SIMPLE ORTHOGONAL PITCH WITH IOI SYMBOLIC MUSIC MATCHING

Iman S. H. Suyoto
Broadmeadows, Vic, Australia
ishs@ishs.net

Alexandra L. Uitdenbogerd
RMIT University
Melbourne, Vic, Australia
SandraU@rmit.edu.au

ABSTRACT

In this submission we resurrect the Fanimae implementation of *n*-gram symbolic matching, as well as a second technique that combines alignment of pitch strings with inter-onset-interval string alignment — a technique that was shown to be marginally better than pitch matching alone for searching collections of polyphonic MIDI files.

1. INTRODUCTION

Our submission this year consists of a couple of matching techniques, the first of which is the *n*-gram matching baseline [1] that we provide for continuity to allow MIREX submissions to be compared across the years.

The second is a technique that builds on our dynamic programming-based approach by adding rhythm information to the matching process. We have found that it does improve effectiveness slightly but significantly in the context of polyphonic symbolic matching, but the weighting for the rhythm component of the similarity measurement must be kept much smaller than that of pitch [2].

The two techniques are implemented in the Fanimae MIREX 2010 Edition [1] (*FM10*) software package.

2. METHOD

Our method is based on the *Three Phase Music Matching Method* [4, 6], consisting of:

1. *Melody Extraction*, in which the highest pitch notes occurring at each instant are selected from the original polyphonic tracks or channels.

2. *Melody Standardisation*. The techniques submitted for this year's MIREX all make use of the basic *directed modulo* pitch standardisation in which the interval between adjacent notes is represented by a character, and where intervals larger than an octave are reduced to the harmonically equivalent interval that is less than an octave. In addition we use a rhythm representation consisting of rhythm contour [2].

3. *Similarity Measurement* We have used two different similarity measurement techniques: alignment using dynamic programming, and *n*-gram-based matching [5].

[1] See http://fanimae.sourceforge.net.

This document is licensed under the Creative Commons Attribution-Noncommercial-Share Alike 3.0 License.
http://creativecommons.org/licenses/by-nc-sa/3.0/
© 2010 The Authors.

2.1 IOI Standardisation

The IOI standardisation implemented in *FM10* is specified in our earlier work [2]. Using this standardisation, a note is encoded as one of five distinct symbols. Each symbol represents the IOI of the note in relation to the IOI of its directly preceding note. The symbols are `S`, `s`, `R`, `l`, and `L` (for much shorter, a little shorter, same, a little longer, and much longer respectively). The quantisation S is given by:

$$S = \begin{cases} \texttt{S}; & r \leq -2 \\ \texttt{s}; -2 < r \leq -1 \\ \texttt{R}; -1 < r < 1 \\ \texttt{l}; \quad 1 \leq r < 2 \\ \texttt{L}; \quad 2 \leq r \end{cases} \tag{1}$$

where $r = \log_2 I_c - \log_2 I_p$, I_c is the IOI between the current note and the following note, and I_p is the IOI between the previous note and the current note. There are special cases here:

- If the melody contains at least two notes, for the last note, I_c is its duration.
- The first note is never encoded. A melody containing only one note (a corner case) is not representable.

In the case of polyphony, the highest note in an onset time is taken. For example, see our earlier work [2].

2.2 Similarity Measurements

Two matching algorithms are implemented in *FM10*:

1. *NGR5*. This is the algorithm implemented in Fanimae MIREX 2005 Edition. In brief, only the pitch feature is used. The directed modulo pitch standardisation is used for queries and answers. Matching a query and the collection uses coordinate matching on the 5-grams of the sequences. For details, see our MIREX 2005 submission [1].

2. *PIOI*. As an improvement to Fanimae, *FM10* now implements this algorithm. In this algorithm, both the pitch and IOI features are used. For a melody, two sequences are produced: pitch sequence as in *NGR5* and IOI sequence produced by the standardisation described in Section 2.1.

   The pitch similarity score between the query and answer pitch sequences is calculated by the local alignment algorithm with 1 for a match, -1 for a mismatch, and -2 for an insertion/deletion. The IOI similarity score between the query and answer pitch sequences is calculated by the local alignment algorithm with -2 for an insertion/deletion and the match and mismatch scores as the following:

   - Matches:
     - `R-R`: 3.
     - `s-s` and `l-l`: 2.
     - `S-S` and `L-L`: 1.
   - Mismatches:
     - `S-s` and `L-l`: 0.
     - `R`-any other symbol: -2.
     - Other mismatches: -3.

   To fuse the pitch and IOI similarity scores, they are modelled as two orthogonal vectors. Adding the two vectors yields the overall similarity vector. The overall similarity score is the magnitude of the resultant vector. Our previous experiment [2] shows that the highest retrieval effectiveness can be achieved if the pitch vector magnitude is multiplied by 18. Mathematically, the overall similarity score (Σ) is:

$$\Sigma = \sqrt{18^2 P^2 + I^2} \tag{2}$$

   where P is the pitch similarity score and I is the IOI similarity score.

   The ranks of the answers are more important than the actual similarity scores. Therefore, in *FM10*, Σ^2 is internally used for ranking to reduce calculation cost by avoiding calculating $\sqrt{\Sigma^2}$.

3. CONCLUSION

The methods that we have entered in this year's evaluation have been shown to be effective on collections of polyphonic symbolic data with symbolic melody queries. The *n*-gram matching technique, as implemented, would be highly efficient due to its inverted index allowing repeated queries to be rapid.

The pitch plus IOI technique that we have submitted has been shown to be more effective than pitch matching on its own. Future work may involve testing its efficacy when combined with the best technique submitted by Uitdenbogerd to the 2007 MIREX [3], which may improve results further.

4. REFERENCES

[1] I. S. H. Suyoto and A. L. Uitdenbogerd. Simple efficient n-gram indexing for effective melody retrieval. In *Music Information Retrieval Evaluation eXchange*, 2005.

[2] I. S. H. Suyoto and A. L. Uitdenbogerd. The effect of using pitch and duration for symbolic music retrieval. In R. McArthur, P. Thomas, A. Turpin, and M. Wu, editors, *Proc. Australian Document Computing Conference*, Hobart, Australia, Dec. 2008.

[3] A. L. Uitdenbogerd. N-gram pattern matching and dynamic programming for symbolic melody search. In *Music Information Retrieval Evaluation eXchange*, 2007.

[4] A. L. Uitdenbogerd and J. Zobel. Melodic matching techniques for large music databases. In D. Bulterman, K. Jeffay, and H. J. Zhang, editors, *Proc. ACM International Multimedia Conference*, pages 57–66, Orlando, Florida, USA, Nov. 1999. ACM Press.

[5] A. L. Uitdenbogerd and J. Zobel. Music ranking techniques evaluated. In M. Oudshoorn, editor, *Proc. Australasian Computer Science Conference*, Melbourne, Australia, Jan. 2002.

[6] A. L. Uitdenbogerd and J. Zobel. An architecture for effective music information retrieval. *J. American Society of Information Science and Technology (JASIST)*, 55(12):1053–1057, 2004.