

Melodic Pattern Anchoring for Score Following

Using Score Analysis

Özgür İzmirli Robert Seward Noel Zahler

Center for Arts and Technology – Connecticut College
email: nbzah@conncoll.edu

Abstract

Building on our previous work in score following, we suggest that research on musical pattern significance, representation and categorization can be usefully integrated into a score follower to automatically identify unique melodic signatures in a composition. These signatures may then be calculated and analyzed over the entirety of a composition providing anchor points. Anchor points replace what has been, in practice, an arbitrary segmentation of scores with a unique division of a composition based on information that is essential to the operation of a score follower. The machine's understanding of the entire score is enhanced and our algorithm's performance is refined.

1 Introduction

The quest to create an accurate and versatile synthetic performer has continued for almost twenty years. There is little disagreement on what such a cyber-performer should be able to do but most approaches in building such a player have only approximated very small parts of the body of knowledge that human individuals bring to their task in performance. There are exceptions, we have seen implementations of “rehearsing the synthetic performer” (Vercoe and Puckette 1985), a “pre-performance score consultation” (Baird, Blevins and Zahler 1993) and the hidden markov model (Orio 2001, Raphael 2001). All these researchers have attempted to give the machine a broader knowledge of the score and an understanding of the expectations a performer has. To date, none of these attempts has established a standard and we do not pretend to suggest one here. We do, however, continue in our efforts to give our own program (Izmirli, Seward and Zahler 2002) as much knowledge of the score as we can. To this end, we add to our previous work on Noel Zahler's Concerto for Clarinet Chamber Orchestra and Interactive Computer by applying concepts of those involved in the analysis of musical surfaces to our synthetic performer program.

At the micro level score followers tend to be short-sighted and lack awareness of order and

position of larger scale structure in a musical composition with respect to the current predicted position they occupy. This weakness may cause a score follower to drift away, losing synchronization due to local position ambiguity either as a result of input mismatch or repetitive self-similarity in the score. In this paper, we present a new method that deals with larger scale awareness in musical pieces during operation of a score follower in real time. The method involves a pre-performance score analysis process during which significant and relatively unique melodic patterns are determined. These patterns are then set as anchor points that are used to guide, orient and affirm the operation of a score follower at a higher level.

An anchor is defined as a melodic pattern within a given fixed-length window that has low resemblance to any other pattern in the score. Ideally, anchors should be individually unique and evenly spaced in a piece. Although, in general, it is not possible to find anchors that are completely orthogonal to each other, an approximate solution provides useful cues for our problem. The aim of this paper is to outline the method of determination of unique note patterns suitable to be used as anchors. Part of the motivation for this approach has arisen from the fact that score sectioning has been common practice in most score following performances. Sectioning is the process of breaking the score apart into smaller sections and enabling manual cueing in terms of these sections during performance. The presented approach is an attempt to replace what has been a fairly arbitrary partitioning of the score with a method that is relevant to the score follower.

Research on musical pattern significance, representation and categorization (Höthker, Hörnel and Anagnostopoulou 2001,) motif clustering (Cambouropoulos and Widmer 2000,) theme extraction (Meek and Birmingham 2001) and melodic similarity (Hu, Dannenberg and Lewis 2002) has concentrated on significance of patterns from a musical salience point of view. The difference between these approaches and the one presented in this paper is that we treat note combinations as patterns without seeking to musically segment them.

Hence, any pattern that bears discriminative information for our purposes is utilized.

2 Pattern Similarity

A distance measure is employed to quantify the degree of similarity between two windows containing pitch patterns. The method, to be explained in the next section, entails a systematic calculation of distances between windows taken from different sections of the score. A window is characterized by its position in the score. We denote the window as $W(t)$ where t is the relative score time that corresponds to the end of the window. The score time is represented in quantized form with approximately 1000 points per quarter note. The duration of the window is fixed and given by t_w .

In order to calculate the distance between two windows, they are first aligned and then logically ANDed according to their pitch content. That is, there is a match in a region only if both windows contain the same note in their corresponding regions. The similarity is defined as the ratio of the total duration of the regions that contain a match to the duration of the entire window. We denote the similarity between two given windows ending at score times t_0 and t_1 respectively as $S(t_0, t_1)$. Figure 1 demonstrates this calculation.

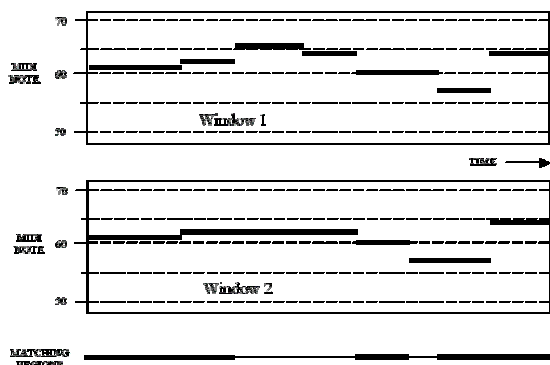


Figure 1. Demonstration of the AND operation to determine the amount of similarity between contents of two windows.

3 Anchors as Signatures

Given any window $W(t_0)$ in the score, the problem of finding other similar note patterns is solved by running the window through the score which corresponds to an operation similar to cross-correlation. This analysis of the score with respect to the window $W(t_0)$ reveals the degree of uniqueness of the melodic pattern in the window. The similarity for a window at t_0 is given by the following:

$$S(t_0, t) \begin{cases} 0 \leq t < t_0 - t_w \\ t_0 + t_w < t \leq t_{\max} \end{cases}$$

Figures 2 and 3 show examples of the similarity for two different windows. It can be clearly seen that in Figure 2 the window has many patterns in other parts of the score that match exactly (with similarity 1.0,) whereas, in Figure 3 there are no patterns that are the same, and furthermore the overall similarity tends to be relatively low. Our aim is to find a collection of patterns in which each one is maximally dissimilar to all other patterns in the collection. For this purpose we define two measures to summarize the degree of uniqueness of a pattern. The first one is the maximum of the similarity:

$$T(t_0) = \max_t (S(t_0, t))$$

The second measure indicates the mean resemblance. It is given in discrete form as the underlying representation is discrete and has a fixed time resolution as mentioned above.

$$M(t_0) = \frac{1}{t_{\max}} \sum_{t=1}^{t_{\max}} S(t_0, t)$$

The maximum similarity and mean resemblance are used as estimates of the degree of uniqueness. When T is close to 1 we know that there is at least one other pattern that is very similar. But, this alone is not informative about the number of similar patterns. M is a measure of overall similarity which is used to differentiate between few or many similar patterns. When T is close to 1, M tells us whether the high similarity is commonly encountered or not. If M is low, then it means that the pattern contained in the window being analyzed is rare in this piece. On the other hand, if M is high, it means the pattern is common in the piece and it would not qualify as a signature. In general, T and M have to be closer to 0 for the pattern to be considered unique.

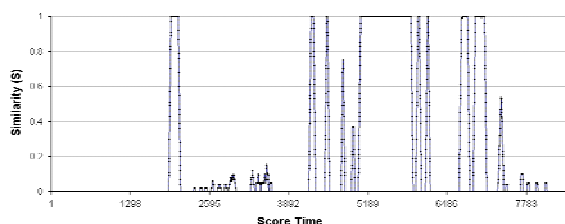


Figure 2. The window has many patterns in other parts of the score that match exactly.

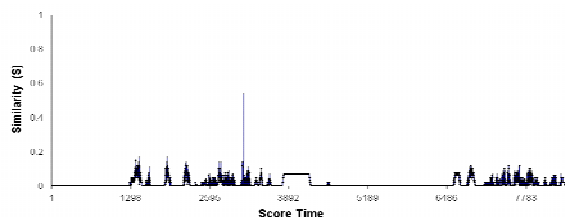


Figure 3. None of the patterns that are the same, and the overall similarity tends to be relatively low.

In order to determine unique patterns, the similarity is calculated for all patterns at note onset boundaries t_i for all notes: (note that although t_i is event based, t has much finer resolution and can be viewed as a continuous slide over the entire score)

$$S(t_i, t) \begin{cases} 0 \leq t < t_i - t_w \\ t_i + t_w < t \leq t_{\max} \\ i = 1.. \max_notes \end{cases}$$

Figures 4 and 5 show the T and M values obtained for the entire score. T and M are plotted as functions of events because they are evaluated at note boundaries. Each value, however, is obtained by continuously sliding the window across the score. We define a combined measure X as the average of T and M. We then perform our search for anchor points on X. The task of finding anchor points is constrained by the number, spacing and uniqueness of the chosen patterns. We formulate the problem on the event axis rather than on the time axis. This means we will have anchors closer together in time where the density of notes is higher.

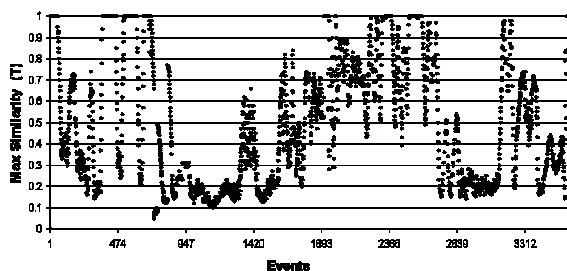


Figure 4. Maximum similarity, T, calculated for the entire score.

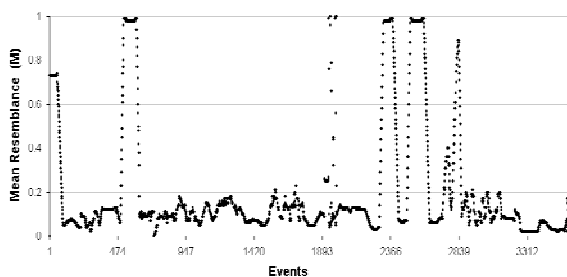


Figure 5. Mean resemblance, M, calculated for the entire score.

Given that we would like to find P anchors in a piece, and assuming that a solution exists for the given sequence X, the problem is to minimize the following expression:

$$\sum_{j=1}^P X(A(j))$$

with respect to $K < |A(j)-A(j+1)| < 3K$ for $j=1..P-1$ and $A(j) < H$ for $j=1..P$. $A(j)$ is the j 'th anchor in terms of event number, K is the minimum inter-anchor distance, again in units of event, and H is the threshold for anchor eligibility. It should be noted that K and P are interdependent.

A fast but approximate algorithm for finding the anchor points using X is given below:

1. $N = \text{event } 1$;
2. Move N , $2K$ events forward. Exit if at end of event list;
3. Find $x = \text{Min } X$ and the corresponding event A in the range $N-K$ to $N+K$;
4. If $x < H$ add A to the anchor list. $N = A$;
5. Goto 2

Figure 6 shows X and the anchor points (shown as triangles on the x-axis) found by the above algorithm. The first four anchors are shown in Figure 7.

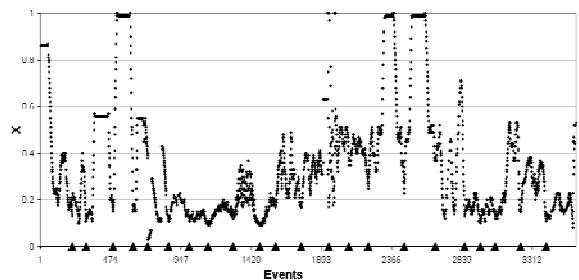


Figure 6. Average of T and M and anchor points found for $K=75$ and $H=0.4$. The anchor points are the triangles on the horizontal axis.

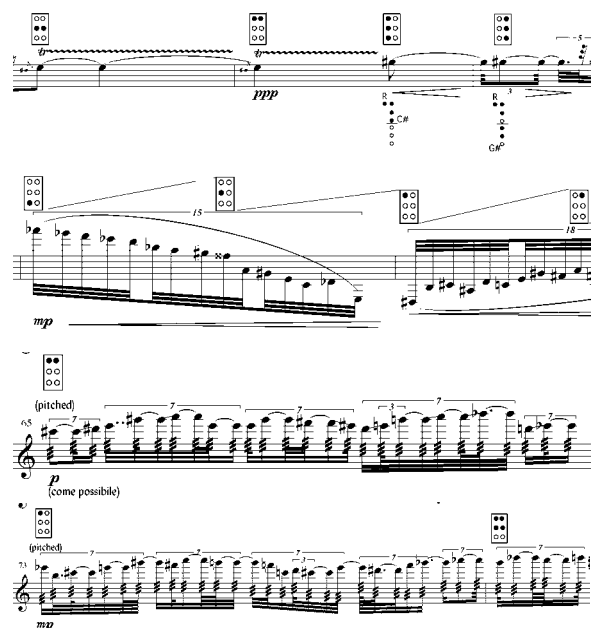


Figure 7. The first four anchors found (G clef.)

4 System Integration

Although the presented method may be integrated with other score following systems, here, we will discuss its integration with our system (Izmirlı, Seward and Zahler 2002.) The score analysis part of the method is implemented following the description given in this paper as an offline module. The output of this module, being a list of anchor points, is used in the real-time score following module to signal the anchor point encounters. The score follower adapts or corrects its operation based on these signals. Once an anchor point is crossed, no local search is carried out beyond that point into the past.

5 Evaluation

To date, we have had the opportunity to perform the Zahler concerto twice in concert. If we count the associated rehearsals, our score follower has executed the composition some eight different times in its entirety. Figures 8 and 9 compare two different performances of the same fragment of the Concerto. The graphs depict the program's analysis of incoming information from the live performer and its estimation of the position in the score on the x axis (SF Response Number) and the program's reaction to that information (spatialization, SF Output) on the y axis. A perfect performance would be evidenced by a straight diagonal from the lower left hand corner to the upper right hand corner of the grid.

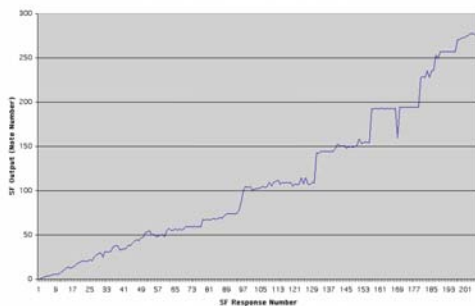


Figure 8. Tracking without program enhancement.

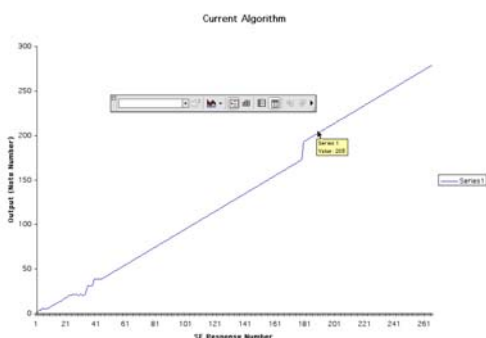


Figure 9. Tracking with program enhancement.

The improvement in the performance should be immediately perceptible.

6 Conclusions and Future Work

We have significantly enhanced the performance of our score follower while simultaneously defining and automating an important structural part of the process for presenting data to the system, score segmentation. We believe that it is possible and preferable to continue to enhance our program's knowledge of the score. While we have been successful in defining historically problematic gestures in score following (repeated notes, trills, fast passages) we would like to have the program automatically recognize these modes of performance and go further by incorporating the capability of capturing polyphonic music in the form of multiphonics. Further, we still need to give the program rehearsal memory and learning, as well as adapt it for accompaniment purposes.

References

- Baird, B., D. Blevins, and N. Zahler. 1993. "The Artificial Intelligence and Music: Implementing an Interactive Computer Performer." *Computer Music Journal* 17.2: pp. 73-79.
- Cambouropoulos, E., Widmer, G. 2000. "Automatic Motivic Analysis via Melodic Clustering." *Journal of New Music Research* 29(4), 303-317.
- Höthker, K., Hörnel, D., Anagnostopoulou, C. 2001. "Investigating the Influence of Representations and Algorithms in Music Classification." Rolland, P.Y., Cambouropoulos, E., Wiggins, G. (editors). *Pattern Processing in Music Analysis and Creation. Special Issue of Journal: Computers and the Humanities* (Kluwer Academic Publishers) Vol. 35, pp. 65-79.
- Hu, N., Dannenberg, R., Lewis, A.L. 2002. "A Probabilistic Model of Melodic Similarity." *Proceedings of the International Computer Music Conference (ICMC2002)*, September 17-21, 2002, Gothenburg, Sweden.
- Izmirlı, O., Seward, R and Zahler, N. 2002 "Compositional Imperatives for Implementing an audio Alignment Program in MAX/MSP", *Proceedings of the International Computer Music Conference (ICMC2002)*, September 17-21, 2002, Gothenburg, Sweden.
- Meeck, C., Birmingham, W.P. 2001. "Thematic Extractor." *Second International Symposium on Music Information Retrieval*. Bloomington, Indiana University, pp. 211-218.
- Orio, Nicola. 2001. "An Automatic Accompanist Based on Hidden Markov Models." *AI*IA 2001, LNAI 2175*, pp. 64-69.
- Raphael, C. 2001. "Music Plus One: A System for Flexible and Expressive Musical Accompaniment." *Proceedings of the Int. Computer Music Conference Havana, Cuba*.
- Vercore, B. and Puckette, M. 1985. "Synthetic Rehearsal: Training the Synthetic Performer." *Proceedings of the 1985 ICMC*. San Francisco: International Computer Music Association, pp. 275-289.