The Yamaha MIBURI MIDI jump suit as a controller for STEIM's Interactive Video software Image/ine

Lindsay R. Vickery

Classical Music Department, Western Australian Academy of Performing Arts,
Edith Cowan University
email: l.vickery@cowan.edu.au

Abstract

The desire on the part of composers to extend the scope of interactive electronics into the visual domain has resulted in the development of a number of systems that use gestural information as a means for controlling video. This paper will discuss the combination of two of these technologies - the Yamaha MIBURI MIDI jump suit and STEIM's Interactive Video software Image/ine - to provide real-time control of images via a performer.

The paper includes an introduction to the Yamaha MIBURI MIDI jump suit and an overview of STEIM's Interactive Video software Image/ine. A number of different MIBURI/Image/ine interface paradigms developed by the author at STEIM are examined in light of the advantages and limitations of the two systems.

1. Introduction

> *In the mid-1990s digital cameras became more widely available to the average user, a step some regard as the 'cinematization' of video. Accompanied by more sophisticated editing equipment, especially digital non-linear systems..video production attained closer links to cinema. (Rush 1999 p.164)*

These advances have opened the doors to a profusion of different systems involving video, sound and movement. Many of these developments – Eric Singer: *Cyclops*, *Palindrome*: *EyeCon*, David Rokeby*: Very Nervous System* etc - have focussed on the use of video as a motion sensing input device for controlling sound. The development of the opposite paradigm – using movement to control video – has been driven by Techno/Dance/VJ-oriented systems such as *Arkaos* and *U&I software*'s *Videodelic*. The 'art' music/dance world has also responded to these developments with software such as *NATO (0f0003.MASHIN3KUNST), Isadora (Troika Ranch)* and *Image/ine (STEIM).*

The other side of the movement/video equation is the method in which information is gathered from dancer(s)' movements to control interactive video. Gestural controller innovator Axel Mulder in his 1994 Technical report *Human Movement Tracking Technology* defines three types of systems for motion sensing:

- inside-in - sensor(s) and source(s) that are both on the body;
- inside-out - on-body sensors that sense artificial external sources;
- outside-in - external sensors that sense artificial sources on the body (Mulder 1994).

Cyclops, *EyeCon* and *Very Nervous System* are clearly outside-in type systems. My current work with interactive video follows the 'inside-in' paradigm, where the dancer's gestures are transduced via Yamaha's MIBURI MIDI jump suit – a product of the company's Tokyo-based experimental division, and the Video is processed by STEIM's Image/ine software.

2. Yahama's MIBURI MIDI jump suit

The MIBURI was released commercially by the Yahama Company's Tokyo-based experimental division in 1994. It conforms to what Todd Winkler refers to as the 'body sensor' group of controllers (the other are spatial sensors, acoustic models and 'new instruments') (Winkler p. 315-8).

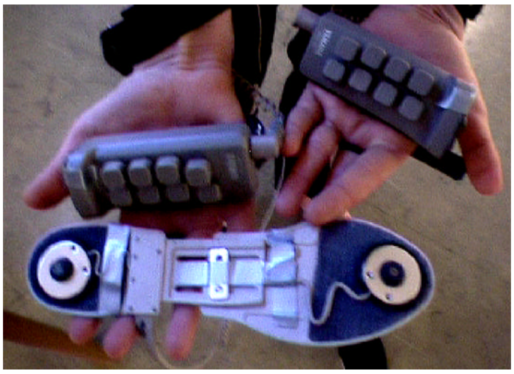

Figure 1. MIBURI handgrips and foot-sensor

The MIBURI system comprises a vest with embedded flex-sensors, two hand-grips, shoe inserts with pressure sensors, and a belt-worn signal distribution unit joined by a cable to a small synthesizer/MIDI converter. A wireless version, conforming to Japanese wireless frequency regulations was available within Japan only.

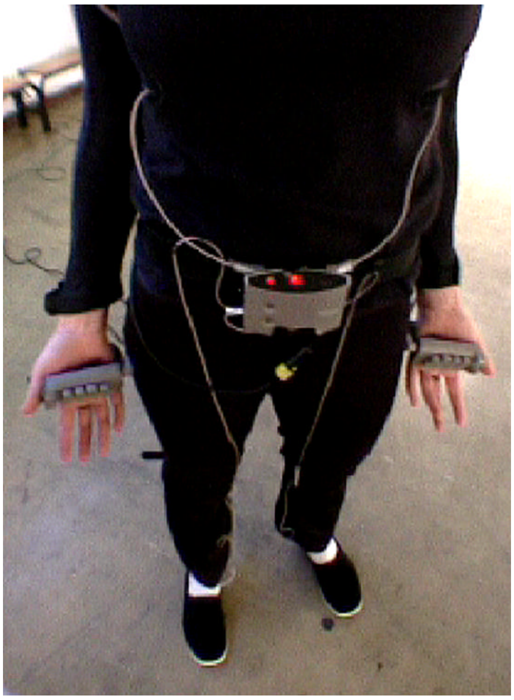

Figure 2. The MIBURI

The MIBURI's belt unit processes data from the sensors into MIDI pitch and velocity information. The unit can be programmed to interpret the data using three 'trigger' modes: 'Cross-point' mode; 'Stop' mode and 'All' a combination of both modes. 'Cross-point' mode measures the speed of the tranducer's flexion as it traverses its zero point (when the flex sensor is straight). 'Stop' mode sends note and maximum velocity values at the conclusion of a gesture. 'All' interprets sensor data in both modes simultaneously (Yamaha MIBURI Manual p.41).

Sensor type	**No.**	**trigger**	**MIDI**
pressure sensors	8	finger tips	Notes (8) / Velocity
pressure sensors	4	heel and toe	Notes (4) / Velocity
'flex' sensors	6	arm joints	Notes (12†)/ Velocity
'mod-wheel' style benders	2	thumbs	Pitchbend/ Controller
buttons	2	thumbs	Program change +/-

Figure 3. The MIBURI 's sensor array

† The six 'flex' sensors send 12 notes – this is because they measure inward and outward movement of each joint as separate notes.

The mapping of each sensor is highly programmable. Each sensor can be mapped on the synthesizer unit to any MIDI note, interpreted in any of the three modes outlined above according to 48 different response modes. The response modes (preset by Yamaha) define the manner in which the sensor's output is graphed to velocity. All the above definitions are components of a single Map 'Preset', there are 32 programmable preset positions available.

These features make the MIBURI extremely effective as a controller. However the MIBURI's synthesizer unit is limited in its possibilities as a sound source and more importantly is only able to process gestures in a direct relationship to the sounds they produce.

Teresa Marrin in her 1996 MIT Master's thesis *Toward an Understanding of Musical Gesture* assesses these shortcomings:

> *'One place where the Miburi might be strengthened, however, is in its reliance on triggers; the synthesizer unit which plays the music has no "intelligence" other than to provide*

a one-to-one mapping between actions and sounds. In fact, the Miburi's responses are as direct as the Theremin's. But because it has so many more controls available for the user than the Theremin it might be advantageous to make use of those controls to create or shape more complex, autonomous musical processes

(Marrin 1996).

For this reason, I have chosen to combine the MIBURI with more sophisticated sound sources and software-based interactive mapping in MAX/MSP and Image/ine. (Note: MAX/MSP mapping will not be discussed in this paper.) The need to 'tether' the MIBURI to its synthesizer box is also clearly a drawback for movement detection and a restriction for the dancer. However, the MIBURI has the robust design, and very predictable sensor output that might be expected from one of the principal electronic instrument manufacturers.

The Miburi represents a rare excursion by a large, commercial organization into the uncharted waters of interface design.

(Marrin 1996)

3. Image/ine

Image/ine was developed at STEIM in the mid-1990s by Tom Demeyer. It is designed to allow the interactive manipulation and processing of live video, QuickTime movies, still images and text in real-time. Image/ine allows control of a number of parameters via a range of controllers including MIDI, the computer ASCII keyboard, mouse position (also the output of a WACOM table) and audio-in.

Image/ine allows three layers or channels of images to be separately manipulated. The layers are called *Foreground*, *Background* and *Displace Source* and each is capable of supporting all of the image sources listed above. The image input sources are viewed via the formats Video Signal, QuickTime Movie, Buffer, Text, Frame or Drawing.

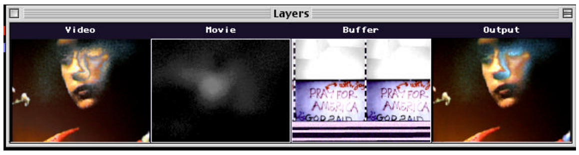

Figure 4. Image/ine's Layer Window from the author's *your sky is filled with billboards of the sky*

Still images are stored in an image buffer. The dimensions of the buffer are user-assignable. It can be navigated in two ways:

- **Frame Mode** divides the buffer into distinct frames (the example below is a 3x3 frame buffer) and allows the user to jump to any particular frame

- **Buffer Mode** in which a Frame with assignable dimensions is navigated around the buffer using the x/y co-ordinates of the Frame's centre.

The strength of *Image/ine*, however, lies in its ability to manipulate and blend these layers in real-time. Real-time processing is determined in the Mapping Window which defines more than 60 editable parameters and what is controlling them. These parameters include the input sources for the three layers, frame selection and buffer frame size and position as described above, as well as parameters controlling the mix between the layers, QuickTime movie playback, Drawing, Text and a number of Photoshop-like processes.

Unsurprisingly, considering its development at STEIM and developer Tom Demeyer, Image/ine is highly programmable. Each Mapping Window parameter has the option of 14 bit mapping of the control source to the parameter source via editable tables. Multiple arrangements of performance parameters can be saved via *Presets* and *Display States* that can create a 'snapshot' of the current state of parameter assignments values.

Figure 5. Image/ine's Buffer Window
from the author's Image/ine interface for Amy Knoles' *Men in the Cities* 1990/2002
(images by Robert Longo*)*

The development of *Image/ine* came to an abrupt halt in 2000 when Demeyer resigned from STEIM, and has only recently begun to proceed with external development on a voluntary basis. This hiatus has resulted in the software being burdened with some difficulties. The most important of these are that the screen resolution has not kept up with rapid developments in digital video - *Image/ine* still operates best at 480x320 pixel resolution and therefore suffers from a pixelated 'jaggy' appearance – it is also not optimized for multi-tasking and is intolerant of other software running simultaneously.

4. Examplar Mappings

In December 2001 I had the opportunity to be in-residence at STEIM, working with Daniel Schorno on Image/ine interfaces. The pieces I have made since then focus on three of the software's strengths: the rapid access to still images available in the *Buffer* and *Frame* Modes; the use of QuickTime video in a similar way and the ability to process and blend live video.

4.1 Men in the Cities and Scan

My first foray into *Image/ine* programming was for Californian percussionist Amy Knoles' *Men in the Cities* (1990), a work she wrote in response to an invitation to perform a computer/electronic realization of the work of artist Robert Longo at the LA County Museum of Art.

The interactive video component for *Men in the Cities* used a relatively straightforward method in which nine drawings by Longo were arranged in the buffer (Figure 5). Two *Presets* corresponding to the principal sections in the work displayed the images in different ways. In *Preset* 1, MIDI note numbers simply selected a particular frame in *the Foreground Layer*, and in *Preset* 2 this same process is blended with a *Background Layer* comprising the *Buffer Frame*. The *Buffer Frame*'s position (both x and y) was determined by note's velocity and its size by pitchbend.

The resultant images (Figure 6 is an example), dynamically combine the two *Layers*, 'keying' them so that the *Background Layer* appears 'through' the white regions of the *Foreground Layer*. These two modes corresponded to sections of Knoles' composition, in which the musical material was sparse and clear (*Preset* 1) or overlaid and dense evoking the textural complexity of city life (*Preset* 2).

Figure 6. *Men in the Cities: Preset* 1 blended *Foreground* and *Background* Layers

Use of the rapid transitions possible from the Buffer also featured in my most recent Image/ine interface for the *skadada* Dance company's dancework *Scan* (2002). Like many 'New Media" works *Scan* had a number of authors: the images were provided by *Tissue Culture and Art* (Oron Catts, Ionat Zurr and Guy Ben-Ary), audio samples by John Patterson and choreography by *skadada* (Jon Burtt and Katie Lavers).

The interactive component of scan comprised only the first six minutes or so of the performance. The 25 microscopic images were accessed using a method roughly opposite to that of *Men in the Cities*. This time MIDI note number determined the *Buffer Frame* size with a result that cueing consecutive note numbers zoomed-in or out from a particular point in the *Buffer* a situation that was factored into the choreography. The zoom point was defined by the MIBURI's thumb operated toggles: left thumb gave 'x' coordinates and right thumb gave 'y' coordinates.

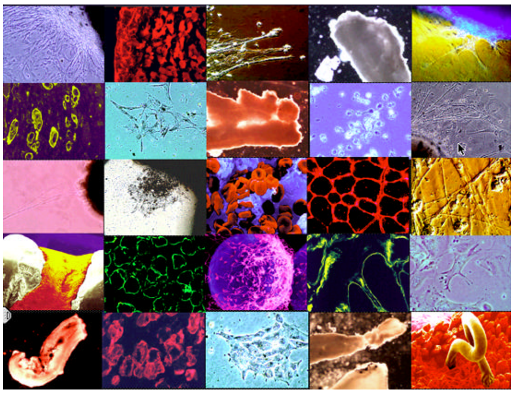

Figure 8. *Scan* Buffer
(Images by Tissue Culture and Art)

The extreme speed of the response time of the MIBURI/Image/ine combination gave the allusion of a hyper-accelerated examination of slides by an electron-microscope. The theme of surveillance/ forensic examination was also a feature of similar sections in the *Image/ine* components of two other works of mine. *Delicious Ironies (Noir)* where a down-and-out private eye dancer manipulates images and sounds from my opera *Rendez-vous* (1995) and dance piece Noir (2000). (*Noir* itself used the MIBURI to interactively control sound and a MIDI lighting desk) The other was Perth digital artist Vikki Wilson's dark meditation on the Claremont Serial Killer *a throw of the dice can never abolish chance* with images by Vikki Wilson and text by Vikki Wilson and Erin Hefferon.

4.3 a throw of the dice

Vikki Wilson initiated this project concerning the serial killer 'still at large' in the urban family paradise of Perth. The work highlights the contrast between the city's naïve sun-bathed exterior and the dark reality of continued disappearance and murder of young women. It also seeks to highlight the ambiguities of modern society's simultaneous repulsion and fascination with serial killing (as evidenced by 100s of Hollywood movies).

The final version of the work is planned to be a non-linear video installation using Finale Cut Pro as the engine to randomly juxtapose text, images and sound. Vikki had allowed me to experiment at STEIM with an interactive version of the piece where the juxtaposition of images is made by a live performer using the MIBURI as a kind of 'all-body' VJ console.

In the Image/ine component of *a throw of the dice* QuickTime videos provided by Vikki Wilson were manipulated by the dancer. The MIBURI's right hand buttons defined the current clip and the pitchbender became a fast-forward/reverse toggle causing the frame-rate to advance when pushed up and decrement when pushed back. Releasing the toggle to its centre position caused the video to stop on a particular frame.

The dancer's other hand defines the arrangement of *Layer*s – principally whether the video is keyed against images in the *Buffer*. Note number determined *Buffer Frame* position, and velocity determined *Buffer Frame* size as well as the degree of keying between the Video and Buffer. In essence the dancer transitions between victim and voyeur through her control of the medium and non-control on its content.

We plan to expand the work considerably in the future, creating a hypertext-like non-linear maze of pathways in which the dancer can access only material according to her previous choices.

4.4 your sky is filled with billboards of the sky

your sky [2002] was developed for the REV festival at the Brisbane Powerhouse in April 02. Thematically, it is a sibling work to my song cycle (for soprano and DVD) **songs of virtual love+war** [1998] – exploring in a more abstract format similar issues of defining identity in the context of a world increasingly comprised of simulated experiences. In **your sky** both the sound

(MAX/MSP) and video elements are manipulated by the dancer, giving the impression that she exists in a loop in which she is called upon to respond authentically to an environment entirely under her own control.

Figure 9. Live Video Image displacement in *your sky*

The interactive nature of the technology used in **your sky** opens the possibility for an open-ended non-linear formal structure, but in this version a mix of improvisation and predetermined structure was applied. The determined formal structure rested on three principal sections: Breath, Text and Exterior each a different *Image/ine Preset*. **your sky** used a live video feed for the bulk of the performance. A brief explanation of the *Breath* section will serve to demonstrate the degree of image control achieved.

In *Breath*, the *Layers* were arranged: Foreground: video, Background: QuickTime Movie and Displace Source: buffer. A large number of parameters were assigned to the MIBURI: Note number: *Buffer* displace angle (the maximum amount the image can be twisted on a central axis); Note Velocity: Buffer Frame pan ‘x’; Controller 1 : Buffer Frame Size as well as image hue/saturation and the degree of Keying between Layers; Pitchbend: Buffer Frame pan ‘y’ and its displace amount (how much the image can is twisted on a central axis within the ‘displace angle’). The bender also determined the luminousity of the Video Input.

Figure 10. Live Video Image displacement, hue/saturation and keying against a QuickTime video in *your sky*

5. Conclusion

Interactive video is still in its infancy. There are currently a number of competing systems and it is likely that as technologies like Cinema, Television and the Internet converge the desire for reliable, high resolution software will continue. Large, well-resourced companies such as Avid and Adobe are also likely to extend their products further in this direction. There are still a number of limitations inherent in the combination of the MIBURI and Image/ine, most importantly question marks about the continuing development of both products. However together, and especially in combination with interactive audio capabilities (which I have not discussed here) they create an effective and responsive system for the development of work in this field

References

Demeyer, T. 1997. Image/ine Manual. Amsterdam, Holland: STEIM Foundation

Marrin, T.A., 1996. *Toward an Understanding of Musical Gesture: Mapping Expressive Intention with the Digital Baton*, Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of Master of Science in Media Arts and Sciences at the Massachusetts Institute of Technology. (Available at web.media.mit.edu /~marrin/Thesis.htm)

Mulder, A., 1994. *Human movement tracking technology*, Hand Centered Studies of Human Movement Project, Technical Report 94-1 http://www.cs.sfu.ca/~amulder/personal/vmi/ HMTT.pub.html

Rush, M., 1999. New Media in late 20th-Century Art. London, United Kingdom: Thames and Hudson

Yamaha 1996. MIBURI R3 Manual. Tokyo, Japan: Yamaha Corporation. (In Japanese, no author listed)

Winkler, T., 1998. Composing Interactive Music: techniques and ideas using Max. Cambridge Massachusetts, MIT Press.

Other Internet References

0f0003.MASHIN3KUNST: Nato
www.bootsquad.com/nato/nato02.html

Arkaos
www.arkaos.net/site/en/index.html

Palindrome: EyeCon
www.palindrome.de/

U&I Software: Videodelic
www.uisoftware.com/videodelic/index.html

David Rokeby: Very Nervous System
www.interlog.com/~drokeby/home.html

Eric Singer: Cyclops
www.ericsinger.com/

STEIM: Image/ine
www.image-ine.org

Troika Ranch: Isadora
www.troikaranch.org/isadora.html