Similarity Measurement Method between Two Songs by Using the Conditional Euclidean Distance

Min Woo Park and Eui Chul Lee*
Department of Computer Science, Sangmyung University
7 Hongji-Dong, Jongno-Gu, Seoul, Republic of Korea
*E-mail: eclee@smu.ac.kr

Abstract: - Since numerous songs have recently been released increasingly, the genre of the song clustering is reasonably more important in terms of the audience's choice. Also arguments for plagiarism are continuously being raised. For this reason, similarity measurement between two songs is important. In previous works, although similarity measurement has been actively researched in the field of query by humming, they only focused on quite partial matching for input humming. To solve this problem, we proposed a novel similarity measurement method between two songs. The proposed method has several advantages compared with previous works. Firstly, it is possible to measure overall similarity between two different songs. Secondly, overall region of a song is represented as 1-dimensional signal which can be obtained by run-length representation of 2-dimensional note information ((pitch, duration)). Thirdly, by sequentially adopting median filter, average filter, Z-score normalization into the 1-dimensional signal, we obtain the overall flow without noise feature such as the eccentric note of the song. Lastly, a new distance metric namely the conditional Euclidean distance is used by combining two distance concepts such as the Euclidean distance and the Hamming distance. To perform the feasibility test, several famous songs by the Beatles and the MIREX`08 dataset were used for our experiment. Also, by applying our method into a comparison between two songs with a plagiarism issue, we confirmed that very high similarity score between the two songs was measured.

Key-Words: - Music similarity measurement, Music clustering, Music plagiarism

1 Introduction

Since multimedia contents have recently been distributed in large quantities, people have the ability to access numerous multimedia contents more easily. Music is also distributed into a digital format such as MP3 files, so numerous songs have been released and allowing people to listen easily. However, due to too much music, selecting the desired music is becoming more difficult. Also, due to the increased amount of music, there is a lot of controversy regarding plagiarism. To solve this controversy, a method is necessary to express quantitatively, compare and classify music data. Since the pitch and the duration of notes can be expressed as the number, it is possible to apply mathematics based data mining and pattern recognition techniques.

Firstly, researches for extracting a main melody track from a polyphonic MIDI file were performed which can be categorized into four classes such as the rules based methods, statistics based methods, methods using the largest repeat pattern, and methods using artificial mark. The meaning of the rules is for determining the main track. The statistics case use traditional methods of pattern recognition. In addition, the largest repeat pattern is identifying the largest repeat pattern as the main track. Lastly, artificial mark case is that mark the main track by artificial ways. In these studies, the related with rules, a numeric vectors are extracted based on note pitches and note duration in MIDI files to automatically find melody track [1]. To determine melody channels, they were clustered depending on average pitch frequency and pitch histogram and the experiments by using Best-k channel and skyline algorithm in case of statistics [2]. They have proposed a dictionary-based to find repetitive patterns based on Lempel-Ziv 78 [3] related with the largest repeat pattern case. Lastly, they have proposed an algorithm for melody extraction from MIDI files by using melody similarity and a hamming retrieval system related with artificial mark [4].

Secondly, query by humming method has been researched which is important meaning for music analysis. The query by humming required pitch contour extraction process based on previous voice signal processing methods [5-8]. For that, pitch contours of query were extracted by using variable-scale windowing and wavelet transformation [9]. In a previous research of query by humming, they have performed Query-by humming system through the use of G.729 feature extractor for selecting the features of humming data and DTW matching using 48 MIDI files and 2,797 humming queries wav files

[10]. Also for improving the retrieval accuracy and retrieval speed, they have proposed multiple distance measurement fusion mechanisms for both note-based matching and frame-based matching with 2,797 sung queries from 118 persons by using Earth Mover`s Distance and Dynamic Time Warping and achieved 92.9 accuracy [11]. They used energy difference, two phases of cepstrum with zero-process, pitch contour and cepstrum peak value curve for doing rough note segmentation, tracking pitch from non-whistle inputs, doing note segmentation and dealing with note insertion respectively [12].

Thirdly, by using these methods for both MIDI files and Query by Humming in the field of music, many studies have been conducted about music similarity. They have approached the similarity of polyphonic music segments by using feature-driven clustering and contrapuntal similarity of the segments. Also, through the Naïve Bayes classifier they have evaluated the classification process, and as a result they have achieved almost 80% accuracy [13]. Furthermore, they have proposed a similarity matching technique based on continuous melody contour by using melody slope alignment and melody contour similarity measure, consequently their system has achieved 88% correct retrieval [14]. In the other work, they performed music genre classification through a time function expressing as n-dimensional pitch and duration vectors with 50 MIDI formats from 3 categories such as classical, pop, and traditional Japanese music [15]. However in previous researches, most music analysis was only focused on local matching. Since the local matching method is not suitable for overall matching between two songs, a new overall matching method should be researched.

In this paper, a new music similarity measurement method is proposed that can measure the similarity between each song by analyzing music data. In previous researches for query by humming, those were the ways to find the most similar song by matching the features of the query which is a part of a song. However, in our research, the correlation between two types of music of different tempo and chord is entirely measured by comparing overall regions of the music.

This paper is organized as follows. In section 2, we explain the proposed method in which generation method of 1D numeric string for overall region of a song and a new distance metric namely conditional Euclidean distance are explained in sections 2.1 and 2.2, respectively. Then in section 3, the difference among three distances measuring metrics are shown by using the most famous artist songs by the Beatles in 3.1. Furthermore, the verification of the proposed method is explained by using MIREX`08 dataset in 3.2. After that, the proposed method is applied into a case regarding a controversial plagiarism issue in 3.3. Finally, conclusion and future works are provided in section 4.

2 Proposed Method

In this paper, the procedure of our proposed method is shown in Fig. 1.

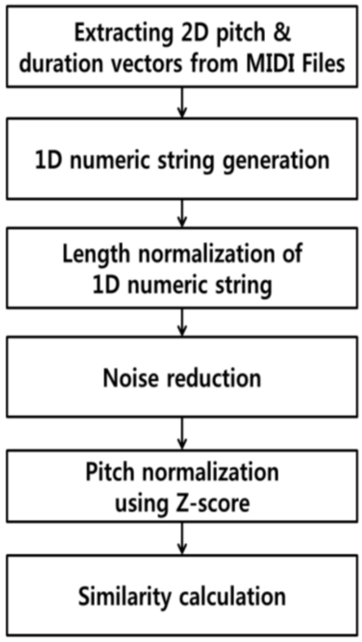


Fig. 1 A flowchart of proposed method.

First of all, 2D numeric vectors for pitch and duration of notes are extracted from music scores. Then, a 1D numeric string is generated by combining the pitch value with the duration value from the previously extracted successive vectors. After that, length normalization of the 1D numeric strings is performed by using the nearest neighbor interpolation. Next, singular points of the numeric signals are eliminated by sequentially using median and average filters. Then, Z-score normalization for amplitude of the 1D numeric strings is performed in order to independently compare the average pitch of the songs. Finally, the normalized numeric string signals are compared by measuring similarity on the basis of several distance measures such as the Euclidean distance, the Hamming distance and the conditional Euclidean distance.

2.1 1D numeric string generation

Firstly, 2D numeric vectors for pitch and duration are extracted from music scores. Typically, pitch and duration information are represented as the height and length of notes in the elements which make up the music. Actually, the MIDI file contains the information of note as a type of positive integer number. Therefore, accurate pitch and duration vectors are extracted from music scores based on MIDI file instead of using acoustic wave file and then arranged as a sequence of the successive 2D numeric vectors form of (pitch, duration).

$$p = 69 + 12\log_2 \frac{f}{440\,Hz} \qquad (1)$$

After extracting the above mentioned vector sequence, the 1D numeric strings are generated by combining the pitch values with the corresponding duration values. In the MIDI file, pitch value (P) can be calculated by using the specific frequency (f) of the pitch as shown in Equation (1). Consequently, the calculated specific pitch values are represented as shown in Fig. 2 in which a semitone offset has 1.

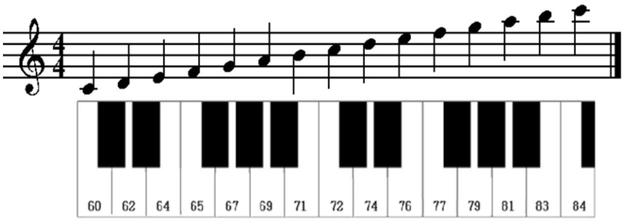


Fig. 2 The pitch values used in MIDI files.

In case of duration values extracted from music scores, they can be relatively changed according to the speed of the music. Therefore, they should be normalized as pre-defined duration values because our proposed similarity measurement method is not dependent upon the speed of music. For that, we defined the length of the sixty fourth note as 1 under the assumption that the sixty fourth note is the shortest length in music. In the same way, the thirty second, the sixteenth, the eighth, the quarter, and the half notes can be respectively represented as 2, 4, 8, 16, and 32 as shown in Fig. 3.

Fig. 3 The duration values assuming that the duration of the sixty fourth note is 1.

Next, 1D numeric string is generated by combining the pitch and the duration values as shown in Fig. 4, which means that the 2D numeric vectors for one note is dimension reduced as a 1D numeric string without a loss of any information. Fig. 4 shows a part of "Yesterday" and its 1D numeric string generation result. For example, if a note is an eighth note for a pitch value of 67, it is converted as "67, 67, 67, 67, 67, 67, 67, 67". Such conversion into 1D type makes the signal processing and similarity measurement easy compared with the 2D case.

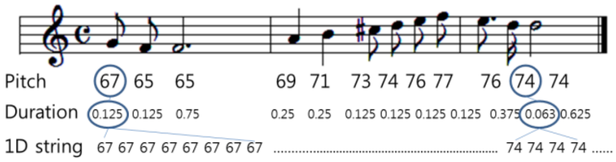


Fig. 4 1D numeric string generation method by combining pitch with duration.

After obtaining 1D numeric string from each different song, the length normalization should be performed in order to compare two strings entirely by using the nearest neighborhood interpolation. By considering the longest length of the used 1D numeric string, the normalized target length is defined as 5000. Here, empty positions after length normalization should be interpolated. To maintain original pitch values, the empty positions are interpolated by using the nearest neighborhood interpolation method. Consequently, two songs can be matched without any difficulties caused by length variation. In this paper, some eccentric notes are defined as noise from 1D numeric string. Since our method is proposed for overall matching, the eccentric notes should be removed. Numerous eccentric notes are happened which occurred up and down frequently. Even though the eccentric notes are features of original song, they make difficult and complex to overall similarity matching between two songs. Therefore, a median filter is applied into the length normalized 1D string [16]. In here, the window size of median filtering was 101.

At result of applying median filter, the shape of 1D string has discrete characteristic which is not suitable characteristic for overall matching. To acquire the smoothed string, average filtering is applied into the median filtered string in order to eliminate the discrete characteristic.

As previously performed length normalization, pitch normalization is progressed by using the Z-score normalization method [17]. The Z-score ($Z(x)$) is calculated by subtracting the average (μ) from value (x) and dividing the standard deviation (∂) as shown in Equation (2).

$$Z(x) = \frac{x - \mu}{\partial} \qquad (2)$$

The reasons for using the Z-score normalization method are solving key variation problems and comparing outlines without detail matching between notes. For example, if the key of a song is transformed or the song is arranged from the original songs, the automatic similarity measurement system based on numeric vector cannot be successfully operated even if people feel high similarity. Also, our method regards a similar changing pattern as a similar song. For example, a song, “Do-Re-Mi-Re-Do” is regarded as the same with a song of “Do-Mi-Sol-Mi-Do” in our proposed method because of using the Z-score normalization method. Therefore, we obtained the result as shown in Fig. 5.

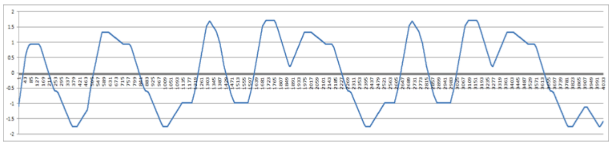


Fig. 5 The result by performing the Z-score normalization for the string.

2.2 Introducing the three distance metrics using two different songs

In this step, three distance metrics are performed such as the Euclidean distance, the Hamming distance and the conditional Euclidean distance.
First of all, the Euclidean distance between two 1D strings is defined as the sum of differences between two pitch values at same position as conceptually shown in Fig. 6 and equation (3) [18][19].

$$ED = \frac{1}{n}\sum_{i=0}^{n}\sqrt{(S1_i - S2_i)^2} \qquad (3)$$

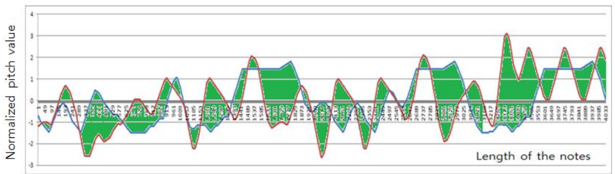


Fig. 6 The conceptual diagram of the Euclidean distance (Red line: “Hey jude”, Blue line: “All you need is love”, Green region: differences between two songs.

Next, the Hamming distance is calculated via dividing the number of different signed values by the number of total samples as equation (4) which is conventionally used for comparing two bit strings.

$$HD = \frac{\left\| \mathrm{sgn}(S1) \otimes \mathrm{sgn}(S2) \right\|}{n} \qquad (4)$$

In our method, since two 1D strings were already normalized by subtracting the mean value, the frequencies of two signs such as + and - in the 1D string were similar. Therefore, we recognized that the sign based on the Hamming distance was reasonably adopted for comparing two songs [20]. Fig. 7 shows the conceptual diagram of the Hamming distance in case of comparing two 1D strings.

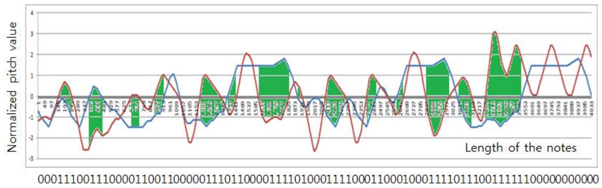


Fig. 7 The conceptual diagram of the Hamming distance (Red line: “Hey jude”, Blue line: “All you need is love”, Yellow region: opposite signed region of two songs, Below bit string: the result of bi exclusive OR).

Finally, we proposed a new distance measure, namely the conditional Euclidean distance (CED) which intends a concept of combining the Euclidean distance and the Hamming distance. That is, the CED can be obtained by calculating the Euclidean distance only in sections of different signed regions as shown in equation (5) and Fig. 8.

$$CED = \frac{\sum_{i=0}^{n}\left| S1_i - S2_i \right|}{\left\| \mathrm{sgn}(S1) \otimes \mathrm{sgn}(S2) \right\|} \qquad (5)$$

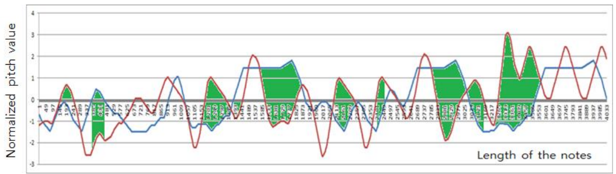


Fig. 8 The conceptual diagram of CED (Red line: “Hey jude”, Blue line: “All you need is love”, Green region: opposite signed region of two songs).

3 Experimental Results

3.1 Results by comparing songs of Beatles

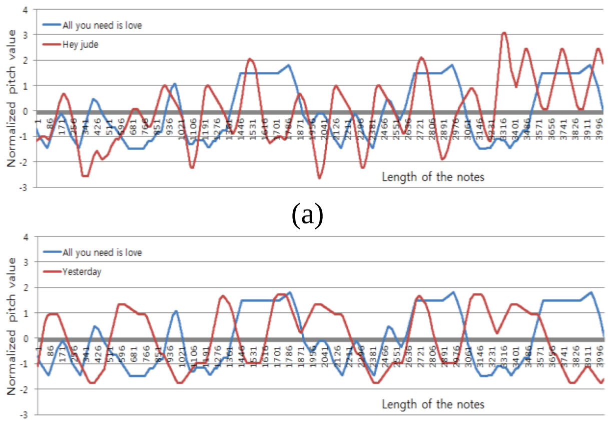
In this section, we compared four Beatles songs [21] by using the above explained distance measuring method such as the Euclidean distance, the Hamming distance and the conditional Euclidean distance.
Firstly, we measured the Euclidean distances of four Beatles songs as shown in the first row (red tone) of each cell on Table 1. Surely, the lower distance value means high similarity between two songs. The result of measuring the Euclidean distance for the comparison between “All you need is love” and

"Hey Jude" showed the minimum distance (1.212). In contrast, the comparison between "All you need is love" and "Yesterday" showed the maximum distance (1.4826). That is, the minimum distance means that the overall layouts and flows of two songs are the most similar, and vice versa. In addition, the Euclidean distance measure has the discretion ratio of 1.223 which is calculated by dividing the maximum distance value by the minimum one (1.4826 / 1.212).

Table 1. A results of respectively measuring the Euclidean distance (1st row (red tone) of each cell), the Hamming distance (2nd row (blue tone) of each cell), and the conditional Euclidean distance (3rd row (green tone) of each cell) between two songs among four Beatles' songs.

Song #2 / Song #1	All you need is love	Hey Jude	Let it be	Yesterday
All you need is love		**1.212**	1.4769	**1.4826**
		0.3954	0.5684	**0.6304**
		0.7734	1.1132	**1.3031**
Hey Jude			1.3303	1.2821
			0.4264	0.4963
			0.8022	0.9761
Let it be				1.2568
				0.4159
				0.7775

Secondly, the Hamming distances were measured as shown in the second line (blue tone) of each cell on Table 1 which was performed by the same scheme of the above calculated Euclidean distances. Similar with the meaning of Euclidean distance, the lower distance value means high similarity between two music strings. Especially, the range of the Hamming distance is restricted from 0 to 1 because the metric is calculated through dividing the number of different bits by the total number of bits. From the result of measuring the Hamming distance, the comparison between "All you need is love" and "Hey Jude" showed the minimum distance (0.3954), on the other hand the comparison between "All you need is love" and "Yesterday" showed the maximum distance (0.6304). Since the comparison cases of having maximum and minimum distances are the same with the case of using the Euclidean distance, the Hamming distance can be regarded as a reasonable distance measurement metric for comparing two music strings. In case of using the Hamming distance, the discretion ratio was 1.594 (= 0.6304 / 0.3954) which is greater than the Euclidean distance's one (1.223). Consequently, we recognized that the Hamming distance is better than the Euclidean distance in terms of comparing two music strings.



(a)

(b)

Fig. 9 The comparison of two music strings in cases of having the maximum and the minimum distances, respectively. (a) "All you need is love" vs "Hey Jude", (b) "All you need is love" vs "Yesterday".

Finally, the conditional Euclidean distances were measured as shown in the third line (green tone) of each cell on Table 1 which was also performed by a similar scheme of the above calculated distances. Likewise, the lower distance value means high similarity between two songs. Especially, since the meaning of the conditional Euclidean distance is the score level fusion of the Euclidean distance and the Hamming distance, we expect that the conditional Euclidean distance may show better performance than the Euclidean distance or the Hamming distance. From the result of measuring the conditional Euclidean distance, the comparison between "All you need is love" and "Hey Jude" showed the minimum distance (0.7734), on the other hand the comparison between "All you need is love" and "Yesterday" showed the maximum distance (1.3031). This result also coincided with the cases of using the Euclidean distance and the Hamming distance. By the way, in case of using the conditional Euclidean distance, the discretion ratio was 1.685 (= 1.3031 / 0.7734) which is the best ratio compared with the ones of using the Euclidean distance and the Hamming distance (1.223 and 1.594, respectively). Consequently, we confirmed that our proposed distance metric such as the conditional Euclidean distance showed the best discretion performance. Fig. 9 shows the comparison of two music strings in cases of having the maximum and the minimum distances such as "All you need is love" vs "Hey Jude" and "All you need is love" vs "Yesterday" from the Table 1.

3.2 Result by using MIREX`08 dataset

For the next experiment, MIREX’08 (Music Information Retrieval Evaluation eXchange 2008) dataset is adopted for detecting the most similar combination among many songs by using our proposed similarity measurement method [22].
MIREX`08 dataset consisted of 48 MIDI files. Firstly, extracting 1D music strings of the MIDI files is performed. Then, we calculated the conditional Euclidean distances of 1,128 combinations (= ${}_{48}C_2$) among 48 MIDI files.

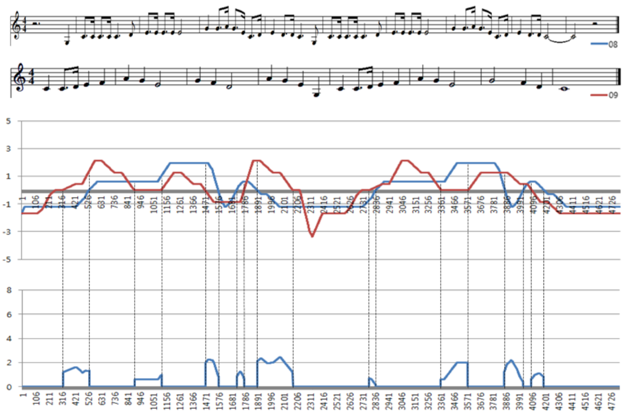


Fig. 10 Two 1D music scores, strings, and the diagram of conditional Euclidean distance in case of having the closest similarity (MIDI #08 vs MIDI #09) (0.38248).

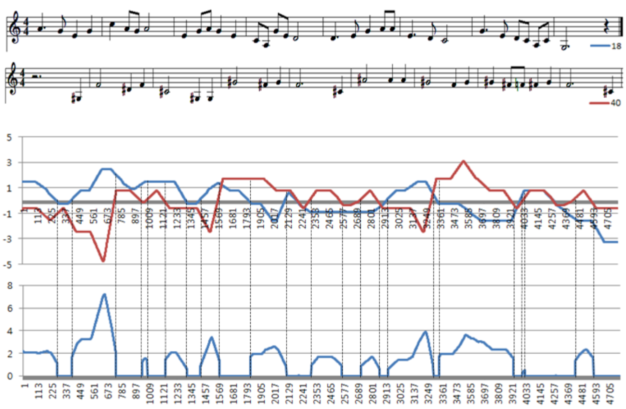


Fig. 11 Two 1D music scores, strings, and the diagram of conditional Euclidean distance in case of having the farthest similarity (MIDI #18 vs MIDI #40) (1.9025).

The result of analyzing 1,128 combinations, the distance between MIDI #08 and MIDI #09 data was the closest having a value of 0.3825 by comparing their 1D music strings, and the distance between MIDI #18 and MIDI #40 data is the farthest having the value of 1.9025. In detail, the flow of the 1D music strings are represented with score and the diagram for conditional Euclidean distance as shown in Fig. 10 and Fig. 11.

At a result of comparative visualization for two cases such as the closest and the farthest ones, we found that distances were appeared from 0 to 2.5 in the closest case as shown in Fig. 10. In contrast, we confirmed that distances of the farthest case were appeared from 0 to 8.0 as shown in Fig. 11. In terms of music score, two songs in Fig. 10 show very similar flow according to the change of notes. However, we recognized that two songs in Fig. 11 might intend for an almost opposite flow in terms of pitch variation.

3.3 Applying for proposed method into controversial plagiarism

Next, the proposed method was applied into a controversial plagiarism issue. The case is between “Don’t look at my eyes” by Sam Harris and “Don’t make me cry” by Shin Seung-Hun. Since the release of Shin’s song in 1990, they had received complaints regarding plagiarism. After all, Shin’s song turned out to be plagiarized. To validate the issue, two songs were compared by using our proposed method. Therefore, the plagiarism issued regions of 2 songs were formed by 1D music strings for similarity matching between “Don`t look at my eyes” and “Don`t make me cry” respectively as shown in Fig. 12.

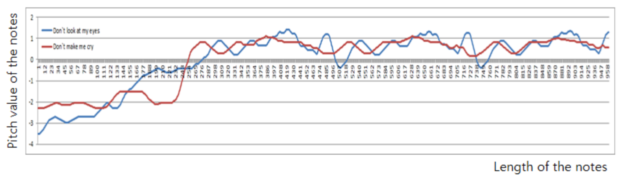


Fig. 12 The replaced sequence of 1D music strings between “Don’t look at my eyes” and “Don’t make me cry”.

Table 2. A results of three distances estimation among 2 songs between “Don’t look at my eyes” and “Don’t make me cry”.

	“Don`t look at my eyes” vs “Don`t make me cry”
Euclidean Distance	0.425
Hamming Distance	0.066
Conditional Euclidean distance	0.049

From the results of calculating distances as shown in Table 2, there were represented as the values of 0.425, 0.066, and 0.049 in terms of Euclidean, Hamming and conditional Euclidean distance, respectively. Consequently, we recognized that the conditional Euclidean distance was reasonably applied into the comparison for the issue of plagiarism.

4 Conclusion

In this paper, similarity measurement method for estimating music similarity was proposed by comparing 1D music strings based on the conditional Euclidean distance. Because our proposed method performed overall matching between two songs, the purpose was clearly different from the previous partial matching method for query by humming. The discretion performance of the new metric used such as the conditional Euclidean distance was validated through several kinds of experiments. From the result, we found that our proposed method could be reasonably applied when two songs had similar structures. Also, we confirmed that the method was adopted for determining the issue of plagiarism.

In future work, we will develop a structure analyzing method in order to overcome the structural variation problem of our method. Also, the calculated distances will be compared with human qualitative feeling for similarity.

Acknowledgement

This research was supported by a 2012 Research Grant from Sangmyung University.

References:

[1] R. David, J. P. Pedro, P.-S. Carlos, P. Antonio, and M. I. Jose, “A Pattern Recognition Approach for Melody Track Selection in MIDI Files,” *Proceedings of the 7th International Conference on Music Information Retrieval*, 2006, pp. 61-66.

[2] G. Ozcan, C. Iskhan, and A. Alpkocak, “Melody Extraction on MIDI Music Files,” *Proceedings of the Seventh IEEE International Symposium on Multimedia*, 2005, pp. 414-422.

[3] H. H. Shih, S. S. Narayanan, and C.-C. J Kuo, "Automatic main Melody Extraction from MIDI Files with a modified Lempel-Ziv Algorithm," *Proceedings of the 2001 International Symposium on Intelligent Multimedia, Video and Speech Processing*, 2001, pp. 9-12.

[4] L. Li, C. Junwei, W. Lei, and M. Yan, “Melody extraction from Polyphonic MIDI Files Based on Melody Similarity,” *Proceedings of the International Symposium on Information Science and Engineering*, vol. 2, 2008, pp. 232-235.

[5] R. Shimokura, L. Tronchin, A. Cocchi, and Y. Soeta, “Subjective Diffuseness of Music Signals Convolved with Binaural Impulse responses,” *Journal of Sound and Vibration*, Vol. 330, Issue 14, 2011, pp. 3526-3537.

[6] T. Hidaka, L. L. Beranek, and T. Okano, “Interaural Cross-Correlation, Lateral Fraction, and Low- and High-Frequency Sound Levels as Measures of Acoustical Quality in Concert Halls,” *Journal of the Acoustical Society of America*, Vol. 98, No. 2, 1995, pp. 988-1007.

[7] L. Tronchin, “The Emulation of Nonlinear Time-Invariant Audio Systems with Memory by Means of Volterra Series,” *Journal of Audio Engineering Society*, Vol. 60, Issue 12, 2012, pp. 984-996.

[8] A. Kupryjanow, and A. Czyzewski, “Real-Time Speech Signal Segmentation Methods,” *Journal of Audio Engineering Society*, Vol. 61, Issue 7/8, 2013, pp. 521-534.

[9] W. Jeon, and C. Ma, “Efficient search of music pitch contours using wavelet transforms and segmented dynamic time warping,” *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2011, pp. 2304-2307.

[10] G. P. Nam, K. R. Park, S.-P. Lee, E. C. Lee, M.-Y. Kim, and K. Kim, “Intelligent Query by Humming System,” *Proceedings of the 4th International Conference on Ubiquitous Information Technologies & Applications*, 2009, pp. 1-6.

[11] L. Wang, S. Hang, S. Hu, J. Liang, and B. Xu, “An effective and efficient method for query by humming system based on multi-similarity measurement fusion,” *Proceedings of the International Conference on Audio, Language and Image Processing*, 2008, pp. 471-475.

[12] L. Gao, and Y. Wu, “A system for melody extraction from various humming inputs,” *Proceedings of the IEEE International Symposium on Signal Processing and Information Technology*, 2006, pp. 680-684.

[13] M. Chan, and J. Potter, “Recognition of Musically Similar Polyphonic Music,” *Proceedings of the 18th International Conference on Pattern Recognition*, vol. 4, 2006, pp. 809-812.

[14] Y. Zhu, M. Kankanhalli, and Q. Tian, “Similarity matching of continuous melody contours for humming querying of melody databases,” *Proceedings of the IEEE workshop on Multimedia Signal Processing*, 2002, pp. 249-252.

[15] A. Ruppin, and H. Yeshurun, "MIDI Music Genre Classification by Invariant Features," *Proceedings of the International Conference of Music Information Retrieval*, 2006, pp. 397-399.
[16] C.-T. Lu, K.-F. Tseng, and C.-T. Chen, "Reduction of residual noise using directional median filter," *Proceedings of the IEEE International Conference on Computer Science and Automation Engineering*, vol. 3, 2011, pp. 475-479.
[17] F. N. Julia, and K. M. Iftekharuddin, "Detection of Emotional Expressions in Speech," *Proceedings of the IEEE SoutheastCon*, 2006, pp. 307-312.
[18] A. Rebelo, J. Tkaczuk, R. Sousa, J. S. Cardoso, "Metric Learning for Music Symbol Recognition," *Proceedings of the 10th International Conference on Machine Learning and Applications*, vol. 2, 2011, pp. 106-111.
[19] U. Simsekli, "Automatic Music Genre Classification Using Bass Lines," *Proceedings of the 20th International Conference on Pattern Recognition*, 2010, pp. 4137-4140.
[20] J.-H. Su, C.-W. Wu, S.-Y. Fu, Y.-F. Lin, W.-Y. C, I.-B. Liao, K.-W. Chang, and V.S. Tseng, "Empirical analysis of content-based music retrieval for music identification," *Proceedings of the International Conference on Multimedia Technology*, 2011, pp. 3516-3519.
[21] "Free Site on MIDI Files" http://rock.freemidis.net/index.php
[22] "MIREX 2008" http://www.music-ir.org/mirex/wiki/2008:Main_Page