

AN ANALYSIS OF THE MONGEAU-SANKOFF ALGORITHM FOR MUSIC INFORMATION RETRIEVAL

Carlos Gómez Soraya Abad-Mota Edna Ruckhaus

Dept. of Computing and Information Technology
Universidad Simón Bolívar, Venezuela

ABSTRACT

An essential problem in music information retrieval is to determine the similarity between two given melodies; there are several melodic similarity measures that have been proposed, among others, the Mongeau-Sankoff measure. In this work we implemented a modified version of the Mongeau-Sankoff measure. We conducted an experimental study to compare the implemented measure with other similarity measures; this evaluation was done in the context of the 2005 edition of the MIREX symbolic melodic similarity competition. The most relevant result of our work is an implementation of the Mongeau-Sankoff measure that presents greater effectiveness when compared to other current melodic similarity measures.

1 PROPOSED APPROACH

Throughout the development of music information retrieval systems, several melodic similarity measures have been proposed. A melodic similarity measure is a function that given two melodies, estimates the degree of identity that can be established between them, expressing it as a real number. This similarity measure is ideally a “continuous” function that is expressed in musical terms.

This work studies the Mongeau-Sankoff measure [2] in the context of music information retrieval, and compares it with some of the existing measures. The Mongeau-Sankoff measure is based on the alignment of musical sequences, where the elements of the sequences are notes with a specific pitch and duration. The similarity between the sequences is expressed in terms of the edit distance between them, which depends on the series of edit operations of least cost that is needed to transform one sequence into the other. In addition to the traditional edit operations (insertion, deletion and substitution), the Mongeau-Sankoff measure defines two operations that apply specifically to music: *fragmentation* and *consolidation*. The measure was designed in 1990, and even though it has been applied previously to music information retrieval systems, to the best of our knowledge, no recent implementation of the measure has been developed with the specific purpose of comparing it to the algorithms evaluated

in a Music Information Retrieval Evaluation eXchange (MIREX) edition, a specific goal of our project.

In this work we propose variants to the measure that are related to several aspects: achieving transposition invariance when the key of queries and the key of documents is not known, computing the interval weight according to its distance, and representing the relative cost of the operations.

Mongeau and Sankoff originally applied their measure to the study of musical variation; therefore, the measure was not conceived specifically for retrieving musical documents in a collection.

The original measure assumes that the key of the two melodies that are being compared is known; based on this knowledge, the measure attains transposition invariance. In the general case of a music information retrieval system, both the key of the query and the key of the documents could be unknown, therefore, it is necessary to adapt the measure in this aspect. In this work, we adopt the approach of making all the possible transpositions of the query with respect to the target, and consider the similarity result as the best value between all the transpositions. Other researchers in this area have applied this approach to Mongeau-Sankoff, and also to other measures.

Another variant of the measure pertains to the weight of the operations that involve the substitution of a note. In the original measure, this weight depends partially on the interval between the note that is being substituted and the new note, and it is defined based on the consonance of the interval. In this work we adopt another approach, which consists of assigning a weight to the interval in terms of the distance between the two notes. This approach has been applied in previous works [1].

Finally, a change was made to the cost functions of the measure. The weight of an operation is the sum of two quantities, that represent the weight associated to the pitch difference between the notes involved in the operation, and that of the duration difference. A parameter of the measure determines the proportion in which the duration difference contributes to the weight, versus that of pitch difference. This constant was originally the same for all operations. In our work, we associate a separate constant to the weight formula of insertions and deletions with respect to the other operations. This allows to vary the relative weight of insertions and deletions with respect to these other operations. Thus, in this work we propose a

Alignment	ADR	NRGB	Average Prec.	R-Prec.	Total time
Local*	0.7337	0.6422	0.5298	0.5361	6.882
All-to-some*	0.6940	0.6044	0.4734	0.4817	5.893
Local	0.6763	0.6046	0.4601	0.4800	6.251
All-to-some	0.6720	0.5771	0.4420	0.4501	5.444
Global	0.6486	0.5586	0.4364	0.4369	5.494

Table 1. M-S with absolute representation. *These configurations implement the proposed variation on the relative weight of the operations.

more general version of the cost formulas that could result in an improvement on the quality of the algorithm.

In our work, we implemented global and local alignment, as well as the additional alignment types described by Meek [1].

2 EXPERIMENTAL STUDY

We conducted experiments to evaluate the effectiveness of the Mongeau-Sankoff algorithm as a similarity measure for music retrieval, using the document collections and the ground truth from the symbolic melodic similarity task of MIREX 2005. The experiments were divided into a parameter optimization stage, and an algorithm evaluation stage. The goal of the first stage was to find a set of parameters that maximizes the quality of the algorithm, using the test collection from MIREX 2005. The goal of the evaluation stage was to assess the quality of the algorithm using the evaluation collection of the contest and the parameters determined in the first stage. The purpose of separating parameter optimization from evaluation was to emulate the conditions of the 2005 contest, in which only the test collection was known by the participants before the contest, so as to avoid algorithm overfitting with respect to the evaluation collection.

We studied several possible configurations of the algorithm according to the alignment type, and the kind of representation (absolute or modulo-12). Experiments were conducted with three types of alignment: global, local, and *all-to-some*.

For each configuration, the optimal value of the algorithm parameters (the parameters of the weight functions) was determined using exhaustive search. The objective function to be maximized was the average of the four quality measures used to evaluate the algorithms in the melodic similarity task of MIREX 2005; a previous use of this objective function can be found in one of the algorithms participating in this contest.

The Mongeau-Sankoff algorithm was then applied to the evaluation collection from MIREX 2005 with 11 queries. The collection was composed of 558 documents. Two of the applied quality measures take into account the order of the results in the ground truth for the computation of the quality of the algorithm (ADR and NRGB), whereas the other two correspond to traditional quality measures and consider all the documents in the ground truth as equally relevant for a query.

Alignment	ADR	NRGB	Average Prec.	R-Prec.	Total time
Local	0.6219	0.5393	0.3919	0.4135	3.869
All-to-some	0.6030	0.5098	0.3728	0.3709	3.233

Table 2. M-S with modulo-12 representation

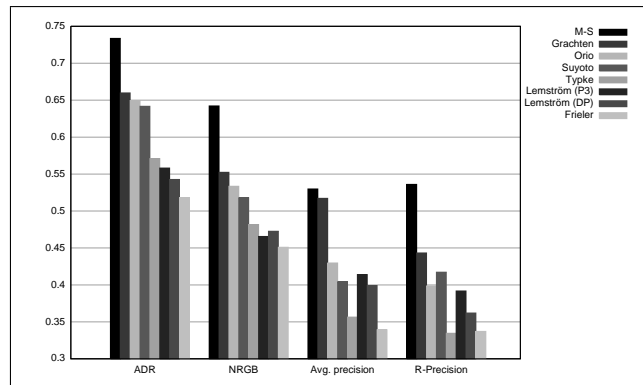


Figure 1. M-S vs. algorithms in MIREX 2005

Tables 1 and 2 show the results of the evaluation. In the following discussion, the ADR was used as a criterion to compare the results. The best quality was obtained with a local alignment and an absolute representation, for the algorithm variant with the proposed formula change; the ADR was 0.7337. In the MIREX 2005 evaluation, the algorithms obtained an ADR between 0.52 and 0.66. Thus, the results of Mongeau-Sankoff with an absolute representation, for the basic algorithm and also for the version with the proposed formula change, are superior to all the results obtained in MIREX 2005 by the participating similarity measures.

Figure 1 shows the comparison between the best Mongeau-Sankoff configuration, and the results obtained by the algorithms in the contest.

Tables 1 and 2 report the total time (in seconds) that was required by the Mongeau-Sankoff algorithm to answer the 11 queries (with an AMD Opteron of 2.2 GHz). The time results show that the algorithm can have an acceptable response time for collections of similar size than the one used. For larger collections, it would be possible to combine the algorithm with an indexing method.

Our implementation with the proposed variants is more effective than all the other similarity measures evaluated in MIREX 2005. Thus, it can be concluded that the Mongeau-Sankoff measure is among the best existing melodic similarity measures.

3 REFERENCES

- [1] Meek, C. *Modelling error in query-by-humming applications*, Ph.D. dissertation, University of Michigan, 2004.
- [2] Mongeau, M. and Sankoff, D. "Comparison of musical sequences", *Computer and the Humanities*, vol. 24, pp. 161–175, 1990.