Real Time Art Engines: Sound in Games
Adam Nash
Novelty, fidelity, utility

Sound and real time 3D graphics share some interesting qualities as media. Sound has already been through the cycle, which I identify as 'novelty, fidelity, utility'. This is the cultural phenomenon of a new technology being initially surprising – novelty; then used to do or recreate something that can already be done – fidelity; then finally integrated into culture on its own terms and used to create something original - utility. In the second half of the twentieth century the medium of sound played out this cultural cycle first with the synthesiser. Initially surprising, even shocking, the sounds of synthesisers were quickly corralled into an existing vocabulary of instrument names and timbre descriptions ('violin', 'reed'). Accordingly, a cultural expectation emerged of these new machines that they produce faithful renditions of the 'real' instruments that the retrofitted vocabulary was describing. Finally emerged artists who were prepared to use the synthesisers as bona fide instruments on their own terms, producing unique and original sounds in the company of conventional instruments.

It would appear that computer gaming (or, specifically, the real time 3D graphics on which the computer gaming industry is crucially dependent) is currently still within the fidelity stage of this cycle, although various indicators suggest it is near the end of this phase. For sound, fidelity meant ‘sounding like physically extant musical instruments’, but for computer gaming it often means, 'looking like physically extant spaces'. This implies ‘sounding like physically extant spaces’, but in today’s competitive marketplace there is also the demand to ‘sound like a movie’. Movies, which have gone through their own novelty, fidelity, utility cycle, often sound like a big orchestra playing a dramatic, emotion-tweaking score and therefore so do games. There is no denying the extra enjoyment derived from playing World of Warcraft while an entire symphony orchestra pumps you along, seemingly playing the score just for you immersed in your adventure. Similarly, the impact of environmental sounds, crisply reproduced in surround sound is undeniable in its ability to add to the feeling of engagement in the game world.

However, the desire for this type of fidelity has set a precedent leading to the widespread, but under examined belief that the goal of all audio within games (and again this belief has its roots in cinematic convention) is to be ‘heard but not noticed’, that the role of sound and music is only to reinforce the visual narrative of the game. This belief is perhaps most startlingly evident in this statement by well-known game audio practitioner and academic Leonard Paul, ‘if the music is too closely tied to events which the player can control, the player may discover how to ‘play’ the music and subvert its normally passive influence’. (Paul, 2003)

The fidelity phase can also be critiqued as one in which artistic progress made during the novelty phase is forgotten. Artists often do their best work in restricted environments and in this way, the so-called 8-bit soundscapes and effects of the early video games from the likes of Nintendo, Atari and Sega in the 1970s and 1980s can be seen as much more original and inventive sonically than the over-orchestrated cinematic scores and slavishly ‘realistic’ sound effects in current computer games. Rob Bridgett and Leonard Paul discuss this in their article Establishing an Aesthetic in Next Generation Sound Design. (2003)
Identifying the danger of a sonic ‘grey goo’ caused by too many sounds being used simply because the new generation gaming hardware can accommodate it technically, they essentially advocate a formalist approach to computer game sound design based on imitating the technical limitations of the pre-digital era in order that the underlying sonic and/or musical idea guide the creation. While theirs is an important and interesting point, it is also important that an approach to sound design in games is based on an evaluation of the real time 3D medium on its own terms, rather than only its superficial, often misleading, relationship to cinema.

Real time 3D environments offer the ability to spatialise not only the sounds, but also the composition. It is this area I have spent several years exploring, and which led to igloo inviting me to contribute to the sound design of SwanQuake. A real time 3D environment allows a user unprecedented access to the form and experiential reception of composed sound and emancipation from a strictly linear perception of sound or music. Each of my works, Memory Plains Returning, Scorched Happiness and Pale Shining Winter, explores this notion by placing sounds that I created within an environment which users navigate independently to experience the soundscape. The interesting challenge of the 3D environment is that all the visuals that define the space are simply visuals, without mass or acoustic properties. Therefore, all sounds need to be separately placed and spaced. Yet these ‘schizophrenic’ sounds, or objets sonores (sounds without context) are the building blocks of a cohesive sonic atmosphere that helps create an overall context. If this works, the result should be different every time for each user, while my underlying compositional values and intentions are maintained.

This means the user takes an active role in the composition, in that the sonic ingredients all exist within the space (some in static positions, some on animated paths) and the user largely determines how these ingredients are put together, by navigating through the space. Often, the user is not in
complete control of their experience because they are disoriented spatially and unable to return to exact locations at will (these are abstract visual spaces without conventional orientating markers of up, down, in, out, etc), and this disorientation is heightened by the mobile nature of the animated sounds. As with any environment, familiarity and exposure breed a kind of physical memory that enables the user to more confidently navigate the space as an audio space to be ‘played’ in an active manner. Ultimately, such environments can be seen as both audiovisual compositions and instruments at the same time, and this is how I fundamentally see the medium, as an audiovisual space where the composition and the performance converge.

Within this concept of composition and performance converging is the exciting possibility of in-world sound synthesis. Currently the accepted, and pretty much only, method of getting sounds into a real time 3D space is by inserting digital audio files prepared externally into the environment. Current game engines can place certain effects upon the sound, such as echo, delay and reverb, chiefly by using the capabilities of higher-end consumer sound cards. However, none of these change the fundamental structure of the sound dynamically. Spatialisation can be achieved through the ‘surround sound’ capabilities of the consumer’s sound card and speaker system, with the intention to increase a user’s sense of immersion in place by having sounds ‘come from the side’ or ‘behind’ the user, but again these are also ‘effects’ that do not change the fundamental structure of the sound.

UK-based digital sound programmer Andy Farnell explains that the current state of audio in games can be compared to game visuals had real time 3D rendering never been invented and all games were presented as a series of static 2D images. (Farnell, 2016) In other words, all sounds must be prepared as a static snapshot of an audio event (eg, footsteps), and the game/sound designers must anticipate all possible contingencies beforehand (eg, footsteps on snow, footsteps on metal, footsteps on interior corridor, footsteps on exterior cobblestone plaza, etc). Due to the often prohibitively large filesize of digital audio, often a single sound will be made to serve several purposes for which it is not optimally suited, resulting in the well-known audio situation of the annoyingly repetitive footstep. This brute force approach becomes less and less viable as visual production values increase, since the gap between visual ‘realism’ and aural ‘realism’ widens by sticking with this approach. As Farnell explains:

Native digital synthesis is the future for multimedia and games. Instead of recording a bunch of sounds like photographs the synthesist creates code to be executed at runtime. These are the sounds, expressed as procedures or formulas which will be run by the client hardware. Synthesis is a direct analogy to the 3D rendering engine used for most modern games, only difference is that it’s a sound rendering process. (Farnell, n.d.)

The advantages of this approach include not only a substantial, almost revolutionary, reduction in filesize required for audio within a game, but also a massive improvement in the level of integration within the environment that audio will be able to achieve, based as it will be on the physical simulation code governing the other aspects of the game experience such as visuals and animation. Once this native sound synthesis has been achieved, then real time
3D environments will have entered the phase of utility, at least from an audio perspective. There is a significant chance, of course, that native sound synthesis itself will be forced through its own cycle that perhaps misses the first phase of novelty altogether since it will be being developed with the express aim of replacing digital audio sampling in the role of fidelity. Nonetheless, native sound synthesis will present an extraordinary opportunity for designers of real time 3D environments to create dynamic, real-time aural compositions.

The acousmatic approach for SwanQuake

When Ruth Gibson and Bruno Martelli of igloo first asked me to compose the sound and music for SwanQuake, we were working together on Sentient Space with Company In Space in Melbourne (I was joining the projects’ other sound artist Luke Pither). They presented me with a series of visual mood boards for each environment of SwanQuake, along with verbal descriptions and some of the level designs, specifically the nuclear power station and some parts of Mars. They asked me to go away and come up with sonic and musical signatures for each environment, providing only the broadest brush strokes regarding the content of the sound. It was clear that they were not expecting a linear or cinematic approach to the sound in SwanQuake, but were interested in having me engage with the project (and its medium) in my fullest capacity as an artist and to push our understanding of what might constitute ‘sound’ or ‘music’ in this context.

I chose to use an acousmatic approach, one I have been using for many years, to composing and constructing the sound and/or music for the project. This is sample-based, but one that refers to the composition process itself, rather than how that will be manifested within the engine. This is a ‘musical’ decision, and is not unique to composing for 3D environments. There is always the choice available to a composer between samples, synthesisers and real world recording. Most often, a composer will take an approach that combines all three approaches. Further complicating the explanation is the fact that samples will often be samples of synthesisers and/or real world sounds or musical instruments.

The acousmatic approach can be seen as a natural convergence and evolutionary result of musique concrete and electronic music. According to Jonty Harrison’s definition, acousmatic composition...

...admits any sound as potential compositional material, frequently refers to acoustic phenomena and situations from everyday life and, most fundamentally of all, relies on perceptual realities rather than conceptual speculation to unlock the potential for musical discourse and musical structure from the inherent properties of the sound objects themselves. (Harrison, 1999)

This approach seemed appropriate for SwanQuake for a variety of reasons. First, it seemed that one of the goals of the visual and experiential aspects of the project was to achieve a kind of immersion on the part of the user that may be described as another place. This is best thought of in terms of later memories of the user experience, when the feeling is of ‘having gone somewhere’. To aid this sonically, it is probably important that all sounds experienced within this ‘place’ share a qualitative commonality at the physical level. In
audio production terms, this commonality is often applied at the end of the production process, during mastering, where some spatial quality such as reverb is applied along with a uniform compression of all sounds to marshal them into a semblance of uniform provenance, since the individual sounds themselves have often been sourced from widely varied origins.

A method that I have found useful, is to build all sounds from the same sample. A ‘sample’ in this case refers to a recording of an actual or realworld sound. This sample then forms the base material that is twisted, effected and manipulated according to formal rulesets or empirical audioaesthetic decision sets. Since the sounds originate from the same physical source, they continue to share a quality that seems to justify their coexistence in the sonic space. By using the same sample, or groups of samples, there is guaranteed a kind of fundamental integrity to the sounds that hopefully is retained regardless of the amount or manner of manipulation that is performed on the sound. This integrity is born from the fact that it is actually the same ingredients that are being manipulated every time. It is possible, then, to ask the ears to follow along a sonic narrative that seems to progress logically step by step, without any shocking sonic event to cut the thread of comprehensibility. It’s not as simple as that though. Since the 3D visuals are presenting a cohesive, apparently physical, space to be navigated by the user, the user subconsciously expects the sonic space to match it. However, this is a visual space that doesn’t actually exist, and often couldn’t actually exist (SwanQuake’s Piranesi environment for example), or if it could it would not be possible for the user to actually experience (SwanQuake’s Mars environment for example) therefore it is impossible for the user to know what it should sound like. In such situations, there should be a shock to the user, the moment when he or she becomes conscious of the sonic space.

It is this moment, this bringing into consciousness a consideration of the nature of the relationship between the user and the sonic space, as mediated by the visual space and the experiential space, that offers the most exciting potential to the sound artist working in this environment, because the user is invited to consciously consider. This consciousness flies, wonderfully, in the face of the accepted (largely cinematic) wisdom that the sound should be unnoticed or subconscious. In my view it is impossible at this point for the sound to be unnoticed by the user as they enter the space, and the users’ experience of the space will be all the better because of this. It is in this moment, as the user consciously asks of the remediated space, ‘what does it sound like?’ that a request is made of the user by the sound artist. It is a specific request, in exact contradiction of Leonard Paul’s proposal, for the user to ‘play’ the soundtrack, to ‘play with’ the soundtrack, to engage in a conceptual back and forth with the sound artist to see how it can unfold.

SwanQuake: environment descriptions
For all the environments of SwanQuake, the sounds broadly fall into one of two categories: ‘foley’ (or sound effects) and ‘soundtrack’. Foley is traditionally for making objects sound like they look; a door creaking and slamming, footsteps, car engines roaring, etc. Whilst there are many points where Igloo specifically asked for a linear visual/sonic mapping as found in the real world, there is no guarantee that the foley effects employed in SwanQuake will always serve to
1.2 Real Time Art Engines 2: Sound in Games

Adam Nash

Fig. 1 (above) House environment from SwanQuake

Fig. 2 Mars environment from SwanQuake
reinforce the visual message provided; i.e., a door may not always sound like a real world door. The ‘soundtrack’ sounds are usually much longer and freer in their experimenting with the kind of concepts I’ve been discussing so far.9

The opening environment, the House, Fig.1 provides a natural opportunity to introduce the user to the kind of sonic journey they are being asked to undertake. Since the environment is based on an existing space (igloo’s office warehouse in London, UK), it was simple to capture the establishing sounds by using a portable recorder and microphone to record the sounds naturally occurring in and around this actual place: the traffic outside, the water dripping in pipes, the rain falling, doors opening, etc. Recording all sounds with the same recording equipment in the same location on the same day guarantees the kind of internal relationship fidelity and sonic integrity discussed above.

Once these had been captured, the sounds were placed in their appropriately corresponding position within the environment. To this I added a long piece of music built from the sound of the rain falling (with accompanying drips and drops), along with some distant thunder. The environment starts with the background sound of the rain. Rain is one of those sounds that is likely to be almost immediately placed at the bottom of the priority list for conscious monitoring on the part of the listener, for it doesn’t significantly change within a known set of parameters. However, to create the music for this environment, the harmonics inherent in the sound of the rain are slowly brought to the foreground, resonating and merging with each other to create a far more ‘musical’ experience. The resonance of the distant thunder provides an underlying bass harmonic portending some possibility that this space is not as readable as it first seems – the resonances within the rain-as-music seem to harmonise with the small dark spaces found all around as the user explores the larger warehouse space.

It is my hope that the user is consciously engaged with the development of the rain sound into harmonic composed sound, and is aware of its relationship with the visual space the user is exploring. However, it must be assumed that many users will not engage so consciously in this process, rather surrendering to the spectacular experience (the cinematic), with its accompanying relegation of sonic perception to the background, and therefore the sounds and music must be able to function, at least somewhat, in this regard as well.

Assumedly, the sound of the rain will fulfil a narrative role, and the ‘music’ it evolves into will provide at least a signifier of the kind of narrative score that users may be used to, or even expecting, from cinema and cinema-style games. Ideally, even those users assuming a conventional spectacle/cinema style role of the soundtrack will notice some foregrounding or activity on the part of the soundtrack as it attempts to capture the user’s conscious, rather than unconscious, attention, since this is the only environment of SwanQuake that lays out the sonic menu so explicitly in its relationship to the precedents of cinematic style soundtrack, music and foley10, found in conventional real time 3D environments.

In sharp contrast to the House, the Mars environment Fig.2, presented a fascinating challenge for a sound artist. What does Mars sound like? Nobody knows! Theory tells us that everything in the universe has
a resonant pattern. (Greene, 2003) However, it is impossible for humans to actually hear the resonant pattern of things on Mars. Even if we could get to Mars, the atmosphere would need to be very similar to Earth’s in order that air could be moved around so that our ears could demodulate the sound waves into information for our brain. Of course we could get a machine to go there for us and record the resonant patterns on Mars so that we can listen. But this raises a problem that is not only limited to sound art, but extends to both philosophy and science. If the resonant patterns happen to be outside of our aurally perceptible range of 20Hz to 20KHz (and a lot of them will be), then we’ll need to modulate them up or down accordingly, but that act of modulation changes both the quality of the resonant pattern itself and the mode of perception. This is the difference between perception and conception; it is impossible for us to perceive the resonant pattern of Mars, but it is relatively simple for us to conceive of it.

The act of modulation, however, involves some conceptual decisions on the part of those performing the modulation. It could equally be said that this is true even here on Earth (ie, that most resonant patterns occur outside of the narrow range of human hearing), and therefore if we could actually stand open-eared on the surface of Mars, our sonic experience would be just as rich as we consider it to be on Earth. Nonetheless, nobody ever has stood open-eared on the surface of Mars, and therefore we haven’t a clue what it sounds like. Since we completely lack the perception side of the equation regarding the sound of Mars, we rely completely on the conception when constructing possible soundscapes for it.

**Summary**

The remaining environments of SwanQuake (e.g., Japanese Garden, Piranesi and the Power Station), whilst continuing the basic acousmatic approach outlined previously, all take different approaches in form, concept and structure. Decisions arose out of conversations with igloo regarding the intention and quality of each environment, with the broad sets of samples I had prepared with each environment acting as a base for growth and evolution into the final sonic implementation. As the user moves through and around all these environments, different aspects of the sonic environment emerge as a function of location within the 3D space. It is envisaged that the user will recognise this as an opportunity to consciously navigate the space using their ears as well as their eyes, thereby creating their own sonic experience. This ability hopefully also increases the ‘replayability’ of SwanQuake as much as the gorgeous visuals, the bewitching and dynamic motion of the sprites and the beguiling story/concept world. I would encourage all users to close their eyes while navigating, for extended periods, to gain an alternative perceptive model of the work.

**Notes**

1. See Mark Pendergrast (2001) for a good overview of the development of synthesisers in music in the twentieth century.

2. Subsequently, in the late 1990s and early 2000s emerged a retro movement celebrating the work of these synthesiser artists, but that is slightly beyond the scope of this argument, although it is interesting to ponder the coming retro movements of computer gaming in general. The current retro taste is for 1970s and 1980s arcade and Nintendo games; assumedly this will grow ‘forward in history’ to the point at which this essay is being written.
(2006) with pop culture-sanctioned movements and patterns emerging that may or may not be actually present in the current gaming landscape.

3. Games such as Darwinia use the processing power of cutting edge real time 3D graphics to produce a decidedly non-‘realistic’ 3D environment.

4. See, for example, Jonathan Miller (2006) paragraph 1: ‘While good sound in movies and games often goes unnoticed’

5. R. Murray Schafer calls sounds divorced from their original context ‘schizophrenic’ and considers it a negative phenomenon. But Pierre Schaeffer, the inventor of musique concrete, called them objets sonores, a kind of ‘aural photograph’, and considered them to be the essential building blocks of sound as art. See David Toop (2004). See also R. Murray Schafer (1994); and Michel Chion (1994).

6. The commercial production of real time 3D environments is often aimed at the economic end of consumption, so there has been little product that targets this ability within the games industry. The 2001 game Rez (Mizuguchi & SonicTeam, 2001), essentially an old-school first person shooter with tunes and beats instead of blood and guts, is notable for privileging music in its gameplay. It probably can’t be said of Rez that its primary aim is the user’s interactive perception of sound, although in practice it can be enjoyably used as such to a limited extent.

7. In whatever format, but .wav, .ogg and .mp3 are the currently favoured formats

8. For an interesting discussion of alternatives, see Philip Brophy (1989)

9. Frustrating as it is to use vocabulary sourced from film, the basic definition of soundtrack (i.e., music serving to reinforce the visual/emotional/conceptual message) holds up, as long as the above considerations of an active role on the parts of both composer/sounds and user/listener are kept in mind.

10. ‘Foley’ is the term used in film production to describe incidental sound effects, perhaps best illustrated by the coconut-shell-as-horses-hooves example.

References

Pendergrast, Mark (2001) / The Ambient Century / Bloomsbury
Leonard, Paul (2003) / Audio Prototyping with Pure Data / GamaSutra / May 28th
Farnell, Andy (n.d) / Practical Synthetic Sound Design / chapter 1, section 1, para. 3
Harrison, Jonty (1999) / 'Diffusion: theories and practices, with particular reference to the BEAST system' / eContact 2:4
Chion, Michel (1994) / Audio-Vision: Sound on Screen / Columbia University Press
Mizuguchi & SonicTeam (2001) / Rez, Playstation 2, SonicTeam/Sega
Brophy, Philip (1989) ‘Film Narrative/ Narrative Film: Music Narrative/ Narrative Music’ / Cinema Papers No. 71 / Melbourne

All links: http://www.swanquake.com/usermanual/AdamNash

Biography

Adam Nash is a new media artist, composer, programmer, performer and writer. He works primarily in networked real-time 3D spaces, exploring them as live audiovisual performance spaces. His work has been presented in galleries, festivals and online in Australia, Europe, Asia and The Americas, including peak festivals SIGGRAPH, ISEA, and the Venice Biennale. He is currently undertaking a Master of Arts by Research at the Centre for Animation and Interactive Media at RMIT University, Melbourne, researching multi-user 3D cyberspace as a live performance medium. He is also a Lecturer in Computer Games and Digital Art in the School of Creative Media at RMIT University. He has been a writer and reviewer for Digital Media World magazine, and editor of the Computers and Internet department at LookSmart. He was also a Project Officer at com.IT, a community charity he helped to establish that recycles computers and redistributes them for free to NFPs domestically and overseas.