

Towards Multimedia Orchestra: A Proposal for an Interactive Multimedia Art Creation System

Kazushi Nishimoto, Kenji Mase
ATR Media Integration & Communications
Research Laboratories
email: {knishi/mase}@mic.atr.co.jp

Sidney Fels
UBC Dept. of Electrical and Computer Engineering
email: ssfels@ece.ubc.ca

Abstract

In this paper, we propose a novel multimedia orchestra system named MusiKalScope-2. From our past experiences with multimedia art creation systems, we are aware of two issues: cognitive overload due to the simultaneous creation of music and graphics by a single performer, and finding a way to provide expert knowledge to a system. To address these issues, we propose two approaches: one, use attribute based role allotment and two, the direct use of experts' knowledge. We describe the construction of MusiKalScope-2 and illustrate the realization of the proposed methods.

1. Introduction

We have been studying systems for art creation that allow expert artists as well as ambitious amateurs to create novel and wonderful works. We have already proposed several art creation (support) systems, including, the Iamascope[1], mainly for image creation and RhyMe[3], for music creation. Furthermore, by combining these two systems into one, we constructed MusiKalScope version 1 (MKS1, hereinafter)[4] for simultaneous image and music creation. This led further to a vision-to-music processing technology incorporated into a new version of the Iamascope[2]¹.

These systems gave rise to several problems as discovered through experimentation. In this paper, we first describe the problems in detail and next propose two solutions. Further, we illustrate a prototype system named "MusiKalScope-2" (MKS2, hereinafter), which resulted from the implementation of the solutions.

The remainder of this paper is organized as follows: Section 2. overviews our previous art creation systems and the associated problems. Section 3. describes the

methods employed in MKS2 in detail. Section 4. illustrates the internal workings MKS2. Section 5. discusses the features of the employed methods. Section 6. concludes this paper.

2. Previous Systems and Problems

2.1. Overview of our previous systems

The goals targeted by our previous art creation systems are these: 1) to let novices easily achieve a reasonable quality of work, 2) to obtain a good balance of quality between individual media in the case of multimedia art creation, and 3) to allow an upward path for skill acquisition so that users can become more expressive. Therefore, the systems are not mere support systems, but new musical/graphical instruments.

The *Iamascope* is an immersive electric kaleidoscope. Figure 1. shows a block diagram of the Iamascope. A video camera captures the performer's image. The system extracts a "pie" slice from the captured image. Then, by reflecting the slice and rendering it sequentially in a circle, a kaleidoscopic image is generated in real-time. The image changes responding to a performer's motions.

RhyMe is an intelligent musical instrument designed especially for jazz improvisation. Figure 2. illustrates a block diagram of RhyMe. The analysis module analyzes a given chord progression of a music piece using the Berklee theory², obtains the available note scale data, and determines the function of each note at each point in the musical piece. The mapping module dynamically changes the mapping of notes onto the positions on a musical instrument (e.g., the keys of a keyboard) in real-time, synchronizing with MIDI³ timing clock messages. We employed a "fixed-function" map-

¹In this paper, hereafter, the "Iamascope" refers to the older version of the Iamascope.

²The Berklee theory was established at the Berklee college of music based on the 40's BeBop jazz.

³MIDI stands for "Musical Instrument Digital Interface".

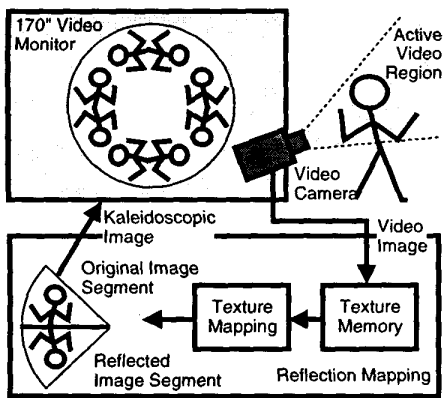


Figure 1: Block diagram of the Iamascope

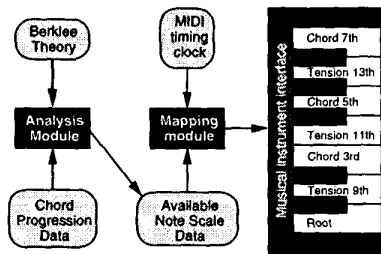


Figure 2: Block diagram of RhyMe

ping method. With this method, a note having a specific function is always mapped to a specific position. Therefore, the performer is always able to access a note to provide a specific function without needing music-theoretic knowledge.

MKS1 is an integrated system of the Iamascope and RhyMe. To integrate them, the Graphical Musical Instrument Interface (GMII) has been added which provides a virtual drum set. Figure 3. shows the block diagram of MKS1. A performer wears one receiver of a position sensor on each hand. When the performer hits the space in front of her/him as if striking a drum pad, the GMII detects when and where the hit is struck. A fixed-function mapping relates spatial regions and notes. While standing in front of the camera mounted in the Iamascope, the performer now plays the virtual drum set. As a result, music and a kaleidoscopic image are simultaneously generated by the same performer.

MKS1 achieves a good balance between the quality of individual media involved. Moreover, because of its support, the MKS1 succeeds in reducing the performer's cognitive load. Because of these features, the MKS1 makes it possible to easily create reasonable

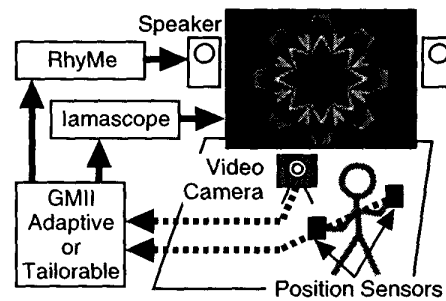


Figure 3: Block diagram of MKS1

quality multi-media artwork.

2.2. Problems within the previous systems

These previous systems generally worked well as expected. However, we encountered two main problems. While performers were able to generate music and graphics of a reasonable quality simultaneously using MKS1, most concentrated on either music or graphics alone. This suggests that it is difficult or perhaps even impossible for people to pay attention to multiple media simultaneously, even if the system reduces the cognitive load for the creation of individual media.

The second problem involved RhyMe. RhyMe requires two kinds of knowledge: knowledge of jazz theory and knowledge of the chord progression of the musical piece to be played. The second factor reduces improvisational flexibility. A performer using RhyMe cannot deviate from the *a priori* chord progression.

3. Employed Methods

3.1. Attribute based role allotment

From past experiments, we have found it useful to allot roles to multiple performers. The question remains, how best to allot them? Here, we would like to briefly review several previous attempts.

In the multimedia opera "Mada Umarenai Kamigami he" by K. Furukawa, and in "Music Plays Images x Images Play Music" [5, 6], music, computer graphics and lighting of the stage all affect each other mutually and interactively. In "Brain Opera" [7, 8], art works are created not only by integrating music generated by the performers, but also by the performers' body actions as well as sounds generated by the audience itself.

These are ambitious and interesting attempts. However, we believe that, in essence, role allotment is done mostly *ad hoc* with no obvious guiding principle in

these examples.

Hence, we propose an “attribute based role allotment” method. Several common attributes exist independent of the specific media. Accordingly, we employ an allotment method whereby each performer is in charge of performing an attribute across different media. For instance, a “color player” decides the “color attribute” of all of the media.

3.2. Direct use of experts’ knowledge

To solve the problem of how to provide direct knowledge to systems, we attempt to employ human-experts’ knowledge without modification. There are two kinds of relevant knowledge: creative knowledge and mechanical knowledge. For example, so-called “reharmonization” requires creative knowledge. It cannot be mechanically determined how a chord progression should be reharmonized. However, when given a chord progression, available note scales can be mechanically obtained. Therefore, we feel that the system should be responsible for *only* mechanical knowledge. Further, creative knowledge should be extracted from experts directly and provided to the system and to other performers.

3.3. Compatibility of the two methods

The above two methods are compatible. Furthermore, their integration has a multiplicative effect. If an expert “color” performer takes part, his/her expert knowledge becomes available to all other performers. This enhances the performance of the whole orchestra. Consequently, we can expect that experts’ knowledge raises the total quality of the performance.

4. Inside of MKS2

4.1. Overall structure

Like the MKS1, the MKS2 consists mainly of the Iamascope with musical components. However, to implement the two methods described above, the interconnecting parts (GMII for the MKS1) were modified.

Figure 4. illustrates a block diagram of the MKS2. Currently, the MKS2 is equipped with five performing instruments: a MIDI keyboard (Roland A-99), an electric drum set (YAMAHA electric drum system), a wearable musical instrument (YAMAHA MIBURI MSU-3⁴), a motion capture system with a CCD cam-

⁴YAMAHA MIBURI consists of a jacket and a sound module. The jacket is equipped with six bend sensors on arm joints, two hand-held keyboards with eight buttons each and one seesaw

era, and a set of tap sensors (our own design). Currently, four modules are executed on three workstations (SGI INDY x 1 and SGI O2 x 2). All of the modules are connected via unix sockets and an ethernet.

4.2. Modules

4.2.1. Tap-module

Tap sensors are simply thin touch switches attached to the bottom a performer’s right and left shoe. When the sensor is pressed, a tap-on signal is transmitted as an audio signal to the receiver unit of a wireless microphone system and is finally received by the INDY. If the input signal exceeds a specific threshold, the tap-module running on the Indy sends the appropriate tap-on event (right or left) to the MKS2-module.

4.2.2. Iamascope-module

The Iamascope-module closely resembles the Iamascope system. However, the modes to generate a kaleidoscopic image (e.g., simulating two mirrors or three mirrors) and so on can be controlled by the commands from the other modules.

4.2.3. Speed-module

The speed-module detects the Iamascope performer’s speed of motion. This module extracts a pie slice from the captured image from the CCD camera, as does the Iamascope, and divides it into 10 segments. Then, the module generates a differential-vector by calculating the temporal difference of image intensity of each segment. This vector is sent to the MKS2-module.

4.2.4. MKS2-module

This module maps all of the input data to the output.

From musical data to music The data from the keyboard, the electric drum, and MIBURI are transmitted on different MIDI channels: channels 1, 2, and 3, respectively. Currently, the data on ch. 1 is assumed to be chord data. The MKS2-module first decides what chord it is, based on the input MIDI note numbers from ch. 1. Then, available note scales are decided in real time by using a simplified method.

The data from MIBURI (ch. 3) decides octaves and provides a sense of tension. When the MIBURI performer is extending both arms, the tension level is switch, and a waist belt with an interface box. With MIBURI, a performer plays music as s/he dances by striking poses.

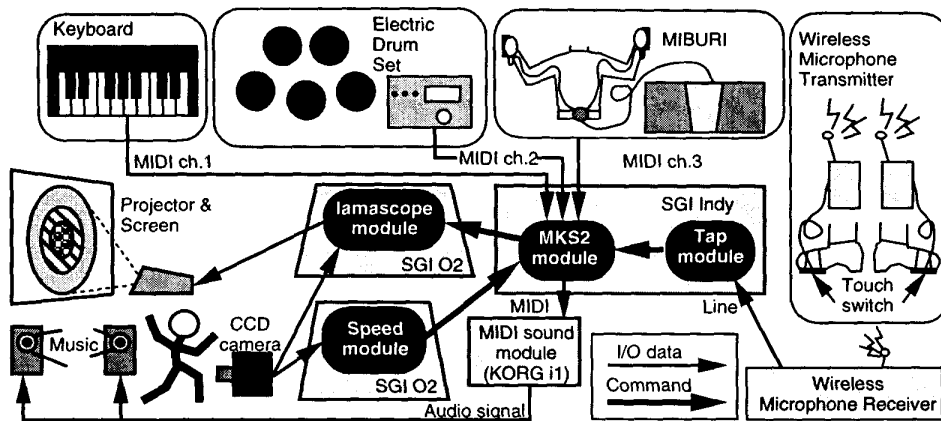


Figure 4: Block diagram of the MKS2

“lowest,” the more s/he bends her arms, the higher the tension level becomes. The octave is decided by the hand-held keys. The index finger key corresponds to the lowest octave ascending to the little finger key which corresponds to the highest.

The data from the drum set (ch. 2) composes the melody. When a pad is struck, corresponding pitch data is sent. However, the pitch assigned to a particular pad changes according to the current scale, as well as the tension level and octave data. If the tension level is at its lowest, only the chord tones are mapped onto. However, the higher the tension level gets, the more dissonant notes are also mapped. The octaves of the mapped notes are changed by the “octave” data.

The data from the tap-module composes a simple bass-line. The tap data from the performer’s left foot plays the root note of the current chord, while the tap data from the right foot plays the 5th note.

From musical data to graphics The input chord affects the color balance of the kaleidoscopic image. Major-type chords restore normal color balance, minor-type chords change the color balance towards blue, and the other chords (e.g., major dominant 7th, diminished and so on) change the color balance randomly.

The viewing mode of Iamascope is also changed when a modulation is detected.

The image flashes when a drum pad is struck harder than some threshold.

The tension-level data changes the rotation speed of the kaleidoscopic image. The higher the tension level is, the more quickly the image rotates. The octave level determines the angle of the pie-sliced captured image. The higher the octave level becomes, the narrower the slice angle becomes.

From graphical data to graphics The input image data simply generates the graphics (the kaleidoscopic image) the same way the Iamascope does.

From graphical data to music Data from the speed-module composes an accompaniment like arpeggio. The pie-slice image is divided into 10 segments along the radius. A specific chord note is always mapped on a specific segment in the pie slice of the captured image: e.g., the most inside segment always corresponds to the root note, and the most outside segment corresponds to the 5th note. By moving more rhythmically and variously, arpeggio-like accompaniment can be performed.

4.3. How the whole system works

Table 1. summarizes the relationships among the inputs and the outputs. The term, “Attribute,” describes the role of the attribute for each instrument performer. As is shown in the table, each instrument is dedicated to a certain attribute role (only the Iamascope and the tap sensors are combined). In this sense, for example, the keyboard is not a musical instrument: it becomes an *attribute* instrument of multimedia art.

The color performer controls the overall color of the work. The musical color as well as the graphical color are simultaneously controlled by the same input. Furthermore, the performed color affects the performances of all remaining parts. They become, in combination, a realization of the proposed two methods.

The mood performer decides the feeling of tension or relaxation within the whole work, under the global color decided by the color performer. The mood performer affects the performances of the remaining two

Table 1: Relationships among inputs and outputs, and the attribute that each instrument performer is in charge of

| Attribute | Instrument | Music | Graphics |
|-----------------|---------------------|--|--------------------------------|
| Basic color | Keyboard | Chord progression and backing | Color balance and display mode |
| Mood | MIBURI | Tension feeling and octave | Rotation speed and slice angle |
| Upper structure | Drum set | Melody line and beat | Flashing |
| Lower structure | Iamascope with Taps | Arpeggio-like backing, bass, and tempo | Base of image |

parts as well.

The mood and color performances are only heard in the context of the remaining two components of the system. Their performances cannot be heard or seen in isolation. Hence, they create the hidden background structure of the work. The color performer constructs the global structure, while the mood performer constructs the local structure.

The remaining two parts therefore provide the foreground of the work. Both of these performances can be directly listened to and watched by the audience. The upper structure performer creates the lead. The lower structure performer generates the base, including, the tempo set by his/her rhythmical motions, and its accompaniment. These outputs are affected by the hidden parts.

5. Discussion

Dynamic interaction and collaboration are always important for creation of art by multiple performers. However, in typical orchestra performances, for example, even if one performer makes a mistake, a reasonable quality of music may still be achieved. Our proposed system is much more delicate. If one performer interacts poorly, the total quality of the performance can be vastly reduced.

The difference derives from structure. A typical orchestra has a vertically divided structure. A performer in the orchestra plays something by her/himself. The orchestra is constructed by bringing together such self-contained parts. Our system has a horizontally divided (layered) structure (see Figure 5). Conceptually, even to play one note effectively, all the performers must play their allotted roles. This feature can be a strength as well as a weakness. Because each individual performance directly affects the total work, there is very little redundancy. Therefore, higher concentration is required which can lead to higher density work.

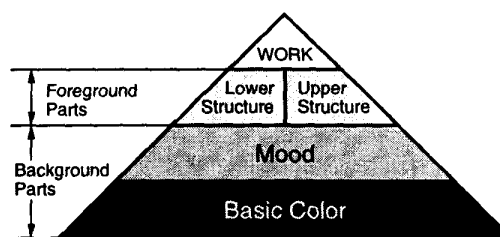


Figure 5: Concept of an orchestra with a horizontally divided structure

6. Concluding Remarks

To construct a good multimedia art creation system, we proposed two methods: attribute based role allotment, and the direct use of experts' knowledge. Based on the proposed methods, we have constructed a prototype system as a multimedia orchestra, named the MusiKalScope-2 (MKS2). We described the inside of the MKS2 and showed the mapping between music and graphics, as well as the layered structure of the MKS2. We also discussed the system's characteristics for the interaction among performers. In the very near future, we would like to attempt to create an art work with the MKS2 to further evaluate its potential.

References

- [1] Fels, S., Reiners, D. and Mase, K., "Iamascope: An Interactive Kaleidoscope", *proc. HCI International'97*, p.35, San Francisco, August 1997.
- [2] Fels, S., and Mase, K., "Iamascope: A Musical Application for Image Processing", *proc. The Third International Conference on Automatic Face and Gesture Recognition (FG'98)*, pp.610-615, 1998.
- [3] Nishimoto, K., Watanabe, H., Umata, I., Mase, K. and Nakatsu, R., "A Supporting Method for Creative Music Performance — Proposal of Musical Instrument with Fixed Mapping of Note-functions", *Trans. of Information Processing Society of Japan*, Vol.39, No.5, pp.1556-1567, 1998 (in Japanese: a brief summary of RhyMe is included in the literature [4]).
- [4] Fels, S., Nishimoto, K. and Mase, K., "MusiKalScope: A Graphical Musical Instrument", *IEEE Multimedia Magazine*, July-September, pp.26-35, 1998.
- [5] <http://www.aec.at/prix/1997/97gnI-images.html>
- [6] <http://sakamoto-iwai.aec.at/>
- [7] <http://brainop.media.mit.edu/>
- [8] Paradiso, J. A., "Electronic music: new ways to play", *IEEE SPECTRUM*, Vol. DECEMBER, pp.18-30. 1997.