Architecture as the Computer Interface: 4D Gestural Interaction with Socio-Spatial Sonification

Kirsty Beilharz
Faculty of Architecture, University of Sydney

Abstract. *Architecture today extends far beyond designing building shells and material, peripheral boundaries. Arguably, it has always been, and shifts increasingly in contemporary environments towards, designing space and interaction with space. Hence, the role of the designer includes integration of computing in architecture through ambient display and non-tactile interaction. This paper explores a framework in which the architecture is the computer interface to information sonification. (Sonification is automatically generated representation of information using sound). The examples in this paper are Emergent Energies, demonstrating a socio-spatially responsive generative design in a sensate environment enabled by pressure mats; Sensor-Cow using wireless gesture controllers to sonify motion; and Sonic Kung Fu which is an interactive sound sculpture facilitated by video colour-tracking. The method in this paper connects current information sonification methodologies with gesture controller capabilities to complete a cycle in which gestural, non-tactile control permutes and interacts with automatically-generated information sonification. Gestural pervasive computing negotiates space and computer interaction without conventional interfaces (keyboard/mouse) thus freeing the user to monitor or display information with full mobility, without fixed or expensive devices. Integral computing, a blurring of human-machine boundaries and embedding communication infrastructure, ambient display and interaction in the fabric of architecture are the objectives of this re-thinking.*

Keywords. *Interactive sonification, gesture controllers, responsive spaces, spatial sound*

Introduction

This paper proposes a framework for gestural interaction with information sonification in order to both monitor data aurally and, in addition, to interact with it, transform and even modify the source data in a two-way communication model (Figure 1). Typical data sonification uses automatically generated computational modelling of information, represented in parameters of auditory display, to convey data in an informative representation. It is essentially a one-way data-to-display process and interpretation by users is usually a passive experience. In contrast, gesture controllers, spatial interaction, gesture recognition hardware and software, are used by musicians and in augmented reality

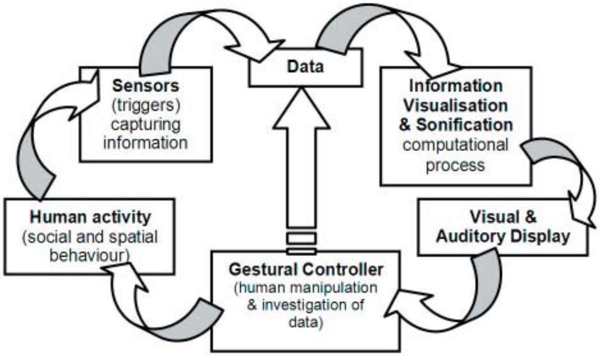


Figure 1. Knowledge flows from socio-spatial activities to sensors that capture data, through a computational process generating a visualisation/ sonification in real time. This loop is completed when gestural controllers are used for spatial interaction to manipulate or investigate this data.

systems to affect, manipulate and perform with sounds. Numerous installation and artistic works arise from motion-generated audio. These technologies are conflated into a single environment in which gestural controllers allow interactive participation with the data that is generating the sonification, making use of the parallel between spatial audio and spatial (gestural) interaction. Converging representation and interaction processes bridge a significant gap in current sonification models.

Social Contexts for Responsive Environments

Existing sonification often focuses on interpretation of abstract data such as meteorological, stock market trends, Internet traffic. These sonifications are removed from the data source: context, place and occasionally time. In contrast, the following examples and the Sensor Cow project focus on real time sonification in which the outcome and input are experienced simultaneously and co-locationally. Hence, the sonification is intended to help people understand their social and spatial activity and interaction (with other people and with space).

Emergent Energy (Figure 2) by Beilharz, Vande Moere and Scott is an iterative, reflexive system of interaction in which motion, speed, number of users and position in a space (triggering pressure sensitive floor mats) determine the growth of a visual design drawn with a Lindenmayer (L-system) generative algorithm. The design provides both an informative monitor of social and spatial behaviour and invokes users to interact with their space to influence their artistic surrounds. The design artefact is an embedded history of the movements, interactions and the number of people who produced it (Figure 2 & Table 1).

The sensate room configuration is explained in figures 3 and 5 which show the grid of pressure sensitive mats installed underneath the carpet connected via Making Things Teleo in/out modules (MakingThings, 2003) to Max/MSP + Jitter (IRCAM, 2003) programming environment.

Enabling buildings with responsive feedback capabilities facilitates flexibility and accessibility to assist environmental comfort, navigation for the visually impaired, building awareness, gerontechnology (technologies assisting the elderly), automated and augmented tasks for the physically disabled. Nanotechnologies - embedding minute

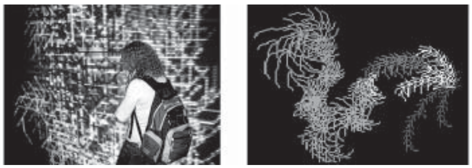
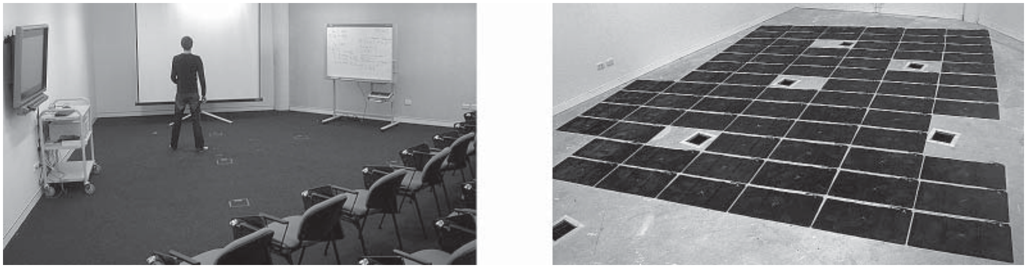


Figure 2. Lindenmayer System generator patch in Max/ MSP & Jitter used to create branched visualisations on screen. Different behaviours modify the algorithmic process of design generation. In the corresponding sonification, the number of people relates to dynamic intensity, position to timbre (tone colour) and speed to frequency (pitch).

Figure 3. The author's Sensate Lab (2 views) showing the "invisible" pressure sensitive floor mats embedded underneath the carpet, triggering the visual and auditory sound system and (before carpeting) the grid of pressure mats, networked to the Teleo modules.



sensor technologies in furnishings, surfaces and pre-fabricated building materials (Figure 4) - facilitate localised sensate regions and unobtrusive (wireless) distributed networks for data collection.

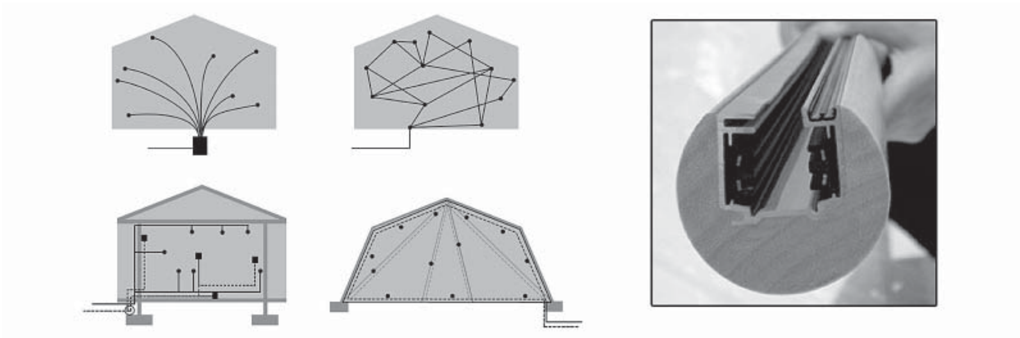
Mapping Towards Aesthetic and Engaging Ambient/Interactive Display

Scientific sonification or visualisation of abstract data is usually designed for the purpose of illuminating and augmenting our understanding of abstract (non-visual) data. There are contexts in which sonification is more helpful than visualisation: utilising the human auditory capacity for detecting subtle changes and comprehending dense data; and to avoid overload on visual senses, e.g. during surgery, anaesthesiology, and aircraft control. These applications of visualisation and sonification contribute to our understanding of well-known issues, particularly in regard to sonification: "orthogonality (Ciardi, 2004; Neuhoff, Kramer and Wayand, 2000) (i.e. changes in one variable that may influence the perception of changes in another variable), reaction times in multi-modal presentation (Nesbitt and Barrass, 2002), appropriate mapping between data and sound features (Walker and Kramer, 1996), and average user sensibility for subtle musical changes (Vickers and Alty, 1998)." There is also evidence to suggest that bimodal (visual and auditory) display has synergistic benefits for information representation (Song and Beilharz, 2004; Song and Beilharz, 2005).

Sound is even more integrally tied to space than light: "in a natural state, any generated sound cannot exist outside its context" (Pottier and Stalla, 2000) – space is a parameter of sound design, just as is pitch or timbre.

The extension to the (normally passive) process of automatically-generated real time auditory display by introducing gestural interaction with spatial

Figure 4. Media House in Barcelona developed by MIT media lab (left) shows the transformation from house with a computer, to house is a computer, to house with a network, to house structure is the network infrastructure and (right) building materials are being developed that embed network cable invisibly in the materials of the construct.



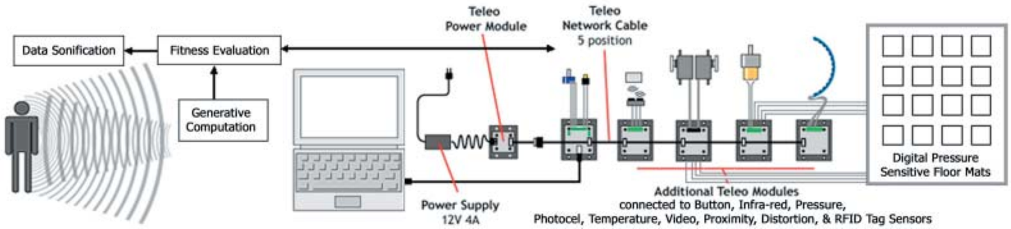


Figure 5. Sensate system: input from digital pressure sensor mats, infra-red, piezo pressure detection, temperature, light-sensitive photocells, proximity, RFID tags.

information to enable examination, transformation, and ultimately manipulation of the source information through 3D spatial and temporal control (the 4th dimension) (Bongers, 2000; Choi, 2000; Pottier and Stalla, 2000; Rovan and Hayward, 2000). For some kinds of information, auditory acuity is ideal for discerning time-based trends, patterns, recurrences and large data structures.

Ambient Display

Ambient visualisation and sonification in buildings merges informative information display with entertainment (infotainment or informative art) bringing a new versatility and purposefulness to graphical and auditory art in our homes and public spaces (e.g. Figure 6). "Ambient displays [including plasma, projection, touch screen and audio] normally communicate on the periphery of human perception, requiring minimal attention and cognitive load" (ADG, 2004). As perceptual bandwidth is minimised, users get the gist of the state of the data source through a quick glance, aural refocus, or gestalt background ambience. In relation to architecture, ambient representation that responds to the building (lighting, airflow, human traffic) as well as to social elements such as human clustering (flocking) patterns, divergences and task-specific data, adds a dimension of responsiveness to the spatial habitat.

4D Spatialised Audio representation and 4D gestural interaction

Both information visualisation and sonification employ a computational process for scaling data, converting it through an algorithmic process of representation, to produce the outcome (passively) received by the user. As the objective is to develop a greater rapport with the information, especially socio-spatial data (Beilharz, 2004) (how people move, relations in spaces, proximity to objects, clustering, eccentric behaviour, velocity and level of traffic), gestural interaction with the sonification enhances the ability to transform and manipulate the data by interacting with its representation. Because socio-spatial data is directly related to positional or spatial axes, 3D spatialised auditory

Sonification	Visualisation	Activity / Trigger
Pitch (frequency)	Length/scale/scope of graphic display on screen	Distance between activities / motion
Texture/density	Density of events / number of branches or iterations of generative algorithm (embeds history by amount of activity)	Volume of activity, number of users and social threshold
Rhythm/tempo of events	Proximity and rapidity of display (animation)	Speed of actions, punctuation of triggering events, tied to velocity of events
Intensity/dynamic loudness	Heaviness and distinction of on-screen drawing	Intensity/magnitude of triggering events
Timbre (tone colour)	Colour and distribution on visual display (screen)	Region/spatialisation – topology, zoning
Harmony	Design artefact	Multi-user manipulation

Table 1. Sonification schema of mapping correspondences

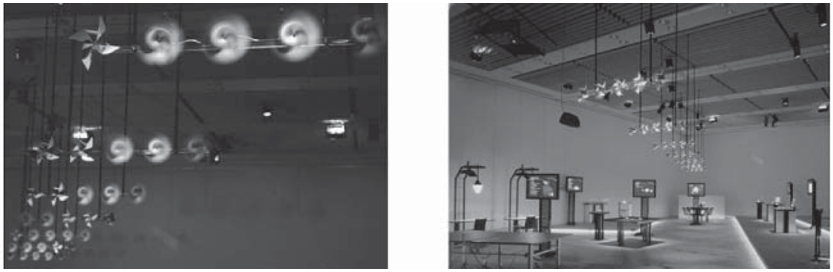


Figure 6. 'Pinwheels': visualising information flow in an architectural space - Hiroshi Ishii, Sandia Ren, and Phil Frei, 2001.Information usage activates the rotation of pinwheels as a visual and visceral indicator of traffic.

representation in physical space over time (i.e. 4D) provides an ideal relation with the source data. Human gestures and sound operate in three dimensional spaces and in the temporal domain (also 4D). Four dimensional interaction (without the traditional mouse and keyboard interfaces) underlies the paradigm of non-tactile, seamlessly-integrated, pervasive, immersive computing in which the computing hardware becomes invisible and more integral to the building structure and space. This is aesthetically beneficial in architectural design.

Using Gestural Controllers and Spatial Interaction to Engage with Data

Introducing gestural controllers as a mechanism for interacting with the 3D spatial auditory and visual representation of information takes this process one step further. The barrier between humans and information, between humans and the smart building are disintegrated while computation and sensing are conflated into a single organism: the intelligent building. Interactive sonification has been used in the past to provide a tangible means for users to negotiate and manipulate the display (Hunt, 2004). In this framework, the interaction can affect and manipulate the data source (determining the display), not only its representation.

The science fiction film, Steven Spielberg's Minority Report (Maeda, 2004) forecasted a kind of interface that is already now achievable: spatial and gestural manipulation of video and computer data on a transparent screen suspended in 3D space (Figure 7). The notion behind gestural information access is an important one: dissolving the hardware and unsightliness of computer interfaces. As computing moves towards people acting in spaces, deviating from our currently sedentary desk-bound lifestyle, the importance of the spatial interaction and experience design, the way in which information is represented, becomes essential. Building architecture and information architecture become one (Figures 4, 7 & 8).

Sonification and Mapping in the Sensor-Cow Project

A distinctive timbre (tone colour) is attributed to each sensor to distinguish the sounds arising

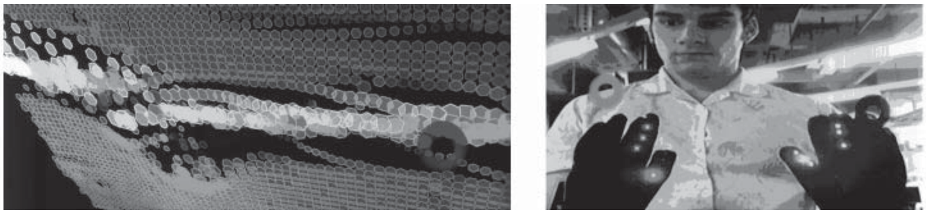


Figure 7. Justin Manor's Manipulable Cinematic Landscapes (Maeda, 2004) is a glove-controlled cinematic landscape interface in 3D space.

Figure 8. Haptic (tactile) manipulable cubes in Reed Kram's Three Dimensions to Three Dimensions (left) are creative tools for expression while sensors attached to digits and limbs can be used as gestural controllers for music (right) (Choi, 2000; Pottier and Stalla, 2000; Rovan and Hayward, 2000).

from each sensor. The rhythm, acceleration and velocity of action is directly realised in real time. The correspondence between rapid gestures and rapid sonification is literal. For both the acceleration and gyroscopic sensor, extremes of motion away from the median, drives the pitch in directional extremes away from a central pitch region. The direction of pitch, ascending and descending away from the mean, corresponds to the x-axis direction of motion. Changes in direction are audible and circular motions of the ear and head produce sweeping auditory gestures that reinforce the audio-visual connection between activity and sonification (Figure 9). The sonification patch was programmed in Max/MSP (+Jitter) using La Kitchen's Kroonde Gamma object (Henry, 2004) to receive data from 4 active sensors. The wireless hardware is recognised using CNMAT Berkeley's Open Sound Control (CNMAT, 2004) object.

Monitoring Spatial Activity in Architecture

While the sensor-cow project is a monitor of calf motion, and it is unlikely that the calf understands the affects of its actions in contributing to the sonification, socio-spatially generated sonification has the potential in a human context to invoke interaction with the outcome. Seamless integration of spatial experience and computational response is a paradigm shift that is essential to the future of designing spaces.

Sonic Kung Fu – Web-cam Colour Tracking and Gestural Interaction with Sound

'Sonic Kung Fu' by Jakovich and Beilharz (exhibited at Sydney Esquisse exhibition, March 2005) is a sonic art installation in which participants wear coloured gloves to perform gestures that produce a real time responsive audio sound-scape (Figure 10). A web cam receives the visual gesture information. The Max/MSP patch responds to the motion of a specific colour (calibrated to match the glove being worn), responding with auditory variation across a range of x and y- axis values. The immediacy and mapping of this work was intentionally as simple and intuitive as possible for recognition to encourage interaction by passers-by in a gallery setting. The result was that users spent considerable time with the 'instrument' learning to understand and control its performance.

Gesture Affected Computation: Completing the Interaction Loop

Finally, translating gestural interaction in 3D space into affectors (software commands) that manipulate the source data demonstrates a complete cycle in which social activities and movement throughout a room produces the sonification that, in turn, can be transformed by the participant. Affectors, in programming terms, are gestures that

Figure 9. Sensor-Cow: bi-directional (mercury) motion sensors are attached to the calf's front legs, a gyroscopic sensor on the forehead and accelerometer on his right ear. The pouch hanging around his neck contains the radio frequency transmitter that sends the real time data to the (La Kitchen) Kroonde Gamma wireless UDP receiver (LaKitchen, 2004). It is connected by Ethernet to the computer running the data sonification with Max/MSP object-oriented programming environment.

Figure 10. Gestural interaction with auditory display created in response to colour tracking of the spatial glove motion.

trigger a change in information, e.g. motion acceleration thresholds, direction, velocity. Spatialised audio display (e.g. using IRCAM's multi-channel SPAT) locates sound attributes in 3D space, making it easier to identify, distinguish, then manipulate specific sounds. Sound represents data through the mapping process, so moving the sound or interacting with it gesturally is essentially a reverse-mapping procedure that alters the data set (Figure 11).

Conclusion

This paper sets out a framework for linking the use of gesture-controlled audio with traditionally passive information sonification. The bridge provided by affecting change in a data set achieved through gestural manipulation of sound completes a loop in the cycle of human-computer interaction. Importantly, the proposed method of transforming data also provides a 4D spatial mode of interaction that is suited to 4D interaction environments, such as Virtual Reality and Augmented Reality as well as pervasive (or ubiquitous) computing in physical architectural spaces. The use of auditory display increases immersion, broadens attentiveness and especially suits information assimilation in already visually-rich environments or those situations where auditory acuity is superior (time-based patterns or low-visibility conditions). Gestural interaction moves towards a seamless integration of architectural (living and working) space and intuitive computing and towards socio-spatially informative architecture.

References

ADG 2004 Berkeley Ambient Display Group http://www.cs.berkeley.edu/projects/io/ambient/

Beilharz, K. (2004). "(Criteria & Aesthetics for) Mapping Social Behaviour to Real Time Generative Structures for Ambient Auditory Display (Interactive Sonification)", in E. Edmonds and R. Gibson (eds), INTERACTION - Systems, Practice and Theory: A Creativity & Cognition Symposium, Creativity and Cognition Press, The Dynamic Design Research Group, UTS, Sydney, pp. 75-102.

Bongers, B. (2000). "Physical Interfaces in the Electronic Arts. Interaction Theory and Interfacing Techniques for Real-timePerformance", in W. W. Wanderley and M. Battier (eds), Trends in Gestural Control of Music, IRCAM - Centre Pompidou, Paris.

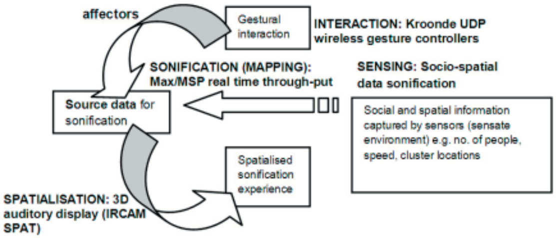


Figure 11. Gestural interaction using gesture controller devices can be used to affect (change) the source data that produces the information sonification in real time.

Choi, I. (2000). “Gestural Primitives and the Context for Computational Processing in an Interactive Performance System”, in W. W. Wanderley and M. Battier (eds), Trends in Gestural Control of Music, IRCAM (Institute of Research in Computer and Acoustic Music) Centre Pompidou, Paris.

Ciardi, F. C. (2004). “sMAX: A Multimodal Toolkit for Stock Market Data Sonification”, in S. Barrass and P. Vickers (eds), Proceedings of ICAD 04 -Tenth Meeting of the International Conference on Auditory Display, Sydney, Australia.

CNMAT 2004 Open Sound Control (OSC) http://cnmat.cnmat.berkeley.edu/OSC/Max/#downloads

Henry, C. (2004). La Kitchen Kroonde Gamma - User Manual & Max Patches, La Kitchen, Paris.

Hunt, S. P. A. (2004). “A Toolkit for Interactive Sonification”, in International Conference of Auditory Display (ICAD), Sydney, Australia.

IRCAM (2003).Max/MSP, Cycling 74: Max/MSP is a graphical object-oriented programming environment for real time control of music, audio, and multimedia objects, capable of hardware and software integration, USA. http://www.cycling74.com/products/maxmsp.html

LaKitchen 2004 Kroonde Gamma 16-sensor Wireless UDP Interface http://www.cycling74.com/products/kroonde.html 03/08/04.

Maeda, J. (2004). Creative Code, Thames and Hudson, London.

MakingThings 2003 Teleo Modules and Sensor Devices http://www.makingthings.com/products/products.htm 1/1/04.

Nesbitt, K. V. and Barrass, S. (2002). “Evaluation of a Multimodal Sonification and Visualisation of Depth of Market Stock Data”, in R. Nakatsu and H. Kawahara (eds), International Conference on Auditory Display (ICAD), Kyoto, Japan.

Neuhoff, J. G., Kramer, G. and Wayand, J. (2000). “Sonification and the Interaction of Perceptual Dimensions: Can the Data Get Lost in the Map?” in P. R. Cook (eds), International Conference on Auditory Display, Atlanta, Georgia, USA.

Pottier, L. and Stalla, O. (2000). “Interpretation and Space”, in W. W. Wanderley and M. Battier (eds), Trends in Gestural Control of Music, IRCAM (Institute of Research in Computer and Acoustic Music) - Centre Pompidou, Paris.

Rovan, J. and Hayward, V. (2000). “Typology of Tactile Sounds and their Synthesis in Gesture-Driven Computer Music Performance”, in W. W. Wanderley and M. Battier (eds), Trends in Gestural Control of Music, IRCAM (Institute of Research in Computer and Acoustic Music) - Centre Pompidou, Paris.

Song, H. J. and Beilharz, K. (2005). “Time-based Sonification for Information Representation”, in World Multi-Conference on Systemics, Cybernetics and Informatics, Orlando, USA, pp. in press 28/01/05.

Vickers, P. and Alty, J. L. (1998). “Towards some Organising Principles for Musical Program Auralisations”, in A. Edwards and S. A. Brewster (eds), International Conference on Auditory Display (ICAD), Glasgow, Scotland, U.K.

Walker, B. N. and Kramer, G. (1996). “Mappings and Metaphors in Auditory Displays: An Experimental Assessment”, in International Conference on Auditory Display (ICAD), Palo Alto, California, USA.