International Journal of Fuzzy Logic and Intelligent Systems, vol. 8, no. 3, September 2008 pp. 192−195

Feature Transformation based Music Retrieval System

Jungim Heo*, Jinmo Yang*, Dong-hyun Kim*, Kyoungro Yoon and Wonil Kim***

*** College of Electronics and Information, Sejong University, Seoul, Korea**
**** College of Information and Communication, Konkuk University, Seoul, Korea**

Abstract

People have tendency of forgetting music title, though they easily remember particular part of music. If a music search system can find the title through a part of melody, this will provide very convenient interface to users. In this paper, we propose an algorithm that enables this type of search using feature transformation function. The original music is transformed to new feature information with sequential melodies. When a melody that is a part of search music is given to the system, the music retrieval system searches the music similar to the feature information of the melody. Moreover, this transformation function can be easily extended to various music recognition systems.

Key words : Feature Transformation, Based Music

1. Introduction

In this paper we propose Music Retrieval System using feature transformation. The query of the proposed system is given by a short humming melody. The humming melody is transformed into a unique sequence of strings and compared with data in music library. The transformed sequence of strings represents the difference between two consecutive note pairs including the note height and duration differences. The simulation results show that the system is able to find the intended musical note with relatively short sequence for specific note. Moreover, it finds several candidate notes for randomly generated note. We will briefly discuss the previous researches in section 2. The detailed proposed system will be explained in section 3, after which the simulation result follows in section 4. The conclusion is given in section 5.

2. Previous Research

M. Anand Raju and Preeti Rao [1] proposed signal processing for melody retrieval system. Pitch tracking and note segmentation are typically carried out based on the estimation of the instantaneous fundamental frequency and amplitude of the acoustic signal. Since many instruments produce only discrete pitches, they especially focus on pitch contour and rhythm.

Pitch contours and rhythm are extracted automatically for all the queries by amplitude and magnitude waveform separately. They show that contours of humming and whistling are identical as a prototype. At the same time they discuss enhancing the scale to solve problem.

Scoring algorithm [2] computes the score of a song and query to evaluate how well they match. A higher score indicates a better match. The results of the algorithm are importance of rhythmic information and finding optimal numbers of level contours. It is supported by histogram of interval changes.

A method for indexing polyphonic music data files using the pitch and rhythm dimensions of music information is introduced by Shyamala Doraisamy and Stefan M Rüger [3]. Their study uses n-grams that are consisted of interval ratio.

3. The Proposed System

When people are humming any part of a song, they tend to hum that part with difference. In most cases, the note sequence generated by each user's humming is little lower or higher than the original note sequence of the music. For example, Figure 1 shows the score of original song and Figure 2 shows the score of the same song hummed by a user. This difference of two note sequences prevents direct comparison between the query and the stored music. In order to solve this problem, we propose music matching method in which the difference of note including height and duration, not the note value itself of the given music, is used.

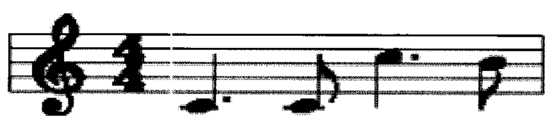


Fig. 1. Score from the real music

Manuscript received Apr. 17, 2008; revised Jun. 4, 2008.
Author for correspondence : Wonil Kim
This Paper is supported by Seoul R&BD Program

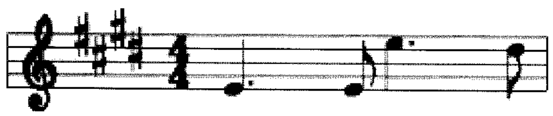


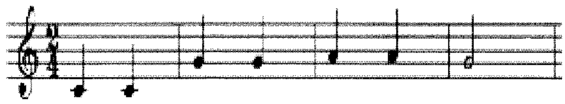
Fig. 2. Score from the user's humming

Music in the database is transformed into sequences of two dimensional feature vectors from the note sequence in two steps. In the first step, the note sequence is represented by a sequence of two-dimensional vectors of integer values. One dimension is allocated to represent the note height of thirty six levels of three octaves. The other dimension is allocated to represent the relative note duration of twelve levels from a thirty second note to full note with variants of point. Hummed query (i.e. humming or whistling) is also transformed to another code vector sequence using the same transformation rule as music in the database. When transforming the hummed query into the code sequences, we give a preset length to the full note, not considering the rhythm, as the relative note duration is used in the second step.

The note height and duration of the actual music and user generated query can be quite different and the possibility of the closely matching the code sequences of actual music and the hummed query generated in the first step is very low. However, the people still tend to follow the melody line or the pattern of the note sequence of the actual music very closely. Using such characteristics, we can create a music retrieval system which can perform quite accurate music retrieval. To employ such characteristics in the system, we perform second step of feature transformation. In the second step, the difference of between two consecutive code vectors is calculated, which is used in calculating the similarity between the music in the database and the hummed query. The similarity between the music in the database and the hummed query is calculated using the weighted distance of a (height difference) + (1-a) (duration difference), where 0<= a <=1.

4. Simulation

By employing the second step of transformation, the overall system performance is improved by neglecting the variations such as pitch or rhythm shift coming from the user's unique singing pattern. Figure 3 is a score, original code sequence, and the vector sequence after the second transformation of a music in the database, whereas Figure 4 is a score, original code sequence, and the vector sequence after the second transformation of the same song hummed by the a certain user.



Original Code Sequence	Note Height	13	13	20	20	22	22	20
	Note Duration	7	7	7	7	7	7	9
Transformed Code Sequence	Height Difference		0	7	0	0	0	-2
	Duration Difference		0	0	0	0	0	2

Fig. 3. Score of music and its corresponding code sequence

Original Code Sequence	Note Height	8	8	15	15	17	17	15
	Note Duration	6	6	6	6	6	6	8
Transformed Code Sequence	Height Difference		0	7	0	2	2	-2
	Duration Difference		0	0	0	0	0	2

Fig. 4. Score of user's voice and its corresponding code sequence

4.1 Query using the songs in the music database

The system successfully retrieves all the songs when original music of the humming music is indeed in the music data base. Table 1 shows 5 songs in top ranks. Other songs included in the top five have the difference measure of about double in average. In average, the difference measure of matched songs 3.383263. The longer user's humming is, the smaller the difference measure of the matched song is.

Table 1. Result of query for a song "Bicycle"

Title	Difference Measure
Bicycle	3.80278
Fall street	8.85591
Children's song	8.86839
You can do it very well	9.26434
Grab	9.27987

4.2 Query using the songs not in the music database

This simulation is for the retrieval result of humming queries, when the hummed songs are not in the music database (e.g. testing music plagiarism). We generated random humming and compared the humming with the music in the database. When the generated random humming do sound familiar to the participants of the experiments, the average difference measure of the best matched song is 4.83013. On the other hand, when the generated random humming does not sound familiar to the participants of the experiments, the average difference measure of the best matched song is more than 10.0 as shown in Table 2.

Table 2. Results of query for a song "Sting" that is not in the database

Title	Difference Measure
Cinderella	11.7585
Hello	11.9373
Pine	12.1456
Children's day	12.2878
My grandmother	12.3233

According to the experimental results, if difference measure is lower than 5.0, we can assume that the two songs are indeed identical. If difference measure is larger than 10.0, we can safely assume that the songs are different.

5. Conclusion

In this paper, we proposed a music retrieval system using feature transformation. According to the result of simulation, we found that the retrieval performance is enhanced when more weight is given to the note height than the note duration. Moreover, this system can be easily extended to the practical case of testing for music plagiarism. According to the simulation, two songs with difference measure under 5.0 are an identical music, whereas over 10.0 are different songs. Any values between them cannot be clearly concluded and need detailed investigation.

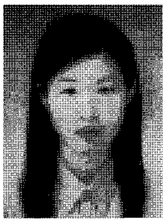
References

[1] M. A. Raju, P. Rao, "Building a melody retrieval system," *Proc. National Conference on Communications (2002)*

[2] Y. E. Kim, W. Chai, R. Garcia, B. Vercoe, "Analysis of a contour-based representation for melody," *Proc. International Symposium on Music Information Retrieval (2006)*

[3] S. Doreisamy, S. M Rüger, "An approach towards a polyphonic music retrieval system," *Proc. International Conference on Music Information Retrieval (2001)*

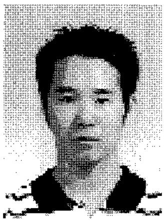
[4] Y. Choi, K. Ko, S. Lim, "Interactive Multimedia Class," Vol. 492, No. 5 (2003) 173-215



Jung-im Heo

She received B.S. in the Digital Contents at Sejong university, Korea. Her main research interests include multimedia and ubiquitous computing. Since 2007, she works for Hewlett-Packard

Phone : +82-2-2640-5917
Fax : +82-2-2640-5953
E-mail : fuke22@yahoo.co.kr



Jinmo Yang

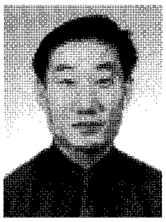
He is studying the B.S. in Computer Engineering from Sejong University, now. His main research interests include ubiquitous

Phone : +82-10-2087-4046
Fax : +82-2-3408-4339
E-mail : djmyang@sju.ac.kr

Dong-hyun Kim

He received the B.E in Architectural Engineering from Yonsei University, Seoul, Korea in 1983. He received the M.E. and Ph.D. in Environmental Engineering from Osaka University, Japan in 1988, 1991.

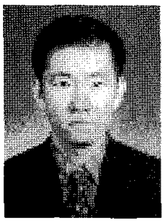
Phone : +82-2-3408-3796
Fax : +82-2-3408-4339
E-mail : mustache@sejong.ac.kr



Kyoungro Yoon

He received the B.S. in Electronic and Computer Engineering from Yonsei University, Korea in 1987, M.S.E in Electrical Engineering Systems from University of Michigan, Ann Arbor, U.S.A., in 1989, and Ph.D. in Computer and Information Science from Syracuse University, U.S.A. in 1999.

Phone : +82-2-450-4129
Fax : +82-2-454-3768
E-mail : yoonk@konkuk.ac.kr



Wonil Kim

He received the B.E in Metal Engineering from Hanyang University, Seoul, Korea in 1982. He received the B.S., M.S in Computer Science from Southern Illinois University, U.S.A. in 1988, 1990 respectively. He received Ph.D. in Computer and Information Science from Syracuse University, U.S.A. in 2000.

Phone : +82-2-3408-3795
Fax : +82-2-3408-4339
E-mail : wikim@sejong.ac.kr