Music Perception and the Computer Aided Analysis of Music

Michael Schutz, University of Virginia
schutz@virginia.edu

Introduction

The increasingly mathematical nature of contemporary analytical techniques have inspired many creative applications of computing to the analysis of music (Cope, 1996; Castine, 1994). In an effort to incorporate interpretive decisions on the part of the performer, some have advocated basing analyses on digital audio files (Dannenberg, 2002). Others such as HUMDRUM are based on abstract representations of music (Huron, 1994). Fully realizing the potential of computer assistance in the analytical process requires finding an appropriate balance between the power of abstract representations and the ability to represent expressive decisions made by musicians during actual performance. Developed specifically for this project, the VirtualScore (Figure 1b) was designed to meet these requirements by allowing for the ability to represent both interpretive decisions and literal information in an abstract manner (Schutz, 2002).

Figure 1a Traditional Notation of a Musical Excerpt

Figure 1b VirtualScore Representation of Excerpt

The VirtualScore format is modeled on a piano-roll style of representation in which vertical position corresponds to pitch height, and horizontal position to time. Start and end times for a given note can be based on actual performance practice.

Each note in the VirtualScore is an independent object, able to represent any number of musical properties (e.g. volume, timbre, etc). Note onset and offset values can be taken from an actual performance, permitting rhythmic flexibility and allowing expressive variations of note length. In the performance captured by the VirtualScore in Figure 1b, the notes D3 (beat two) and C3 (beat three) were performed shorter than notated; the notes D3 and F3 on the "and" of four were delayed. This approach offers all the computational flexibility of abstract representation while preserving aspects of expressive performance.

The following two analyses were designed as a proof of concept and are based on VirtualScores derived from purely abstract data (e.g. without expressive information). While they were successful in demonstrating the potential of this analytical approach, the unexpected volume of data generated by these analyses raised complex issues of saliency and segmentation that must be addressed before moving on to more advanced analyses taking full advantage of the VirtualScore's abilitiy to represent expressive musical information.

Demonstration 1: Arnold Schonberg Piano Piece, Op 11. No. 1

An initial search identified 1605 instances of the interval class vector 101100, corresponding to trichord (014). Note groupings were limited to those notes which occurred within one musical beat. As the (014) trichord was identified as the basis for pitch content in Op 11. No. 1 by Straus (pp 40-41), it was expected to be discovered frequently.

However, the results of an exhaustive search of all trichords identified 10968 groupings, showing that while (014) was the most prominently used, several others were observed nearly as frequently. These findings are summarized graphically in Figure 2.

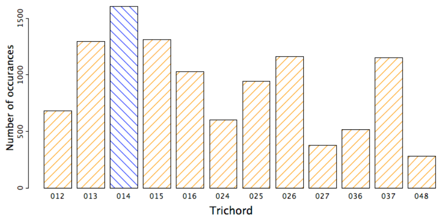

Figure 2 Observed Trichords

While the IC vector in blue corresponding to the (014) trichord was used the most, others were found frequently in the 10968 identified groupings

Demonstration 2: Anton Webern 5 Pieces for String Quartet Op. 5, movement 4 (1909)

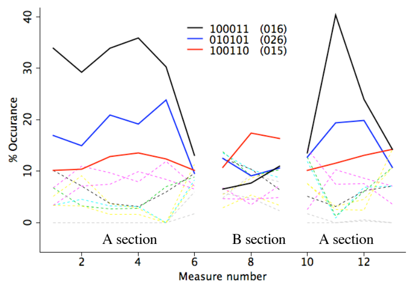

Figure 3a Trichord Use with Original Segmentation

The most frequently used trichords are plotted with solid lines; Others with dashed. Frequency of trichord varies by section, with (016) and (026) used predominantly in the A section (mm 1-6; 10-13)

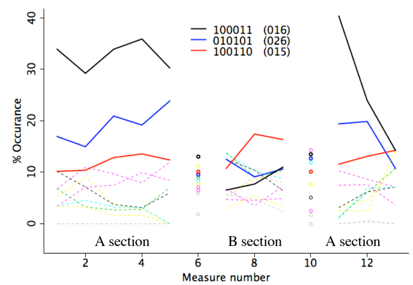

Figure 3b Alternative Segmentation

Treating mm. 6 and 10 as transition points results in a more convincing ternary segmentation with clearer sectional differences

In his analysis of Webern Op 5, No. 4 Joseph Straus comments the ternary form of the piece is demarcated by shifts in tonal structure between the A (mm 1-6; 10-13) and B (mm 7-9) sections (p 79-84). While he generally favors set class sizes of 7 notes, for the sake of clarity Figure 2a summarizes an exhaustive analysis of 3 note patterns. It is clear the trichords (016) and (026) are used more prominently in the A sections (1-6; 10-13), with a great deal of harmonic ambiguity in the B section (7-9). While the analysis is in agreement with the traditional segmentation (Figure 3a), the alternate segmentation shown in Figure 3b treating mm 6 and 10 as transition points results in a cleaner sectional division. Automated analysis provides a valuable tool for exploring such alternate interpretations.

This example demonstrates the potential for analysis using the VirtualScore to exhaustively examine all potential musical patterns. However, while this approach offers tremendous possibilities, it is likely only a small fraction of the 3573 discovered patterns spread over 13 measures are perceptually salient, suggesting interpretation of the multitudes of analytical information generated by such an approach must be interpreted in light of the pre-conscious filtering of auditory information in music perception.

Conclusions

The VirtualScore possesses enough flexibility in its representation to allow for the coding of expressive performance characteristics, yet is abstract enough to allow for complex analyses on a variety of dimensions. In two separate demonstrations, algorithms searching a VirtualScore were able to identify, sort, and catalogue thousands of potentially meaningful musical patterns. Such an analytical approach would not be feasible without automation, suggesting computer aided analysis holds potential to augment current analytical approaches.

However, the process of sorting and identifying which of these patterns are perceptually salient is a complex task which requires an understanding of more than just music theory - it requires a working knowledge of music perception. Many thousands of readily identifiable patterns are disregarded without conscious awareness. Therefore even an approach taking advantage of the analytical power of abstract representation while capturing the nuances of expressive performance cannot fully "analyze" musical information without accounting for the role of the perceptual system in transforming "acoustic sound" into "music."

While the results of these analyses are promising and suggest the use of software to aid music analysis may yield many interesting applications, a better understanding of the characteristics of saliency is needed in order to realize these possibilities. As music has been composed within the context of the perceptual system, a firm understanding of this system is necessary not only to continue pursuit of automated analytical tools, but to better understand the music which has evolved in response to it.

References

Castine, Peter. (1994). Set Theory Objects. Frankfurt am Main: Europäischer Verlag der Wissenschaften.

Cope, David. (1996). Experiments in Musical Intelligence. Madison, Wisconsin: A-R Editions, Inc.

Dannenberg, Roger. (2002). Listening to 'Naima': An Automated Structural Analysis of Music from Recorded Audio. Proceedings of the 2002 International Computer Music Conference. pp. 28-34.

Huron, David. (1994). "Humdrum News." Issue No. 1 (Sept 23, 1994)

Schutz, Michael. (2002). Computer Aided Analysis of Post-Tonal Music. Pittsburgh Undergraduate Review. Fall 2002.

Straus, Joseph N. (1990). Introduction to Post-Tonal Theory. London: Prentice-Hall.