, “An Industrial Strength Audio Search Algorithm.”, ISMIR, **(2003)**, pp. 7-13.
[15] Smith III, O Julius and J S. Abel, "Bark and ERB bilinear transforms." Speech and Audio Processing, IEEE Transactions on, vol.7, **(1999)**, pp. 697-708.
[16] S. Xi, C. Xu and S. Mohan Kankanhalli, "Unsupervised classification of music genre using hidden markov model.", Multimedia and Expo, 2004., ICME'04. 2004 IEEE International Conference on. vol. 3. **(2004)**, pp. 2023-2026.
[17] Chaudhuri, Surajit and Gautam Das, "Automated ranking of database query results." In CIDR. , (2003).
[18] Jeongseok Jo and Jeongweon Kim, “Improvement of a music identification algorithm for time indexing “, Advanved Science and Technology Letters, DCA 2015, vol.120, **(2015),** pp.765-769.
[19] Chaudhuri and Surajit, "Probabilistic ranking of database query results." Proceedings of the Thirtieth international conference on Very large data bases, VLDB Endowment, vol.30, **(2004)**, pp. 888-899.
[20] O. Lartillot and P. Toiviainen, “A Matlab toolbox for musical feature extraction from audio”, Proc. the 10th Int. Conference on Digital Audio Effects, **(2007)**, pp. 10-15.

Authors

Jeongseok Jo, he received his B.S. degree in Computer Science and Intellectual property from Sangmyung University, Korea, in 2015. He is currently pursuing the M.S. degree in Copyright Protection, Sangmyung University, Korea. His research interests are digital watermarking, 3D model retrieval, 3D model identification, music identification, digital forensics.

JongWeon Kim, he received the Ph.D. degree from University of Seoul, major in signal processing in 1995. He is currently a professor of Dept. of Contents and Copyright at Sangmyung University in Korea. He has a lot of practical experiences in the digital signal processing and copyright protection technology in the institutional, the industrial, and academic environments. His research interests are in the areas of copyright protection technology, digital rights management, digital watermarking, and digital forensic marking.

