# GESTURE, SOUND AND PLACE

Roger Alsop, Faculty of the VCA/MCM, University of Melbourne, Vic, Australia
E-mail: ralsop@unimelb.edu.au

Mark Pedersen, Faculty of the VCA/MCM, University of Melbourne, Vic, Australia
E-mail: markp@chailight.com

**Abstract**

This paper discusses two multichannel interactive audiovisual artworks, Action A/V and SoundLabyrinth, that explore approaches to the experience of gesture, sound and place. Both works were situated in a geodesic dome frame and built within the Max, Ableton and Max for Live computer programs, and produced ostensibly similar outcomes, however the approaches taken by the two authors differ in intention, processes, and philosophy.
These approaches were presented and discussed in workshops delivered at ISEA2013, on Sunday June 9 and Monday June 10 2013. In these workshops participants improvised with the two systems, both through moving in the dome and by operating the related software, and discussed approaches and understandings of the three terms listed in the title of this paper.

**Keywords:** sound, space, place, interaction, gesture, interactive art, workshop

## Introduction

*Gesture, Sound and Place* was an interactive workshop held at *ISEA2013* in which Mark Pedersen and Roger Alsop introduced their approaches to creating interactive audiovisual processes and artworks that engage with and interrogate the concepts Gesture, Sound, and Place.

The breadth of human gestures, as Adam Kendon posits, "can play a role in how interactants regulate each others' patterns of attention; participants in non-speaker roles … indicate their assessment or understanding of another's utterance; … [are] incorporated into discourse as objects of deictic reference, and … used in alternation with spoken elements in discourse, partnering words as syntactic elements"[1].

This understanding is related to a physical action that accompanies a (usually spoken) utterance. However, gestures can encompass much more. They may be physical, such as a wave of the hand; ineffectual, 'just a gesture'; empathetic, 'a kind gesture'; and spontaneous or considered. Cadoz and Wanderley offer a review of various definitions of gesture in a number of contexts [2], and Clynes offers an interesting position on gesture in conducting scored music [3]. There are also mental gestures [4], the gestures that cause affect through " a multitude of perceptual and cognitive mechanisms" [5] or the thoughts that cause physiological results, as explored by Eaton and Miranda [6].

Gestures are also considered to be present in music, painting, narrative, thought, speech and just about every other human action or interaction, and can be either deliberate or unconscious. *Gesture, Sound and Place* focuses primarily on deliberate physical gestures that the gesturer can recognise as having an effect, such as those of a dancer or other performer, or simple movements from one place to another.

Like Kendon's conceptualisation of gesture, sound is a physical event that can be objectively observed and mapped. It has a range of functions, including: transmitting information (through speech), evoking emotion (through music), alerting to danger (it has fewer barriers than sight, and can be registered from all directions), indicating place (through sonar, or simply listening to reverberations), and so on. The understandings/conceptualisations/theorisations of the sound gesture have developed to include more than merely what can be heard; Kim-Cohen [7], Kahn, [8] and LaBelle [9], for example, consider it from conceptual, philosophical and social perspectives.

There have also been a vast number of tools developed regarding the sonifaction of physical gesture, as seen in the work of Rokeby [10] and Winkler [11], and the sonification of data [12].

The same can be said for place, which, like sound, can be considered a natural, geographic event, one that can be objectively considered. It can also be considered experientially, as seen in the discussions of Tuan [13], Bachelard [14], Seaman [15] and Griesinger [16], the last of whom blends objective and subjective experience of place in his discussion of inclusion. Pedersen and Alsop consider the objective place as harbouring and generating a subjective experience; in the geodesic dome created for the *ISEA2013* workshop, the physical gestures of participants generated and mediated the sounds they heard.

## Processes

In *Gesture, Sound, and Place*, Alsop and Pedersen contrast switch-based and continuous control interaction paradigms. These two systems are comparable to the piano keyboard, a fixed set of switches which offer depth of variation through velocity, sensitivity and combinatoric complexity, and strings, which offer continuous pitch and timbre variation through such factors as bow pressure and speed, and finger position.

In practice, Alsop uses digital camera input analysed via Max/MSP [17] to make a grid of 16 zones (the switches), which respond to the degree of light change within each zone (analogous to velocity), to generate a set of variable sine tone frequencies output through a surround-sound speaker array.

In contrast, Pedersen uses a Kinect depth-sensing camera to continually track the positions of the participants within the space, and their hand gestures, to create a shifting soundscape of field recordings. Pedersen's tracking devices and speakers are shown in Fig. 1, while Alsop's grid-based input can be seen in Fig. 2 and his Kinect-based tracking system is detailed in Fig. 3, and Alsop's control software interfaces are shown in Fig. 4.

**Fig. 1. Sensor and speaker placement in the dome**

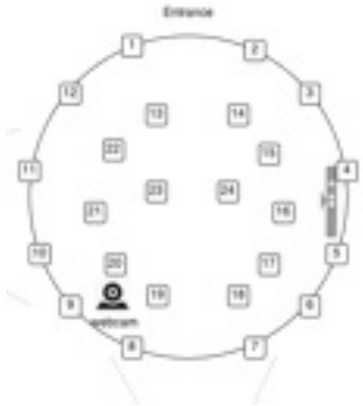


**Fig. 2. 16 quadrant camera image, showing the degree of light and light change as numbers in each quadrant.**


These systems were developed independently, and are both able and intended to be installed in a variety of environments. Alsop's systems are generally created to fit any environment and technical specifications, while Pedersen's are oriented around a bespoke 24-channel audio system he

designed. Pedersen's system is currently optimised for a domelike environment, in which sound may be easily dispersed horizontally and vertically, as all speakers can be equidistant; however this environment is not a prerequisite to the use of this system.

Integrating the two systems within the dome required adjusting Pedersen's dispersion processes to accommodate Alsop's processes. In the *ISEA2013* workshop Alsop's sound was dispersed horizontally and vertically through the speakers to the left of the dome, and Pedersen's through all speakers, shown in Fig. 1. This fitted the requirements of both approaches to interactive sound dispersion.

## Aesthetic approaches

Alsop and Pedersen have complementary approaches to gesture, sound and place, drawn from both their individual aesthetic philosophies and aspirations, and also their respective approaches to the role of sound in art making and its relationship to the environments in which it is presented.

Alsop's position is based on simple approaches to sound art presentations. He recognises that these are made in environments which heavily influence the reception, or perception, of sound, and also subsequently the ways in which participants (those who engage with the work – audience, gesturer, etc.) experience engagement, envelopment [16] [18] [19] and agency. In *ActionA/V* (a continually developing work/program), he explores processes that allow participants to develop a sense of personal engagement, envelopment, and ownership of the artwork that results from their intended and non-intended interactions with *ActionA/V* [20].

This approach has certain core requirements: that the interface between gesture and heard/seen result is robust; that the result of a gesture be easily acknowledged and recognisable by the gesturer; that interface learning comes through play; that it be extensible; and that it is able to create responses which range from completely predictable to unexpected (yet are recognisable as related).

The result is a system similar to a typical musical instrument and music conductor, affording similar gestural control; that is, physical gestures cause and influence timbre, amplitude and frequency. This results in the gesturer creating, and the participants experiencing, rhythmic, timbre, and frequency gestures, and their consequent aesthetic/conceptual interpretation of those gestures.

The *ISEA2013* workshop version of *ActionA/V* used sine waves. Here the heard frequencies ranged from about100 to 16 KHz; this meant that the 16 'players' seen in Fig. 4 (below) could generate 16 different sine waves at different frequencies and amplitudes.

This provides a number of options; for example, if the frequencies to be propagated are set close together, with a variation range of 2 Hz, the gesturer/listener would experience a form of additive synthesis produced timbres as the different sine waves faded in and out, and the location of the gesturer created different phase relationships between the waves. A more traditional 'instrument' would have each player set to frequencies similar to those of a well-tempered scale, allowing the possibility of creating melodies and harmonies similar to those of traditional western music.

The two approaches outlined above indicate the symbiotic relationship between the place, gesture, gesturer and listener that creates the sonic environment. In the first, the slight differences between frequencies (2 Hz as suggested above) creates a subtle sonic, physical environment. Here the interaction between the sonic environment and a participant's attention is perhaps not captured or excited to the same degree as it would be by the comparatively dramatic changes of the second approach, in which there are larger differences between frequencies. This approach creates both introverted and extroverted senses of psychological 'place': one may not overtly expect or inspire the audience's attention, while the other does.

The interface for this version of *ActionA/V*, shown in Fig. 4 (below), offers the opportunity for gesture in the creation and dispersion of the audio. This shows the aspects of the program that can be altered by the participants and, when used in a performance setting, the users/performers.

Alsop sees this kind of interaction as a gesture fitting the understandings and interpretations of the term 'gesture' listed earlier. There is the gesture required to adjust and respond to those adjustments of the controls in the oblong bubbles in Fig. 4, and using a mouse to draw in the matrix to the left of the screen in order to spatalise the sounds.

Alsop's underlying aesthetic requires that *ActionA/V* respond to the physical and mental gestures of the participants as possible gestures, in such a way that learned sociocultural bases are transparent. It should interact with the participant's deliberate, unconscious, composed, and improvised gestures; all are brought to the attention of the participant in a way that inspires further interaction, and develops understanding of the effect of their interactions through play.

## Aesthetics and Cognition

In *SoundLabyrinth*, Pedersen focuses on place through the use of evocative field recordings, and the relationship between physical gesture and sound. This approach is inspired/informed by the Embodied Sound Cognition framework developed by Godøy and Leman [21] and others. For Godøy and Leman, knowledge emerges out of a need to act in the environment, not just to collect information for its own sake; hence the focus of the embodied cognition approach is on action. Key to the embodied perspective is evidence of a close coupling between the cognitive processes that underpin movement and perception. Leman [22] provides extensive discussion of this evidence, including the behavioural observation of infants' innate ability to perceive gestures and replicate them, and the neurobiological observation that some of the same neurons which are fired to create a gesture such as grasping-with-the-hand (the so-called 'mirror neurons') also fire when the subject observes another person performing the same action.

The tight coupling of movement and perception at a cognitive level gives rise to the idea that, just as our movements arise from intentions (simulation of the movement), so perceptions of the external world map back to intentions because of the trace left by the shared cognitive processes. This action-oriented ontology suggests that even at the social level, the actions of others are understood in terms our own intentions, that is, our own simulated actions; the moving sonic forms of music, created by a concatenation of sonic gestures, are likewise attributed with intentionality because of the coupling of perception and movement. Thus, because individuals develop their own action-oriented ontology in a similar way by virtue of their common human physiology, if not common culture,

semantic communication is possible through sound and music [23].

Pedersen's particular goal with *SoundLabyrinth* is to create a space in which it is possible to experience an embodied encounter with the sublime, through sound [24]. The intersection of the abstract/virtual and the embodied/actual is the space upon which *SoundLabyrinth* focuses.

The compositional approach is one of gradual intensification of this intersection. Elements of the installation's soundscape respond to simple movement within the space, thus gradually progressing from a natural environment to an aural landscape that becomes more abstracted as the participant delves deeper into the space of the dome (desert winds, snatches of poetry, synthetic drones).

Along the way, the participant may discover another mode of interaction: one that responds to gesture, rather than just location. Through linking physical gesture and sound, there is an intensification of the inherent connection between sound and movement posited by Leman.

Given the casual nature of the relationship, this temporary fusing of participant and installation as a combined performative system only works if the mapping between gesture and sound is grounded in the embodied intelligence of the participant; without an awareness of the relationship between gesture and sound, the sounds generated in response to gesture could feel alien and disconnected, and fail to invite further exploration.

## Physical Infrastructure

In *SoundLabyrinth*, the Kinect provides input into a Max/MSP patch which uses the Kinect's skeleton tracking algorithm to provide participant locations as XY coordinates to the main sound management and ambisonic spatialisation patch.

In contrast, *ActionA/V's* quadrant-based analysis of the camera input was mapped directly to a subset of speakers (10–12 and 21–23), which allowed gestures in various quadrants to be mapped to specific speakers, providing spatialisation both horizontally and vertically; **Error! Reference source not found.** provides an overall view of sensor and speaker placement, while Fig. 3 provides a more detailed view of the analysis and spatialisation subsystems.

For *SoundLabyrinth*, a person's position within the dome is used to smoothly transition between different soundscapes. The Max/MSP nodes object provides a zone-based trigger system which can also provide weighted output from each proximal trigger zone, allowing interpolation between triggered soundscapes. This approach provides a natural mechanism for exploration of the *SoundLabyrinth*, as triggered material fades in and out of hearing in response to position, the primary form of agency within the *SoundLabyrinth* component of the system.

The Kinect is also used for gesture tracking of a single participant at a time, with large-scale hand gestures enabling sound objects, or discrete sonic events, to be moved about within the spatialisation system. In addition, these hand gestures apply simple effects, such as reverb or delay amount and feedback, to be applied. The effective range of the Kinect means that participants are only tracked in the left half of the dome closest to the Kinect. The Kinect gesture tracking subsystem is able to automatically recognise and calibrate to any humanoid shape within range, without the need for a specific calibration gesture; hence engagement of the gestural interaction is seamless with respect to the rest of the installation.

## In the workshops

Alsop and Pedersen held six workshops, each lasting about an hour, over two days. Their aim was to expose and interrogate the artists' processes for multimedia interaction, to facilitate participants' understanding of the processes, and to discuss and expand on these processes.

Each workshop had between five and twelve participants, and often led to conversations which continued long after the workshop had finished, with participants experimenting with *SoundLabyrinth* and *ActionA/V*, discussing their own processes, and networking.

The loose format for the workshops involved the participants first introducing themselves, discussing their experience in interactive art, their practice and affiliations, and explaining what interested them about the workshop. Participants were then asked to reflect on gestures they considered particularly meaningful and to explain why, and to consider how these gestures might be observed and measured. They were next introduced to concepts of gesture, sound and place, and embodied cognition, as Alsop and Pedersen considered them in their respective art making.

The responses were unique to each group, and to try to distill the responses would make it impossible to accurately represent them. However, responses to the questions asked leaned towards a gesture being a deliberate physical movement, sound being musical when in an art context, and place being a contrived or purpose allocated/specific environment.

When it was suggested that a gesture, sound or space could include those listed above and more, participants questioned and considered those possibilities and acknowledged that such interpretations of the terms could be advantageous, particularly when creating artwork and developing conceptual frames. The contrasting interaction paradigms of *SoundLabyrinth* and *ActionA/V* were explained, in terms of both the conceptual paradigms and a high-level description of the practical implementation.

The six workshops had differing cohorts and interests: some participants wanted to experience the sonified geodesic dome, while others developed or wanted to share understandings of, and approaches to, gesture recognition/mapping processes and systems, and a third group attended just to explore and play.

Often the responses ranged through: waving to someone, proximity to others, dance-like movements, body language, and expressive gestures such as jumping-for-joy. After discussion, a broader sense of what a gesture might be arose, and this was, in some cases, reflected in the gestures participants made in the dome and while using the programs.

Participants' approaches to *ActionA/V* were usually exploratory, as can be expected when any new object or experience is encountered. The physical gestures were initially timid, mostly slowly waving arms. As familiarity and experience grew, however, rapid bodily motion, movement around the dome, and a sense of play became prevalent. After this it was suggested that participants move closer to and farther away from the camera, try to perform as ensembles, and use small torches provided to activate sound.

Participants were also asked adjust any of the parameters seen in Fig, 4, and the approach was similar. When doing so, the activities of the participant(s) in the dome and those adjusting the *ActionA/V* parameters created a two-way

improvisation, as the actions of one influenced the other. The image in Fig. 4 does not clearly show what the variable parameters are, or what they affect. The program, which has a built-in pop-up help and a description of what each parameter does, is available on request from Roger Alsop.

Participants' interaction with the *SoundLabyrinth* components of the installation was initially unconscious, as the system reacted to their position within the space without them needing to actively engage. As participant awareness of the relationship between sound and position developed, more conscious explorations of the sound space that overlaid the physical space occurred.

Participants with a knowledge of gestural systems, particularly those who recognized the Kinect sensor, used arm-waving gestures to explore the gestural affordance of the system. Some participants noted that they would not normally use such gestures in an installation setting, and did not discover these aspects of the system until prompted. Participants remarked on a sense of satisfaction with gestures which mapped to obvious changes in sound, such as the 3D panning of a sound, but were frustrated by other more subtle effects, such as reverb changes, as they did not get a clear sense of agency or relationship.

## Developments beyond the workshop

There is much discussion regarding the sonification of gesture which can be seen in the work of the authors listed, and many more not listed. Alsop and Pedersen did not intend to offer commentary on, or a history of, interactive designs, processes, or technologies in *Gesture, Sound, and Place*. Rather, they sought to develop processes within current technologies, and share and develop paradigms and concepts whilst doing so.

The workshops and subsequent works made by Alsop and Pedersen simply build on this work in idiosyncratic ways, and it is hoped that they add to the increasing body of knowledge and of ways of working with interactive technologies, and the opportunities that these provide. In terms of *SoundLabyrinth*, one element emerging from the workshop was the importance of relational gestures, that is, gestural interactions between human participants, rather than gestural interactions between the participant and the system. S*oundLabyrinth* had been conceived as a solo experience for the participant, with interactive elements designed to heighten the participant's individual experience of the sonic space of the installation.

Following on from the workshop, greater emphasis has been placed on enabling the *SoundLabyrinth* system to respond to the gestural relationships between two or more participants. Simple mappings, such as linking the volume of a sound to the distance between two participants, with other variations on the sound being controlled by individual gestures, produced greater sustained interest in the interactivity of the system, compared to the solo gestural mode.

Since the workshops, *ActionA/V* has been developed as an approach to gesture-generation of audio, and is now considered a tool for sonifying physical gestures (whatever they may be). The current goal in developing the work further is to create a process that transparently allows the integration of physical and mental gestures in the creation of an audio and visual environment.

While this may well be a difficult aspiration, as mental gestures are unseen until articulated externally, it is possible to develop a system for understanding and responding to intention, for example: sustained, slow gestures that signify contemplation; gestures in a small space that signify intimacy; sustained gestures on the horizontal plane that indicate expansive intentions; and sustained gestures on the vertical plane that indicate contracting intentions. These interpretations are subjective and arbitrary at best, but can lead to a model from which to develop more accurate interpretations of mental gestures.

Altavilla, Caramiaux and Tanaka see that "gestural-sonic affordances may provide insight into designing future interactive and gestural music systems that balance the morphological characteristics of the sound with its potential cultural identification" [25], and may equally contain morphological characteristics of gesture and place.

As those three authors imply, the field of gestural music systems, while very active, is still nascent and lacking the maturity of traditional instrumental gesture sonifying systems. In linking interactive audiovisual installation practice to traditional instrumental practice, such as cello or piano, through common underlying gestural paradigms, Alsop and Pedersen hope to further develop both the accessibility and the potential for virtuosity in future instances of their work.

**Fig. 3. *SoundLabyrinth* spatialisation and sensor system detail**

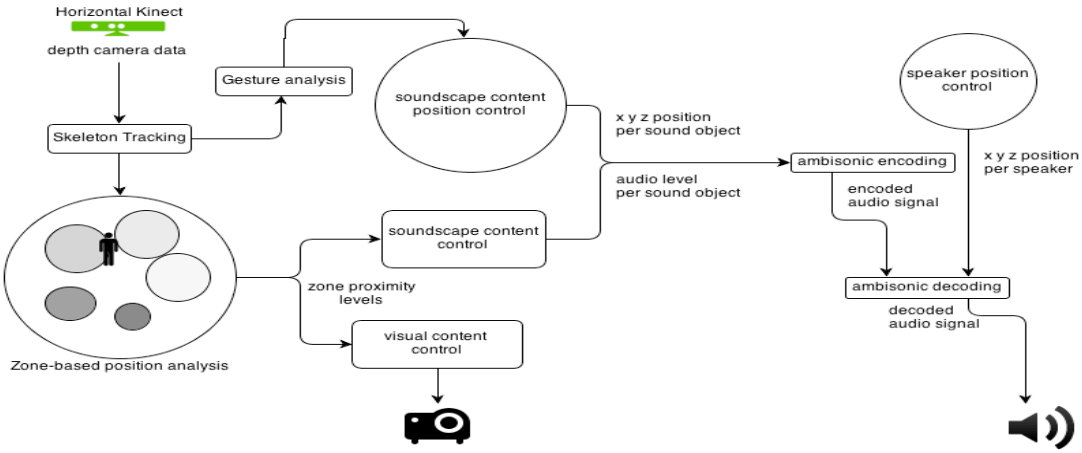
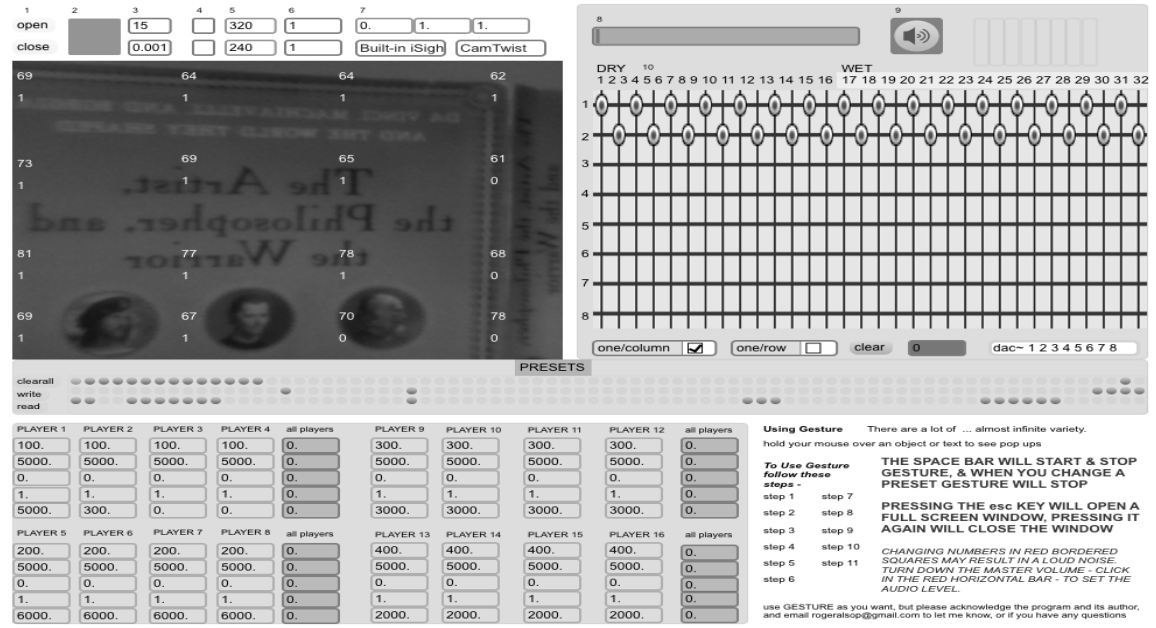


**Fig. 4. Developed control surface for *ActionA/V*, as used in *IESA2013* workshop**




## References and Notes


**1**. Adam Kendon, 'Gesture', *Annual Review of Anthropology*, 26 (1997), pp. 109-28.

**2**. Claude Cadoz and Marcelo M. Wanderley, 'Gesture - Music', in M.M. Wanderley and M. Battier (eds.), *Trends in Gestural Control of Music* (Paris: Ircam - Centre Pompidou, 2000), pp. 71-94.

**3.** Manfred Clynes, 'Living Micro-Microstructure: Instant Rehearseless Conducting Communicates Fine Structure of Musical Intention', in *The Inaugural International Conference on Music Communication Science* (Sydney, Australia: 2007), pp. 27-30.

**4.** Fernando Iazzetta, 'Meaning in Musical Gesture', in Wanderley and Battier*Trends in Gestural Control of Music*, ed. by M.M. Wanderley and M. Battier (Paris: Ircam - Centre Pompidou, 2000).

**5.** Christopher Tough, 'Can Gesture Recognition Repair the Disembodiment of Live Electronic Music?', (2013).

**6.** Joel Eaton and Eduardo Miranda, 'Real-Time Notation Using Brainwave Control' (2013).

**7**. Seth Kim-Cohen, *In the Blink of an Ear: Toward a Non-Cochlear Sonic Art* (New York: Continuum, 2009).

**8**. Douglas Kahn, *Noise, Water, Meat: A History of Sound in the Arts* (MIT Press, 1999).

**9**. Brandon LaBelle, *Background Noise: Perspectives on Sound Art* (New York: The Continuum International Publishing Group, 2006).

**10.** David Rokeby, 'The Construction of Experience: Interface as Content', in *Digital Illusion: Entertaining the future with high technology* (1998), pp. 27-48.

**11.** Todd Winkler, 'Creating Interactive Dance with the Very Nervous System', in *Proceedings of the 1997 Connecticut college symposium on art and technology, New London, Connecticut* (1997), pp. 212-17.

**12.** David Worrall, Michael Bylstra, Stephen Barrass and Roger Dean, 'Sonipy: The Design of an Extendable Software Framework for Sonification Research and Auditory Display', in *Proceedings of the 13th International Conference on Auditory Display, Montréal, Canada* (2007).

**13**. Yi-Fu Tuan, *Space and Place: Humanistic Perspective* (Springer, 1979).

**14**. Gaston Bachelard, *The Poetics of Space*. Vol. 330 (Beacon Press, 1994).

**15.** David Seamon, 'A Way of Seeing People and Place', in *Theoretical Perspectives in Environment-Behavior Research* (Springer, 2000), pp. 157-78.

**16**. David Griesinger, 'The Relationship between Audience Engagement and the Ability to Perceive Pitch, Timbre, Azimuth and Envelopment of Multiple Sources' http://www.davidgriesinger.com/IOA2011/IOA2011_preprint.doc [Accessed June 6 2012]

<http://www.davidgriesinger.com/IOA2011/IOA2011_preprint.doc> [Accessed June 6 2012].

**17**. Max. cycling74, San Francisco.

**18**. Fred Geil and David J. Weinber, 'Concert Audience Engagement - the Perception of Closeness: Engage Us with Sound', 2010. http://www.enjoythemusic.com/magazine/bas/0810/ [Accessed June 6 2012].

**19**. Guy Harries, 'Embodied Subjectivity and Audience Engagement Via Drawing in the Project Shadowgraphs' (2012). http://process.arts.ac.uk/content/embodied-subjectivity-and-audience-engagement-drawing-project-shadowgraphs [Accessed May 5 2013].

**20**. The program *ActionA/V* is available by contacting rogeralsop@gmail.com

**21**. Rolf Inge Godøy and Marc Leman, *Musical gestures: Sound, Movement, and Meaning*, Taylor & Francis (2009).

**22.** M. Leman, *Embodied music cognition and mediation technology* (MIT Press, 2008), pp 77-102**.**

**23.** M. Pedersen, `Sound Labyrinth: Exploration Of The Embodied Sublime Through An Immersive Audio/Visual Installation', in *Proceedings of the 2013 International Computer Music Conference* (Perth, 2013)**.**

**24** Alessandro Altavilla, Baptiste Caramiaux and Atau Tanaka, 'Towards Gestural Sonic Affordances', in *NIME'13, May 27-30, 2013, KAIST, Daejeon, Korea* (Korea, 2013).