Using the Shape of Music to Compute the Similarity between Symbolic Musical Pieces

Julián Urbano, Juan Lloréns, Jorge Morato and Sonia Sánchez-Cuadrado

University Carlos III of Madrid
Avda. Universidad, 30
28911 Leganés, Madrid, Spain
jurbano@inf.uc3m.es   llorens@inf.uc3m.es   jorge@ie.inf.uc3m.es   ssanchec@ie.inf.uc3m.es

Abstract. We describe a new approach for computing the similarity between symbolic musical pieces, based on the differences in shape between the interpolating curves defining the pieces. We outline several requirements for a symbolic musical similarity system, and then show how our model addresses them. Results obtained by an early prototype, tested with the MIREX 2005 Symbolic Music Similarity collections, show that this approach performs well. This new model opens a new and promising line for further research on the melodic similarity problem.

Keywords. Music information retrieval, melodic similarity, interpolation.

1   Introduction

The problem of Symbolic Melodic Similarity has been approached from very different points of view [1]. Some techniques are based on geometric representations of songs [2], [3], others rely on representations based on n-grams to calculate similarities [4], [5], and some others use editing distance algorithms, as well as local- and global-alignment algorithms [6].

Some of these techniques have proven effective for retrieving short incipits melodically similar to a query, given a reduced number of candidates. Evaluations in this way were carried out by the International Music Information Retrieval Systems Evaluation Laboratory (IMIRSEL) group in three editions of the Music Information Retrieval Evaluation eXchange (MIREX) [7], from 2005 to 2007, using ground truths based on partially ordered lists or similar [8].

We adapted the local alignment approach to work with n-grams instead of with single symbols, and the corresponding substitution score function between n-grams was also adapted to take into consideration a new geometric representation of musical sequences.

In the next section we outline several problems that a symbolic music retrieval system should address, and then we discuss the general approaches given in the literature to these requirements. Next, we introduce our representation model, which compares two musical pieces by their shape, and see how this model addresses the requirements discussed. We then give some early experimental results, and finish the paper with conclusions and lines for future research.

2 Melodic Similarity Requirements

Due to the nature of the information treated in Music Information Retrieval (MIR) [9], there are some requirements that have to be considered from the very beginning when devising a retrieval system. Byrd and Crawford [10] identified some requirements that they consider every MIR system should meet, such as the need of cross-voice matching, polyphonic queries or the clear necessity of taking into account both the horizontal and vertical dimensions of music. Selfridge-Field identified three elements that may confound both the users when they specify the query and the actual retrieval systems at the time of making comparisons. These elements are rests, repeated notes and grace notes [9]. Mongeau and Sankoff addressed repeated notes and refer to these situations as fragmentation and consolidation [11].

We list here some general requirements that should be common to any MIR system since we consider them as basic for the general user needs. The requirements are divided into two categories: vertical (i.e. pitch) and horizontal (i.e. time).

2.1 Vertical Requirements

Vertical requirements regard the pitch dimension of music: octave equivalence, degree equality, note equality, pitch variation, harmonic similarity and voice separation.

2.1.1 Octave Equivalence

When two pieces differ only in the octave they are written in, they should be considered as the same one in terms of melodic similarity. Such a case is shown in Fig. 1, with simple versions of the main riff in *Layla*, by *Dereck and the Dominos*.

Fig. 1. Octave equivalence

It has been pointed out that faculty or music students may want to retrieve pieces within some certain pitch range such as C5 up to F#3, or every work above A5 [12]. This type of information need should be easily handled with metadata or a simple traverse through the sequence. However, we argue that queries issued by users without such a musical background will be interested on the recognition of a certain pitch contour, and such cases are much more troublesome because some measure of melodic similarity has to be calculated.

2.1.2 Degree Equality

The score at the top of Fig. 1 shows a melody in the F major tonality, as well as the corresponding name and tonality-degree for each note. Beneath, Fig. 2 shows exactly the same melody shifted 7 semitones downwards to the B$^{\flat}$ major tonality.

Fig. 2. Degree equality

The tonality-degrees used in both cases are the same, but the resultant notes are not. One would consider the second melody a version of the first one, because they are the same in terms of pitch contour. Therefore, they should be considered as the same by a retrieval system, which should also consider the possibility of a key signature change somewhere throughout the song.

2.1.3 Note Equality

We could also consider the case where exactly the same melodies, with exactly the same notes, are written in different tonalities and, therefore, each note corresponds to a different tonality-degree in each case. Fig. 3 shows such a case, with the same melody as in Fig. 1, but in the C major tonality.
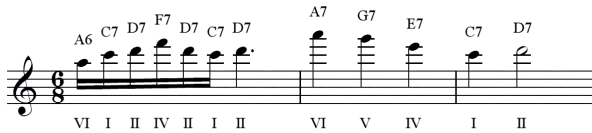

Fig. 3. Note equality

Although the degrees do not correspond one to each other, the actual notes do, so both pieces should be considered the same one in terms of melodic similarity.

2.1.4 Pitch Variation

Sometimes, a melody is altered by changing only the pitch of few particular notes. For instance, the first melody in Fig. 1 might be changed by shifting the last note from D7 to A6 (which actually happens in the original song). Such a change should not make a retrieval system disregard that result, but simply rank it lower, after the exactly-equal ones. Thus, the retrieval process should not consider only exact matching, where the query is part of a piece in the repository (or the other way around). Approximate matching, where documents can be considered somewhat similar to a query, should be the way to go.

2.1.5 Harmonic Similarity

Another desired feature would be to match harmonic pieces, both with harmonic and melodic counterparts. For instance, in a triad chord (made up by the root note and its third and fifth intervals), one might recognize only two notes (typically the root and the fifth). However, some other might recognize the root and the third, or just the root, or even consider them as part of a 4-note chord. Fig. 4 shows the same piece as in the

top of Fig. 1, but with some notes added to make the song more harmonic. These two pieces have basically the same pitch progression, but with some ornamentation, and they should be regarded as very similar by a retrieval system.
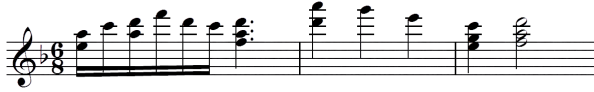

Fig. 4. Harmonic similarity

Thus, a system should be able to compare harmony wholly and partially, considering again the Pitch Variation problem as a basis to establish differences between songs.

2.1.6 Voice Separation
Fig. 5 depicts a piano piece with 3 voices, which work together as a whole, but can also be treated individually. Indeed, if this piece is played with a flute only one voice could be performed, even if some streaming effect might be produced by changing tempo and timbre so that two voices might be perceived by a listener [13].
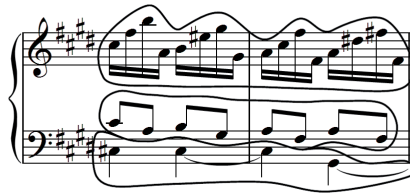

Fig. 5. Voice separation

Therefore, a query containing only one voice should match with this piece in case that voice is similar enough to any of the three marked in the figure.

2.2 Horizontal Requirements

Horizontal requirements regard the time dimension of music: time signature equivalence, tempo equivalence, duration equality and duration variation.

2.2.1 Time Signature Equivalence
The top of Fig. 6 depicts a simplified version for the beginning of *op. 81 no. 10* by *S. Heller,* with its original 2/4 time signature. If a 4/4 time signature is considered, like in the bottom of Fig. 6, the piece would be split into bars of duration 4 crotchets each.

Fig. 6. Time signature equivalence

The only difference between these two pieces is actually how intense each note should be played. However, they are in essence the same piece, and no regular listener would tell the difference. Therefore, we think the time signature should not be considered when comparing musical performances in terms of melodic similarity.

2.2.2 Tempo Equivalence

For most people, the piece at the top of Fig. 6, with a tempo of 112 crotchets per minute, would sound like the one at the top of Fig. 7, where notes have twice the length. However, the whole staff is played twice as fast, at 224 crotchets per minute, resulting in the same actual time.

Fig. 7. Tempo equivalence

On the other hand, it might be considered a tempo of 56 crotchets per minute and notes with half the duration, like the bottom of Fig. 7 depicts. Therefore, actual note lengths cannot be considered as the only horizontal measure, because these three pieces will sound the same to any listener.

2.2.3 Duration Equality

If the melody at the top of Fig. 6 were played slower or quicker by means of a tempo variation, but maintaining the rhythm, an example of the result would be like the score in Fig. 8.

Fig. 8. Duration equality

Even though the melodic perception does actually change, the rhythm does not, and neither does the pitch contour. Therefore, they should be considered as virtually the same, maybe with some degree of dissimilarity based on the tempo variation.

2.2.4 Duration Variation

As with the Pitch Variation problem, sometimes a melody is altered by changing only the rhythm of a few notes. For instance, the melody in Fig. 9 maintains the same pitch contour as in Fig. 6, but changes the duration of some notes.

Fig. 9. Duration variation

Moreover, the tempo can change somewhere in the middle of the melody, and therefore change the actual time of each note afterwards. Variations like these are common and they should be considered as well, just like the Pitch Variation problem, allowing approximate matches instead of just exact ones.

3   General Solutions to the Requirements

Most of these problems have been already addressed by the research community, although not all together in the same system. Next, we show the most used solutions.

3.1   Vertical Requirements

The immediate solution to the Octave Equivalence problem is to consider octave numbers with their relative variation within the piece. Surely, it is not the same a progression from G5 to C6 than a progression from G5 to C5. For the Degree Equality problem it seems to be clear that tonality degrees must be used, rather than actual pitch values, in order to compare two melodies. However, the Note Equality problem suggests the opposite.

The solution so far for these three vertical problems seems to be the use of relative pitch differences as the units for the comparison, instead of the actual pitch or degree values. A good discussion for modeling pitch can be found in [6] and [5].

3.2   Horizontal Requirements

Although the time signature of a performance is worth for other purposes such as pattern search, it seems to us that it should not be considered at all when comparing two pieces melodically.

According to the Tempo Equivalence problem, actual time should be considered rather than score time, since it would be probably easier for a regular user to provide actual rhythm information. On the other hand, the Duration Equality problem requires the score time to be used instead. We consider that both measures have to be taken into account. The actual time is valuable for most users without a musical background, whilst the score time might be more valuable for people who do have it.

However, when facing the Duration Variation problem it seems necessary to use some sort of timeless model. The solution could be to compare both actual and score time, or to use relative differences between notes, in this case with the ratio between two notes' durations. Besides others, this approach has been show in [5] and [6].

4   A Model based on Interpolation

We developed a geometric model that represents musical pieces with curves. Each note is placed in a plane with positions relative to their pitch and duration differences.

Then, we define the curve $C(t)$ as the interpolating curve passing through each note (see Fig. 10). Should the song have multiple voices, each one would be placed in a different pitch-time plane, sharing the same time dimension, but with a different curve $C_i(t)$ (where the subscript i indicates the voice number).
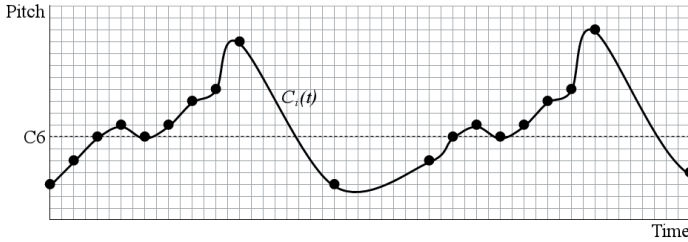

Fig. 10. Melody represented as a curve in a pitch-time plane

With this representation, the similarity between two songs could be thought of as the similarity in shape between the two curves they define. Thus, every vertical requirement identified in section 2 would be met with this representation: a song with an octave shift will keep the same shape, if the tonality changes the shape of the curve is not affected either, and if the notes remain the same after a tonality change, so does the curve. The Pitch Variation problem can be easily addressed analytically, and different voices can be compared individually in the same way because they are in different planes.

Same thing happens with the horizontal requirements: the Tempo Equivalence and Duration Equality problems can be solved analytically, because they just imply a linear transformation in the time dimension. For example, if the melody at the top of Fig. 6 is defined with curve $C(t)$ and the one at the top of Fig. 7 is denoted with curve $D(t)$, it can be easily proved that $C(2 \cdot t) = D(t)$. Moreover, the Duration Variation problem could be addressed analytically as the Pitch Variation problem, and the Time Signature Equivalence problem is not an issue because the shape of the curve is independent of the time signature.

4.1 Measuring Dissimilarity with the Change in Shape

Having musical pieces represented this way, each one of them could be defined with a polynomial of the form $C(t) = a_n t^n + a_{n-1} t^{n-1} + \ldots + a_1 t + a_0$. The first derivative of these curves measures how much the songs are changing at a particular point in time (i.e. how the shape changes). To measure the change of one song with respect to another, the area between the first derivatives could be used.

Note that a pitch transposition would mean just a change in the a_0 term. As it turns out, when calculating the first derivative of the curves this term disappears, which is why the vertical requirements are met: transpositions in pitch are not reflected in the shape of the curve, so they are not reflected in the first derivative either.

Therefore, the song is actually defined by the first derivative of its interpolating curve, $C'(t)$. The dissimilarity between two songs, say $C'(t)$ and $D'(t)$, would be defined as the area between them (measured as the integral of the absolute of their difference):

$$diff(C, D) = \int |C'(t)\text{-}D'(t)|dt \tag{1}$$

This dissimilarity measure turns out to be a metric, because it has the following properties:

- Non-negativity, $diff(C, D) \geq 0$: because of the absolute value, there will never be a negative distance.
- Identity of indiscernibles, $diff(C, D) = 0 \Leftrightarrow C = D$: because calculating the absolute value, the only way to have no difference is with the same curve.
- Symmetry, $diff(C, D) = diff(D, C)$: because we are integrating the absolute value of their difference.
- Triangle inequality, $diff(C, E) \leq diff(C, D) + diff(D, E)$:

$$\int |C(t)\text{-}E(t)|\,dt \leq \int |C(t)\text{-}D(t)|\,dt + \int |D(t)\text{-}E(t)|\,dt$$

$$\int |C(t)\text{-}D(t)|\,dt + \int |D(t)\text{-}E(t)|\,dt = \int |C(t)\text{-}D(t)| + |D(t)\text{-}E(t)|\,dt$$

$$\int |C(t)\text{-}D(t)| + |D(t)\text{-}E(t)|\,dt \geq \int |C(t)\text{-}E(t)|\,dt$$

Therefore, many indexing and retrieval techniques, like vantage objects [14], could be exploited if using this metric.

4.2 Interpolation with Splines

The next issue to address is the interpolation method to use. The standard Lagrange interpolation method, though simple, is known to suffer the Runge's Phenomenon [15]. As the number of points increases, the interpolating curve wiggles a lot, which is not acceptable because it does not maintain the same shape as the melody. Moreover, a very small difference in one of the points could translate into an extreme variation in the overall curve, which would make virtually impossible to handle the Pitch and Duration Variation problems (see top of Fig. 11).
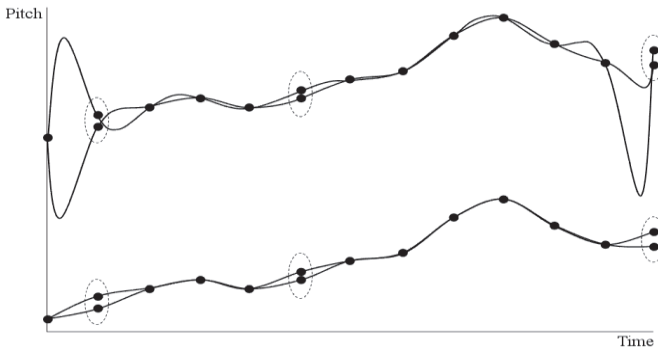

Fig. 11. Runge's Phenomenon

Instead, splines should be used, because they are usually easy to calculate too and they do not suffer the Runge's Phenomenon (see bottom of Fig. 11). Besides, splines

are defined as piece-wise functions, which come in handy when addressing the horizontal requirements. We saw above that the horizontal problems could be solved, as they implied just a linear transformation of the form $D(t) \Rightarrow D(k{\cdot}t)$ in one of the curves. However, the calculation of the term k is anything but straightforward, and the transformation would apply to the whole curve and so would not allow us to measure differences for the Duration Variation problem. The solution would be to split the curve into spans, and define it as

$$C_i(t) = \begin{cases} c_{i,0}(t) & t_{i,0} \le t \le t_{i,1} \\ c_{i,1}(t) & t_{i,1} \le t \le t_{i,2} \\ \vdots & \vdots \\ c_{i,n_i-1}(t) & t_{i,n_i-1} \le t \le t_{i,n_i} \end{cases} \tag{2}$$

where $t_{i,j}$ denotes the onset time of the note $m_{i,j}$ (the j-th note of the i-th voice), and n_i is the length of the i-th voice. With this representation, linear transformations would be applied only to a single span without affecting the whole curve. Moreover, the duration of the spans could be normalized from 0 to 1, making it easy to calculate the term k.

Most spline interpolation methods define the curve in parametric form (i.e. with one function per dimension). In this case, it results in one function for the pitch and one function for the time. This means that the two musical dimensions could be compared separately, giving more weight to one or the other. Therefore, the dissimilarity between two spans $c(t)$ and $d(t)$ would be the sum of the pitch and time dissimilarities as measured by (1):

$$diff(c,d) = K_{pitch} \cdot diff\left(c_{pitch}(t), d_{pitch}(t)\right) + K_{time} \cdot diff\left(c_{time}(t), d_{time}(t)\right) \tag{3}$$

where $c_{pitch}(t)$ and $c_{time}(t)$ are the corresponding parametric functions for pitch and time, respectively, and K_{pitch} and K_{time} are fine tuning constants. Different works have shown that pitch is much more important than time for comparing melodic similarity [5], [6], so more weight should be given to K_{pitch}.

5   Implementation and Experimental Results

Various approaches for melodic similarity have used a textual representation of musical sequences to which an editing distance algorithm was applied in order to calculate the dissimilarity between two pieces. This representation has been improved in recent years, and sequence alignment algorithms have proved to perform better than editing distance algorithms [6] [16].

To practically apply our model, we followed a basic n-gram approach. We used the Smith-Waterman local alignment algorithm [17], with the two sequences of spans as input, defined as in (2). Therefore, the input symbols to the Smith-Waterman algorithm are actually the pitch and time functions of a span, which we consider an n-gram. The function used to calculate the substitution score between two symbols in the sequences (two n-grams in our case), is the function defined in (3), which

measures the shape similarity of the two spans. We used Uniform B-Splines of degree 3, which need spans of 4 notes. Therefore, the n-gram length is 4, noted in [5] as an effective length.

The n-grams were modeled as follows. The pitch of each note was represented as the relative difference to the pitch of the first note, and the duration was represented as the ratio to the duration of the first note as well. For example, a span with absolute pitches $\langle 74, 81, 72, 76 \rangle$ and durations $\langle 240, 480, 240, 720 \rangle$ milliseconds, would be modeled as $\langle 7, -2, 2 \rangle$ in relative pitch and $\langle 2, 1, 3 \rangle$ in relative duration. Note that the first note is removed as the relative difference with itself would be the unit.

We tested our approach with the training and evaluation test collections of the Symbolic Melodic Similarity task of MIREX 2005 [8], measuring the Average Dynamic Recall score [18]. Both collections consist of about 580 incipits and 11 queries each, with their corresponding ground truths.

5.1 Evaluation with MIREX 2005 Test Collections

Results in [5] and [6] suggest the use of both the pitch and time dimensions in the n-gram representation. More specifically, $K_{pitch}=0.75$ and $K_{time}=0.25$, ignoring rests. We tested first considering only pitch, and then pitch and duration. We checked two more possibilities: retrieve the relevant results from all the about 580 incipits, and also assume there is a fictional previous filter which makes us retrieve relevant documents only from the about 50 incipits that were manually-ranked when constructing the ground truth. We did this to examine how our method works when retrieving and ranking documents, or just ranking. Table 1 outlines the results.

Table 1. ADR scores for the Splines method proposed, with and without filter, for the MIREX 2005 training and evaluation test collections.

Tunning	Collection	Filter	Avg.	Min	Max
$K_{pitch}=1, K_{time}=0$	Training	No	0.639	0.271	0.864
		Yes	0.837	0.696	0.963
	Evaluation	No	0.709	0.314	0.911
		Yes	0.795	0.383	0.911
$K_{pitch}=0.75, K_{time}=0.25$	Training	No	0.643	0.312	0.864
		Yes	0.841	0.724	0.963
	Evaluation	No	0.710	0.314	0.911
		Yes	0.809	0.520	0.919

We found the improvement of considering the time dimension completely incidental. Only one query benefited when modeling time, and the average improvements were under 0.02 in the ADR score in all four cases. The average ADR score for the test collection was 0.643, and for the evaluation collection it was 0.710.

Table 2 shows the detailed results of our approach for the evaluation test collection of MIREX 2005, along with the results of the seven systems that participated in that edition. We can see that our model seems to perform better than all systems, obtaining the best score for 5 of the 11 queries. On average, the best system obtained an ADR score of 0.660, which is outperformed by our approach, with 0.710. However, the

difference was not found to be significant between our approach and the top 3 systems, according to 1-tailed Mann-Whitney U tests.

Table 2. ADR scores for the Splines method proposed (without filter), compared with the systems that participated in MIREX 2005 with the evaluation test collection. GAM = Grachten, Arcos and Mántaras; O = Orio; US = Uitdenbogerd and Suyoto; TWV = Typke, Wiering and Veltkamp; L(P3) = Lemström (P3), L(DP) = Lemström (DP); FM = Frieler and Müllensiefen. Best scores per query appear in bold face. Best scores per system appear in italics. * for significant difference at the 0.10 level, ** at the 0.05 level and *** at the 0.01 level.

Query	Splines	GAM	O	US	TWV	L(P3)	L(DP)	FM
190.011.224-1.1.1	0.803	0.820	0.717	**0.824**	0.538	0.455	0.547	0.443
400.065.784-1.1.1	**0.879**	0.846	0.619	0.624	*0.861*	0.614	*0.839*	0.679
450.024.802-1.1.1	**0.722**	0.450	0.554	0.340	0.554	0.340	0.340	0.340
600.053.475-1.1.1	*0.911*	*0.883*	*0.911*	*0.911*	0.725	0.661	0.650	0.567
600.053.481-1.1.1	**0.630**	0.293	0.629	0.486	0.293	0.357	0.293	0.519
600.054.278-1.1.1	0.810	0.674	0.785	**0.864**	0.731	0.660	0.527	0.418
600.192.742-1.1.1	0.703	**0.808**	**0.808**	0.703	**0.808**	0.642	0.642	*0.808*
700.010.059-1.1.2	0.521	0.521	0.521	0.521	0.521	**0.667**	0.521	0.521
700.010.591-1.4.2	0.314	**0.665**	0.314	0.314	0.314	0.474	0.314	0.375
702.001.406-1.1.1	0.689	0.566	**0.874**	0.675	0.387	*0.722*	0.606	0.469
703.001.021-1.1.1	**0.826**	0.730	0.412	0.799	0.548	0.549	0.692	0.561
Average	**0.710**	0.660	0.650	0.642	0.571*	0.558***	0.543**	0.518***

6   Conclusions and Future Work

We have proposed a new approach for modeling musical pieces and compute their melodic similarity. We consider songs as curves in the time-pitch space, which allows us to compute the melodic similarity between two songs by the similarity in shape of the curves they define. We have implemented this model using Uniform B-Splines and the Smith-Waterman alignment algorithm. This approach is shown to perform at least as well as the systems that participated in the MIREX 2005 edition of the Symbolic Melodic Similarity task.

These results indicate that this line of work is promising. Future research should still address several points. The alignment algorithm to use, n-gram representation and length, interpolation method, weights for the different tuning parameters, take rests into consideration and normalization of the incipits lengths are the most immediate points for further research. Also, this model should be extended to allow harmony and polyphony, as the results given in this paper concern only melodic pieces.

Acknowledgements

References

1. Typke, R., Wiering, F., Veltkamp, R.C.: A Survey of Music Information Retrieval Systems. International Conference on Music Information Retrieval. pp. 153-160 (2005).
2. Ukkonen, E., Lemström, K., Mäkinen, V.: Geometric Algorithms for Transposition Invariant Content-Based Music Retrieval. International Conference on Music Information Retrieval. pp. 193-199 (2003).
3. Typke, R., Veltkamp, R.C., Wiering, F.: Searching Notated Polyphonic Music Using Transportation Distances. ACM International Conference on Multimedia. pp. 128-135 (2004).
4. Uitdenbogerd, A., Zobel, J.: Melodic Matching Techniques for Large Music Databases. ACM International Conference on Multimedia. pp. 57-66 (1999).
5. Doraisamy, S., Rüger, S.: Robust Polyphonic Music Retrieval with N-grams. Journal of Intelligent Systems. 21, 53-70 (2003).
6. Hanna, P., Ferraro, P., Robine, M.: On Optimizing the Editing Algorithms for Evaluating Similarity Between Monophonic Musical Sequences. Journal of New Music Research. 36, 267-279 (2007).
7. Downie, J.S., Ehmann, A.F., Bay, M., Jones, M.C.: The Music Information Retrieval Evaluation eXchange: Some Observations and Insights. In: Zbigniew, W.R. and Wieczorkowska, A.A. Advances in Music Information Retrieval. pp. 93-115 Springer (2010).
8. Typke, R., den Hoed, M., de Nooijer, J., Wiering, F., Veltkamp, R.C.: A Ground Truth For Half A Million Musical Incipits. Journal of Digital Information Management. 3, 34-39 (2005).
9. Selfridge-Field, E.: Conceptual and Representational Issues in Melodic Comparison. Computing in Musicology. 11, 3-64 (1998).
10. Byrd, D., Crawford, T.: Problems of Music Information Retrieval in the Real World. Information Processing & Management. 38, 249-272 (2002).
11. Mongeau, M., Sankoff, D.: Comparison of Musical Sequences. Computers and the Humanities. 24, 161-175 (1990).
12. Isaacson, E.: Music IR for Music Theory. The MIR/MDL Evaluation Project White paper Collection, 2nd. edition. pp. 23-26 (2002).
13. McAdams, S., Bregman, A.S.: Hearing Musical Streams. In: Roads, C. and Strawn, J. Foundations of Computer Music. pp. 658-598 The MIT Press (1985).
14. Bozkaya, T., Ozsoyoglu, M.: Indexing Large Metric Spaces for Similarity Search Queries. ACM Transactions on Database Systems. 24, 361-404 (1999).
15. de Boor, C.: A Practical guide to Splines. Springer (2001).
16. Hanna, P., Robine, M., Ferraro, P., Allali, J.: Improvements of Alignment Algorithms for Polyphonic Music Retrieval. International Conference on Computer Music Modeling and Retrieval. pp. 244-251 (2008).
17. Smith, T.F., Waterman, M.S.: Identification of Common Molecular Subsequences. Journal of Molecular Biology. 147, 195-197 (1981).
18. Typke, R., Veltkamp, R.C., Wiering, F.: A Measure for Evaluating Retrieval Techniques based on Partially Ordered Ground Truth Lists. IEEE International Conference on Multimedia and Expo. pp. 1793-1796 (2006).