Consciousness Reframed 2004

Presence and Sound; Identifying Sonic Means to "*Be there*".

Paul Doornbusch & Sarah Kenderdine

Summary:
The concept of "presence" is a major part of the discourse of VR. Creating a sense of presence can be crucial to the effectiveness of museum exhibits and VR presentations. Sound can play various roles in technology-based presentations, by creating or enhancing a sense of presence, and providing a continuity of experience. This paper discusses these points and how presence and sonic identification are important to such multi-disciplinary works, with examples drawn from the creation and implementation of the sound for a VR presentation of the Angkor ruins in Cambodia.

Keywords: Sound; Presence; Virtual Reality (VR).

Introduction:

Virtual reality presentations usually strive for the greatest sense of "presence" or "being there" possible. Thus observers can "suspend their disbelief" and believe they are completely engaged in another world, allowing for the possibility of internal dialogue and reflection, because of the confluence of sensory perception, maximising the engagement with the work.

Sound is indispensable for creating complete engagement with virtual environments of almost all kinds (Doornbusch 2002). Particularly with photorealistic visual imagery, similarly realistic sonic elements are required to fully engage the observer, avoiding cognitive dissonance associated with sonic elements that mismatch the fidelity of the image.

The crux of this issue is what Pierre Schaeffer termed the source-cause, where the listener identifies the source or cause of the sound. While initially this may seem a trivial matter, there are many subtle effects at play. Our hearing is very adept at quickly analysing the sonic environment and identifying elements within it such as the cause or source of a sound, the setting or environment of a sound, and the space of a sound such as its location and surrounding physical environment. The extension of Schaeffer's concept to source-identification, setting-identification and space-identification, is discussed at length below.

Of particular interest with VR presentations is that the sonic source-identification, setting-identification and space-identification are appropriate to the image displayed. While this may be easily managed with less than realistic images, as photorealistic images become the norm in VR presentations then the degree to which the sound must match is similarly more demanding.

How sound is disregarded

Hearing is a generally neglected sense, particularly in our general environment, despite the richness of information which sound delivers. That we are largely visually dominated is indisputable. However, in our reductionism, we overlook many things deemed non-essential and this is true of the sound milieu. Although we disregard sounds unconsciously, the subconscious surely takes note and maintains a coherent and consistent representation of our environment. We concentrate on obvious sounds and take for granted the thousands of details that give a setting for the sounds we hear everyday. Quoting Bryson, "Perhaps one may draw on the distinction made by Charles Stirling between 'megalography' and 'rhopography'. Megalography is the depiction of those things in the world which are great – the legends of the gods, the battles of heroes, the crises of history. Rhopography (from *rhopos*, trivial objects, small wares, trifles) is the depiction of those things which lack importance, the unassuming material base of life that 'importance' constantly overlooks. The categories of

megalography and rhopography are intertwined. The concept of importance can arise only by separating itself from what it declares to be trivial and insignificant; 'importance' generates 'waste', what is sometimes called the preterite, that which is excluded or passed over." (Bryson, p.41) While focusing on the "important" we overlook the "trivial", yet in a complex and rich soundscape it is exactly those "trivial" sounds - the midground and background and spatial sounds - which make the setting so compelling and informative of the environment. It is similar with images; the details in a high-resolution photograph are what make it compelling compared with a low-resolution rendering.

Bateson makes a strong case for the use of multiple types of information to more completely understanding any situation. "… various ways in which the combining of information of different sorts or from different sources results in something more than addition. The aggregate is something more than the sum of the parts because the combining of the parts is not a simple adding but it is the nature of a multiplication or fractionation, or the creation of a logical product. A momentary gleam of enlightenment." (Bateson pp80-81)

The inclusion of sound with an image, what I call "augmentation", provides an enhanced experience for observers due to the input of other information from another sense. For example, a piece of skilfully edited and selected music underneath an animated or film scene, while clearly unrealistic, can provide the observer with an enhanced experience and continuity by evoking both a context aurally while also providing a continuity of experienced in one of the senses.

However, the inclusion of accurate sounds, with appropriate presentation of sonic source-identification, setting-identification and space-identification, goes beyond "augmentation" to create a powerful synthesis of sensory input such that the engagement is a qualitatively different experience. We unconsciously use sophisticated sonic cues to understand our environment and the inclusion of realistic soundscapes with the appropriate images informs the understanding.

Image, sound, and other sensory input, are different languages to understand the world. They provide us with separate pieces of information, but each is limited in different ways. Hayakawa argues that language is both a means of communicating ideas as well as limiting what is said. "Whatever may be the language one happens to inherit, it is at once a tool and a trap. It is a tool because with it we order our experience […] What is true of verbal language is also true of visual 'languages': we match the data from the flux of visual experience with image-clichés, with stereotypes of one kind or another, according to the way we have been taught to see. […] If the abstractions, the words, the phrases, the sentences, the visual clichés, the interpretative stereotypes, that we have inherited from our cultural environment are adequate to their task, no problem is presented. But like other instruments, languages select, and in selecting what they select, they leave out what they do not select. The thermometer, which speaks one kind of limited language, knows nothing of weight. […] Every language, like the language of the thermometer, leaves work undone for other languages to do." (Hayakawa, in Kepes 1969)

Additionally, Shklovsky discusses the pervasiveness of everyday experience and how we ignore our environment: "We live in a poor and enclosed world. We no more feel the world in which we live than we feel the clothes we wear. We fly through the world like Jules Verne characters […] in our capsule there are no windows. […] The Pythagoreans used to say that we do not hear the music of the spheres because it goes on uninterruptedly. In the same way those who live by the sea do not hear the noise of the waves. […] We live as if coated with rubber. We must recover the world." (Shklovsky, 1923)

Although sound is largely disregarded, there is much information about the environment within the soundscape. We correlate this information with our other senses to build a complete map of the space and world. Even if we do this somewhat unconsciously, it undoubtably takes place and sound plays a vital part in our understanding and sense of the environment in which we are immersed.

Sound, reality and identification

The source-identification of a sound is of greatest significance when determining how "real" a sound is. Source-cause theory seeks to relate sounds to their physical causes, which are profoundly connected to reality. Schafer describes the increasing variety in our environmental soundscape, saying that the meaning within it is decreasing because with standard signals for particular messages (alarms etc.) there is reduced need to evaluate sounds. Despite this, the skills to identify the source of sounds are not far below the surface of human consciousness, as proved by the numerous aircraft spotters in WWII who would accurately tell the type and number of approaching aircraft by the sound.

The context in which a sound exists, the setting and space, also contribute to the sense of a sound belonging to reality. Regardless of how accurately we may recognise the source of a sound as real, the setting of the sound and the space it occurs in contribute greatly to the sense of reality or not. While the sighted navigate space visually, the blind will navigate a space aurally, counting lampposts and doorways on the street as they pass them, and aurally appraise a space to determine its size and shape, the surface coverings etc. Thus the unsighted make a mental map from the sounds to navigate by. For the sighted, the mental maps from the visual and aural input match, but in a VR presentation there may be discontinuity between these two maps that causes cognitive dissonance. While contemporary society and media practice may make people forgiving of such mismatches, a great sense of reality is achieved when they match to a high degree of fidelity. While this is demonstrable in practice, like much of psychoacoustics and psychology itself, it is difficult to "prove" scientifically, which is not attempted in this work.

There are several crucial elements for sound to enhance the sense of presence:

- Source-identification - the source(s) of sounds can be authentically identified.
- Setting-identification - the environmental ambience matches the visual setting.
- Space-identification - the acoustic space faithfully reconstructs the expected acoustic image of the presented visual image and the location of the sounds.

These elements are interlinked, yet individually crucial. Achieving a high degree of fidelity in these areas creates a strong continuity of experience, where the aural and visual "maps" converge, which is required to suspend disbelief. These allow for identification of the source of sound because the fidelity is sufficient to define it in our mind. The setting is cued because the ambient sounds that accompany the source can be heard with sufficient clarity, and authenticity, to convince the listener that it is a real environment. The space the sounds occur in can be identified because the early and late reflections are accurately captured or synthesised, such that there is a high degree of congruence between the reflected sound and the visual presentation of the space.

Angkor, a practical example

In the Angkor project, high-resolution, stereographic, photographic panoramas were taken from several locations within the Angkor complex (Kenderdine 2004). This visual imagery, especially when presented on a large screen and in full scale, is stunning in its lifelike richness and detail. When presented in the Virtual Room (VROOM, an octagon of screens with a four-channel surround-sound system per screen: refer to http://www.vroom.org.au), there exists sufficient visual and audio playback resources for a semi-immersive experience. Initial

experiments were undertaken where various sounds were used to augment the images, such as traditional music, or various sound effect library soundscapes of jungles and so on. What became apparent was that while this approach may have been suitable for a film or television presentation, and it was clearly significantly better than having no sound, it was insufficient to achieve the sense of presence that was desired.

Source recordings were required from the temples themselves to achieve the required fidelity of source-identification and setting-identification. In the Angkor project, sound is used to create a strong sense of presence and a sense of continuity of experience. It was found that source recordings from the sites, despite their limited flexibility, were the most useful sound sources because of their strong source-identification and setting-identification properties. However, much more than simple recording playback is required, as the recordings need to be modified substantially to achieve the relevant presence. Murray Schafer documented how the soundscape changes throughout the day, and year, based on environmental activity, and this needed to be taken into consideration, often modifying the source sounds (editing, filtering, processing) to make them suitable in terms of setting-identification and space-identification, as perfect recordings from appropriate places were not always possible. An example of the subtle effects that greatly enhanced the sense of presence was the use of early-reflection modelling for sounds near walls. As the panorama rotated, the sound (played back through a 4-channel system), tracks the rotation with early reflections from nearby walls coming from the appropriate locations.

There was a requirement for dialogue in the soundscape, but it was discovered that this introduced voice quickly destroyed the illusion of presence because of its artificial nature. A solution was found with the use of bluetooth headsets for the dialog, which people could choose to use or not. A summary of the findings is:

- Source recordings worked best, but were only a basis.
- Dialogue interrupted presence or the continuity of presence.
- The sound provided a stabilising element for the movement or slow rotation of the image, which previously, without the sound, caused seasickness.
- Visual cues were not always needed, eg birds, but could help.
- Narrative/illustration is possible with careful sonic augmentation, eg water sounds to illustrate water features that are dry, but setting and space must be carefully maintained.

In the Angkor project, several points emerged during the construction of soundscapes as being crucial to providing the greatest sense of presence:

- Source-identification, where source(s) of sounds can be authentically identified as from specific visual elements.
- Setting-identification, where the environmental soundscape accurately matches the visual setting.
- Space-identification, where the acoustic space faithfully reconstructs the expected acoustic image of the visual image and the location of sounds within that environment.

One of the most significant discoveries was that the strong sense of presence achieved with the best sound encouraged an internal dialogue for the observers, such that contemplation could be achieved similarly to the real site.

Conclusions

The source-cause model for sonic identification, first proposed by Schaeffer and enhanced by others, can be extended from source-identification with the concepts of setting-identification and space-identification, as a means of providing a way to build soundscapes that maximise the sense of presence. The accurate matching of visual and aural maps and the lack of cognitive dissonance causes a confluence of events in the observer such that the presence achieved can be greater than the sum of the parts - a holistic experience is attained. In the Angkor project the great sense of presence, of belief in being there at the site, allowed an internal dialogue of the contemplation of ruins to develop which was not dissimilar to that experienced at the actual site.

Acknowledgments:

We would like to gratefully acknowledge the following organisations for their generous support in enabling this research: APSARA-Authority, Cambodia; Museum Victoria; the Virtual Room project; I-Cubed (RMIT).

References:

Bateson, G. 1979. *Mind and nature: A necessary unity*. New York: Bantam Books, pp. 80, 81

Bryson, N. 1990. *Looking at the Overlooked: Four Essays on Still Life Painting*. London: Reaktion Books, p. 41

Doornbusch, P. 2002. *Sound and Reality*. Melbourne, Australia. SimTech Conference Proceedings CDROM.

Field, A. 2000. *Simulation and reality: the new sonic objects*. in Emmerson, S. (Ed) *Music, electronic Media and Culture*. Hants, England: Ashgate Publishing Limited.

Hayaka in Kepes, G. (Ed) 1969. *The Language of Vision*. Chicago: Paul Theobald, p. 8

Kenderdine, S. 2004. *Avatars at the Flying Palace – Stereographic Panoramas of Angkor Cambodia*. The International Conference on Hypermedia and Interactivity in Museums (ICHIM) 2004 Berlin: CDROM Museum Archives and Informatics.

Schaeffer, P. 1966. *Traité des Objets Musicaux*. Paris: Editions du Seuil.

Schafer, R. M. 1980. *The tuning of the world*. Philadelphia: University of Pennsylvania Press

Shklovsky, V.B. 1923, *Form and material in art.* Internet: Center for Book Culture.org. Viewed 20 May 2004, http://www.centerforbookculture.org/context/no2/shklovsky.html

Paul Doornbusch (pauld@iii.rmit.edu.au) is a composer and sonologist at the Interactive Information Institute of RMIT University, Melbourne, Australia. Mr Doornbusch is also President of the Australasian Computer Music Association, and he is responsible for the sound infrastructure and content for the Virtual Room of the Melbourne Museum. http://www.iii.rmit.edu.au/

Sarah Kenderdine (skender@museum.vic.gov.au) is project manager for The Virtual Room of the Melbourne Museum. Miss Kenderdine trained as a marine archaeologist and museum curator, but since 1994, she has designed digital environments for cultural institutions. Miss Kenderdine is a Director of the Virtual Heritage Network.