ORB3 – MUSICAL ROBOTS WITHIN AN ADAPTIVE SOCIAL COMPOSITION SYSTEM

Dan Livingstone
Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
dlivingstone@plymouth.ac.uk

Eduardo Miranda
Computer Music Research

School of Computing, Communications and Electronics, University of Plymouth, Drakes Circus Plymouth PL148AA United Kingdom
eduardo.miranda@plymouth.ac.uk

ABSTRACT

Gesture capture, motion tracking and 3D visualisation technologies have generated many new musical forms, often extending the mannerisms or behaviours of a given performer or discipline, providing new compositional frameworks for real time synthesis in response to action.
In many cases these approaches are presented within a single domain, a live stage performance, a site specific installation, a shared networked visualisation of collaborative composition. The reality is that these 'interactivating spaces' [1] whether haptic, [5] tactile [9] or ubiquitous [11] is that they manifest new forms of interaction, between people, systems and the medium of sound.
Free Sound can be understood to be an extension of the 'open work' where the base materials for a compositional process are created through a model of exchange, interaction and resynthesis. The resulting output of these activities can be broadcast and disseminated through a range of technologies to both social and private spaces. This research suggests that there are new interaction models and social compositional frameworks to be found in these cybrid spaces, a previously intangible location often dominated by the broadcast and publishing industry. A marketing model defined by revenue streams and a value chain. In the case of socially mediated composition or 'free sound' there is still a value chain, it's investors and beneficiaries are the open source community, the collaborators and participants within such mediated systems and the resulting free sound.

Keywords

Adaptive System, Sound Installation, Smart Interfaces, Music Robots, Spatial Music, Conscious Subconscious Interaction, Interaction models.

1. INTRODUCTION

In the design of new interfacing methods [1] for sound manipulation and control it is often the case that the primary focus is the point of tactile interaction, the exploration of new gestural controllers or methods for mapping and transforming data to create sound material [2]. This approach has led to the development of numerous novel and individual interfaces [3], in many cases the interaction mode is learnt by the user, in order to complete the feedback loop, thereby achieving dynamic results through an exploratory model of interaction.
With a modular adaptive systems approach the emphasis is on providing an interface framework for different types of interaction that can be initiated by both users and 'smart' interfaces, ie new interaction behaviors can be identified by the system independently, in response to users actions, whether direct tactile control or simple movement, location, gesture or position. With a modular adaptive systems approach the emphasis is on providing an interface framework for different types of interaction that can be initiated by both users and 'smart' interfaces.

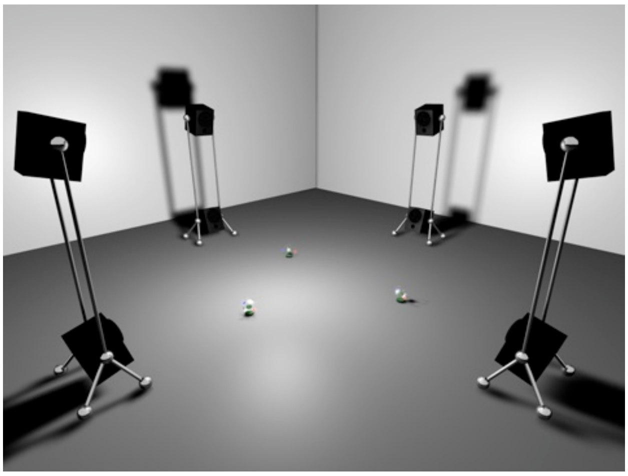

Figure 1. *Auditory Sphere. 8 active speakers angled to provide versatile software controlled diffusion. Diffusion and synthesis generated from environment/interaction data collected by each Orb. Software developed in MAX/MSP running on G4 Apple laptop with M-Audio 410 Firewire mobile multi - channel interface, custom built 'composer – listener' objects (Orb3 interfaces)*

The goals and aspirations of many researchers and educators in the field of musical interfaces for social composition as opposed to dedicated instrument controllers for accomplished performers was clearly expressed during an engaging keynote speech given by Gil Weinberg:
"…to have a musical response accentuated by the player who sent the original call, to plant a musical "seed" that would be picked up by the group in various manners, etc. An effective network would therefore promote interpersonal connections by encouraging participants to respond and react to evolving musical behaviors in a social manner of mutual influence and response."[9] A

significant observation during performances by children using the Beatbug system developed at MIT was described at the International Computer Music Conference in Miami 2004; the children made exaggerated swooping motions with the Beatbugs as they ‘passed’ sounds while interacting with the controller. At the time the Beatbugs were not equipped to react to this *emergent behavior*, although neither the audience or children were aware of this at the time. This observation led to the next refinement of this social network of interfaces, using blue tooth technologies and motion detection to refine and utilize this interaction, this anecdote reinforces the value of an adaptive systems approach which is a continually evolving field of applied research for novel interfaces and interactive music systems.

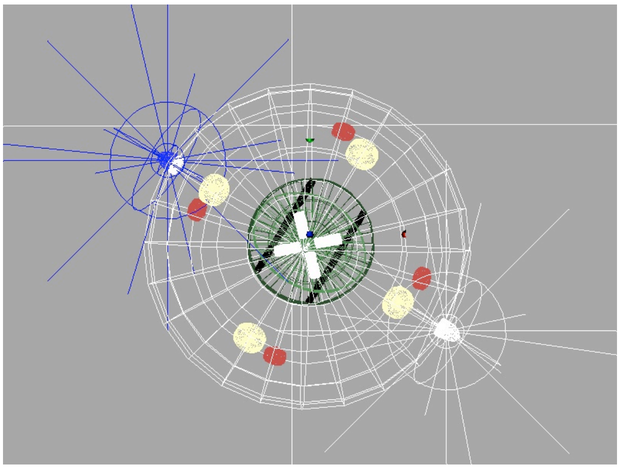

Figure 2. *Orb3 Design Constraints. Design aesthetic can be achieved by designing custom PCB’s for sensor placement, careful selection of plastics for manufacture and careful consideration of tactile properties for intuitive interaction.*

2. ORB3 DEVELOPMENT PROCESS

The Orb3 interface design was developed through observation of interaction with wired ‘composer and listener’ objects[i]. These original objects were static spheres housing a cluster of analogue sensors (Light dependant resistors, bead thermistors, vibration and tilt switches etc) for measuring ambient light, ambient heat, general motion and orientation. The original system comprised four such spheres which could be placed and relocated to generate and vary data used to synthesize sound material for 7.1 sound diffusion controlled by a gesture and motion based video tracking system[ii]
Each sphere contained a total of 8 sensors, wired to a control voltage to midi converter (Infusion Systems Icube) this method worked effectively for developing software and refining synthesis and sound design for prototyping a large scale adaptive system. Local interaction was less successful due to restricted movement of wired objects and unexpected behaviors and reactions of participants. For example; using the prototype system the shadows cast between spheres as participants moved around the room were recorded by a drop in light values sensed by the sphere affected, causing subtle changes in base sound materials generated for the sound scape, this was an intended compositional element of the system but on realizing this process, many participants could not resist the temptation to explore further, initially cupping or shielding areas of the spheres and inevitably moving and repositioning them, anticipating a direct response. It was immediately apparent that the simplicity of the sphere encouraged a series of interactions that could further inform sound design for socially mediated sound spaces. It also led to the realization that the software techniques applied to the vision system for adding new data relationships based on symbol recognition could be migrated to the interface design for each sphere developing more expressive tactile control, and more significantly, using the relative position and orientation of each sphere as a compositional parameter that could be heard in the diffused sound-scape, that was also registered by visual or tactile feedback on the interface itself. Other observations were that often participants chose to work collaboratively, taking a sphere each, influencing a parameter passing it on, this worked particularly effectively in groups of three, where patterns of motion and exchange had the potential to create rhythm and flow, some general experiments were done with different numbers of spheres to see if this affected interaction modes, it is speculated that providing an odd number of interfaces provides more movement through transfer and exchange and encourages turn taking. It was also noted that during periods of inaction or when participants were more passive different listening modes were reported, this in turn has influenced the sound design of the refined system, incorporating different ‘play states’ or modes - some further controlled experiments are in data to support these assertions. The logical development of these passive and active modes mediated by participants is to add simple robotics to each sphere to allow each one to move and interact with other spheres independently.

2.1. Design for Collaboration

Having established some significant refinements from the initial prototypes a specification for a more robust adaptive interface was resolved. Primarily a wireless approach was required, high performance with reliable transfer of digital and analogue data from sensors, in addition a wireless microphone embedded in each unit for live sampling. Internal lighting was added to indicate interaction modes and force feedback in response to interaction. These features introduced new design challenges, as the revised design needed onboard power for wireless operation and ideally solar charging to extend session times. A final addition was the inclusion of lasers and proximity sensing to enable quick alignment and event triggering between spheres. A simple method for overhead positional video tracking (max/msp Jitter) using a single fixed camera provides an effective method for documenting movement and behavior of each orb during a live session through time lapse imagining.

A mobile robotic element has been prototyped for each sphere, allowing them to move and reposition themselves autonomously or in 'collaboration' by integrating positional tracking and proximity triggers. This dynamic motion provides a visual element that reveals the compositional potential of the system, while demonstrating some of the synthesis and diffusion properties that are influenced by the interaction between or with each sphere. When each Orb is collecting data to influence sound synthesis and diffusion, or being followed or manipulated by participants this collaborative process can be displayed from a top down perspective, using either projection or plasma screen display.

"Most of the systems that allow the creation of sound and image in real-time don't have the capability for organizing events at a global level. This is however, required if the aim is to allow the composition of a piece that involves feedback from events sonic and visual, in the construction of interactive audiovisual compositions." [3] (Franco et al 2004)

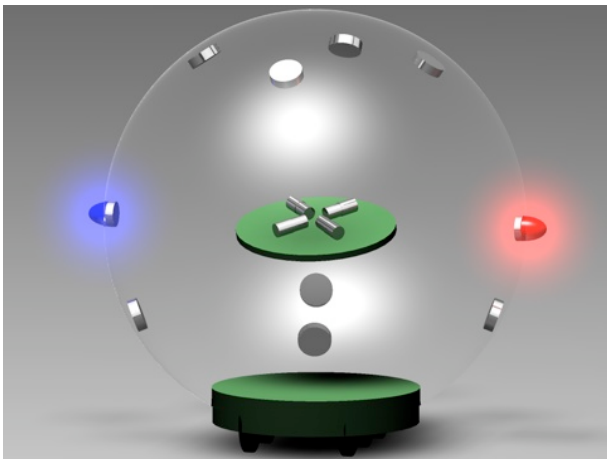

Figure 3. *Wireless Mobile Orb v2.0 in Absorb mode.*

(OrbV2 features data collection through light, temperature, orientation, motion sensing, laser alignment, microphone, mobility, rotation, Exploratory interaction transforming data into sound material)

For the purpose of this paper emphasis has been placed on the Orb3 interfaces, the key features are interaction modes and social composition, simply expressed as 'play states'. Sound is the primary medium but in order to make visible the transformative processes underpinning the compositional output ways are being sought to create a visual aesthetic from both the data and interaction of people, making visible behavior and interactions, effectively creating graphical transcription as real time feedback to participants.

Developers of collaborative musical interfaces with tactile, graphical and sensory feed back are developing new terminologies to describe the design process for these systems in terms that begin to articulate their compositional and social modalities. Collaborative interactive music systems, such as 'Block Jam' (Newton-Dunn, Nakano, Gibson 2002) where interconnecting blocks are collaboratively assembled to organise musical phrases and sequences begin to identify new musical forms:

"By creating both a tangible and a visual language, we are able to create endless meaningful musical structures in a novel and intuitive way that predisposes itself to collaboration and exploration, face to face or via a network, pushing interactive music towards the casual user."

[6] (Newton Dunn et. al. 2002)

Other collaborative works such as ToneTable [2] (Bowers J. 2001) use interactive visual elements as an integral interface element, in this case participants manipulate 4 trackballs, 'disturbing' a projected fluid surface with associated textures and diffused sounds, again it is the observation of improvisation and collaboration with a realtime composition system that distinguishes this emerging musical form. The author discusses emergent behaviors and extended engagement as a development of the system design.

"we have tried a number of design strategies for addressing such settings. We have explored notions of 'collaboration through a virtual medium', 'collaborative added value', 'layers of notice ability', 'structures of motivation'. These are all concepts intended to suggest ways for orienting design for variable participation."

[2] (Bowers, J. 2001)

A highly refined table top tactile control surface for two or more participants has been developed by Patten and Brecht, 'Audiopad' [8], which has been extensively exhibited. The system provides a graphically dynamic projected overlay oriented around electronically tagged tracked physical objects or 'pucks' for real-time control of preprogrammed electronic music, moved by hand with fingertip control.

"Audiopad not only allows for spontaneous reinterpretation of musical compositions, but also creates a visual and tactile dialogue between itself, the performer, and the audience."

[8](Patten J. Brecht B. 2003)

The design and installation for the Orb3 system forms an auditory sphere (fig. 1.) using an 8 Channel sound diffusion through which participants move, view, listen and reconstruct the compositional process through social interaction within it. The audiovisual feedback in response to these varied interaction modalities is an active process, one of content driven collaboration.[iii]

2.2. Communications & Parameters

Each Orb sends data via a 2.4ghz wireless RF interface to a G4 laptop running Max/MSP, a combination of analogue and digital data can be sent and processed by the control software created in max. The software itself is not simply a parameter mapping utillity, it is designed to correlate different data against previous interactions, a form of compositional memory where environmental parameters of previous sessions are compared with current ones to identify repeated behaviors of the system and actions of participants. The software is designed to be adaptive, previously un-recorded or new data configurations are identified and used to compose new sound events or objects. The software sends data to each Orb to indicate it's state and trigger visual or tactile feedback, ie; activate laser/proximity sensing for positioning, activate status leds, activate force feedback. Each Orb has two compositional states – *Absorb* and *Adapt*. In *Absorb* mode an Orb is autonomous and located on the floor, it's sensors are calibrated to collect environmental data, ambient light, ambient temperature, relative position and orientation, it can also live sample sound for processing - the software controls this calibration which is activated through Orb alignment – each Orb is fitted with proximity sensor, a laser and LDR - placing the three Orbs in a triangle and directing each Laser to the next Orbs locating LDR activates this mode, which is part of the initial setup process.

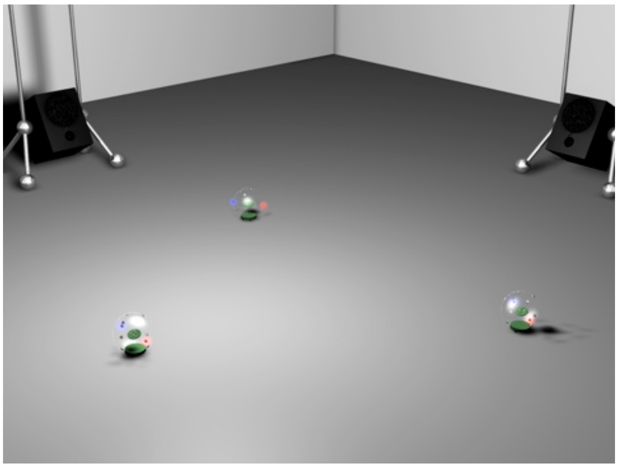

Figure 3. *Laser alignment - Triangulation.*

(Alignment –view of each Orb, lasers are activated, two Orbs are shown in listening 'Absorb' mode after calibration, one (lower right) is about to move out of alignment in response to parameter changes, autonomous – may move or rotate to attract participants, system responds by panning a sound in relation to it's movement).

Environmental data changes are usually slow in interior environments so these elements are mapped to the timbre and color of sounds created with larger fluctuations affecting diffusion, thus providing an overall structure for the real time composition that is responsive to ambient light, temperature and general movement. *Adapt* mode is activated when the alignment of Orbs is disturbed, either by walking between them, interrupting the laser tracking or by picking them up which also activates vibration sensors and initiates orientation mapping - angle and orientation of each Orb in this state directly influences panning and diffusion rates of synthesized sounds. During *Adapt* mode the laser is deactivated and the ambient temperature measurement is recalibrated to respond to body heat through hand contacts on the Orbs lower surface. Bead thermistors with fast response times are used so as an Orb is passed from one hand to another, or between users, it registers and marks these changes. This data combined with orientation data allows for a range of subtle and dynamic sound events to be initiated by each participant in collaboration with both the system and with other people.

2.3. Emergent Behavior

As an adaptive portable system, the Orb3 environment creates an opportunity for observing and recording forms of emergent behavior in relation to spatial sound interaction, this provides researchers in this field with a structured framework to inform the design of mobile and autonomous interfaces, such as musical robots or adaptive social composition systems.

" we should not forget that humble reactive robotic systems capable of sensing and reflecting the complexity of their environments have the capacity for unpredictable and life like behavior that encourages playful somatic interaction." [10] (Woolf & Beck 2002)

The inclusion of play through collaboration is not a by product of this system, it has been developed explicitly to motivate different responses through consideration of ergonomics and human factors, developing from the considered observations of researchers and practitioners in related fields. The 'play states' or modes titled *Absorb* and *Adapt* have been designed with consideration of both composer/listener object interaction and the listening process or perceptual triggers to motivate participants.

In the 'play state' *Absorb* the Orbs are programmed to activate when certain parameters or sound events are captured, or when conditions match previously encountered sequences, the 'intention to listen' is shown through both the status LEDS and motion/rotation in response to stimulus. This modality can also be attributed to the behavior of participants, who move towards the 'Auditory sphere' of course initially their interest is more likely to be the spectacle of the technology or other participants behavior, however moving into the 'Auditory sphere' shows an intention to participate, *to listen.* Participants interaction at this stage can be described as Subconscious, they are not necessarily aware that their presence and orientation is influencing the system.

The 'play state' *Adapt* is active when the triangular alignment of the three Orbs is disturbed, the software reconfigures itself to a more sensitive state, ready to be interacted with, held, passed, moved in relation to sound synthesis & diffusion as perceived and manipulated by a participant. In software terms this is achieved by switching the algorithms mediating data analysis, through patterns stored in short and long term memory (Max objects *capture, decode, funbuff, histo* and *spray* are integrated with *mtr* to record, store and replay streams of data, which are compared against previously collected and live data [short term memory], a form of score following). By picking up an Orb a participant is moving from the *Absorb* state, instead choosing to interact, to explore and through this action perceiving and identifying the source of broadcast sounds, through their manipulation of an Orb. This modality is further reinforced when direct control of sounds are influenced by the participant. Their behavior changes as they *Adapt* to the parameters they have influence over. This can be described as Conscious interaction, a heightened state of attention and engagement, [6] [Newton Dunn et. al. 2002] the intention to collaborate with the system and others using it, improvisation, not simply 'call and response' [4] [Lippe C. 2002] as there are no familiar, formal or structured elements in the form of musical patterns, note sequences or beats inherent in the open nature of this spatial sound environment. A key development with this system is that it continues to adapt while capturing, archiving and broadcasting new behaviors. Several task based experiments using perceptual constructs [12] to establish and refine interaction models to refine compositional processes within the system. Assigning participants simple compositional tasks based on establishing mental models for the relationship between sound objects and their perceived location in the auditory sheper relative to the listener yield to forms of data. Quantitative data can be extracted from the short term memory of the system, noting the start event of an interaction, such as repositioning a sound by manipulating an orb. After this action has been archived Qualitative data from participants reports can be established by comparing transcribed verbal accounts of the set task and system response. Both data types can be considered in context by reviewing the overhead broadcast documented by the system, archiving actual position of participants and orbs against archived positional data.

3. CONCLUSIONS

The compositional approach is not modeled on a *'fixed or even consistent excitation-sonification relationship'* [Paine G. 2004] many elements of the sound-scape generated are through transformative synthesis methods, in this instance the creation of sound through traditionally unrelated real world variables. Neither are the sounds randomly generated; the capture and transformation of variables such as heat, light, proximity, motion and time create values that could be mapped to conventional parameters for musical control of predefined note sequences, loops and formally structured phrases but in this adaptive approach through a process of observation, listening and sound design these parameters are treated as explicit elements of the real time composition environment.

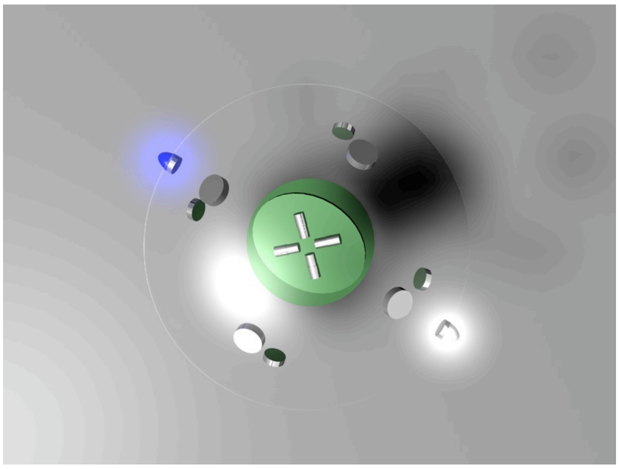

Figure 4. *Orb in adapt mode.*

(Overhead camera view (simulated), As an Orb is picked up sensitivity is heightened by re mapping parameters, accelerometer sensitivity maps motion, angle, orientation to sound diffusion while archiving lifting motion as a new behavior or gesture for the current synthesized sound object)

The system design approach is adaptive, one that aims to create synthesis to express physical real world properties in collaboration with participants through social interaction, sound synthesis and diffusion. *"...we have mentioned that developing a new instrument is not an easy task, and* [we have] *introduced a structured and modular approach. By focusing on the content of the compositorial material rather then the interface, we felt it was possible to perform with the instrument even from its most rudimentary state."[1] (Bongers 2002)*

The emphasis on compositional content rather than purely refining the interface technology has proved to be a significant design methodology, each interface is fairly simple, basic electronics are used, however the combination of participants behavior, adaptive software and 'smart' interfaces creates a new compositional process. Through further observation and refinement of this type of system a deeper understanding of 'play states' and collaborative compositional processes will be described. *"Response to musical stimuli can cause significant changes in both behavior and brain activity"[5] (Machover T. 2004)*

Developing systems that adapt and respond to these essential elements of musical activity is a demanding challenge to this field. Consideration of social interaction through the medium of sound is a core concern of this research; how we perceive and interact with sound environments or interface objects that adapt to our behavior. In this sense the Orb3 interface is 'smart' our social interactions and interplay are part of the 'instrument' but the instrument is not merely a separate controller or extension of an individual performer, it is a socially mediated compositional environment with the potential to adapt to emergent behavior. 'An adaptive systems approach that exhibits process driven collaboration.' [13]

4. REFERENCES

[1] Bongers B. Harris Y. A (2002) Structured Instrument Design Approach: The Video-Organ *Proceedings of the Conference on New Instruments for Musical Expression, May 2002, Media Lab Europe, Dublin Ireland.*

[2] Bowers, J. (2001) TONETABLE: A Multi-User, Mixed Media, Interactive Installation P2, and 4,5.

[3] Franco E. Griffith N. Fernstro E. (2004) " Issues for Designing a flexible expressive audio visual system for real-time performance & composition" *Proceedings of the Conference on New Instruments for Musical Expression, June 3-5, 2004 Shizuoka University of Art and Culture Hamamatsu, Japan*

[4] Lippe, C (2002) "Real time Interaction Among Composers, Performers, and Computer Systems" *Information Processing Society of Japan SIG Notes, Volume 2002, Number 123, pp. 1-6.*

[5] Machover T. Shaping Minds Musically *BT Technology Journal Vol 22 No 4 October 2004*

[6] Newton-Dunn H., Nakanon H., Gibson J. "Block Jam: A Tangible Interface for Interactive Music" *Proceedings of the Conference on New Interfaces for Musical Expression*, Montreal, Canada *2003*

[7] *Paine G. (2004) "Gesture and Musical Interaction: Interactive Engagement Through Dynamic Morphology" Proceedings of the Conference on New Instruments for Musical Expression, June 3-5, 2004 Shizuoka University of Art and Culture Hamamatsu, Japan*

[8] Patten J., Recht B., Ishii H."Audiopad: A Tag-based Interface for Musical Performance" *Proceedings of the Conference on New Interfaces for Musical Expression*, Dublin, Ireland, 2002.

[9] Weinberg G. Key Note Address: Interconnected Musical Networks –Bringing Expression and Thoughtfulness to Collaborative Group Playing" *In Proceedings of the International Computer Music Conference 1-6 November, 2004 University of Miami, Florida, USA*

[10] *Woolf S. Experiments with Reactive Robotic Sound Sculptures. ALife VIII: Workshop proceedings 2002 P2, 3.*

[11] Livingstone D. and Miranda E. "Composition for Ubiquitous Responsive Sound Environments" . *In Proceedings of the International Computer Music Conference 1-6 November, 2004 University of Miami, Florida, USA*

[12] Livingstone D. (1998) "The Space between the Assumed Real and The Digital Virtual". *In Ascott, R. Reframing Consciousness – Art, Mind & Technology Intellect Books pp138-143*

---

i *Composer and Listener Objects detailed here form part of a larger integrated system included in proceedings ICMC 2004* [11]

ii *Details of this symbol based adaptive tracking system are included in proceedings ICMC 2004* [11]