A SIMPLE ALGORITHM FOR RHYTHM SIMILARITY COMPARISON

Shu-Jen Show Hsiao
Department of Computer science
National Chiao Tung University
Hsinchu, Taiwan ROC
Show.cs95g@nctu.edu.tw

Tyne Liang
Department of Computer science
National Chiao Tung University
Hsinchu, Taiwan ROC
tliang@cis.nctu.edu.tw

Hao-Ren Ke
Institute of Information Management
National Chiao Tung University
Hsinchu, Taiwan ROC
claven@lib.nctu.edu.tw

ABSTRACT

Rhythm is one of the music features for man to perceive intuitively, even one without formal music training. So, it's a good idea for Content-Based Music Retrieval System (CBMRS) to employ rhythm information in retrieval process. This extended abstract details our two submissions for MIREX 2008 Query by Tapping task (QBT). These submissions present a rhythm similarity function to compare MIDI extract with user tapping.

Keywords: CBMRS, MIREX, QBT.

1 SYSTEM OVERVIEW

Data processing perform in two phase. Feature extraction and rhythm intermediate form transformation from ground-truth MIDI in Indexing phase. Every query file will perform a matching task run to compare the similarity between all ground-truth intermediate forms in matching phase. Then report top 10 candidates in answer file after each run complete.

Submission 1 uses recorded tapping rhythm in WAV audio format as query files. We use "aubioonset"[1] with the threshold value for the onset peak picking=0.7. There are two execution files in this submission both are implemented by Matlab 7. See Figure 1.

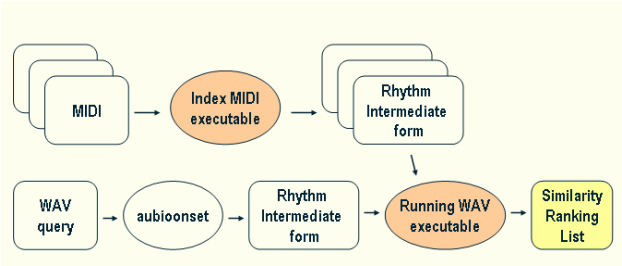

Figure 1 Use WAV query files in submission 1

Submission 2 uses symbolic onset format as rhythm query files in stead of WAV query file in submission 1. See Figure 2.

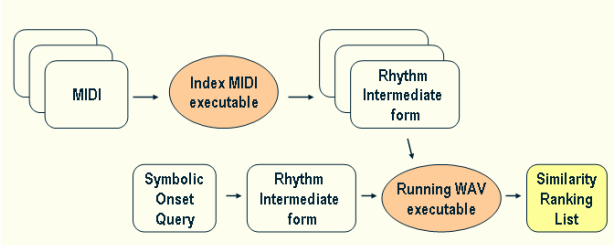

Figure 2 Use symbolic onset query files in submission 2

The rhythm intermediate form will depict the time interval (in ms) between two consecutive Note On events in MIDI. For example the rhythm intermediate form for the first 12 Note On events in "Happy birthday" would be: 450, 150, 600, 600, 600, 1200, 450, 150, 600, 600, 600.

2 SIMILARITY FUNCTION

Most MIR system compares rhythm similarity by distance measure [2]. Such as warping distances [3], edit distance [4], earth mover distance [5] to measure the similarity. This extended abstract propose a new way to measure the rhythm similarity by "Rhythm Quotient Standard Deviation". RQSD compares rhythm directly by numerical form instead of symbolic form. For example, we have a query sequence Q and index document sequence D, RQSD processes by the following step:

A. Cut Q and D into grams by 4-gram method, and every gram with 4 rhythm elements. Then calculate each corresponding gram similarity between Q and D.
B. Suppose a 4-gramm in Q is called QG and the corresponding 4-gramm in D is called DG. Divide each rhythm element in QG and DG to obtain four rhythm quotient values

$$q1=QG(1)/DG(1)$$
$$q2=QG(2)/DG(2)$$
$$q3=QG(3)/DG(3)$$
$$q4=QG(4)/DG(4) \quad (1)$$

C. Calculate the standard deviation of these 4 rhythm quotient values. The smaller the standard deviation is, the more similar QG and DG are.

$$\sigma = StdDev(q1, q2, q3, q4) \quad (2)$$

D. The standard deviation value allows changing in wider range for bigger rhythm quotient values. In other words, for the same similarity degree, bigger q1, q2, q3, q4 will get a bigger standard deviation value. So it is necessary to create a factor δ to reduce this kind of effect

$$\delta = (q1+q2+q3+q4) / 4 \quad (3)$$

E. Gram similarity can be calculated:

$$Gram\ Similarity = (1 - (\sigma / \delta))*100\% \quad (4)$$

F. Finally calculate the global similarity of Q and D. Sums the gram similarity of all grams and divides the summation by the number of gram pairs compared.

$$Similarity = \frac{\sum GramSimilarity}{compared\_gram\_count} \quad (5)$$

For example if a query sequence Q: 313, 343, 641, 656, 688, 1234**.** And a index document sequence D: 450, 150, 600, 600, 600, 1200. The similarity between Q and D is 0.66. See Figure 3.

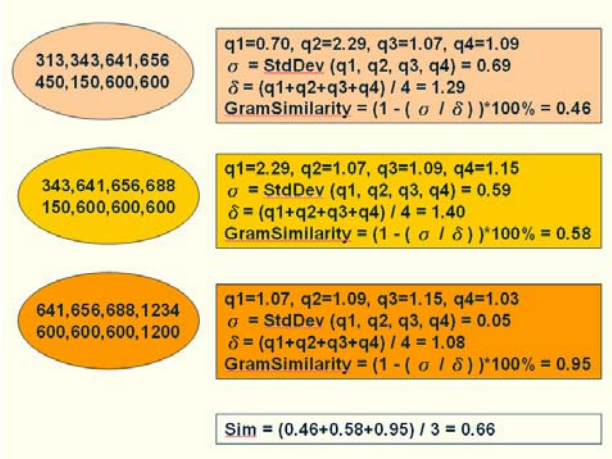

Figure 3 Examples for RQSD calculation

3 EVALUATION RESULT

There are 103 MIDI files in QBT test collection. 481 query files both in WAV query (for submission 1) and onset query (for submission 2). Submission 1 & submission 2 should have 481 task runs to report the similarity ranking list for each query files. Figure 4 and Figure 5 are the running result references from MIREX 2008 QBT result page

	Top 10 hit rate by query	Top 10 hit rate by query group	MRR by query	MRR by query group
Sub1	0.31	0.25	0.20	0.18
Sub2	0.69	0.66	0.57	0.51

Figure 4 Top 10 hit rate and MRR for submission 1 & 2

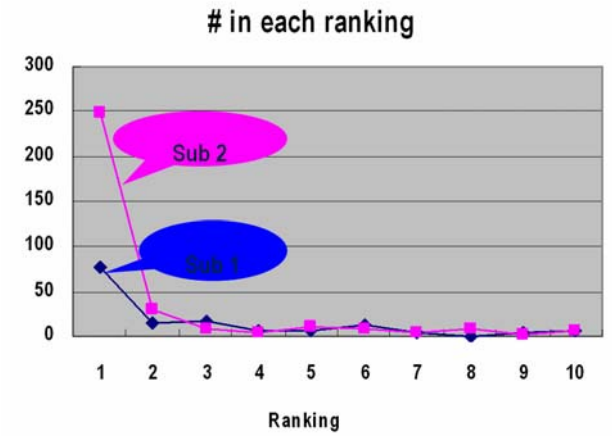

Figure 5 Number of counts in each ranking for 481 task run.

4 DISCUSSION

A. Both submission1 and submission2 is used same comparison algorithm. But the result shows awful difference because of onset detection method. So the "aubioonset" seems not work perfectly in this application (Submission 1). And thanks again for HR Lee to contribute these quality onset files (Submission 2).

B. Onset detection plays important role in QBT task. If we can provide query files which recorded directly from tapping keyboard instead of microphone, the better result could be expected.

C. Number of count in ranking 2 is dramatically decreased from ranking 1. And counts in ranking 2 to ranking 10 relatively seldom. That's why we use the top 10 hit rate measure can't have obvious improvement from MRR.

REFERENCES

[1] http://aubio.org/

[2] Rainer Typke, Frans Wiering, Remco C. Veltkamp: "A survey of Music Information Retrieval System", Proceeding of the Sixth International Conference on Music information Retrieval (ISMIR), London, 2005.

[3] JS Roger Jang, Hong-Ru Lee, Chia-Hui Yeh: "Query by Tapping: A New Paradigm for Content-based Music Retrieval from Acoustic Input", The Second IEEE Pacific-Rim Conference on Multimedia, Beijing, China, 2001.

[4] S. Pauws. CubyHum: "A fully operational query by humming system", In ISMIR Proceedings, pages 187–196, 2002.

[5] Rainer Typke, A.C. Walczak-Typke: "A tunnelling-vantage indexing method for non-metrics, In ISMIR Proceedings, 2008