The Music Table

Rodney Berry, Naoto Hikawa, Mao Makino, Makoto Tadenuma, Masami Suzuki

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

Abstract

The Music Table enables the composition of musical patterns by arranging cards on a tabletop. An overhead camera allows the computer to track the movements and positions of the cards and to provide immediate feedback in the form of music and on-screen computer generated images. This paper describes the basics of the system design and outlines some future directions for the project.

1   Introduction

The Music Table is a system that enables a player to compose musical patterns by arranging cards on a tabletop. An overhead camera allows the computer to track the movements and positions of the cards and to provide immediate feedback in the form of music and on-screen computer generated images. The *Music Table* provides both tactile and visual representations of music that can be easily manipulated to make new musical patterns. It can enable inexperienced music makers to experience their own music as patterns in musical space. This project is the first of a planned series of interface designs under the umbrella title of *The Augmented Composer*. An *Augmented Composer* is our name for a person who creates and learns about music with the help of augmented reality. We see great potential for augmented reality and mixed reality interfaces for music because it allows for a physical representation to be manipulated while being connected to an additional visual representation generated by the computer. (Berry et al, 2002). This representation could be displayed in a head-mounted display or on a screen or projection adjacent to the interaction area.

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"*.

2   Description of the System

Physically, The Music Table consists of a table on which are placed a variety of specially marked cards. A camera is placed directly above the table, so that it can 'see' the entire interaction surface, and video from the camera is sent to the computer. The display
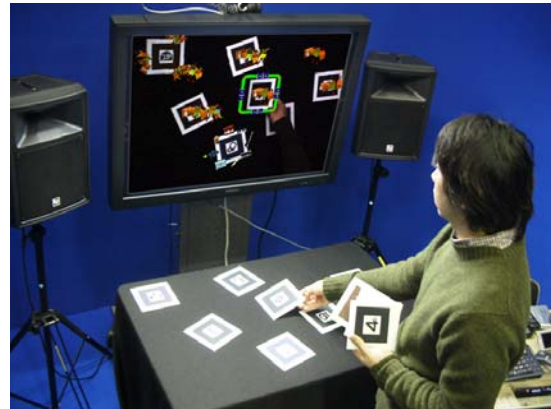

Figure 1. Music through multiple representations

output from the computer is sent to a large display screen directly behind the table (opposite where the user stands).

2.1   Software

The software component of the system is in two main parts. The Music Table program itself handles the camera-based tracking of the cards, the rendering of VRML objects, and the compositing of CG graphics into the original video image. To do this, it uses the Augmented Reality Toolkit programming libraries (Kato and Billinghurst, 1999) devised by Hirokazu Kato and supported by a lively community of enthusiasts. The sequencer for midi events is built in the PD music programming environment (Puckette, 1996) and receives data from the Music table program via MIDI and via UDP sockets. This use of MIDI is a carry over from the Augmented Groove, an earlier project that used the same marker tracking technology to mix live dance music tracks and effects via midi (Poupyrev et al, 2000). At the time of writing, MIDI communication between the applications is being entirely replaced by the use of socket communication in order to make the architecture simpler and more flexible.

2.2   Sequencer

The sequencer in the *Music Table* is really four banks of ten MIDI sequencers, each of which plays just one note. Each of these sequencers is assigned to

one of the cards on the table and that card's associated marker pattern. When a card is placed on the table, and its tracking pattern becomes visible, the music table software recognizes the position of the card on the table. The card's location on the y axis (towards or away from the user or, from the camera's perspective, up and down) produces midi controller values between 0 and 127 that are sent out to the sequencer module. The sequencer then maps this value to a MIDI note number. The range is adjustable according to how many octaves the part being played should cover. At present, this is preset before use, and there is an unavoidable trade-off between range and resolution. The position on the x axis (left to right across the table) is converted to a value from 0 to 127 then sent to the sequencer. The sequencer component takes this value and re-maps it to one of eight equal steps representing eight steps on a looping time-line across the tabletop. When the card is rotated clockwise on the table, it sends a message to increment the sequencer's stored velocity value for that card. If the card is tilted slightly to the left or right, a value is incremented or decremented in the *Music Table* software then sent to the sequencer as the duration value for that note.

The sequencer has a master clock that generates regular pulses, and outputs a repeating count from 0 to 7. When one of the one-note sequencers receives this count, it checks to see if it matches its own card's location on the time-line. If it matches (and if the card is visible at that time), it generates a MIDI note-out using the pitch value from the position of the card and the velocity from its rotation. The note is sustained for an amount of time determined by its duration value. All MIDI info originating from the cards on the table is sent on MIDI channel 1.

The values for each of the sequencers are stored in one array in PD. This array can be copied to one of several other arrays via intermediate storage in a 'clipboard' array. Once a musical pattern is made on the table, a copy card is placed near one of the note cards on the table. This triggers a command in PD to copy the main array to the clipboard array. Then, a phrase card is placed on the table and the copy card is used to copy the clipboard to the array represented by the phrase card. Each phrase card has a corresponding sequencer that is identical to the sequencer connected to the tabletop interface. These three sequencers (one for each phrase card) transmit on channels 2, 3 and 4 respectively.

After copying a phrase to a phrase card, the phrase can be later opened up for editing on the table without the presence of the original note cards. In this case, a *phrase-edit* card is applied to the images of the notes allowing the user to change the timing, pitch, velocity and duration of each note. The changes made are written directly to the array in PD that corresponds to that particular phrase.

An instrument card allows the user to change the instrument sound of either the notes playing from the table, or any of the individual phrase cards' sequencers. The card is placed close to either a note card on the table or the appropriate phrase card. By tilting the card to the left or right, the sequencer cycles through a 'palette' of instrument sounds and sends the appropriate MIDI program-change message to the MIDI output.

2.2 Visual Representation

The Music Table uses the Augmented Reality Toolkit libraries to combine computer generated 3D objects and characters with the camera image of the table and cards. It does this by identifying each card's unique marker and its surrounding black square frame. By tracking the distortion of the square shape of the black frame, the system can identify the orientation of the pattern. Once the pattern's position and orientation in space is known to the system, it can render a 3D VRML object with the appropriate scale, position and orientation and mix it in with the original video signal. This makes the VRML objects appear to be inserted in the 'real' world as shown on the screen.

In general, the *Music Table's* visual design creates a sense of fun and cuteness. This is partly due to our hopes of making a system for children to use, but also, being a Japan-based project, a contemporary Japanese sensibility underpins the design aesthetic. The projects two designers, Mao Makino and Naoto Hikawa each contribute to a design style that is both playful and iconic.

Figure 2 A short and smooth creature.

Figure 3 A long and spiky creature.

Note-cards are coupled with animated computer generated *note-creatures*. When a note is made louder, the creature's body grows more spikes. When a note is made longer, its creature also becomes longer. *Phrase-cards* are coupled with marching circles of creatures.
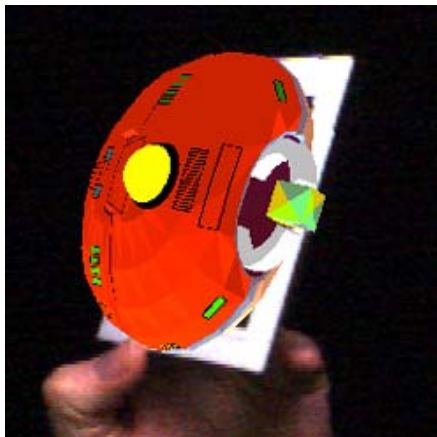

Figure 4 *Phrase-card*

Figure 5 *Copy-card*

Figure 6 *Instrument-card*

The *copy-card* shows a beetle-like machine with pincers that grab material to be copied when activated, and eject material from their rear when writing to a phrase.

The *instrument-card* shows a rotating menu of animated cartoon instruments. The *phrase-edit* card is a frame made of 3D tubing and arrows reminiscent of vintage computer game graphics.
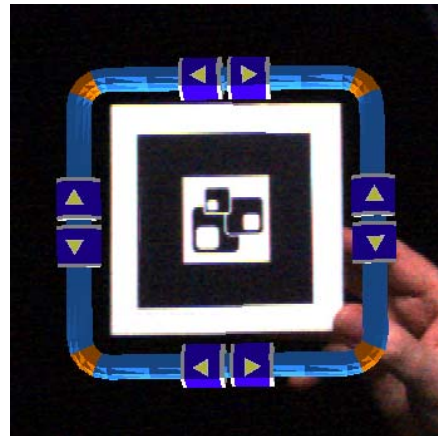

Figure 7 *Phrase-edit* card

3   Limitations and Solutions

Because the marker pattern must be big enough for the camera to recognize, there is a minimum size that a card can be before the system no longer finds it. In the present system, cards must be at least 10% of the total field of view in size. This means that, even though the position detection is quite fine, chords are necessarily limited to the space that two or more cards take up on the table. Because each pattern must be fully visible to the camera, cards making up a chord can not be allowed to overlap. Recent improvements in the tracking process, coupled with special retro-reflective material in the markers can reduce the minimum marker size considerably.

The current version has been designed with the specific purpose of demonstrating in the Emerging Technologies exhibit in SIGGRAPH2003. For this reason, the design leans more towards the requirements of a quick and dramatic technological demonstration with an expected engagement time of about 2 minutes per visitor. In its present form, the user can only make patterns of eight steps in length. By layering several tracks ('phrases') on top of each other, it is easy to produce nice pounding techno grooves that repeat until changes are made via the tabletop interface. To make it into a real composition tool, we must develop a way of concatenating phrases in order to build up new longer musical structures. The phrases will probably remain as basic units referenced by a purely graphical representation so that it is still possible to open and edit them even within larger compositions.

The use of VRML for display of 3D content is currently limited to the loading of ready-made VRML files at program start-up. The VRML models are associated with individual patterns. Changing states are currently achieved by switching between different

models. For example, to change the length of a creature involves switching between four different models of different length. If we multiply that by four different degrees of 'spikiness', we must load 16 separate VRML models just to represent one creature. In the original VRML standard, there are provisions for a variety of scripted behaviors and interactivity that are not supported by the current libraries. We are currently working on some modifications to the underlying software libraries that may enable us to build a more interactive and configurable system.

The *Augmented Composer Project* intends to explore methods of musical representation beyond the *x is time, y is pitch* world of *The Music Table.* To achieve this, the aforementioned modifications should ideally be open enough for a non-programmer to author their own mixed reality environments without having to become a master of C++ beforehand.

The Music Table currently sits somewhere between the world of children and the world of adults and will most probably be split into two separate projects to better meet the needs of each.

4 Conclusions

The Music Table is, within the stated limitations, a very usable and fun music making tool. It enables the integration of physical and visual representations that may help young composers to more intuitively understand patterns in the music they make. If the immediate technological barriers are overcome, augmented and mixed reality has a good chance of becoming a profound vehicle for human expression.

5 Acknowledgments

Thanks to the following people:

Naoto Hikawa who designed the marker patterns and 3D models for the *copy-card* and *phrase-edit* card (and who is probably the first 'power user' of *The Music Table).*

Mao Makino who designed the creatures themselves and characters for other *Augmented Composer* projects.

Takashi Furuya of CSK who wrote the software.

Hirokazu Kato and the ARToolkit community for the underlying technology for this project.

Masami Suzuki head of ATR-MIS Cooperative Learning Media Dept.

References

Berry, R., "Augmented Reality for Music" *Proceedings of the International Computer Music Conference.* International Computer Music Association, pp. 100-104.

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

Puckette, M. 1996. "Pure Data." *Proceedings, International Computer Music Conference. San Francisco*: International Computer Music Association, pp. 269-272.
http://crca.ucsd.edu/~msp/Publications/icmc96.ps
http://crca.ucsd.edu/~msp/software.html

Poupyrev, I., Berry, R., Kurumisawa, J., Nakao, K., Billinghurst, M., Airola, C., Kato, H., Yonezawa, T., and Bald-win, L. 2000. "Augmented Groove: Collaborative Jamming in Augmented Reality". *Proceedings of the SIGGRAPH'2000 Conference Abstracts and Applications* (pp. 77). ACM.