INSIDE THE SCORE

Music and Augmented Reality

Rodney Berry*, Ivan Poupyrev[†], Makoto Tadenuma*, Nobuji Tetsutani*,
Shigeo Imura*

**ATR Media Information Science Laboratories,
2-2-2 Hikaridai, Seika-cho, Kyoto 619-0288
Japan - rodney@atr.co.jp*

*†Interaction Lab., Sony CSL
13-14-13 Higashi Gotanda
Tokyo 141-0022
Japan – poup@csl.sony.jp*

Abstract: The paper is intended to provide a brief prelude to a new research project, the *Augmented Composer*, that seeks to apply augmented reality to music composition. It sketches the conceptual framework for the research and its roots in the preceding *Augmented Groove* project, as well as potentials for near future development.

Key words: Mixed Reality, Music, Composition, MIDI, Education

1. Introduction

The terms *augmented reality* and *mixed reality* represent a philosophy of technology as much as they represent the technology itself. They show a breakdown of the assumption that the virtual and the real must be exclusive of one another. In Jean Baudrillard's 'The Map Precedes the Territory', humans increasingly abandon the territory and inhabit their maps, (that is, if there ever was a territory in the first place)[1]. Maps and simulations can become territories or realities of their own, quite separate from what we call the *real* world.

The original version of this chapter was revised: The copyright line was incorrect. This has been corrected. The Erratum to this chapter is available at DOI: [10.1007/978-0-387-35660-0_65](https://doi.org/10.1007/978-0-387-35660-0_65)

R. Nakatsu et al. (eds.), *Entertainment Computing*

© IFIP International Federation for Information Processing 2003

Augmented and mixed reality form a bridge between the world of experience and the world of symbols and concepts in a way that echoes our natural tendency to attach meaning to objects and make them into symbols. In music, the use of written notation creates a similar bridge between the physical sensation of sound and the formal organisation of sound that allows us to experience it as music. Musical instruments also impose certain structures on the music played upon them through a mixture of physical, ergonomic and cultural constraints. Because of this relationship between sensation, structure and meaning, it is reasonable to assume that there can be some useful applications of augmented and mixed reality in the field of music.

2. Levels of Musical Experience

Music exists on several levels at once. Sound itself exists only as periodic fluctuations in sound pressure. The complex qualities we attribute to music are a property of our perception, cognition and, ultimately, our culture. For example, we perceive evenly repeated fluctuations in pressure as ‘tones’ of a particular frequency. More importantly for music, we can perceive the relative frequencies or ratios between different perceived tones. Experimental evidence indicates that people remember melodies by these relative relationships between the notes rather than by remembering the specific frequencies of the tones themselves [1].

Beyond our perceptual ordering of sensory information, however, some kind of organisation and manipulation must take place on a cognitive level. People appear to refer to some kind of mental representation of the music when listening to it, in fact, the above reference to the way we remember pitch implies some level of abstraction is taking place. Musicians and composers however, use various abstract representations that make it easier for them to perceive and manipulate musical structure and form. In western art music of the last four centuries, some kind of graphic representation or score has played an increasing role in music making.

3. The Changing Role of the Score

Originally, music notation served primarily as a tool to help performers remember the details of a piece of music. Later, the score became a tool of analysis allowing others to discover patterns in the music not immediately obvious through simply listening. Although the notations retained their primary function as performance instructions, they also provided a symbolic

world open to interpretation and manipulation. The structures found through analysis in turn informed subsequent compositions resulting in the massive orchestral forms of the Classical and Romantic periods.

Most representations of music treat music as a space of some kind. Time typically forms one axis and other musical qualities are usually mapped out along the other axes or represented by special symbols. In the score example below, the pitch of the notes is determined by their vertical placement on the horizontal lines. The duration of the notes is shown by various symbols that indicate fractional divisions of a standard longer note.

Figure 1. Western 'Common Practice' music notation.

Japanese notation for the shakuhachi runs down the page in columns from right to left, the same as in Japanese written text. In contrast, the computer *piano roll*, notation shows the time/pitch space as a matrix of squares indicating pitch on the vertical axis and time on the horizontal axis.

Figure 2. Traditional Japanese notation for the Shakuhachi flute, and the PianoRoll' style representation popular in desktop music software.

Computer composition software enables many composers nowadays to rely less on traditional western music notation, but still refer to some kind of graphic representation on the computer. The main attraction of desktop music software for composers, is that it is now possible to directly alter the on-screen representation and simultaneously hear the change in the music. This creates a very different relationship to the score than the traditional case where the composer continually moves back and forth between the piano and the piece of paper. The composer directly experiences the flow of time

through the representation and becomes aware of the effect of small changes to the composition as soon as such changes are made.

4. Augmented Realities for Music

Augmented would be useful to musicians for a variety of reasons:

- the ability to directly experience the structures of the music in conceptual space as tangible structures in real physical space.
- the translation of physical gestures (expressed as paths in physical space) into musical gestures (musical products of the same paths through the system's parameter space).
- the ability to provide direct visual feedback regarding the states of a multitude of variables in real time.
- because the real world is not eliminated by the viewing technology (as in immersive VR), collaborative and shared experiences are easier to achieve.

4.1 The Augmented Groove

In 2000, with the above points in mind, researchers led by Ivan Poupyrev at the former ATR Media Integration and Communications Research Laboratories, built an AR based music mixing system as an Emerging Technologies demonstration for SIGGRAPH2000[1]. The system, *Augmented Groove*, employed camera-based tracking of markers placed on old LP records to control aspects of the music. Movements of the markers (x,y,z rotation and translation) were mapped to output various MIDI continuous controller commands according to their position in space as viewed by an overhead camera. This enabled a number of musical tracks to be mixed and processed with effects in real time simply by moving the records around in the air or across the surface of a table.

In addition, VRML models of strange *gauges*, reflecting the states of the MIDI parameters, were composited into the camera's video output and projected onto a screen for users to see. The tracking and image compositing was achieved using the *Augmented Reality Toolkit* libraries, developed by Dr. Hirokazu Kato of Hiroshima University [1].

4.2 The Augmented Composer

Where the Augmented Groove is focussed on mixing existing music in a performative play context, the *Augmented Composer* seeks to provide a tabletop environment for the composition of the user's own original pieces

of music. The most basic functional part of the system is the *phrase constructor*.

To construct a musical phrase, the user takes a stack of marked cards in his or her hands and lays them out on the table. Via the overhead camera, the computer tracks the position of each card on the table and translates the pattern into a melody. The left to right position of a card determines an individual note's position on a time-line. The front to back (toward or away from the user) position determines the note's pitch. While the cards are arranged on the table, the notes are heard immediately as a repeating sequence.

In addition to the *note* cards, a *modifier* card is used to change the loudness and duration of individual notes in the sequence by selecting that note, then moving along the timeline to lengthen it or toward the camera to make it louder. The *modifiers* can also be used to make global changes to the whole phrase such as transposition, change of instrument and scale mapping.

Figure 3. HMD view: notes represented as toy snakes and a spanner for the modifier card

At the time of writing, we have a very simple version of the *phrase constructor* working. The results are so far encouraging and the system, although quite limited, is fiendishly addictive. When viewed through a head-mounted display or a large screen, animated green cartoon snakes appear to represent the individual notes. When a note is made longer with a *modifier* card, the tail of the snake also gets longer accordingly. The modifier card is visible as a bright red cartoon spanner in the augmented display.

4.3 Planned Development

Larger structures will be built up using *phrase cards*. Once satisfied with a phrase, the phrase can be attached to a *phrase card* for subsequent manipulation as a separate object. A phrase might then be transformed by inverting, reversing, stretching etc., all by the manipulation of physical and virtual objects in the space on and above the tabletop. Several phrases could then be arranged to make larger phrases over time, as well as different concurrent musical tracks. A phrase could then be opened up again for editing, only this time, without a card for each note. Instead, the user would use the *modifier* cards to act upon the graphic objects inserted by the system into the user's field of vision.

A finished composition could be performed using *Augmented Groove* to control the mix. Also, since the placement of the cards provides a schedule of events, a similar system in a remote location could share this schedule and play back the notes, allowing a kind of *jamming* to take place over the net.

5. Educational Potential

The *Augmented Composer* mostly involves larger motor movements making it easier for a young child to play with it. It introduces elements of musical abstraction in a tangible physical form making it ideal for children aged 7-10 who are just starting to develop more abstract modes of thinking¹. Children could work together in twos or threes, sharing and manipulating the musical space in a cooperative learning context. A teacher could scale the complexity of the interface by giving out only a limited subset of the *modifier* cards then adding more later.

6. Conclusions

The Augmented Composer is still very much in its infancy, but may reveal some new methods of making and representing music for young and older users as we well.

- [1] Sloboda, John A.,(1985) The Musical Mind – The Cognitive Psychology of Music Oxford University Press, New York. pp.4-5.
- [2] Poupyrev, I., Berry, R., Kurumisawa, J., Nakao, K., Billinghamurst, M., Airola, C., Kato, H., Yonezawa, T., and Baldwin, L. (2000). Augmented Groove: Collaborative Jamming in Augmented Reality. In Proceedings of the SIGGRAPH'2000 Conference Abstracts and Applications (pp. 77 Poupyrev, I., Berry, R., Billinghamurst, M., Kato, H.,

- Kurumisawa, J., Nakao, K., Airola, C., , and Baldwin, L. (2001). Augmented Reality Interface for Electronic Music Performance Proceedings of HCI International 2001, Volume 1, pp.805-808
pp.805-808
- [4] Kato, H., and Billinghurst, M. (1999). Marker Tracking and HMD Calibration for a Video-based Augmented Reality Conferencing System, Proc. of 2nd Int. Workshop on Augmented Reality, pp.85-94 (1999). In Proceedings of the 2nd Int. Workshop on Augmented Reality (pp. 85-94).
- [5] Sloboda, John A.,(1985) p209