MIREX Symbolic Music Similarity

Pascal Ferraro, Pierre Hanna, Julien Allali and Matthias Robine
LaBRI - Université de Bordeaux 1
351 cours de la Libration
F-33405 Talence cedex, France

`firstname.name@labri.fr`

Abstract

This extended abstract presents a submission to the Music Information Retrieval Evaluation eXchange (MIREX) in the symbolic Melodic Similarity task. This submission is an implementation of an extension of the edit-distance technique to the polyphonic context. Melodies are represented by quotiented sequences. A quotiented sequence is a sequence graph defined with an additional equivalent relation on its vertices and such that the quotient graph is also a sequence graph. The core of the method relies on an adaptation of edit-distance metrics, regularly applied in bioinformatic context.

Keywords: symbolic melodic similarity, edit-distance, quotiented sequence, polyphony.

1. Similarity Scores

To take into account the polyphonic nature of musical sequences, we propose to use a quotiented sequence representation. Formally, a quotiented sequence is a sequence graph with an equivalence relation defined on the set of vertices, and such that the resulting quotient graph is also a sequence. By definition, in a quotiented sequence, quotient graph and support sequence are both sequences. A quotiented sequence can thus be considered as a self-similar structure represented by sequences on two different scales. Note that a quotiented structure can also be viewed as a tree graph.

We propose to extend the representation of monophonic melodies as sequences of pitches and durations [2]. In the context of polyphonic music, notes that occur at the same time are grouped to form a quotiented sequence. Each vertex of the quotiented sequence is labelled by the pitch and the duration of each note. The pitch is coded according to the difference with the tonic (in semitones), so that the algorithms proposed are transposition invariant. The tonic of the piece considered has to be known beforehand. The duration is coded in sixteenth note values.

The score s depends on the pitch range and on the duration difference between two notes x_1 and x_2. One parameter determines the relative weight of the pitch difference with the duration difference. The score due to the pitch difference is thus determined according to consonance: the fifth (7 semitones) and the third major or minor (3 or 4 semitones) are the most consonant interval in Western music. The score associated with the empty symbol is computed according to the score between a note and a rest. This score has been fixed.

This approach leads us to consider any polyphonic sequence as a series of ordered pairs. As for monophonic sequence, each pair is defined by the pitch of the note and its length.

1.1. Edit Score

We propose to consider only three operations (substitution, deletion and insertion), that are usually used to compare musical sequences. Our approach propose to compare quotiented sequences at the more macroscopic scale. Basically, quotiented sequences refer to sequences whose nodes are also sequences. Edit score related to quotient vertices is then defined as an edit score computation between the support subsequences of these vertices. This method is based on an adaptation of an edit distance between unordered quotiented trees based on a comparison of support graph and edit operations that preserves equivalence relations recently introduced by Ferraro and Godin [1].

The main difference between this recursive relation and the computation with the global edit score between sequences lies in the computation of the score of the edit operations between quotient vertices. They are computed as the edit score between the support subsequences corresponding to the quotient vertices. The complexity of this algorithm is $O((|S_1|+|Q_1|)\times(|S_2|+|Q_2|)$, where $|Q_1|$ and $|Q_2|$ represent the respective number of chords in polyphonic musical sequences S_1 and S_2 and $|S_1|$ and $|S_2|$ represent the maximum size of a chord in S_1 and S_2.

1.2. Local Alignment

In many applications, two strings may not be highly similar in their entirety but may contain regions that are highly similar. This is particularly true when long streches of anonymous sequences are compared, since only some internal sections of those strings may be related. In this case, the task is to find and extract a pair of regions, one from each of the two given strings, that exhibit high similarity. This is called *local alignment or local similarity problem* [3] and is defined as follow: given two strings S_1 and S_2, find substrings ρ_1 and ρ_2 of S_1 and S_2, respectively, whose similarity is maximum over all pairs of substrings from S_1 and S_2.

The computation of a local similarity allows one to detect local conserved areas between both sequences and has been generalized to quotiented sequences on the same principle as the global edit score.

Similarly to the computation of an optimal score between quotiented sequences, the complexity of solving the local suffix mapping problem between quotiented sequences is $O((|S_1|+|Q_1|)\times(|S_2|+|Q_2|)$. All optimal local alignments of two quotiented sequences can be represented in two dynamic programming tables and can be found by tracing any pointers back from any cell with the optimal value.

2. Acknowledgment

This work is partially supported by the program *ANR Jeunes Chercheurs* JC07_188930 (projet SIMBALS) from the french Ministry of Research.

References

[1] P. Ferraro and C. Godin. An Edit Distance Between Quotiented Trees. *Algorithmica*, 36:1–39, 2003.

[2] M. Mongeau and D. Sankoff. Comparison of Musical Sequences. *Computers and the Humanities*, 24(3):161–175, 1990.

[3] T. Smith and M. Waterman. Identification of common molecular subsequences. *Journal of Molecular Biology*, 147:195–197, 1981.