Music Ranking Techniques Evaluated

Abstract

Several techniques have been proposed for matching melody queries to stored music. In previous work [2], we found that local alignment, a technique derived from bioinformatics, was more effective than the n-gram methods derived from information retrieval that are used in other work. In this paper we explore a broader range of n-gram techniques, and test them with both manual queries and queries automatically extracted from MIDI files. Our experiments show that alternative - indeed, simpler - n-gram matching techniques than those tested in earlier work can be as effective as local alignment; one highly effective technique is to simply count the number of 5-grams in common between query and stored piece of music. N-grams are particularly effective for short queries and manual queries, while local alignment is superior for the automatic queries.

In our experiments we have used a database of 10 466 standard MIDI files down loaded from the internet. These were processed to extract their melodies as described in [1]. In that earlier paper, we used volunteer listeners to judge a small set of extracted melodies. We discovered that our simple algorithm that chooses notes starting at each instance to be melody notes and that we have named *all mono*, was the most successful in extracting melodies. However, we continued to test all of our melody extraction algorithms in our second paper [2] in the context of melodic similarity measurement. In this experiment we used three methods of melodic standardisation to produce the representations that would be used for matching. They were: *contour*, which uses a three-symbol alphabet to represent where a melody goes up, down or stays the same; *modulo-12 interval* (pitch distance between notes), which keeps the size of intervals and their direction but reduced to one octave; and *exact interval*, which keeps the exact interval size and direction. All these represented pitch information only.

A set of 28 automatically extracted melody queries were used to locate other versions of the same piece of music. The results were evaluated by computing the eleven-point precision average and the precision at a recall level of 20. Again the all mono method was successful, as were the *entropy channel* and *all channels* methods. The entropy channel method applies the all mono technique to each channel and then selects the melody channel based on its first order predictive entropy. The all channels technique also applies the all mono algorithm to each channel and then retains them all for the matching process. In addition, we confirmed that contour was insufficient for good matching, but that the use of interval sequences worked well. Queries of 30 intervals produced good answers. The similarity measurement techniques tested were local alignment, longest common subsequence, an n-gram counting technique that included the frequency of occurrence of each n-gram, and the Ukkonen n-gram measure. Local alignment was a clear winner. However, local alignment is a slow technique. It would be useful to have a good n-gram-based technique as indexes can be built to speed up matches. This led to our current exploration of n-grams.

In this current set of experiments we used the same experimental framework as that used in [2], but also tested a set of manually produced queries. These manual queries were created by a classically trained pianist with perfect pitch, who listened to the source MI DI files and played a melodic fragment to represent each piece. We explored the variables of n-gram length, n-gram counting method, song-length normalisation, melody extraction and query length. The results clearly showed that n-grams of length five work best using the count distinct method, where term

frequency is ignored. Some length normalisation was found to be best, such as dividing by the log of the number of intervals. Query lengths of 20 and 30 each gave similar results, but those of length 10 were much less effective. The all channels technique was found to be far superior to the other methods when manual queries were used. For the automatic queries the results are less distinct.

It is quite clear from our experiments that an index-based approach to melody matching using n-grams can be used to produce good answers to melody queries and we recommend an n-gram length of five for this purpose. Term frequency should be ignored in this approach. The method of melody extraction and standardisation used within the collection is also important in implementing a successful system.

Author Information

Alexandra L. Uitdenbogerd, Justin Zobel
Department of Computer Science, RMIT University
{alu, jz}@cs.rmit.edu.au

Suggested Readings

Downie, Stephen J. and Nelson, Michael. 2000, "Evaluation of a simple and effective music information retrieval method", *Proc ACM-SIGIR Conference*, Athens, Greece.

Uitdenbogerd, Alexandra and Zobel, Justin. 1998. "Manipulation of music for melody matching", *Proc. ACM International Multimedia Conference*, Bristol, England.

Uitdenbogerd, Alexandra and Zobel, Justin. 1999. "Melodic matching techniques for large music databases", *Proc. ACM International Multimedia Conference*, Orlando, Florida, USA.