Supporting Music Interaction and Analysis with Computers

Adam Mikeal
Department of Computer Science
Texas A&M University
Harvey R. Bright Building
College Station, TX 77840 USA
adam@tamu.edu

ABSTRACT
The relationship of computers to music is long and rich, dating back to the earliest years of electronic music experimentation and Musique concrete. Working at Bell Labs in 1957, Max Matthews wrote MUSIC, the first software application for creating sound on a computer, and computer technology has been integral to modern composition ever since. As we approach the 50th anniversary of the invention of computer-created music and computer-aided composition, it is useful to step back and take inventory of the current state of things; where we have been, but more importantly where we still have far to go.

Keywords: Music; computer-aided composition; education; performance; information retrieval.

INTRODUCTION

Music has been a universal form of human expression for as long as records have been kept. Music's ability to speak directly to our emotions, and its incredible expressive faculties have allowed it to transcend traditional barriers to communication, such as culture and language differences. Given that computers and computing technology has gradually taken a leading role in our global society as a communicative technology, it is therefore not surprising that an eventual integration of computers and music would happen. Indeed, the application of new technologies towards music-centered activities is often one of the first developments in the application of said technologies to the outside world.

The relationship of computers to music is long and rich, dating back to the earliest years of electronic music experimentation and *Musique concrete*. Working at Bell Labs in 1957, Max Matthews wrote MUSIC, the first software application for creating sound on a computer, and computer technology has been integral to modern composition ever since. As we approach the 50th anniversary of the invention of computer-created music and computer-aided composition, it is useful to step back and take inventory of the current state of things; where we have been, but more importantly where we still have far to go.

Given the time frame involved in the relationship between computers and music, it is not surprising that there is a significant volume of literature on the subject. Indeed, merely providing some crude structure to the current literature proved a difficult task. Creating a classification system that did justice to the majority of the works was challenging. Ultimately, I settled on 5 broad categories:

Search & Retrieval Research that falls into this category includes classification systems, various mechanisms querying a database of music (such as query-by-humming), or automatic structural analysis of music for the purposes of surrogate creation, which are used for search & retrieval applications. Basically, this category is focused on mechanisms to allow a user to *search* through datasets of music and retrieve relevant results.

Composition This category is fairly straightforward; software that deals with the composition process. There are two major divisions within this category, however: applications that seek to assist the human composer in his creative task, and software that seeks to replace the human in the composition task. In addition, the line between these two divisions has begun to be increasingly blurred.

Music Education This category included research in the fields of educational psychology, and focused on the results of using computer software and tools for learning and pedagogy.

Performance & Experience Items in this category focused on the creative and performance acts integral to music as a shared experience; the applications described tended to focus on the collaborative and interactive potential of computer technology in regards to the performance act.

Analysis & Scholarship The category that has the most direct relationship to the field of digital humanities also has the weakest representation in the literature. Admittedly, these categories are not impermeable; there are certainly aspects of scholarship and analysis that exist in the other four categories (especially in the "Search & Retrieval" category); however, in those samples the analysis occurs as a secondary result to the primary goal of information retrieval, and not as the primary goal of scholarship.

Of these five categories, it is the last two that are of particular relevance to the field of Digital Humanities, since they both deal with the humanist's primary task, and how computer technology can aid or assist in that task.

This paper will provide a brief outline of the current research available in the field in each of the aforementioned categories, and conclude with a brief analysis of the situation, and where I expect to see the research progress in the near future.

SEARCH & RETRIEVAL

The research space for search and retrieval is arguably (by my highly un-scientific measurements) the largest of the five categories that I have laid out, and accordingly, the available research in this category is largest and most easily evident (again, by my own, quite un-scientific measurements). Of course, given the increasing role that digital music plays in our society and economy, it is not surprising that more attention is being paid to improve the mechanism we have to catalog, store, search, and find it.

Traditional information retrieval techniques have been centered around datasets that are text-centric, and music—particularly polyphonic music—present unique challenges that often leave these traditional data mining approaches wanting. This is largely due to the fact that the sheer amount of data conveyed in even a small polyphonic piece—not to mention a larger-scale work such as a symphony—creates a massive scaling challenge for traditional data mining techniques. Most of the research in this category deals with attempts to modify or adapt basic IR approaches to more effectively interact with music as a primary information source, either by creating algorithms and systems that natively understand music as well as traditional systems understand text, or by creating surrogates of the musical data to represent the original information, and caching and indexing those surrogates.

Moerchen, Mierswa, and Ultsch [14] dealt with a related, but different problem: how can one accurately group songs into meta-categories to aid in the task of recommender systems (a growing need in the new digital music economy). Their approach was to model the genre and timbre of a piece by using exhaustive feature generation, and a learning-algorithm approach that trains regression models in order to summarize a subset of those features. Eschewing heuristics and psycho-acoustic analysis, the authors instead opt for a statistical sampling based on short-term features from small time windows. Ultimately, they produce a semantic description of the song, which allows them to classify the piece with a genre to a high degree of accuracy when compared to traditional methods.

Sapp [16] used a similar approach to solve another classification problem; recognizing the musical key of a piece.[1] In this paper, a method is described which uses a statistical algorithm to identify which pitches occur most often in a piece relative to the others. This method is very context-sensitive, in that a larger or smaller time window will significantly affect the outcome. In order to expand this method to provide a visual indication of when key shifts occur *within* the piece, the author color-maps all possible key interpretations for a particular piece onto a keyscpae plot, providing a visual indication of the harmonic structure of the piece (see Figure 1).

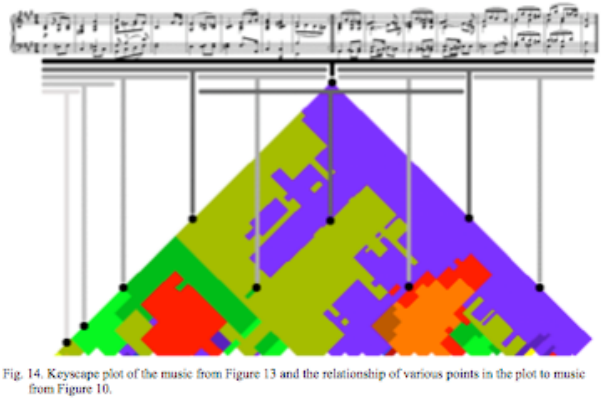

Figure 1: Example of the author's keyscape plot with a correlation to the associated music score.

In another piece about music analysis with the goal of indexing and retrieval, Chai and Vercoe [4] turned their attention toward determining a piece's overall structure (the larger-scale repetitive patterns). By recognizing repetitions of each segment of a fixed length, the algorithm is able to summarize the repetitions and infer a structure based on heuristic rules. Finally, based on the structural analysis report, the authors propose an approach for music thumbnailing. The authors used a corpus of Beatles music to perform preliminary testing on their methods, and reported favorable results.

Finally, to examine a final work in this category, Ferrara and his colleagues [8] describe an ontology to provide a Semantic Web compatible representation of any given piece based on its content. They utilize an OWL representation that describes music information, as well as rules to build a flexible genre classification scheme. Additionally, their approach is not restricted to a rigid taxonomy, but can adapt to fuzzy metadata, and imprecise correlations between a piece and any corresponding musical genres (see Figure 2).

COMPOSITION

Composition is often thought of as the quintessential creative act for a musician—one conjures up an image of Beethoven hunched over a desk penning his famous four-note theme from his fifth symphony. Nevertheless, as with almost every other aspect of the musical experience, here too computers are making their presence felt. Within this category I've labeled "Composition" there are really two major subdivisions: research that deals with computational tools to aid in the compositional task, or *computer-aided composition*, and research that aims to use a computer to do the composition itself, or *algorithmic composition*.

[1] In western music, pieces generally follow a predictable progression of harmonic development, with a single note in a scale being the dominant tone, or *key*. Typically, the piece will begin with pitches most closely related to its key, move away from those pitches (the tonic, or root or the key) creating tension, then resolve with a gradual move back to the dominant pitches of the key. This pattern can be seen in a single monophonic melody, and in more complex polyphonic scores.

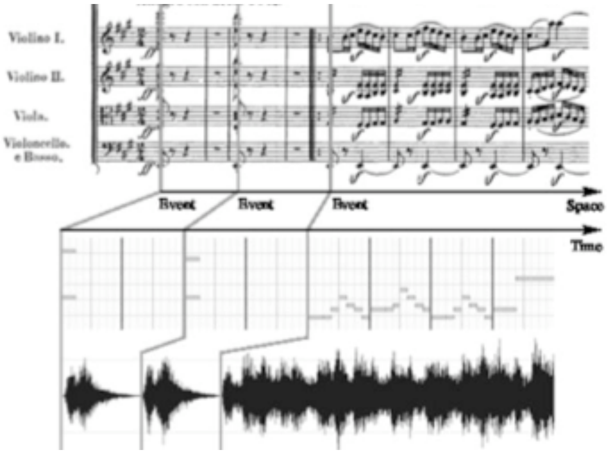

Figure 2: Spine: showing the relationships between notational, performance, and audio layers.

Figure 3: Students begin each lesson with verbal instructions from Beethoven.

An example of the former is the application QSketcher described by Abrams, et al [1] Basically, QSketcher is like an IDE for music composition—specifically, it aims to aid in the process of creating music for film. The creators paid particular attention to the early stages of the compositional workflow; from idea conceptions to realization, as opposed to a particular focus on the linear order or the musical fragments in the film. This same approach is also illustrated by a 1983 paper by Alan Talbot [17], *Finished Musical Scores from the Keyboard: An Expansion of the Composer's Creativity*, in which an intelligent music typesetting program is described to free the composer from rote mechanical duties.

MUSIC EDUCATION

Perhaps this category would be better titled "Music *and* Education", because I found at least one paper that while it dealt with using music and computers in an educational setting, rather than using the computer technology to teach music, the author described the results of an experiment in which he used *music* to teach *computer science* [9]. Specifically, he used music theory and composition to introduce college freshman to design patterns and object-oriented programming. According to his results, after a trial semester he reports promising findings, and said that the universal nature of music—its ability to transcend language and culture—proved to be a great benefit to the experiment.

On a more traditional note (at least, what you would think of as typical within a "Computer and Music Education" category), there was the class of publications that described software whose purpose was to instruct, inform, and educate. Usually geared toward younger users, these applications use multimedia, interactivity, and immersion to create a rich learning environment for children in which to experience more complex music. Ian McKinnon described one such application, entitled *Children's Musical Journey* and the various immersive environments that are available to be explored [12]. In Figure 3, the user is receiving instructions from Beethoven prior to beginning the lesson.

PERFORMANCE & EXPERIENCE

Items in this category focused on the creative and performance acts that are integral to any musical art as a shared experience. The applications and systems described in these papers tended to focus on the collaborative and interactive potential of computer technology in regards to the performance act. Obviously, there is a heavy emphasis on the approaches found in disciplines like computer-supported cooperative or collaborative work.

For example, Roger Dannenberg and Christopher Raphael [5] discuss the challenges in creating a computer accompanist that can follow—and sometimes even learn from—a live human performance. Their goal is to achieve a correspondence between the symbolic representation of music that we call a score, and the audio performance of that same score. Figure 4 shows a spectrogram of a complex polyphonic piece by Gustav Holst; the vertical lines indicate the computer-predicted measure demarcations based on the score alignment algorithms. By frequently sampling the audio, they are able to calculate pitch and beat estimates, which provide the basis for a computer-based accompaniment. Their system is able to respond to a human soloist, by extrapolating the expected arrival time of the next beat based on previously detected note offsets.

Other authors take different approaches to the challenges of integrating computers and the performance arts. Oshima, et al [15] proposed several concepts for facilitating musical expression during performance utilizing a prototype musical instrument they dubbed CiP. CiP consists of a MIDI keyboard, a music database, a function for note-number substitution, and a tone-generator. They have experimented with using this system in both performance and educational (i.e. lessons) settings. Tanaka [18] develops a novel approach to collaborative musical expression and performance between large groups of disconnected individuals over a wide-spread computer network, such as the internet.

ANALYSIS & SCHOLARSHIP

Finally, the last category we will consider consists of the overlapping arenas of analysis and scholarship; specifically,

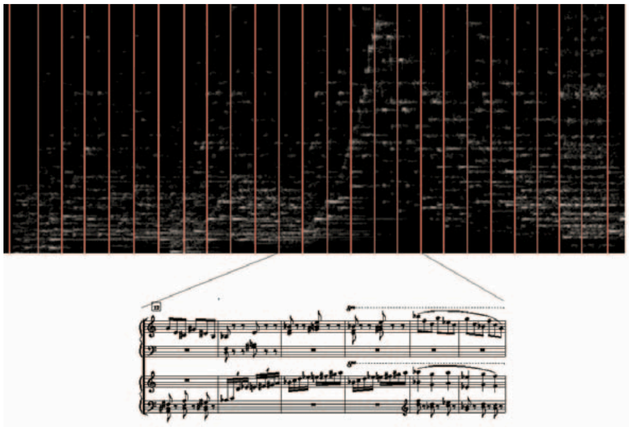

Figure 4: Spectrogram of the opening of Holst's *The Planets*. Vertical lines indicate the automatically-generated bar lines for score alignment.

using computational technology to aid or assist in the traditional humanistic tasks of music analysis. Unfortunately, the category that has the most direct relationship to the field of digital humanities also has the weakest representation in the literature. As mentioned earlier, these categories are certainly not impermeable; there are aspects of scholarship and analysis that exist in the other four categories (especially in the "Search & Retrieval" category [6][11]); however, in those studies the analysis occurs as a secondary result to the primary goal of information retrieval, and not as the primary goal of scholarship.

For this final category, it is that body of research which concentrates on tools to directly aid the researcher in this *research task* that most directly interest us. While computers have been employed in the aid of humanists for many years [7], most of that research has focused on textual materials as its primary subject matter, and only recently, video sources. The existing literature applying this approach to music is rather limited.

Illustrating the difficulties in locating this genre of scholarship, one of the papers I found actually dates back to 1979—James Meehan's *An Artificial Intelligence Approach to Tonal Music Theory* [13]. In this work, he cites the lack of a computer model of tonal music theory, and blames the lack of such a model on a preoccupation with syntax that comes from the traditional linguistics communities, and specifically, transformational grammar. Instead, he uses the concepts of semantic primitives from AI and natural language processing research to provide a correspondence to larger structures in the music.

Carol Krumhansl develops a well-reasoned analysis of polyphonic tonality from a geometric point of view [10]. Using traditional tools of music theory such as the circle of fifths and fourths, she maps the relationships present between pitches and key signatures onto a geometry space that allows for novel analysis. Particularly fascinating are the 3-dimensional graphs of keys and pitch she includes from older sources, such as Arnold Schoenberg, and the comparison to similar visualizations produced using modern, computer-aided tools and approaches.

A final example to consider from this category would be the IBM Glass Engine [3], as am example of a unique, interactive visualization technique. This interface into almost all of Philip Glass' compositions allows the user to explore and interact with his pieces in a variety of ways, and discover connections that would otherwise have been impossible or difficult to extract. This is an excellent example of a unique visualization and navigation tool that would only be possible on a computer, and whose primary purpose is to aid the scholar or student of Philip Glass in experiencing his repertoire.

CONCLUSION

Since Bush first introduced us to the concept of the computer as an integral component of the knowledge worker's toolset [2], the computer has held great promise for assisting in the creation, analysis, categorization, storage, and retrieval of musical information. While the application of computer technologies to the field of music and musical analysis has produced many results, the majority of these results are centered around a very narrow range of research interests. Information search and retrieval dominates the field, and many of those publications discuss research that is geared toward commercial music retrieval or recommender systems—obviously research that is driven my economic realities as the music industry finds itself increasingly forced to adapt to a digital world.

There is certainly a gap in the research, however, where analysis and scholarship is concerned. Traditional digital humanities tools—visualization methods and systems to analyze pieces and make connections in ways that are difficult or impossible for humans—are few and far between. It seems as though the digital humanities and humanities informatics communities have allowed themselves to be trapped into a largely text-centered focus.

Hopefully, this state of affairs will not persist in the long-term. As noted previously, the economies of the music industry is rapidly catching up to our digital world. With luck, this movement toward digitization will prompt more and more music scholars and researchers into collaborations with computer scientists, and we will see new, interesting tools and systems that will challenge our conceptions of music—what it is, how it is made, and how it is understood.

REFERENCES

1. Steven Abrams, Ralph Bellofatto, Robert Fuhrer, Daniel Oppenheim, James Wright, Richard Boulanger, Neil Leonard, David Mash, Michael Rendish, and Joe Smith. Qsketcher: an environment for composing music for film. In *C&C '02: Proceedings of the 4th conference on Creativity & cognition*, pages 157–164, New York, NY, USA, 2002. ACM Press.

2. Vannevar Bush. As we may think. *Interactions*, 3(2):35–46, 1996.

3. IBM T. J. Watson Research Center. The ibm glass engine, 2001.

4. Wei Chai and B. Vercoe. Structural analysis of musical signals for indexing and thumbnailing. In *Proceedings of the 2003 Joint Conference on Digital Libraries (JCDL)*, pages 27–34, 2003.

5. Roger B. Dannenberg and Christopher Raphael. Music score alignment and computer accompaniment. *Commun. ACM*, 49(8):38–43, 2006.

6. Daniel P.W. Ellis. Extracting information from music audio. *Commun. ACM*, 49(8):32–37, 2006.

7. D. C. Engelbart. Augmenting human intellect: A conceptual framework (afosr-3223). Technical report, Stanford Research Institute, Menlo Park, California, 1962.

8. Alfio Ferrara, Luca A. Ludovico, Stefano Montanelli, Silvana Castano, and Goffredo Haus. A semantic web ontology for context-based classification and retrieval of music resources. *ACM Trans. Multimedia Comput. Commun. Appl.*, 2(3):177–198, 2006.

9. John Hamer. An approach to teaching design patterns using musical composition. *SIGCSE Bull.*, 36(3):156–160, 2004.

10. Carol L. Krumhansl. The geometry of musical structure: a brief introduction and history. *Comput. Entertain.*, 3(4):1–14, 2005.

11. Namunu C. Maddage, Haizhou Li, and Mohan S. Kankanhalli. Music structure based vector space retrieval. In *SIGIR '06: Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 67–74, New York, NY, USA, 2006. ACM Press.

12. Ian McKinnon. Children's music journey: the development of an interactive software solution for early childhood music education. *Comput. Entertain.*, 3(4):1–10, 2005.

13. James R. Meehan. An artificial intelligence approach to tonal music theory. In *ACM 79: Proceedings of the 1979 annual conference*, pages 116–120, New York, NY, USA, 1979. ACM Press.

14. Fabian Moerchen, Ingo Mierswa, and Alfred Ultsch. Understandable models of music collections based on exhaustive feature generation with temporal statistics. In *KDD '06: Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 882–891, New York, NY, USA, 2006. ACM Press.

15. Chika Oshima, Kazushi Nishimoto, Yohei Miyagawa, and Takashi Shirosaki. A concept to facilitate musical expression. In *C&C '02: Proceedings of the 4th conference on Creativity & cognition*, pages 111–117, New York, NY, USA, 2002. ACM Press.

16. Craig Stuart Sapp. Visual hierarchical key analysis. *Comput. Entertain.*, 3(4):1–19, 2005.

17. Alan D. Talbot. Finished musical scores from the keyboard: An expansion of the composer's creativity. In *ACM 83: Proceedings of the 1983 annual conference on Computers : Extending the human resource*, pages 234–239, New York, NY, USA, 1983. ACM Press.

18. Atau Tanaka, Nao Tokui, and Ali Momeni. Facilitating collective musical creativity. In *MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia*, pages 191–198, New York, NY, USA, 2005. ACM Press.