Searching musical incipits by means of sequence alignment

Anonymized for peer reviewing

Introduction

Various methods for symbolic melodic search have been proposed, from string matching based approaches to geometric models [3]. A study on melodies in the Dutch Song Database (www.liederenbank.nl) has shown that sequence alignment methods are quite powerful at dealing with melodic variability [2]. We investigate the potential of such methods for retrieving RISM incipits, comparing 27 different alignment methods.

Relation to music encoding: this work is part of our ongoing endeavour to make encoded music searchable

Aligning incipits

Informally, in musical sequence alignment, the notes of two melodies are arranged in such a way that corresponding notes are vertically aligned. The quality of an alignment is expressed in a score. Identical notes in the two melodies receive the maximum benefit (typically 1); insertion of gaps and substitution of one note by another receive lower benefits, or even penalties. The highest possible alignment score of two melodies (computed by Dynamic Programming) can be regarded as a model of melodic similarity. By aligning a query incipit to all items in a collection, we can create a ranked list of similar items.

A crucial step is the computation of the substitution cost by means of *raters* for the relevant properties of the notes. We look into three such properties: pitch (represented in Base 40 [1]), metric weight (computed using Inner Metric Analysis [5]), and duration. For each property, three different raters are defined (Table 1). Rater scores are multiplied to create the final score.

Rater	abbreviation	benefit/penalty calculation and remarks
Pitch		transposition invariant through pitch histogram comparison
Kranenburg	pi1	1 for unison, linear decrease to 0 for fifth, -1 for larger intervals (modulo octave) [2]
Exact	pi2	1 for unison, 0.5 for same pitch class, otherwise -1
Zigzag	pi3	linear function of interval class, between 1 (unison) and -1 (triply augmented fourth)
Metric		
(none)	mw0	
IMA weighted	mw1	IMA scaled to average 1; benefit = average of resulting note weights
IMA combined	mw2	absolute difference of note weights
Duration		
(none)	dur0	
Fixed	dur1	duration as notated; benefit = overlap
Scaled	dur2	compares duration histograms to determine corresponding durations; scale and benefit = overlap

Table 1: raters

Evaluation

All 27 combinations are evaluated on the complete RISM data set. For 11 incipit queries, a ground truth is available [4] which is however based on a much smaller version of the RISM data set. The ground truth was extended using the pooling method. For each incipit, the top-50 results for all methods were combined. Subsequently, all incipits not appearing in the original ground truth were evaluated. The extended ground truth is published at http://anonymo.us.

The evaluation metric we employ is the Area Under Curve (AUC), the curve being the Precision-Recall curve calculated for each method. All AUC values are shown in Table 2.

method	AUC
pi1mw0dur0	0.30
pi1mw0dur1	0.31
pi1mw0dur2	0.32
pi1mw1dur0	0.06
pi1mw1dur1	0.12
pi1mw1dur2	0.03
pi1mw2dur0	0.24
pi1mw2dur1	0.27
pi1mw2dur2	0.26
pi2mw0dur0	0.35
pi2mw0dur1	0.41
pi2mw0dur2	0.38
pi2mw1dur0	0.31
pi2mw1dur1	0.41
pi2mw1dur2	0.39
pi2mw2dur0	0.39
pi2mw2dur1	0.42
pi2mw2dur2	0.38
pi3mw0dur0	0.33
pi3mw0dur1	0.36
pi3mw0dur2	0.35
pi3mw1dur0	0.22
pi3mw1dur1	0.31
pi3mw1dur2	0.24
pi3mw2dur0	0.29
pi3mw2dur1	0.35
pi3mw2dur2	0.33

Table 2: AUC for all methods

Averaged over all combinations, the best raters are pi2 (exact pitch, AUC=0.38), mw0 (no weight, AUC=0.35) and dur1 (fixed duration, AUC=0.33). If we look at individual combinations, the best performing ones are pi2mw2dur1 (AUC=0.42), and pi2mw0dur1 and pi2mw1dur1 (AUC=0.41).

Case study

RISM contains 31 incipits entitled *Roslin Castle*. Twenty-nine incipits clearly refer to the same tune, though sometimes with substantial variants. One of these was used in [4] as a ground truth query. Unfortunately, this is an atypical version, possibly lacking a key signature of one sharp (Example 1). Table 3 shows the number of retrieved items to this query in various experiments.

Example 1: query

method	retrieved items	comment
RISM OPAC	17	key signature added
Earth Movers Distance	12	[3] (17 items known)
Proportional Transportation Distance	16	[3] (combining 3 different queries)
pi2mw2dur1	23	
pi2mw0dur1	25	
pi2mw1dur1	25	
pi2mw2dur2	26	
pi2mw0dur2	27	
pi2mw1dur2	27	

Table 3: Retrieval results for various methods on *Roslin Castle*

Interestingly, the best performing methods are only 4th and 6th in Table 2. Specifically, they pick up 2 items that methods with fixed duration ignore (Example 2).

Example 2: item found by scaled duration

Example 3: item with octave error

All methods find items that contain local rhythm and pitch errors (Example 3), contain insertions and/or deletions, split or merged notes, and ornamentation. The price one pays is that sometimes seemingly unrelated melodies may also be highly ranked (for example, using pi2mw0dur0, one finds instances of the chorale *Jesus meine Zuversicht* ranked quite highly). Yet this property could also be relevant for tracing melodic topics and allusions.

Conclusion

Alignment search methods show a good performance on RISM incipits, clearly outperforming the current string based method of RISM OPAC. We believe that the utility of RISM OPAC could be increased by integrating one or several alignment methods in addition to the current search method.

A full RISM search demo with all matching methods is available at http://anonymo.us. Searches typically take only a couple of seconds, which is well within the limits of acceptability for research purposes. The alignment software can be downloaded from http://anonymo.us.

References

[1] Hewlett, W.B. (1992). A Base-40 number-line representation of musical pitch. *Musikometrika* 4, 1-14.

[2] Kranenburg, P. van (2010). A computational approach to content-based retrieval of folk song melodies. PhD thesis Utrecht University

[3] Typke, R., Giannopoulos, P., Veltkamp, R. C., Wiering, F., & Van Oostrum, R. (2003). Using transportation distances for measuring melodic similarity. In *ISMIR*.

[4] Typke, R., den Hoed, M., de Nooijer, J., Wiering, F., & Veltkamp, R. C. (2005). A ground truth for half a million musical incipits. *Journal of Digital Information Management*, 34-39.

[5] Volk, A. (2008). The study of syncopation using inner metric analysis: Linking theoretical and experimental analysis of metre in music. *Journal of New Music Research*, 37(4), 259-273.