Augmented Reality for Music

Rodney Berry, Makoto Tadenuma

ATR Media Information Science Laboratories
email: rodney@atr.co.jp

Abstract
Augmented reality presents a number of opportunities for musical exploration. Using a camera-based marker-tracking system, a number of interesting avenues have been explored and many more are revealing themselves as the research develops. Issues of representation and meaning relating to symbols and objects provide fertile ground for artistic investigation. An older music system, Augmented Groove, is briefly described, as is the current project, Augmented Composer.

1 Introduction
The first time I experienced an augmented reality system in action, it triggered in me an immediate, instinctive response that inspired me to investigate the musical possibilities of such technologies. Also referred to as mixed reality, this family of technologies enables us to see virtual objects mixed together with real objects in our field of view. This is achieved by a variety of techniques usually involving partially see-through head-mounted displays or semi-transparent projection screens. In addition, camera based systems exist that show the viewer the direct input from the camera, as well as the computer-generated parts of the scene. It is the latter kind of system I will be talking about in this paper, as it is the kind that I have used for my recent experiments.

2 How it works
The augmented reality system used in my current research is based on the Augmented Reality Toolkit, a set of programming libraries developed by Hirokazu Kato of Hiroshima City University (Kato, and Billinghurst 1999). Using a camera, a computer is taught to recognise certain patterns. When a pattern is recognised, the computer also recognises the size and orientation of the thick black border framing the pattern. The border is a uniform square shape, the size of which is known to the computer. Therefore, if the square becomes larger, the system knows that the marked object is moving toward the camera. If it becomes smaller, the square must be moving away from the camera.

Figure 1 Typical tracking Pattern.
The system also knows which way up the learned pattern should be, and how much it has been rotated. When the object is tilted in relation to the camera, the square border becomes more trapezoid in shape so, by measuring the relative lengths of the sides of the square, the system can track the amount and direction of tilt.

Figure 2 Tilted card showing distortion of the square shape.
As a whole, this is enough information for the system to accurately and quickly track the movement, position and orientation of the marked object, making it very interesting as a
3 Basic ideas

3.1 Representation

Our structural understanding of music owes a lot to the kinds of representations we employ. Traditional music notations provide ingenious ways of indicating where events should occur on a timeline, and what kind of events they should be. We learn to see time as one of the axes of a musical space, through which we pass as we hear or perform the music. Since the advent of computer music, time is no longer simply a linear trajectory through musical space but can become complex non-linear path taken between various nexes of decisions. When we make a symbolic representation of a series of sound events, changes we make to that representation can be reflected in the music. For example, a reversal of the notes written on a musical score gives us a correspondingly backwards version of the original melody. Given the capability of the system described above to associate 3D computer graphics with objects in the real world, it is reasonable to suspect that such a system might be able to give a musical space a renewed physicality not possible with a mouse, keyboard and screen interface. It should be possible to create a manipulable physical representation of the music coupled with a mutable virtual representation.

3.2 Reality, virtuality, objects and symbols

Augmented reality makes it possible to bring the virtual and the real together. In some cases, it could perhaps even make the two terms superfluous. I think it was this feeling of a silent collision between the world of physical experience and the world of symbols that first piqued my interest in augmented reality as a vehicle for musical and artistic expression. It also brought back to the surface, many things I had been thinking about the relationship between symbols and objects. If we consider the symbols we use to be a part of our thinking process, then at least a part of what we call our mind resides in such symbols. Symbols do not exist only as abstractions flickering around inside our brain however. These words as you read them or hear me speak them show that symbols regularly pass outside the brain and become part of our environment. Objects in the everyday world often serve dual roles as symbols. By manipulating these object/symbols, a kind of externalised thinking takes place, giving some substance to the idea that, as our body may extend through the tools we use, so may our mind extend outward with the symbols we project, absorb and manipulate. An augmented reality based artwork might be able to express some of this line of thought in ways not possible by other means.

3.3 Past project - Augmented Groove

Dr. Ivan Poupyrev and myself developed the Augmented Groove (Poupyrev, Berry et al 2000) along with a team of designers, programmers and musicians at ATR Media Integration and Communications Research Laboratories in 2000. It was a popular demonstration during the SIGGRAPH emerging technologies exhibition in New Orleans that year. This body of work is to be more fully documented elsewhere by Dr. Poupyrev, so I will only make a brief description here. The Augmented Groove was designed with modern forms of repetitive dance music in mind and loosely applied the metaphor of a D.J., mixing and effecting pre-existing musical material in real time. Several old vinyl LP records were marked with patterns such as the ones in figure 1 and figure 2. A camera was placed above a table on which the records were placed. When a particular LP’s pattern was visible to the camera, MIDI continuous controller messages were sent to several MIDI synthesiser modules. The different controller numbers were mapped to the height of the record from the table (closeness to the camera), the position on the left/right axis, the position on the front/back axis, the amount of tilt in relation to the camera, the rotation of the record, as well as the amount the user shook the record about. Several pre-prepared tracks were played from a MIDI sequencer subject to the MIDI continuous controller messages coming from the Augmented Groove system. The various controller messages were assigned to the operation of various filters and effects parameters on the synthesisers and the mixer. In this way, a user could wave the record around in the air and simultaneously affect a number of different aspects of the mix at one time. On a
screen behind the table were displayed images of the table as viewed from the overhead camera and 3D virtual objects associated with the patterns on the records. In this case, the representations used were to give some feedback about the movement of the records and their effect on the MIDI parameters.

### 3.4 Current project - Augmented Composer

Augmented Composer is an attempt to extend some of the ideas explored in the Augmented Groove and to enable the construction of musical phrases from scratch. There is more emphasis on using the system as a composition tool than as a continuous musical controller. The Augmented composer is also aimed at helping young people develop a more intuitive understanding of musical structure by partially immersing the body in a musical space. A deck of marked cards is used as an interface. When the user places a card on the table, a repeated single note is heard. When another card is added, a second note is heard in sequence with the first. A card’s Left/Right position on the table determines its place on a continuously looping time-line. The card’s position from front to back determines its pitch.

When viewed through a head-mounted display (HMD) or on a screen, 3D representations of each individual note (in this case animated green toy snakes, figure 3) can be seen. Individual notes can be changed (altered velocity or note length, for example) by applying a modifier card next to the note card to be edited. In figure 3, the modifier card is shown as a red spanner.

![Figure 3 HMD view showing snakes and spanner card.](image)

When a note is lengthened, the tail of the snake also gets longer. Once a phrase is complete, it can be encapsulated and copied to a phrase card. Then the note cards can be re-used to make up a new phrase. Phrase cards can be organised side by side to make longer phrases or, when placed one in front of the other on the timeline, separate tracks. They can also be copied to a new phrase card. In this way, large compositions are possible. Going back and editing a phrase after it has been made is one of the more difficult design challenges. So far, we intend to ‘open’ the phrase and place the snakes in the same positions as if the original cards were on the table. Then, the modifier cards can be used to alter aspects of the representation by moving the elements around in the same manner as the original cards. In future we hope to find a more elegant way of moving between action upon the physical representation versus the virtual representation. Once that is achieved, it will be possible to do very intuitive manipulations of phrases. Imagine achieving a retrograde by simply grabbing the phrase and flipping it backwards. The same gesture could be used for inversions, compression and stretching of the phrase.

### 4 Future variations

As a medium for broader artistic expression, I think this kind of system will naturally suggest its own language as we get used to its workings. I have planned a series of art works that deal with aspects of representation and the relationship between objects and symbols in our memory and thought. The Augmented composer will undergo a number of changes, chiefly, I think a move to a less linear attitude to time is important with more objects that allow logical jumps on conditions and other old computer music chestnuts. The system could also be used as an interface for music synthesis of various kinds and might provide a fun way for a young person to learn about synthesis by arranging components with 3D extended information and hearing the results of their actions as they go.

A networked version is also on the agenda. The arrangement of the cards does not generate events in real time but merely schedules the events for playing by the sequencer. If two remote systems were connected, latency would not be a problem, as each party would have their own clock with which to interpret the schedule.

A more distant goal is to allow the user to edit the representation on a more low level. It may be possible to edit the 3D objects from within the application instead of reaching for the mouse and building them up on another system.
5 Conclusion

Augmented reality systems based on the Augmented Reality Toolkit offer potential for exploration in the area of music and media art. As a gestural controller and as a composition/education tool it also shows promise. By considering the role of symbols, representation, metaphors, objects and images, perhaps even a new artistic genre may be possible.

The research reported here was supported in part by a contract with the Telecommunication Advancement Organization of Japan entitled, "A Study of Innovational Interaction Media toward a Coming High Functioned Network Society".

References


AR Toolkit libraries available for download:
http://www.hitl.washington.edu/projects/shared_space/download/