

Special Section

Transactions

Guest Editor: Ernest Edmonds

Transactions, edited by Ernest Edmonds of Creativity and Cognition Studios in Sydney, Australia, publishes short refereed papers and provides a fast track to publishing key new results, ideas and developments in practice. Researchers and practicing artists are encouraged to report on new work and new concepts through Transactions.

Papers are restricted to two pages of published material and must be submitted electronically in final camera-ready form at <www.leonardo-transactions.com>, where formatting instructions and a template may be found.

Leonardo Transactions are published camera ready as provided by the author and approved by the peer reviewer(s). The author is fully responsible for any errors in the text or imperfections in form or language.

The first Call for Papers is for contributions that address one of the following two areas: (1) The impact of the interactive arts on audience experience; (2) New art practice that also advances science or technology.

Special thanks to our Transactions reviewers: Paul Brown, Computer Arts Society; Dave Buraston, Creativity and Cognition Studios; Linda Candy, University of Technology, Sydney; Sean Cubitt, University of Melbourne; Ross Gibson, University of Technology, Sydney; Sue Gollifer, University of Brighton; Janis Jeffries, Goldsmiths College, University of London; Mary Lou Maher, National Science Foundation, USA; Lev Manovich, University of California, San Diego; Eduardo Miranda, University of Plymouth; Bonnie Mitchell, Bowling Green State University; Kumiyo Nakakoji, University of Tokyo; Frieder Nake, University of Bremen; Jack Ox, University of New Mexico; Doug Riecken, IBM; Cynthia Beth Rubin, Rhode Island School of Design; Stephen Scrivener, Chelsea College of Art and Design; Jeffrey Shaw, University of New South Wales; Mike Stubbs, FACT Liverpool.

Transactions staff: *Managing Editor:* Sarah Moss. *Webmaster:* Damian Hills.

HARNESSING THE ENACTIVE KNOWLEDGE OF MUSICIANS TO ALLOW THE REAL-TIME PERFORMANCE OF CORRELATED MUSIC AND COMPUTER GRAPHICS

Ilias Bergstrom and R. Beau Lotto,
University College London, Institute of
Ophthalmology, 11-43 Bath Street,
London, EC1V 9EL, United Kingdom.
E-mail: <i.bergstrom@ucl.ac.uk>.

Keywords

Real-time performance, enactive interfaces, visual music, new media, digital musical instruments, mapping, colour organs, programming as art.

Artists and scientists have a perpetual interest in the relationship between music and art. As technology has progressed, so too have the tools that allow the practical exploration of this relationship. Today, artists in many disparate fields occupy themselves with producing animated visual art that is correlated with music (called ‘visual music’). Despite this interest and advancing technology, there still is no tool that will allow one to perform visual music in *real-time* with a *significant level of control*. Here we propose a system that would enable a group or individual to perform live ‘visual music’ using the musical instrument(s) itself as the primary source of control information for the graphics. The hypothesis driving this choice of interface is that, by connecting musical control data (i.e. scales, notes, chords, tempo, force, sound timbre and volume) to graphical control information (a process called *mapping* [6]), a performer will be able to more readily transfer his/her enactive knowledge [1] of the instrument to creating visual music. The term *enactive knowledge* refers to knowledge that can only be acquired and manifested through action. Examples of human activities that heavily rely on enactive knowledge include dance, painting, sports, and per-

forming music. If our hypothesis is correct, this will enable a mode of musical/visual performance different from current practice, which is likely to enhance the experience of both the performer(s) and audiences.

The outcome of this process will not simply be music visualization, with the graphics being subordinate. Rather, we believe that images controlled directly by the same physical action that generates music will feedback to shape what is actually played musically by the performer. Furthermore, many people will be able to collaboratively perform a complex musical and visual experience live – in a manner that live musicians are already used to, because each performer will be influenced by the performance of the others in the group.

Background

The immediacy with which music can communicate emotion has been envied by many visual artists, most notably Wassily Kandinsky [2], who set out to recreate it in painting. The first known machine for exploring the relationship between music and visual art was Louis-Bertrand Castel’s “Clavecin oculaire” (1734); Castel implemented a modified version of the note-to-color correspondence proposed by Isaac Newton [2]. Many such systems have since followed, made to either accompany music with colour or provide a form of visual music – named “Lumia”. The term *Lumia* was coined by Thomas Wilfred, developer of the “Clavilux” color-organ (1922) [2], who, rejecting the notion of an absolute correspondence between sound and image, concentrated on generating visual compositions that were meant to be viewed alone, i.e., without musical accompaniment.

Though lacking an entirely rigid definition, what qualifies as Visual Music is sufficiently well described by Brian Evans [3] as: ‘time-based visual imagery

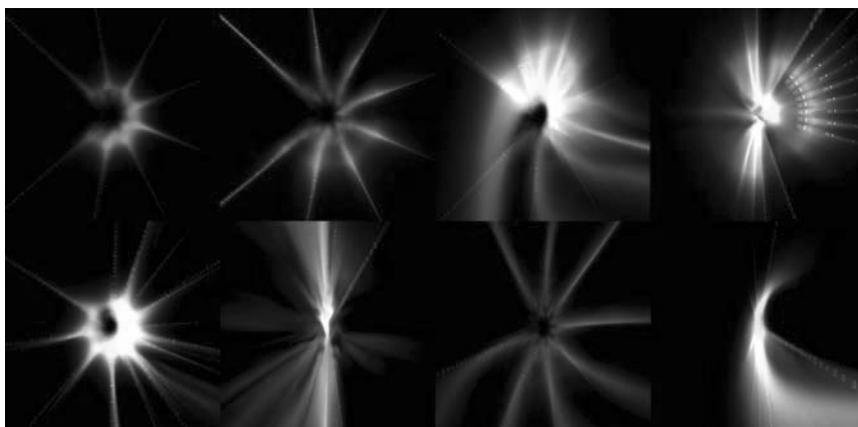
that establishes a temporal architecture in a way similar to absolute music. It is typically non-narrative and non-representational (although it need not be either). Visual Music can be accompanied by sound but can also be silent’.

In modern times, analogue video synthesizers, laser shows and more recently computer graphics have all been employed to accompany music. For instance, at live music concerts and at clubs with music played by a DJ, there are often live graphics performed by a VJ (Visual Jockey), who mixes pre-recorded video clips together, while altering the playback parameters of the individual clips, as well as processing these clips using real-time video effects. A further advance has been the recent development of computer programming environments, such as Processing [4] and Cycling74’s Max/MSP/Jitter combination, that allow the description and performance of real-time procedural graphics by non-expert programmers, and as such their use has begun making its way out of the avant-garde and into VJ sets and other ‘new media art’ performances. Despite the intense interest in visual music, and advances in its associated technologies, current tools/practices are limited in three ways: (i) highly constrained, linear mappings, (ii) limited controllability, and (iii) overly complicated process for preparing new performances.

Existing Limitations

In current practice the mapping between sound and graphics is typically reduced to the beat and amplitude of the music. What is more, such mappings quickly become highly predictable, thus restricting the use of suspended expectations [5], as well as tension and release [3], which are crucial aspects of artistic/esthetic narratives. It is known that the human perceptual system is apt at detecting correlated stimuli across modalities and fusing these into a single percept before their interpretation [8]. More detailed correlation may thus further encourage the unified experience of music and image. Furthermore, visual music performances are controlled using often highly *unintuitive* interfaces on computer screens and/or external hardware controllers (with knobs, sliders and buttons, etc., also referred to as non-musical controllers). Together these limitations mean that controlling visual music is a far cry from the level of control that musicians have of their instruments.

Fig. 1. Sequence of images produced using prototype. (© Ilias Bergstrom)



Finally, existing multimedia programming environments prevent a program created by one performer from being easily used by other performers for achieving their individual esthetic goals, making such programs very limited in their usability.

Proposed System

The system proposed here overcomes these three limitations, while simultaneously introducing several novel concepts for the real-time visual music performance using computers. Here we will describe the proposed system and the concepts that it engenders.

The system consists of three modules: one to host the graphics rendering, a second module to gather musical control information from musical gesture controllers [6], and a third to adapt the ‘mutable mappings’ between the parameter spaces of the control data input and of the graphics rendering.

Although mixing multiple layers of moving graphics, as in VJing, will be retained, these graphics are not constrained to pre-rendered video clips, but use the output of real-time visual synthesizers (‘synths’) that run in parallel within the main host application. Each visual synthesizer is a program that renders a particular visual effect, the control parameters of which are all accessible on-line during a performance, so that the appearance of the visual is animated over time. ‘Synths’ can be created as Processing ‘Sketches’ [4], using a processing library provided that enables the sketch to work within our host application. Alternatively, synths can come from a collection of sketches provided with the host, or from other users. In this way we facilitate the performance of advanced real-time computer graphics by non-programmers by encapsulating the programming complexity in flexible modules that are easily exchanged and managed.

More importantly, the parameters of the visual synthesizers are not primarily controlled using a screen and/or non-musical controllers more generally. Both visuals and music are instead controlled using the instruments themselves: using, for instance, scales, notes, chords, tempo, force, sound timbre and volume. Additional non-musical controllers, such as banks of sliders, knobs, switches, light sensors, touch-screens, etc., may still be used as further sources of control data. Possible input sources include musical and non-musical data from MIDI or OSC [6], as well as multi-channel audio.

MIDI technology is a well-established standard for transmitting musical information, while OSC is its recently developed successor. There are few instruments whose output cannot be translated to MIDI/OSC, while many non-musical controllers also produce data in these formats. In addition, any monophonic sound source (e.g. the human voice) can be used to produce musical information using pitch-identification [6]. Using the musical instrument as a user interface will allow a previously unachievable level of control over real-time computer graphics, since the method takes advantage of the enactive knowledge [1] that playing a musical instrument affords.

Of particular importance for the proposed system is that there will be *no hardwired mapping* between the musical control information and the control parameters of the visual synthesizers. Instead the mapping(s) will be mutable: created and modified before and even during a performance, thus enabling performers to construct a non-linear, dynamic and free-form narrative in the correlation between sound and vision, thereby introducing a further means of expression.

Given mutable mappings and the retention of non-musical controllers in the system, an additional role in the performing group, dubbed the ‘visual mixing engineer’, will be created, who is responsible for designing the system’s initial state: (i) the visual synthesizers, (ii) animations to be controlled, (iii) the mutable mappings, as well as gradually altering connections between the streams of control input and control parameters of the visual synthesizers during performance. At any instant, one, a subset, or even all connections may evolve. In this sense, his/her role is analogous to that of the audio mixing engineer, who focuses on the overall (visual) music performance, while the musicians concentrate on their individual parts.

Naturally and finally the proposed system will be useable in all contexts where live music is performed, either using traditional direct manipulation instruments or conducting gesture controllers [6].

Discussion

Here we have provided a high-level description of a novel system that would allow the live performance of ‘visual music’ with a level of control significantly greater than is currently possible. In current practice only limited correla-

tions between the aural and visual, and limited real-time control over the graphics drawn are possible. The novel methods introduced here directly address these identified limitations of current practices. The proposed system will allow performers to control real time generated imagery using control information directly derived from the musical instrument(s). We will furthermore enable artists without the skills of computer programming to significantly influence the content of their visual performance, while also keeping the programming of new graphics algorithms accessible to those seeking greater creative freedom.

More generally, our proposed system (and our programme of research more generally) may in the long run help address a shortcoming that computers often have when employed as creative tools: the creative process is rarely in real-time. This is especially true in the field of computer graphics. Recent advances have allowed us to display real-time, interactive graphics of great visual quality, but the process for creating and controlling their animation is still highly time consuming and technical. Recent research clearly shows the advantages of creating user interfaces that acknowledge the understanding of human cognition as being highly embodied in the interaction between humans and computers [1] [7]. This however has not yet been sufficiently explored in practice.

Finally, the conduct of performing by continuously altering the mapping between two parameter spaces is in itself a novel mode of performance, equally applicable to many other contexts.

References

1. Enactive network: <www.enactivenetwork.org>, accessed 7 March 2008.
2. K. Peacock, “Instruments to Perform Color-Music: Two Centuries of Technological Experimentation,” *Leonardo*, vol. 21, 1988.
3. B. Evans, “Foundations of a Visual Music,” *Computer Music Journal*, vol. 29, 2005.
4. C. Reas and B. Fry, *Processing: A Programming Handbook for Visual Designers and Artists*, MIT Press, 2007.
5. N. Steinbeis, S. Koelsch, and J.A. Sloboda, “The Role of Harmonic Expectancy Violations in Musical Emotions: Evidence from Subjective, Physiological, and Neural Responses,” *Journal of Cognitive Neuroscience*, vol. 18, 2006.
6. E.R. Miranda and M.M. Wanderley, *New Digital Musical Instruments: Control and Interaction Beyond the Keyboard*, AR Editions, 2006.
7. P. Dourish, “Where the Action Is: The Foundations of Embodied Interaction”, MIT Press, 2001.
8. P. Larsson, “Virtually Hearing, Seeing, and Being: Room Acoustics, Presence, and Audiovisual Environments,” *Doctoral Thesis*, 2005.

WIDENING UNEQUAL TEMPERED MICROTONAL PITCH SPACE FOR METAPHORIC AND COGNITIVE PURPOSES WITH NEW PRIME NUMBER SCALES

Roger T. Dean, australYSIS, Sydney, and MARCS Auditory Laboratories, University of Western Sydney, NSW 1797, Australia.
E-mail: <roger.dean@uws.edu.au>.

Abstract

I define a new set of microtonal scales based on the prime number series, and containing 41 to 91 pitches spread over the whole audible range, rather than subdividing the octave. I designed these scales for metaphorical purposes, and applied one within my performance piece *Ubasuteyama* (2008), written with Hazel Smith, for speaker, computer sound and digital processing. Simple timbres using partials bearing prime number ratios to their fundamental were also used to embody the scale. The scales and timbres will be amongst the subjects of cognitive studies of pitch combinations, of large and unbroken pitch intervals in melodies, and of the relation between scale and timbre.

“Our senses are piano keys upon which surrounding nature plays, and which often play upon themselves.” [1]

Microtonal scales may be viewed as scales with sequences of pitches in which some or all steps are significantly smaller than the Western semitone [2]. The Western semitone, the core of ‘equal temperament’, is a pitch interval in which the higher pitch has a frequency ratio to the lower which equals the twelfth root of 2, such that twelve such equal ratio steps in succession create a frequency ratio of 2 with the initial note, and constitute the so-called octave [3]. Octaves may be perceived as bearing a special relationship to each other, so that melodic pitch patterns which recur within different octaves are perceived as equivalent, i.e. repetitious.

The psychology of such ‘octave equivalence’ is poorly understood, yet even cross-culturally, virtually all microtonal scales use repetition at the octave. A few have patterns which repeat with a 2-octave cycle, and even fewer again (the so-called Pierce-Bohlen scale in particular) repeat at an octave and a fifth and do not include a pitch corresponding to the octave. Possible exceptions to this rule are found in the music of indigenous Aboriginal Australians from the Western Desert areas [4].

There has been a considerable interest in creating scales involving prime number relationships. Prime numbers are those which are only divisible by one and by the number itself (such as 2, 3, 5,

7, 11, 13, 17, 19, 23, 29, 31 ... 61, 67). The 91st in the series of primes is 467, and the largest numerical jump between adjacent primes up to that point is 14, between 293 and 307. The jumps become generally wider after the first 9 prime numbers.

In the Scala database of c. 3500 largely microtonal scale structures, maintained in the Netherlands by the Huygens-Fokker Institute for microtonal music [5], most if not all the scales involving prime numbers use them as the number of equal ratio frequency steps in the scales to which they are attached, generally subdividing the octave into the chosen prime number of equal ratio steps. Such scales are ‘equal tempered’ scales, like the conventional Western 12-tone scale.

In contrast, I have designed a category of scales in which the chosen number of prime numbers are used in their increasing series successively as multiplier to construct a scale covering the whole of the audible range, from c. 15Hz to c. 4500Hz. These scales have no octave repetitions (because of the definition of prime number), nor any other form of regularly repeating pattern. The scales qualify as ‘just’ tunings, since the definition of this is that every frequency ratio be expressible as a series of powers of prime numbers (in my case, the power need only be 1). The (now non-repeating) scale space is widened from the octave to the whole audible range, corresponding roughly to that of the piano keyboard. In other words, there is just one rendering of the scale across the whole of the audible range, instead of >7 renderings of the Western equal tempered scale.

The scales comprise any chosen and sufficient (perhaps greater than 12) number of primes from the first to the 91st, but the examples discussed here consist of 41, 51, 61, 71, 81 or 91 successive primes. The prime numbers are used as successive multipliers of a base frequency (from 10 to 20Hz), such that the resultant frequency range is as specified above. The first prime number is 2, so the lowest pitch is (2*base-frequency). Note that the multiplication principle (always using the base frequency as one of the two numbers multiplied together) is quite different from that of equal temperament, in which each pitch is the base for the next multiplication, thus achieving consistent pitch ratios between each successive pair of pitches. In my scales, there are several jumps between adjacent prime numbers

which correspond to 2, 3, 4 or 6 times the base frequency, thus giving recurrent pitch differences of 20, 30, 40, or 60Hz for a scale with base frequency 10Hz, and such linear differences thus appear to some extent throughout the whole pitch range. However, there are no octaves, though some ratios between pairs of distal pitches are quite close to 2:1. Thus the scales involve ratios to the base pitch, but generate irregular linear steps over the whole of pitch space: they constitute unequal temperament. The core 6 examples described have been admitted to the Scala database (August 2008).

Performance and Compositional Metaphoric Purposes of the New Prime Number Unequal Temperament Scales in *Ubasuteyama*

I constructed these scales both for compositional and empirical purposes. There is considerable evidence that listeners perceive music partly by means of physical analogies, for example, such that ‘higher’ pitches are taken as higher in physical space [6], and pitch movement may construct virtual spatial movement [7]. There are also perceived relationships with colour, mass and other properties [8].

I envisaged that this pitch-physical height relationship could be perturbed and perhaps extended by the wider pitch jumps in my scales, in comparison with equal-tempered twelve tone systems on which the described perceptions are largely based. The slow increment of pitch as one progresses through the lowest notes of the scales is succeeded by faster rates of increment, and this too might transform the metaphorical relation with physical elevation, and perhaps its perception.

I also anticipated that in making music with such scales, the composer or improviser might choose to make melodic contours covering a much wider range than is common in equal-tempered pitch space, where virtually all melodies are contained either within one-, or occasional two-octave pitch ranges and repeat in the remaining registers. Again, such unusual kinds of melodies might be used for distinct metaphorical and affective intents.

Pioneers of microtonal Western microtonal tuning such as Wendy Carlos [9] and the composer-theorist William Sethares [2] have suggested that those using microtonal scales might also want to mutually adapt the timbres and the scales which they realize. One discussion

concerns the possibility of controlling perceived roughness or 'jangliness' of the scales by adjusting the spectral properties of the timbres. It is notable that a key feature of the microtonal scales embodied in the metallophones of Balinese music is its percussive 'jangle' [10]. There seems to have been little overt systematic use of this idea in pitch-based music, even though spectral morphology is a key concern of many electroacoustic music composers. In my application so far of the new prime number scales, I have developed simple FM and noise-based timbres which were synthesized in real-time using a Max/MSP patch. The spectral components were chosen to emphasise prime number components, for example by using small number primes to control the timbre of a simple FM synthesis patch. Specifically, in this case the prime numbers controlled the degree to which a sound is harmonic (consisting of overtones from the harmonic series) or inharmonic.

These techniques were used in a new piece *Ubasuteyama* (2008), by Hazel Smith and myself. The Japanese word *ubasuteyama* means, roughly, 'grandmother thrown upon the mountain'. The piece involves a poem by Hazel which relates to the Japanese rural and historic tradition in which an old woman is carried up an exposed mountain near to her village by her son, and left there to die. The highly metaphoric text is spoken in performance, and the recorded speech is timbrally and texturally manipulated in real-time, and this process continues after the text's enunciation.

Japanese traditional music uses microtonal features far more than does most Western music. Correspondingly, my new 81-prime scale, and the related pitch and timbral structures discussed above, were used compositionally in the performance score. The purpose of their use was to seek various impressions of progressive motion upwards in height, and of 'evaporation' from a high space into another physical unknown space, a virtual after-world. The scale was realized by several synthesis patches on a computer, and in some cases discrete FM-pitches were played in response to MIDI-input from a keyboard. The piece has been premiered and recorded by my ensemble *austraLYSIS* and released on CD (2008) [11], and an mp3 version is available online.

An interesting future compositional approach will be to form morphing relationships between pre-existent microtonal scales, such as those of Bali or

Japan, and Western ones, such as my prime number scales, and to use these relationships structurally and motivically. An electroacoustic exchange between microtonal and equal tempered might also be effected, as in a sense it is in the more extreme microtonal inflections of post-Coltrane jazz improvisation, such as in the contemporary saxophone improvisations of Evan Parker.

Some Empirical and Cognitive Aspects of the Prime Number Scales

My scales raise questions about perception and recognition of melodic patterns generated over multi-octave pitch space. Even when such melodies are (as is rare) constructed on equal temperament instruments such as the piano, they commonly involve conventions of 'return to the mean', in which after a large jump in one direction several small steps in the opposite direction gradually bring the pitch range back to a central norm [12]. This is less likely or feasible with the non-octave scales for the reasons indicated already, amongst others.

The melodic structures in *Ubasuteyama* are purposely realized with monophonic FM sounds (as described above) and accompanied by noise-derived textures (also bearing imprints of the prime number relationships in their spectra) and sometimes by speech components or derivations. The physical height perceptual implications of large microtonal intervals deserves study. But more generally, the issue of perception of chords constructed with these new and other microtonal intervals is also of interest. We understand little of the possible perceptual impact of acoustic 'roughness' (which may be a physical property of overlapping spectral energies, as originally proposed by Helmholtz [13]). This requires study in relation to more familiar $\frac{1}{4}$ and $\frac{1}{8}$ th tone equal tempered scales, or other tuning systems, just as it does with Balinese pitch combinations and with those from the new scales described here. In view of the common occurrence of prime number structures within physical objects, it will also be of interest to determine whether there are any environmental or speech sounds which express some of the pitch relationships incorporated in my or other microtonal tuning systems.

Such empirical studies will not only be of benefit to composers such as myself, but also provide understanding rele-

vant to music perception by listeners more generally. We are undertaking such work in my laboratory presently, while other collaborative work by Greg Scheimer, Emery Schubert, Richard Parncutt and others is assessing the precision with which singers or acoustic instrumentalists can learn to perform microtonal tunings in music. Whatever the outcome of that work, the feasibility of precise control in computer-interactive music remains and can contribute exciting musical and scientific ideas and understandings in the future.

References and Notes

1. Denis Diderot, *Conversation between D'Alembert and Diderot* (1769); quoted in notes to Fyodor Dostoevsky, *Notes from the Underground*, R. Pevear and L. Volokhonsky, trans. (London: Vintage Books, 1993), p.133, in relation to the 'metaphor of the piano key'.
2. W. Sethares, *Tuning, Timbre, Spectrum, Scale* (London: Springer Verlag, 1998).
3. Carol L. Krumhansl, *Cognitive Foundations of Musical Pitch* (New York: Oxford Univ. Press, 1990).
4. U. Will and C. Ellis, "A Re-Analyzed Australian Western Desert Song: Frequency Performance and Interval Structure," *Ethnomusicology* 40.2 (1996) pp.187-222.
5. The Scala website of microtonal scales (undated), <http://www.xs4all.nl/~huygensf/>, accessed August 2008.
6. S. Ferguson and D. Cabrera, "Vertical Localization of Sound from Multi-Way Loudspeakers," *J. Audio Eng. Soc* 53 (2005) pp.163-73.
7. Z. Eitan and R.Y. Granot, "How Music Moves: Musical Parameters and Listeners' Images of Motion," *Music Perception* 23 (2006) pp. 221-247.
8. N.Sagiv and J. Ward, "Crossmodal Interactions: Lessons from Synesthesia," *Progress in Brain Research* 155 (2006) pp. 259-271.
9. W. Carlos, "Tuning: At the Crossroads," *Computer Music Journal* 11.1 (1987) pp. 29-43.
10. J.Rahn, "Javanese Pélog Tunings Reconsidered," *Yearbook of the International Folk Music Council* 10 (1978) pp. 69-82.
11. R.T.Dean and H. Smith, *Ubasuteyama* (2008), musical performance piece, recording released on CD by Wirripang Wirr011 (2008). See <www.australysis.com> for further information and links to mp3 material.
12. D. Huron, *Sweet Anticipation* (Cambridge, MA: MIT Press, 2006).
13. H. von Helmholtz, *On the Sensations of Tone as a Physiological Basis for the Theory of Music* (translated from the German), (Whitefish, MT: Kessinger Publishing, 2005; original published in 1862).

AUGMENTING CREATIVE REALITIES: THE SECOND LIFE PERFORMANCE PROJECT

Kathryn Farley, Michael Nitsche, Jay Bolter and Blair MacIntyre, School of Literature, Communication and Culture, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332-0165. E-mail: <kfarley6@mail.gatech.edu>, <michael.nitsche@lcc.gatech.edu>, <jay.bolter@lcc.gatech.edu>, <blair@cc.gatech.edu>.

Abstract

This article discusses the development of a new interface that allows for the creation of mixed-reality performances. It details the features of the new technology, charts the ways in which the interface has been used in rehearsal sessions and describes how the technology functions as an innovative tool for creative expression. *Key words:* mixed reality, performance, MMO, Augmented Reality, Second Life.

Massive multi-user online (MMO) worlds, such as *Second Life*, have the potential to act as facilitators of cross-media and cross-cultural exchange, providing locations for newly forming communities. These communities often grow from a wide range of new possibilities beyond the conditions of the physical world. Questions of gender representation, race, social status, and political organization are mirrored but often unfold in new ways in virtual environments. We applaud these new features but are concerned that a complete reliance on the new digitally infused forms of expression might become so self-sufficient that it detaches from the more “physical” history we are surrounded by in the real world.

In addition, the potential of online worlds for performance and artistic expression remains problematic because the digital platforms are still limited. *Second Life* only serves as the most prominent and active platform that illustrates these limitations. Players are confined to existing animation systems, render pipelines, and interfaces [1]. As a result, expression is limited and “real world” topics or practices are often abstracted into the purely digital. It is our intention to focus on the friction between the purely digital and the purely “real” to investigate new forms of performances. This paper describes a new interface between digital MMO worlds and physical spaces that supports mixed media performances.

We have developed an Augmented Reality (AR) interface into *Second Life* that allows us to combine real actors and

virtual ones (avatars) on a shared visual plane. With this setup, we blend the world of *Second Life* with real world artistic practices and their expressive range. Thanks to this interface we can address the friction zone between the ephemeral virtual realms and the “real” physical human body, its history, limitation and power. Our project involves the cooperation of performance artists, such as Georgia Tech’s resident improv troupe, *Let’s Try That!*, who help us explore the range of expression such a cross-media platform allows. Based on the feedback we receive from these performers, we are constantly re-iterating our interface and developing the project further.

Our work is influenced by new advances in video game technology. Largely driven by the community of game players, video games have seen a shift towards games as a platform for self-expression. Players, in these cases, do not play the game to achieve a certain high score but to make an artistic statement and express opinions. This form of emergent play illustrates the complexities of video game textuality. The text in a video game – as in a theater piece – needs to be performed to come into existence [2]. Theorist Espen Aarseth traces this performance back to the ‘ergodic,’ muscular activity of the player during the interaction process [3]. Players are always performers. The question is: how much can they affect the arsenal of expressions for the performance?

The mechanisms of *Second Life* puts it in-between the restricted access of a game world and the freer concept of a game modification. *Second Life* actively encourages expression through its own creative construction, as users – “inhabitants” – of *Second Life* can add their own virtual creations to the existing world. For example, users can model their own 3D objects, modify the space, add their own textures, re-shape and animate their avatar. However, a lot of underlying restrictions still apply. The animation control and render system is fixed, as well as the number of polygons used in a certain location. The use of dynamic lights and detailed avatar appearances is usually very limited, as both put heavy pressure on the render engine. Ultimately, *Second Life* offers an exciting digital performance venue but presents new restrictions of expression at the same time.

Instead of relying solely on in-game adjustments, we modified the *Second Life* client program in the Augmented

Environments Laboratory under the supervision of Blair MacIntyre. Thanks to these changes we can combine AR interfaces with the *Second Life* platform. The system virtually conflates two spaces: the real and the virtual stage. Avatars active on the virtual stage in *Second Life* are combined with the real-world actors in the physical performance space. Certain objects can be present in both worlds, allowing overlap of action that supports a more effective blending of the two worlds. For example, we usually have some virtual invisible chairs in the *Second Life* space positioned at the same location that the camera sees real chairs in the physical space. Thus, in the combined image a virtual avatar seems to sit on a real chair. The mirroring of other objects and entities is used to provide for occlusion.

In practice, the performing player can use the prototype in two ways: either through a Head-Mounted Display (HMD) or using fixed cameras. When the HMD is used, the system allows for precise 6 degree-of-freedom (DOF) tracking within the lab setup of the HMD tracker. This setup depends on the range and functionality of the tracking devices. In case of the fixed camera, the player is less confined to the cumbersome helmet interface and the system can be installed in different locations relatively easily. However, the camera is restricted to one perspective and the actor feels like playing against a fake mirror (a large LCD screen) on which the composite image is projected. In both cases, tracking proved to be very precise and robust. We also record from the real stage and broadcast it into the *Second Life* stage and vice versa to allow for direct audio conversations. Finally, we stream the composited video image back into the *Second Life* environment and project it as live broadcast on virtual cinema screens.

We have run weekly rehearsals as well as special performances with the new technology during the 2007–2008 academic year. Instead of relying on a fixed performance text, we have opted to experiment with improvisation games and exercises that allow the live actors to interact with virtual characters in dynamic ways, such as the creation of scenes using improvised dialogue and movement. We have found that the hybrid performance space acted as a malleable environment in which people from various perspectives and levels of experience can come together to experiment with new options for creative expression. The visiting performance

artists (directors, art directors, actors) all agreed in the value of such a new form. We will continue to use the technology to establish a workshop-like environment for the production of short video pieces, live performances, and mixed media presentations and lectures.

There are a number of technical challenges that remain inherent in the system. For example, audio transmissions have a short delay, which seemed manageable for our actors; video transmissions into *Second Life* have a larger delay (more than 6 seconds) as they depend on the video server used (*Quicktime* streaming video). However, the setup has proven to be stable in numerous conditions as

well as affordable. Especially the single-camera version does not depend on any expensive hard- or software. Once installed, it remains mobile and relatively easy to set up. The biggest challenge is the constant stream of updates for the *Second Life* client itself. In order to follow the development of this client (which happens at times on a weekly basis), we have to recompile our modification whenever a new version is released.

The prototype is a first combination of a MMO world with a physical stage using consumer electronic AR. As such, it offers the necessary combination of the “real” with the virtual and opens up new friction

zones for mixed-media performances. The project is very much intended for ongoing experimentation. We are still expanding this usage further and hope to include pieces of puppetry or dance in future installations.

References and Notes

1. Michael Nitsche and Maureen Thomas, “Play It Again: Film Performance, Virtual Environments and Game Engines,” in Colin Beardon and Gavin Carver, eds., *New Visions in Performance: The Impact of Digital Technologies* (Lisse: Swets & Zeitlinger, 2004).
2. Laurel, Brenda, *Computers as Theatre* (Reading: Addison-Wesley Publishing Company, 1993).
3. Aarseth, Espen J., *Cybertext: Perspectives on Ergodic Literature* (Baltimore: The John Hopkins University Press, 1997).

FLOW— A FLOWING INFORMATION INTERACTIVE ART

Yu-Chuan Tseng, Chia-Hsiang Lee and Ding-Ming Wang, National Chiao-Tung University / Shih Hsin University, National Taipei University of Technology, Hsuan Chuang University, Taiwan. E-mail: <t.yuchuan@gmail.com>, <jasonlee2357@gmail.com>, <dmwang2008@gmail.com>.

Abstract

Flow is an interactive art work creating a context of flowing information which is unable to be held, but is instead in a state of intermix. The practice aims to represent the chaos state of information in which we find ourselves in our present-day society. When confronted with a huge amount of information, we process it differently based on our opinions. We assume that we understand without question, but how can we firmly believe that our cognition is correct?

In the digital era in which we live, with its profound change in culture, arts, science, and ways of living, we live with information overflow. The World Wide Web is not only a tool, but also a medium for social, cultural, and political activity. Virtual ways of living have affected the ability to know people's experiences and have created new methods of perception. On the Internet, time, location, and identity no longer matter. We can exist in both the real and virtual worlds, playing different roles in each simultaneously.

The Internet is also a medium in which humans retrieve and create content. Internet services provide convenient tools to search information and create content in the net. Information can be gathered from across the globe by just a simple click: feelings are conveyed and lives are shared in blogs, on online chat forums, and in online chatrooms. People immerse themselves in large communication nets struggling to grab all the information possible to prevent themselves from being isolated behind the world stream and to strive to fit into the global village. With the huge amount of information available, bits are gleaned according to geographical location, culture background, and political conviction. However, the meanings are uncertain and changeable: perception and conception of meaning are different due to varying viewpoints and manners.

A Chaos State of Information

The meanings we receive are only digital signals created by data transformation. In transmission, digital technology analyzes



**Fig. 1. News pictures blur on the screen. (Photo © Yu-Chuan Tseng)
(Exhibition: “Boom! An Interplay of Fast and Frozen Permutation in New Media,”
Taipei National University of Arts, Kuando Museum of Arts, Taipei, 2007)**

and encodes information into fragments of digital data and then recombines those fragments. At the remote site, the recombined data is received, but what is perceived is only a pileup of fragments, not an entirety. Sokolowski utilizes the perception of a cube to explain a conscious experience: when we look at a cube, we see only three faces at any given time. Through subjective completion we add intentionally to the concept cube [1]. However, in the virtual world, our intention of objectivity is vague and uncertain. The authenticity of information has been exaggerated in digital media, throwing the human being into the dilemma of believing or denying the authenticity of information. Jean Baudrillard in *La guerre du Golfe n'a pas eu lieu*, claimed that the Gulf War was a collective illusion created by mass media [2]. In the digital era, the meanings of information have been regenerated, and a fake reality created. Humans own memories and existing intentions granted by digital apparatus. We fantasize the existence experience and believe the illusion.

Flow aims to represent a chaos state of information, which we conceive in world news by nearly seamless integration of mixed fragments. It grabs the information in the Web and visualizes a chaotic information status. People's images will be unconsciously immersed with information on the screen. *Flow* creates a context of flowing information which is unable to be held, but is in a state of intermix. It indicates the compelling

power of the mass media control system. It also presents the confusion that is faced daily and reminds that rationality must be maintained.

A Confused Information Fluid

There are two parts in this work. The first part is the retrieval of information and the automated system of its conversion into audio. Information is chosen from news portal sites which have the availability of RSS feeds. The selected news portal sites are from different countries: CNN (USA), UDN news (Taiwan) and the BBC (United Kingdom). Prominent news photos with accompanying text were selected from various categories such as politics, entertainment, digital life, the economy, etc. The contents of these parts, reflecting political, social, and economical issues, interact with each other deep inside the mind. The selected news fragments represent the modern congested information status.

At the exhibition site, the installation consisted of a room connected to a network, mounted video projectors, and a camera. At times of non-interaction with the installation, the news photos are projected in a blur (Fig. 1). The voice files heard are jumbled and indistinguishable. Images and audio appear layer by layer and compound with each other. Upon entering the installation, a participant's photo is taken and superimposed into the blurred news images. The images of geographically distributed information combined with that of the participant

represent a unity in the virtual world (Fig. 2).

The flashlight, as a metaphor of a conductor, was an interactive tool, lighting the way to information. Participants select to view the photos and hear the narration clearly by waving the flashlights. However, the hand is unsteady: as the flashlights wobble, the pictures and audio respond to the interaction and immediately cascade. Participants are forced to face a flood of information including images and sounds. As participants wave their flashlights in attempting to view the photos and hear the narration clearly, they encounter distress. Participants choose one picture and initiate the audio message, but news pictures come in quick succession. New images replace the one selected and multiple audio messages play simultaneously. Participants are bombarded with a collage of visual input and only hear noise created by multiple channels of audio. In the controlling process, participants laboriously attempt to comprehend. Multiple attempts are made to choose precisely, but in vain. Participants found themselves lost in the information stream.

The result of interaction is uncertain and fraught with difficulty. Participants are given control over the pictures and sounds, but are not able to control the final appearance and find themselves caught in the subject matter. As a result, the perception is one of being in a state of chaos. Although a photo can briefly be viewed clearly, its story can't be understood lucidly. The participant is left wondering and eager to know its meaning. In reality, no one is able to control and understand what is really happening in the world.

Free/Libre/Open Source Software Implementation

In cross-platform development, Java is chosen as the main programming language, as the Java platform can potentially encompass the choice of programming languages. Five open/free source software are selected to be integrated into the requirements: AJAX uses Microsoft XMLHTTP as an XMLHttpRequest object; TTS, FreeTTS API, a Java-based speech synthesizer or online TTS service; Tomcat server, a Java-based web server; BlobDetection, the programming library for finding blobs on an image; and Processing, an open source programming language to create a flying-like blurry effect (Fig. 3).

In the absence of a participant, the system randomly selects 30 images from the local file directory and displays them layer upon layer on the screen with an alpha transparent value to produce a blurry effect. As new images appear older ones dart off with a velocity of flying. Thus, images are presented in a colorful flying-like blurry mode. The system also selects a sound file from the local file directory, gives it a revert effect, and plays it in loops in the background. Every 15 minutes, the system randomly selects another 10 images to maintain a current status.

When a participant stands in front of the screen, their photo is taken by the web camera and becomes the background of the drifting vision. Participants control the beam of a flashlight into an input coordination signal. The signal mapping into the position acts as the cursor, controlling images and sounds to appear clearly. When the beam is detected and mapped to the cursor on the screen, it is triggered to enlarge the image, stop its movement, allowing clear viewing of it. The system does a loop to check whether the cursor is in the news



Fig. 2. Participant's image is superimposed with blurred news images. (Photo © Yu-Chuan Tseng) (Exhibition: Imaginary Number "i" in Arts, Kao Yuan Art Center, Kaohsiung, Taiwan, 2006)

image; once the image is triggered to enlarge, the same cursor will not be able to trigger another image until the cursor is out of the area of the triggered image. Different detected cursors trigger difference images. When the cursor moves out the image, the image reduces and moves again in a flying-like blur mode.

When the images pop up on the screen, the related sound file plays once in amalgamation. The sound is played in different channels according to the image on the screen, left image to the left channel, right image to the right channel. The more images are triggered, the more sounds are played. The system can deal with multi-users at the same time.

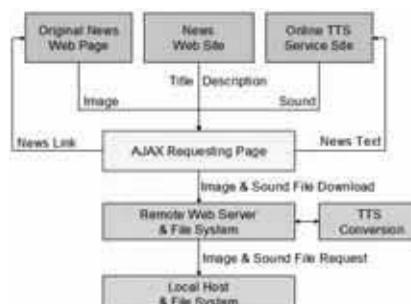


Fig. 3. Flow chart for requestion new items. (© Chia-Hsiang Lee)

A variety of different services are available online: search, map, dictionary, translation, text-to-speech, etc. They can be integrated to obtain information instantaneously with the huge information flow available on the Internet. This effect is mimicked by using AJAX and website technology. Data, information, and online services are sought on the Internet. Further, these websites and their services are integrated to meet many different information requirements.

Conclusion

In *Flow*, the image scenes consist of blurry news pictures that flow by on the network. Viewers use a light source to capture clear news pictures and voice reports from the real-time information field. The simulated news reports are in fact floating in the cyberspace information flow and are digitally converted into new combinations of information topics. This strange noise screened from the incessant chatter of news media has become a perceptual field with which humankind is increasingly familiar but perplexed. *Flow* constructs a fluid reality and indicates a situation faced. The same information is conceived differently based on one's opinions. It is assumed that something is understood without question. But how can one firmly believe one's cognition as being correct? Are the meanings behind the appearance really known, or is one simply caught in a maze? One drifts into the information flow to receive reality. However, what is reality?

Acknowledgment

Many thanks to Prof. Tien Chun Chang, National Chiao-Tung University, for supporting us to complete the practice.

References

1. Sokolowski, Robert. Introduction to Phenomenology. (Chinese edition). (Psy garden Publishing Company. Taipei. 2004). pp. 35-42.
2. Baudrillard, Jean. La guerre du Golfe n'a pas eu lieu. (Chinese edition). (Rye Field Publications. Taipei. 2003). pp. 74-105.