Extended abstract for submissions to MIREX 2006: Three algorithms for symbolic similarity computation

Klaus Frieler
Universität Hamburg
Neue Rabenstrasse 13
20354 Hamburg, Germany
kf (at) omniversum.de

Daniel Müllensiefen
Goldsmiths College, University of London
New Cross Road, New Cross
London, SE14 6NW
d.mullensiefen (at) @gold.ac.uk

Keywords: Symbolic Similarity, Similarity of Melodies.

1. Introduction

In this abstract we describe very briefly the three algorithms that we submitted to this year's MIREX competition for symbolic similarity. All three algorithms only work for the comparison of monophonic melodies, and thus only entered the RISM task of the symbolic similarity competition.

2. Constructing hybrid algorithms from the SIMILE toolbox

In the past we have explored different methods for abstracting information within different musical dimensions from melodies and for comparing, i.e. measuring the similarity, two abstract sequences which may represent monophonic melodies. All abstraction and similarity computation methods are implemented in our software toolbox SIMILE. The individual methods are described in greater detail in [7].

2.1. Melodic dimensions and abstraction methods

The abstraction methods we implemented so far for representing data in different musical dimensions of single line melodies are:

- Pitch: MIDI quantisation, leap/step-quantisation, Parsons Code, see [7].
- Rhythm: Categorisation to five duration classes, representation as 'gaussified' values, see [7], [2].
- Contour: Different methods for smoothing coarse directional movements, Fourier transform, see e.g. [10].
- Implied tonality: Categorisation according to harmonic content as based on Krumhansl's tonality vector, e.g. [5].
- Accent structure: Combinations of Gestalt-like accent rules from psychological literature, e.g. [4].

2.2. Similarity algorithms

Data from any abstraction method can be combined with most of the following similarity algorithms that are employed for doing the actual comparision.

- Edit Distance: e.g. [6]
- n-grams: e.g. [1]
- Geometric distance: [10]; [8]
- Correlation coefficient: e.g. [10]

3. The employed algorithms

We tried three different algorithms on the competition items (RISM incipits) to learn a little bit about the behaviour of different approaches with this particular melodic material (short beginning phrases of classical melodies).

3.1. KF1

KF1 was the name for the algorithm we submitted to last year's MIREX competition. This algorithm was optimised on a set of human experts ratings of short melodic phrases mainly from pop tunes. A detailed description can be found in the extended abstract of our last year's submission [3]

3.2. KF2

KF2 stands for an algorithm which proofed to be the best hybrid combination of abstraction methods and algorithms in our study with pop music tunes [7]. There it was termed `opti3`, and it consists of three different individual algorithms:

$$\begin{aligned} KF2 &= 0.505 \cdot \texttt{nGrUkkon} + 0.417 \cdot \texttt{rhythFuzz} \\ &\quad +0.24 \cdot \texttt{harmCorE} - 0.146 \end{aligned}$$

where

- `nGrUkkon:` measure based on the Ukkonen distance of pitch intervals 3-gram.
- `rhythFuzz:` Edit distance of sequences of categorised rhythm values.
- `harmCorE:` Harmonic measure: The main tonality of the melody is calculated according to Krumhansl's algorithm and the two tonalities are compared with the edit distance.

3.3. KF3

KF3 was one of our new and still very experimental approaches to melodic similarity measurement. It is based on the calculation of melodic accents. To this end, for every note of the two melodies, we calculate binary accents weights according to the following rules:

- `phrasend:` Ending of melodic phrase.
- `phrasbeg` Beginning of melodic phrase.
- `beat13:` Note on beat 1 or 3 of bar.
- `shortpr:` Accent of second note of 2-note phrase.
- `longmod:` Duration longer than mode of all durations in phrase.
- `pextrem:` Melodic contour turning point.
- `jumpaft5:` Note after a jump of at least 5 semitones.

For every rule that evaluates to true an accent value of 1 is allocated to that specific note. All accent values are summed for each note. This general approach and the individual rules in particular are described in greater detail in [9].

For the two melodies we receive therefore two resulting number sequences, which are then compared using the Edit Distance algorithm where each number is treated as a separate symbol and costs for deletion, insertions, and substitutions is always 1.

4. Contest results

As last year's MIREX test set hasn't been published and we weren't able to learn from the particular type of melodic phrases, we submitted three very different algorithms. The failure of KF1 in last year's and this year's MIREX shows clearly the detrimental effects of overfitting. KF1 performed very well in predicting human listener judgements regarding the similarity of a specific set of pop music phrases, when it was tested on a separate test set comprising short phrases from the same repertoire. But its internal complexity and its very bad performance on the RISM incipits suggest that it is heavily overfitted to the specific type of data set that is was constructed from.

We are surprised by the still acceptable performance of KF3, which ignores all specific pitch and duration information and instead relies on some quite abstract melodic accents computation. This could mean that accent structure is indeed an important aspect of a melody's identity. In a future optimisation round, this algorithm should be combined with similarity approaches that consider the melody data (pitch and duration) with less abstraction.

KF2 is the algorithm that has proven to be reliable in a variety of situations in the past (e.g. comparing pop music tunes and folk song phrases). We therefore are quite pleased that without any further modification it worked also well on the competition data set and delivered results which are close behind the two algorithms that scored best.

References

[1] Downie, J.S. (1999). Evaluating a simple approach to music information retrieval. *Conceiving melodic n-grams as text*. PhD Thesis, University of Western Ontario.

[2] Frieler, K. (2004). Beat and Meter Extrraction Using gaussified onsets. *ISMIR Proceedings Barcelona, 2004*

[3] Frieler, K., and Müllensiefen, D. (2005). The Simile algorithm for melodic similarity *Abstract MIREX 2005*. http://www.music-ir.org/evaluation/mirex-results/articles/similarity/frieler.pdf

[4] Jones, M.R. (1987). Dynamic pattern structure in music: Recent theory and research. *Perception & Psychophysics*, 41, p. 621-634.

[5] Krumhansl, C. L. (1990). *Cognitive foundations of musical pitch*. New York: Oxford University Press.

[6] Mongeau, M., and Sankoff, D. (1990). "Comparison of musical sequences". *Computers and the Humanities* 24, 161-175.

[7] Müllensiefen, D., and Frieler, K. (2004). "Cognitive Adequacy in the Measurement of Melodic Similarity: Algorithmic vs. Human Judgements". *Computing in Musicology* 13, 147-176.

[8] O'Maidin, D. (1998). A Geometrical Algorithm for Melodic Difference in Melodic Similarity. Melodic Similarity: Concepts, Procedures, and Applications. *Computing in Musicology* 11. Ed. Walter B. Hewlett & Eleanor Selfridge-Field. Cambridge: MIT Press.

[9] Pfleiderer, M., and Müllensiefen, D. (2006). "The perception of accents in pop music melodies. *Proceedings of the 9th International Conference on Music Perception and Cognition, Bologna, 2006*. http://www.escom-icmpc-2006.org/pdfs/513.pdf.

[10] Steinbeck, W. (1982). *Struktur und Ähnlichkeit. Methoden automatisierter Melodieanalyse.* Kassel: Bärenreiter.

[11] Uitdenbogerd, A. L (2002). *Music Information Retrieval Technology*. PhD thesis. RMIT University Melbourne Victoria, Australia