6 Transportation Distances and Their Application in Music-Notation Retrieval

Frans Wiering, Rainer Typke, Remco C. Veltkamp

Institute of Information and Computing Sciences
Utrecht University
P.O. Box 80.089
NL-3508 TB Utrecht
Netherlands
{frans.wiering,rainer.typke,remco.veltkamp}@cs.uu.nl

Abstract

This article introduces the representation of music notation as a weighted dot pattern and defines two distance measures—the Earth-Mover's Distance and the Proportional Transportation Distance. Their performance is compared in the ranking of melodic-match candidates.

Queries with both distance measures are evaluated and their applicability in music research is discussed. The research is related to the Orpheus project (*http://www.cs.uu.nl/centers/give/multimedia/music/*), which aims to develop efficient methods for content-based retrieval from large collections of encoded music notation producing meaningful results.

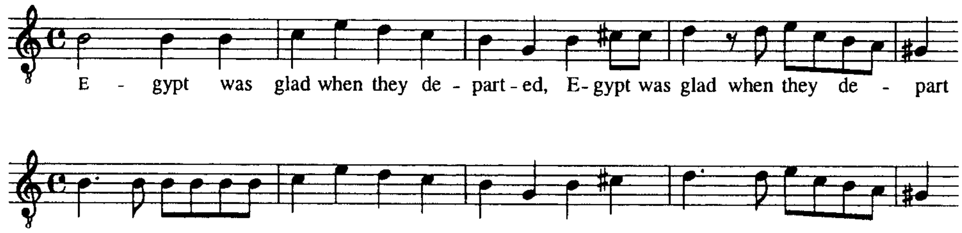
Music Query: Methods, Models, and User Studies (Computing in Musicology 13). Published by CCARH and the MIT Press, 2004.

6.1 Introduction

Suppose we are studying Handel's oratorio *Israel in Egypt* and would like to know more about sources the composer used to create this work. Take, for instance, the chorus "Egypt was glad when they departed". Just entering these words in Google gives the source of the text: Psalm 105, verse 38 (Hebrew/Protestant numbering; retrieved March 2, 2004, from *http://www.sacred-texts.com/bib/jps/psa105.htm*). However, were we to try to search the Internet for the beginning of the music (Example 6.1a), we probably would not be so lucky. If the data are there at all, then an obvious tool for searching them is lacking. To identify the musical source of Handel's chorus, a canzona by Kerll (Example 6.1b), we would depend on paper tools such as catalogs, thematic indices, and scores.

a

b



Example 6.1. (a) G. F. Handel, Israel in Egypt, HWV 54; (b) J. K. Kerll, Canzona 4 from Modulatio Organica super Magnificat (1686).

Human beings perceive the two melodies in Example 6.1 as very similar. There are some slight rhythmic differences in order to fit the text to the notes, but the melodic contour and the intervals are the same. We would expect a useful search engine to deal with musical similarity as humans do.

Despite several decades of research into computer analysis of music and music retrieval, the most frequently used way of searching notation is pattern matching in text representations. Generally, it is possible to search one dimension only: pitch (intervals being another way of looking at the same dimension). This makes it hard to take the relative importance of a note into account, which depends on such factors as duration and metrical position. As a consequence, a musical query may result in false negatives, when relevant items are missed, or false positives, when irrelevant items are retrieved. An example of the former would be not finding Example 6.1b when searching for Example 6.1a; this would happen when the query required the composition to begin with precisely three repeated notes. If, however, one would restrict the number of notes in the query to six (as in Example 6.2a) and allow an arbitrary number of repetitions of the first note, false positives such as Example 6.2b may result (retrieved September 26, 2003, from *http://www.themefinder.org*).

Example 6.2. (a) Query without pitch durations; (b) L. van Beethoven, Symphony No. 6, first movement.

The need for multi-dimensional music searching is widely acknowledged, yet only a few effective implementations exist. We have selected two that most resemble our approach for comparison, PROMS (Clausen et al. 2000) and C-Brahms (*http://www.cs.helsinki.fi/group/cbrahms;* Ukkonen et al. 2003).

In PROMS, the score is represented as a set of notes with two characteristics: MIDI pitch and onset time (but no duration). Queries, which may be polyphonic, are similarly encoded. PROMS tolerates missing notes by default, so long as the time structure is kept intact. Therefore, our two melodies in Example 6.1 would be matched, but it would be impossible to match Example 6.2a with these. One can construct a fuzzy query by specifying pitch alternatives for one or more notes. Such a fuzzy query is interpreted as a set of exact queries. There are indexes that allow fast queries over the test collection of *c.* 12,000 classical compositions. However, there is no way of finding rhythmical transformations of the query. Note also that this method is non-continuous: a small extra deviation from the pattern (e.g. a note that is just a semitone higher) may lead to an entirely different judgment (no similarity rather than similarity).

Ukkonen et al. (2003) represent notes as line segments in a pitch-time space. Example queries are all monophonic, but polyphonic queries appear to be possible too. Three algorithms are described that respectively find exact matches (P1), the largest common subset of onset times between the score and the query (P2), and the longest common shared time between the score and the query (P3). Algorithm P2 can be used to find partial matches with less than a given number of errors. The two melodies in 6.1 have 4 mismatches. One could also match Example 6.1 to Example 6.2b (assuming that 2b is rewritten with note values doubled), with 5 matching notes and 9 mismatches. Algorithm P3 allows another form of incomplete matching, where onset and offset times need not be precisely coordinated. The algorithms were tested in the C-Brahms project on a database of 10 times 279 musical documents. These methods too are non-continuous. In P2, one duration change, however small, will disrupt the match, as will the change of one melodic interval. P3 is somewhat more tolerant of duration change, but has the same problem with interval changes as P2.

Our method shares with the ones just described the "geometric" representation of music notation in a two-dimensional [pitch-time] space: we represent notes as dots with different weights that reflect their musical importance. Other than in PROMS, we do not compare dot patterns by deciding whether they match or not; we apply similarity measures to pairs of dot patterns that calculate their distance on a continuous scale. The hypothesis is that the smaller such a distance is, the more likely the two patterns resemble each other musically. Music retrieval would thus be improved by enabling ranking of the retrieval output. The C-Brahms algorithms already allow ranking, but this ranking is undercut by the algorithms' sensitivity to certain small deformations that are very common in musical practice. Our method is better able to handle these problems.

The similarity measures we discuss in this article are the Earth-Mover's Distance and the Proportional Transportation Distance. To evaluate these, we employed the RISM A/II collection (2002). This database contains bibliographic data about musical manuscripts dating (mainly) from the period 1600-1850. For most compositions, the beginning of one of the voices (an *incipit*) is given (or one for each movement), so that the work is characterized not only by its metadata but also by its initial musical content. Typically, incipits are *c.* 25 notes long, with considerable variation (a minimum of twelve pitches is required, though in practice some incipits are shorter). The data are used in several common musicological tasks, such as attributing a work to a composer, identifying its (physical) sources, or tracing musical relationships between different compositions. There were about 470,000 incipits in the version of the RISM A/II collection we employed.

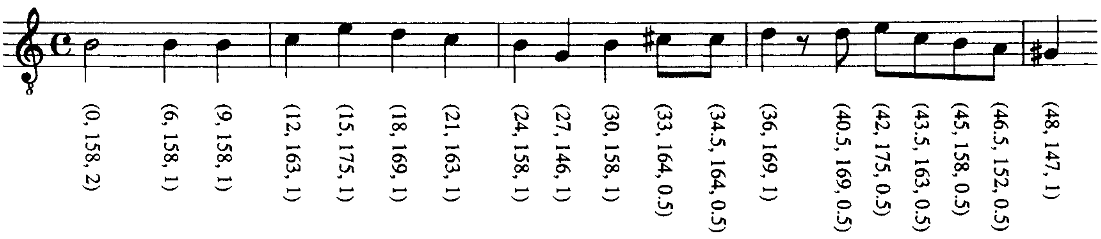
Size was an important reason for selecting this collection. It is also maintained by a professional organization and designed to be used for a variety of representative scholarly purposes, so that both quality and usefulness are guaranteed. A search engine for musical incipits is missing in RISM A/II, although some user interfaces to the collection provide searches of the *Plaine & Easie Code* by which the incipits are represented [see Appendix I]. Finally, the similarity measures discussed here work best when comparing units of roughly the same size. Using the same measures for querying scores of complete works necessitates preprocessing, a step we wished to avoid in the early phases of our research.

6.2 Musical Notation as a Weighted-Dot Pattern

Rendition of a piece of music notation as a weighted-dot pattern works as follows. Each note onset is represented as a dot in a two-dimensional space. The axes are time and pitch; and the weight of the dot is an estimation of the note's importance to the melodic shape of the incipit. So far, we have assumed that a note's importance is proportional to its duration, but ultimately weights should be assigned on the basis of what is known from music perception. Rests are not encoded.

Pitch is encoded using Walter Hewlett's base-40 notation (1992). Like MIDI, this is a number-based notation, but with two additional advantages: enharmonic differences are respected, and enharmonically correct intervals can be calculated by subtracting pitch values. Durations and time coordinates are expressed by means of fractions of quarter notes (for the precise calculation see Typke et al. 2003).

Because of the base-40 pitch representation, the pitch range of an incipit is often much larger than its duration range. An extreme case is Example 6.2, with a pitch range of 46 and a duration range of only 5.5. To compensate for this, we multiplied the time coordinates by 3 for all examples in this article. For Handel's incipit, the values are given in Example 6.3.



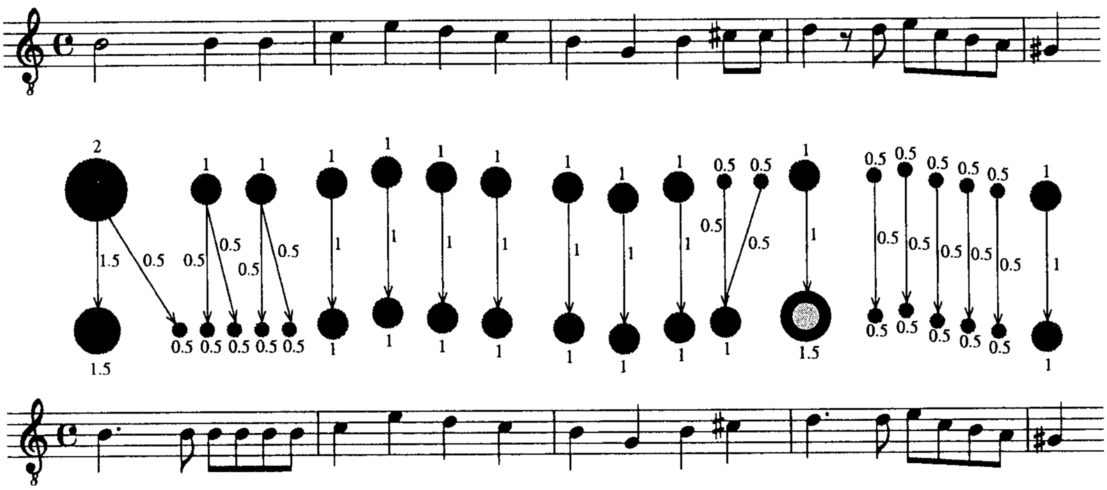
Example 6.3. Handel's incipit as a weighted-point set. The parameters given inside parentheses are time, pitch, and weight.

6.3 Incipit Comparison using the Earth-Mover's Distance (EMD)

The Euclidian distance between the weighted dots acts as the "ground distance" for the calculation of transportation distances. The Earth-Mover's Distance (described in Cohen 1999) calculates the minimum effort needed to transform one dot-pattern into another. One can imagine this to happen as follows. One pattern (Kerll's incipit in Ex. 6.4) is represented as heaps of earth, the sizes of which correspond to the weights of the dots; the other pattern (Handel's incipit in Ex. 6.4) as holes with a certain capacity, likewise corresponding to the dots' weights. The task is to fill the holes with as little effort (that is, ground distance times weight) as possible. For our example, many heaps and holes have the same coordinates and size, so their contribution to the total effort is zero.

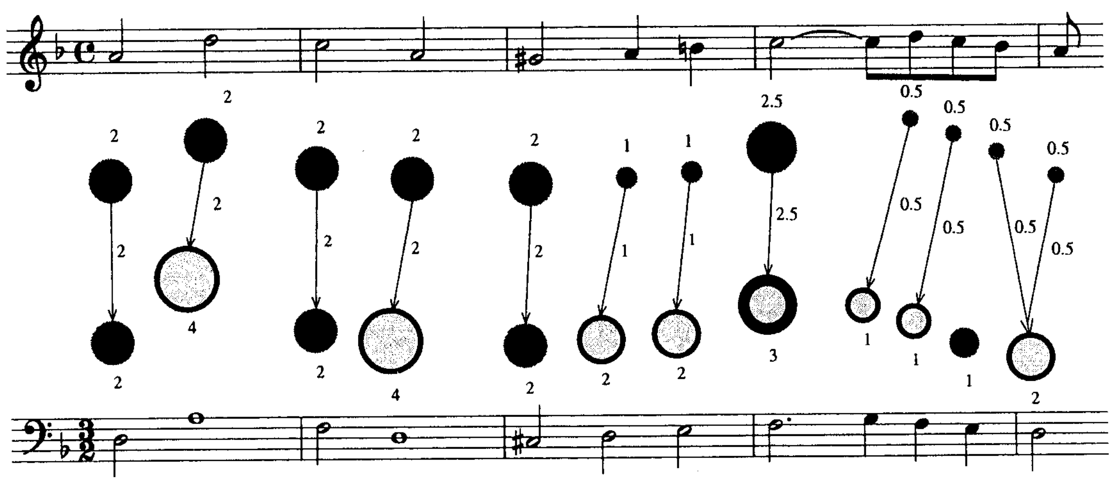
The heap of Handel's first note, however, is too big for Kerll's first note, and the remaining earth is moved to the second note. The effort is 4.5 x 0.5 = 2.25. Likewise, Handel's second and third notes are split up, each resulting in an effort of 1.5 x 0.5 = 0.75. On the other hand, the two C-sharps in Handel's measure 3 merge into one note in Kerll, also resulting in an effort of 0.75. Some capacity remains for the first note in Handel's measure 4, since all earth has been used: this has no effect on the EMD.

The total effort equals 4.5. The effort is normalized by dividing it by the total weight of the lighter pattern. The EMD is thus 4.5/16.5 = 0.2727. The mapping of weights is shown in Example 6.4. Note that for clarity the melodies are shifted apart; in reality, most of the notes are actually at the same coordinates.



Example 6.4. EMD weight flow between Handel's and Kerll's incipits. The rim of Item 15 (lower row) corresponds to the matched part of the weight.

One interesting property of the EMD is that it allows partial matching. Compare for instance the first two measures of Handel's incipit and the complete incipit. The EMD is zero, as all "earth" in the former fits in the "holes" of the latter. That holes remain (or earth) is not an issue. The downside of this property is that the EMD is less suitable for constructing an index. (See Typke et al. 2003.)



Example 6.5. J. S. Bach, Kunst der Fuge, BWV 1080: EMD weight flow between Contrapunctus 1, beginning of soprano (above) and Contrapunctus 12, beginning of bass (below). The EMD is 1.51256. In the lower row of circles, partial matching is indicated by a lighter shade of grey.

Example 6.4 is deliberately kept simple. Weight is only moved in the time dimension. It can also be moved in the pitch dimension, as in Example 6.5. But before computing the EMD, the two incipits need to be aligned. Durations are aligned by stretching the shorter incipit so that the beginnings of their last notes coincide (unless we are looking for a partial match). For pitch alignment we align the weighted average pitch. This is an approximation that may be improved in various ways. The diagram between the two incipits in Example 6.5 shows the optimum weight shifts needed to match them.

6.4 Incipit Comparison using the Proportional Transportation Distance (PTD)

To remedy the EMD's problem with indexing, Giannopoulos and Veltkamp (2002) developed another distance measure, the Proportional Transportation Distance. For the PTD, the total weight of both patterns is normalized to one; subsequently the EMD is applied. In terms of the earth metaphor, this means that all earth is always distributed over the holes proportionally to their capacity. Because of its different mathematical properties, the PTD can be used to construct an index, so that searching is much more efficient than when using the EMD (see Typke et al. 2003 for a discussion of the complexity of the algorithm). The price one pays is that partial matching is not possible with the PTD.

6.5 Applications in Music Research

Composer attribution is one of the three musicological tasks we mentioned in our introduction. The PTD was used in a "de-anonymifying" script that compared *c.* 80,000 anonymous incipits from the RISM (*Répértoire Internationale des Sources Musicales*) A/II database to all other incipits. In 13% of the cases, the system reported candidate attributions. A random sample of a hundred of these were checked manually. Fifty-five cases were matches with other anonymous works. Nineteen works could be attributed to a named composer because the compared incipits are identical. In eleven cases we could attribute the work although the incipits are not identical. Our method thus leads to the identification of about 3.9% (30% of 13%) of the anonymous pieces (see also Typke et al. 2003). Schlichte (1990) managed to identify only 2.08%.

Example 6.6 shows the process of assigning authorship in practice, with two candidate attributions for the anonymous incipit. Note that the incipits are similar, but not identical: there are differences in both rhythm and intervals. The individual character of the melody makes it very likely that we are dealing with closely related pieces, while the metadata enable us to establish that the anonymous work is probably a piano arrangement of Umlauff's aria.

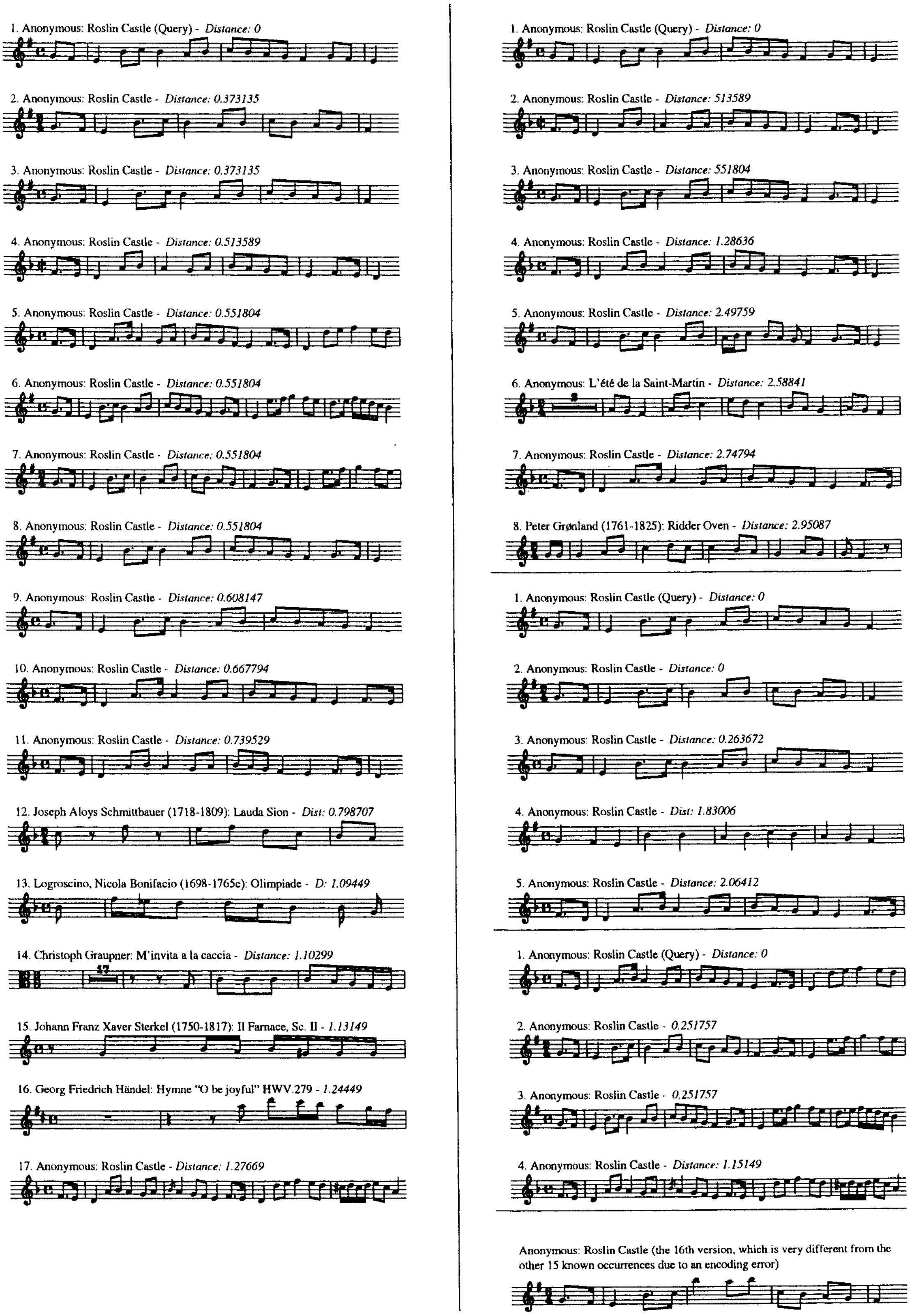
Example 6.6. *(a) Anonymous keyboard piece; (b) I. Umlauff/ A.F.J. Eberl/ W.A. Mozart, Ariette variée; (c) I. Umlauff, Das Irrlicht, aria "Zu Steffen sprach im Traume."*

Whether the piece is identical to the *Ariette variée* ascribed to various composers cannot be established from the available data; for this the complete works need to be compared. The situation is thus no different from working with thematic indices, except that the speed and arguably the reliability are much greater. Also, the idea of being able to link directly to a score is tempting though as yet unrealized.

We tested the second task, the identification of physical sources, by means of the melody of "Roslin Castle," which was previously investigated by Howard (1998). Example 6.7 shows our results with both the EMD and the PTD. Howard, using an earlier version of the database, was able to retrieve 6 out of 13 known occurrences in his test set (46%). With the EMD, we were able to group together 11 out of 15 known correctly encoded occurrences (73%). The 16th occurrence is incorrectly encoded; as a consequence the EMD is greatly increased and the ranking is correspondingly low.

To retrieve the 15 correct occurrences with the PTD, we had to query three times, and the largest of these groups has only 6 incipits (for full results see Typke et al. 2003). The reason for this is that the PTD does not allow partial matching, whereas incipits of "Roslin Castle" in the database differ in length. Queries therefore need to have different lengths too, that is, 13, 16 and 21 notes. For the RISM database to be searched more efficiently by means of the present method, the musical data would need to be encoded according to more strict rules than is done at present. On the other hand, PTD-based searching could be improved by making it less dependent on length. One way to do so would be to add a weight component that is proportional to the inverse note number, thereby making notes progressively less important. Another would be to devise a way of dividing the incipits into chunks that can be better compared, and then to integrate the results.

The EMD Example 6.7 shows a number of false positives that are not at all musically similar to "Roslin Castle." This is caused by the partial-matching property: note that all false positives are relatively short. Here, too, a chunking strategy that produces chunks of similar size may be the solution. Rests within an incipit may cause trouble as well. A possible remedy is to replace note duration with distance-to-the-next-note in the weighting.



Example 6.7. "Roslin Castle," retrieved with one EMD-based query (left) and three PTD-based queries (right) in declining order of relatedness.

The third task, tracing musical relationships, is the hardest to perform. It seems that the EMD and PTD work well when variation is small, as in the case of Handel's borrowing from Kerll; when it is larger, the distances between related melodies are about the same as between unrelated ones. Example 6.8 shows two related themes from J. S. Bach's *Kunst der Fuge*; yet the EMD and PTD between these are greater than 5. In practice, the second theme will not be retrieved using the first as a query because there are too many intervening false positives. One way to get a better match between the principal theme and that of the Canon is to introduce other features into the weighting. Since in this case similarity is determined first of all by the notes that receive metrical stress, this feature is an obvious candidate.

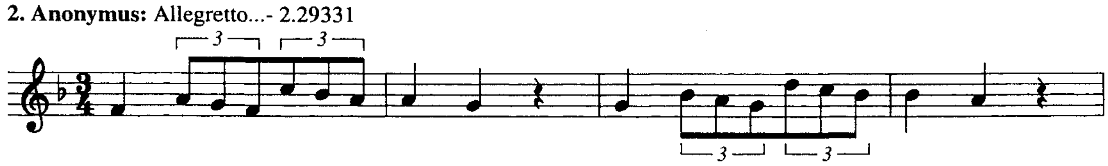
Example 6.8. Two themes from J.S. Bach, Kunst der Fuge, with EMD and PTD distances.

There are many different levels of musical similarity. So far most of the examples have been about a specific kind of similarity, where different instances of one melody are compared. Example 6.8 is the exception: here some effort is required to see—or hear—the similarity between two melodies sharing the same underlying structure. From a historical or music-analytical perspective, searching for laxer forms of similarity is an important task. One may think of inter-textual relationships like allusion and emulation, and of searching for music that shares some generic quality, such as works of one artist, works belonging to the same genre, or works using a similar principle of construction. Listeners (as opposed to music professionals) may want to find music that "sounds like" or has the same emotional quality as a song they like. This type of search brings us to an area that is barely covered by present-day music retrieval techniques. Audio retrieval can be used to make a very general division of music into categories, one of which is often described as *classical music* (McKinney and Breebaart 2003). A finer granularity seems to be very hard to reach. Techniques for notation retrieval are so strongly tied to notes that concepts that abstract from these are hardly an object of retrieval—with the exception of some interesting experiments in harmonic retrieval (Pickens et al. 2002).

So far our approach has been strongly note-oriented, although obviously transportation distances could be used to measure the similarity of any other quantifiable musical property too (Berenzweig et al. 2003). Yet some queries we have performed using the *Orpheus* system suggest that it is possible to search for more generic resemblances too.

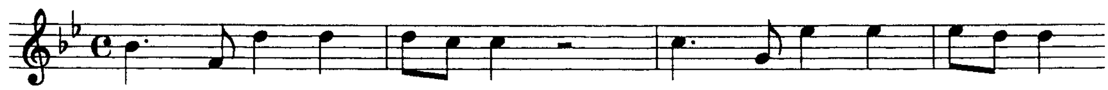
1. Hasse, Joh[ann] Adolf 1699-1783: Artemisia... - 0

2. Anonymus: Allegretto...- 2.29331



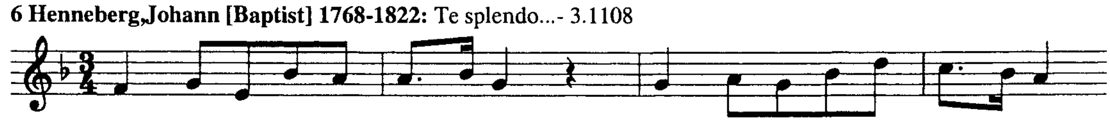
3 Hasse, Joh[ann] Adolf 1699-1783: Romolo ed...- 2.8981

4. Anonymus: Totengrab...2.902



5. Carissimi, Giacomo 1605-1674: Democritus...- 2.94952

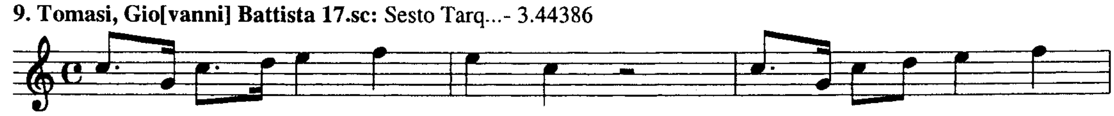
6 Henneberg,Johann [Baptist] 1768-1822: Te splendo...- 3.1108



7. Grétry, [André-Ernest-Modeste] 1741-1813: Fausse mag...- 3.17135

8. Aprile, Giuseppe 1731-1813: Bacio di N...- 3.3104

9. Tomasi, Gio[vanni] Battista 17.sc: Sesto Tarq...- 3.44386



10. Anonymus: LibertÀ...- 3.46176

11. Vranicky, [Pavel]: Oberon Kö...- 3.50097

12. Hofmann, Leopold 1738-1793: Symphonies...- 3.53097

13. Anonymus: Dances fl...- 3.54719

14. Anonymus: Altes Kirc...- 3.55255

15. Scopazzagatti, Michele: Sonatas or...- 3.56192

Example 6.9. Query for a melody by J.A. Hasse (no. 1) and first 15 hits sorted by transportation distance. Multiple hits for the same composition (nos. 1, 3, and 7) were removed.

In Example 6.9, a query was made for a melody from Johann Adolph Hasse's opera *Artemisia* by means of the EMD (duration weights only). Only hits for different melodies are shown, as variants of the same melodies (nos. 1, 3, and 7) show only minuscule differences. Surprisingly, several of the remaining hits show a conspicuous yet generic similarity to the query (nos. 2, 3, 4, 6, 7, 8, and 13; to a lesser extent, also 10 and 11). These melodies consist of two phrases, each beginning with a melodic ascent and closing with a descending second that acts as the resolution of a suspension. The second phrase is typically a second higher than the first. In addition, most of the incipits seem to be based on the very common harmonic framework I-II$_6$-V-I, as illustrated in Example 6.10.

Example 6.10. Inferred harmonic progression for J.A. Hasse's incipit.

Interesting as these results may be by themselves, they are no more than chance hits at this stage. If we inspect weight flows between pairs of incipits, as was done in Examples 6.4 and 6.5, some notes are matched very convincingly, while others may be split up and their parts matched to notes quite some distance apart. Weight flows for false positives such as no. 5 are not too dissimilar from those for the true positives, making it clear that the EMD by itself is not a sufficiently accurate model of human perception of similarity. Yet at the same time we can derive from this example the hypothesis that application of contour and harmony as additional criteria will increase the precision of our retrieval methods by eliminating false hits like nos. 5, 9, 12, 14 and 15.

6.6 Evaluation

Compared to Schlichte's and Howard's experiments with the RISM A/II collection using string-based techniques, we were able to attribute almost twice as many anonymous works to a composer, and to retrieve a distinctly higher percentage of instances of the "Roslin Castle" melody using the EMD. The reason for our better performance in computer attribution is, first of all, that we base our judgments on similarity rather than identity. It is unfortunate that Howard's data pertain to two melodies (from which we chose one) only, so that quantitative comparison of methods is impossible. As the RISM A/II database is large, high-quality, and generally available, it is useful to determine a ground truth for it, that is, a number of queries that have been evaluated by human experts against the whole database. This ground truth can then be used to benchmark the retrieval performance of search engines. We intend to compile and make publicly available a ground truth for RISM A/II in the near future.

Our hypothesis concerning the use of transportation distances for musical patterns is that *the smaller the distance, the more the two patterns resemble each other musically*. In the experiments described above, weights depend on duration only; corresponding query outputs show, in addition to many true hits, too many false positives to accept the hypothesis unconditionally. We still need to improve the way in which weights model the musical importance of notes—and test this by means of an established ground truth. Four factors that may improve weighting have been mentioned so far: note number, distance to next note, metrical stress, and contour. We expect the relative importance of such features to depend on genre. In plainchant, for example, metrical stress plays no role, and duration a minor one; contour is likely to be a major factor. Duration and metrical stress has an obvious place in dance music.

Finally, how does our work compare to other geometric approaches to music retrieval? The main advantage of our approach is that our distance measures have the property of *continuity*. Small changes of point sets (shifts of points and/or differences of weights) have a small effect on the transportation distance; if the changes in the point sets increase, then the transportation distance increases correspondingly, without sudden changes. It is because of this property that matches as in Example 6.9 are possible at all. Continuity is especially an improvement over PROMS (Clausen et al. 2000), where pitch may vary only within a range that has to be specified for each note, and duration may not vary at all. The C-Brahms P3 algorithm has some tolerance against duration change, but as soon as notes no longer overlap, matching is impossible. However, PROMS beats our present solution on two important issues. Searches are much faster due to efficient indexing techniques, and the underlying database consists of complete polyphonic scores rather than monophonic incipits.

6.7 Beyond Incipits

The ultimate challenge is to fully search large numbers of polyphonic scores in an efficient and reliable manner by means of polyphonic queries. Can our approach be used for this? We can only give a tentative answer. In any case, it makes no difference whether the dot patterns are polyphonic or not, as preliminary experiments using the EMD have confirmed. For matching queries against complete scores, the scores, and possibly also the queries, must be divided into chunks. One way to do so is by musical units such as measure, phrase or chord. Another is to set a fixed chunk size and divide scores and queries into overlapping chunks. This way presupposes no musical knowledge, which is an advantage. The application of the EMD for partial matching again raises the indexing problem. We are investigating how the EMD can be used to construct a usable index after all.

So far, all our searching has been melody-oriented. But polyphonic queries may partly or wholly consist of chords, and implicit harmony may have an important place in judging the similarity of melodies (see Example 6.10). Searching chords raises an interesting problem, as each chord can be expanded into a large number of pitch combinations. This relates to pitch representation. Currently we measure interval size, making an octave more related to a seventh than to a unison. If we order intervals by degree of consonance and use this order as the basis of distance calculation, this might offer a solution for dealing with different forms of the same chord, and at the same time better agree with perception.

To make our approach suitable for polyphonic "full-score" retrieval, computational, perceptual and musical issues need to be solved. If we solve these, we may be able to push music notation retrieval past its experimental stage and provide generic tools that help end-users and not just their developers.

References

Berenzweig, Adam, Beth Logan, Daniel P. W. Ellis, Brian Whitman (2003). "A Large-Scale Evaluation of Acoustic and Subjective Music Similarity Measures," in *ISMIR 2003: Proceedings of the Fourth International Conference on Music Information Retrieval*, ed. Holger H. Hoos and David Bainbridge (Baltimore, Johns Hopkins University), 99–105.

Clausen, M., R. Engelbrecht, D. Meyer, and J. Schmitz (2000). "PROMS: A Web-Based Tool for Searching in Polyphonic Music," in *ISMIR 2000: Proceedings of the 1st International Symposium on Music Information Retrieval* (Plymouth, MA); *http://www-mmdb.iai.uni-bonn.de/download/publications/promsIsmir 2000.pdf.*

Cohen, Scott (1999). "Finding Color and Shape Patterns in Images." Ph. D. thesis, Stanford University.

Giannopoulos, Panos, and Remco C. Veltkamp (2000). "A Pseudo-Metric for Weighted Point Sets," in *Proceedings of the 7th European Conference on Computer Vision (ECCV)*, ed. A. Heyden, G. Sparr, M. Nielsen, and P. Johansen (Copenhagen: Springer Verlag), 715–730.

Hewlett, Walter B. (1992). "A Base-40 Number-Line Representation of Musical Pitch Notation," *Musikometrika* 4, 1–14. *http://www.ccarh.org/publications/reprints/base40/*

Howard, John (1998). "Strategies for Sorting Melodic Incipits," *Melodic Comparison: Concepts, Procedures, and Applications (Computing in Musicology 11)*, 119–128.

McKinney, Martin F., and Jeroen Breebaart (2003). "Features for Audio and Music Classification," in *ISMIR 2003: Proceedings of the Fourth International Conference on Music Information Retrieval*, ed. Holger H. Hoos and David Bainbridge (Baltimore: Johns Hopkins University), 151–158.

Pickens, Jeremy, Juan Pablo Bello, Giuliano Monti, Tim Crawford, Matthew Dovey, Mark Sandler, and Don Byrd (2002). "Polyphonic Score Retrieval Using Polyphonic Audio Queries: A Harmonic Modeling Approach," in *ISMIR 2002 Conference Proceedings: Third International Conference on Music Information Retrieval*, ed. Michael Fingerhut (Paris: IRCAM), 140–149.

Répertoire International des Sources Musicales (RISM) (2002). *Serie A/II, manuscript musicaux après 1600*. Munich: K. G. Saur Verlag.

Schlichte, Joachim (1990). "Der automatische Vergleich von 83,243 Musikincipits aus der RISM-Datenbank: Ergebnisse-Nutzen-Perspektiven," *Fontes artis musicae* 37, 35–46.

Typke, Rainer, Panos Giannopoulos, Remco C. Veltkamp, Frans Wiering, and René van Oostrum (2003). "Using Transportation Distances for Measuring Melodic Similarity" in *ISMIR 2003: Proceedings of the Fourth International Conference on Music Information Retrieval*, ed. Holger H. Hoos and David Bainbridge (Baltimore: Johns Hopkins University), 107–114.

Ukkonen, Esko, Kjell Lemström, and Veli Mäkinen (2003). "Geometric Algorithms for Transposition Invariant Content-Based Music Retrieval" in *ISMIR 2003: Proceedings of the Fourth International Conference on Music Information Retrieval*, ed. Holger H. Hoos and David Bainbridge (Baltimore: Johns Hopkins University), 193–199.

Appendix I. The RISM A/II Database

The RISM A/II Database, a work-in-progress for more than 50 years, attempts to provide bibliographical information on all music manuscripts, worldwide, for the seventeenth and eighteenth centuries. In some of the 60 countries which have collaborated to create this corpus of material, attention is now being devoted to manuscripts of the nineteenth century.

RISM (the *Répertoire International des Sources Musicales*) is headquartered in Frankfurt, Germany. The project referred to here (*http://rism.stub.uni-frankfurt.de/*) is currently directed by Klaus Keil. The designation A/II differentiates the music-manuscript project from a large number of text-based projects (printed music, music-theory treatises, et al.) also under RISM's jurisdiction.

Access to the RISM A/II Database *per se* is principally available via a CD-ROM (412,000 compositions with a total of 470,000 incipits) produced by K. G. Saur (Munich) and via library subscription (456,000 records) through NISC (Baltimore; statistics as reported in March 2004). Some national collections of RISM materials offer alternative means of access. For example, the musical incipits displayed for manuscripts in the Cambridge University Library are based on a separate implementation of *Plaine & Easie Code* (which is described in E. Selfridge-Field, *Beyond MIDI: The Handbook of Musical Codes*, The MIT Press, 1997).