Chapter 7

Ashby's Grandmother's Footsteps: an Interaction Installation

> " In the environment, the participant is confronted with a completely new kind of experience. He is stripped of his informed expectations and forced to deal with the moment in its own terms. He is actively involved, discovering that his limbs have been given new meaning and that he can express himself in new ways. He does not simply admire the work of the artist; he shares in its creation." - Krueger (1976) p.84

This chapter opens up the closed generative networks of the last chapter to the real world, experimenting with the use of simple adaptive systems as a means of mediating responsive environments in an interactive and generative sonic game.

Figure 7.1: Visitors treading carefully in Ashby's Grandmother's Footsteps at Artpool, Budapest, 2006.

Ashby's Grandmother's Footsteps is a play on the children's game *Grandmother's Footsteps* where one person stands at the end of the room facing a wall and all the others have to creep up on them. At any point (for instance if they hear movement), grandmother can turn around, whereupon any child caught moving has to return to the far wall and start again from there. This is a cybernetic version, where grandmother is a homeostat, receiving information about people's movement via video analysis and commanding control sonically.

The piece was installed at the Artpool[1] gallery, Budapest, as part of the 'Process Revealed' exhibition in conjunction with the EvoMusArt workshop at EuroGP 2006. Approaching the back of the underground room, visitors to the exhibition hear a faint drone emanating from a pair of wireless headphones hanging on a nail on the whitewashed wall. Fortunately people love being allowed to pick things up in galleries, so this is enough to get them interested. They put them on, and turn to face a previously concealed corridor - Alice-in -Wonderland what's-down-the-rabbit-hole curiosity is again enough to make them approach it. Stepping past an invisible boundary, they suddenly *hear* themselves walking as if on amplified delicate gravel. Walking the corridor, every movement is heard. As they progress down the corridor, the monophonic drone builds and differentiates, developing strange harmonies. A bold step forward triggers a SCREECH - loud enough to halt them for a moment and aversive enough to want to avoid it. Now they must creep slowly to the end of the corridor and turn off grandmother's eyes.

7.1 Power and Play in Responsive Environments

Since the early days of digital, and even electronic arts, people have explored ways of engineering situations in which the audience themselves 'bring forth' a work of art via their interaction with a space, engaging them in a process of co-creation. Often these pieces employ 'invisible' interfaces such as video cameras and sensors to create responsive environments: physical spaces with no obvious exposed technology that respond to visitors movements, gestures or vocalisations. These 'natural' interfaces, bolster a suspension of disbelief in the audience, arguably making them more open to the unlikely events that unfold within.

One of the many devices that gets played out in these environments, is a tipping of the power balance between user and system. Turning the push-button reactivity model on its head, many artists have found ways of engaging audiences in a game where they become manipulated by the system. This may be taken as a social comment of the cultural effects of technology by some, but it also presents a pertinent model of interaction for man-machine performance in demonstrating ways in which digital systems can be made to take the lead and actively induce a response from the human user. This section looks at some historical examples of power and play in responsive environments, and considers how auditory feedback can be given to maximise the audience's engagement.

7.1.1 Responsive Environments

The diffuse and intuitive interfaces of much contemporary European interactive art build on ideas laid down by players in the Art and Technology movement of the late 1960s. As early as 1969, Myron Krueger worked on multi-media 'responsive environments', such as *GlowFlow* (1969) which combined pressure sensitive floor pads with basic surround sound and reactive light elements. Phosphoresent particles were pumped through tubes attached to the walls of a darkened room in such a way as to distort the visitor's perception of the room's shape. The glowing particles and sounds were triggered by users standing on the sensors placed through the room. Others such as Seawright and Rauschenberg explored similar environments (Dinkla (1994)), but it was Krueger who developed the technology to create a more complex dialogue between user and environment.

[1] `http://www.artpool.hu/`

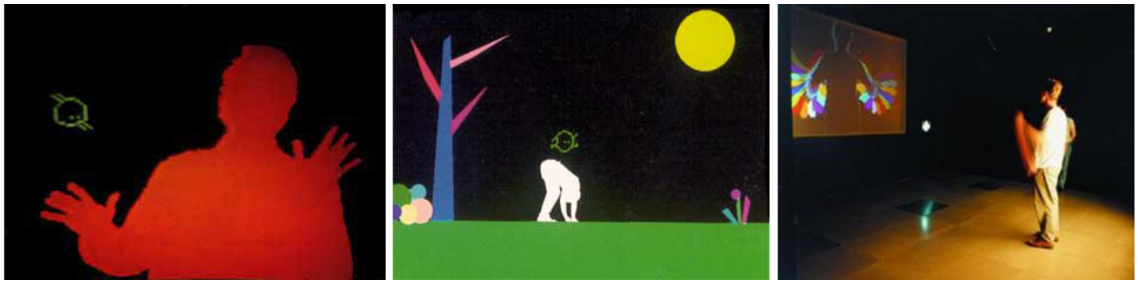

Figure 7.2: Artificial Reality space of Myron Krueger's *Critter* from *Videoplace*.

Krueger was concerned with creating what he described as 'artifical realities', aiming for a full-body participation in computer-mediated events that were so compelling that they would be accepted as real experience. In contrast to the head-mounted displays of nascent VR, or touch-screen/ mouse interfaces popular in the USA at the time, Krueger adopted the closed circuit video technologies that were fashionable amongst his contemporary video artists. The real time projections were combined with responsive, generative software to create these graphical 'artificial realities'. The viewer would enter a space dominated by a projected image which combined their own self-reflection within a world inhabited by computer generated images which were designed to respond to, and often evoke responses in, the participant. Figure 7.2 (right) shows one such projection where a simple delay on the video creates after images of their movements.

Beside his technical innovations, Krueger pioneered the notion of the artist as a 'composer' of intelligent, real-time computer-mediated spaces. Works such as *Videoplace* are presented as sets of composed interactive experiences, in which different forms of gestural analysis interpret, or even anticipate user's actions. In perhaps the most famous piece in the Videoplace collection, *Critter*, the user is taunted by a cartoon like creature. Shown in Figure 7.2 (left and middle). As someone enters the space, they see themselves as a shadowed outline on the screen into which jumps the small green critter. The critter tries to 'make contact' with the visitor, steering toward their outline, landing there and attempting to climb up their arm, shoulder, neck until it reaches their head. The user must try to outwit the critter, to move around in the real space, contorting their projected image to try and shake the Critter off. If they fail, and the Critter conquers the summit, it performs a joyful dance to signify its victory.

This simple game represents an early example of the probing of power distribution between user and system, inverting the paradigm of control and navigation common in other interactive forms. Real-time projections of the user's actual shadow are interlaced with computer-mediated graphics. The user's movements directly affect their shadow and influence the computer mediated graphics. The graphics in turn are heavily contrived to influence the user's physical movements. This sets up a simple but powerful play on the real and the artificial, on leadership and submission.

A sonic analogue of Krueger's closed-circuit video interfaces can be seen in works such as David Rokeby's *The Very Nervous System* (VNS). The various incarnations of VNS also work with video analysis as a diffuse, invisible interface, but rather than operating in a visual space, invite interaction with a synthesiser. Rather than the mixed realities of Krueger's projections, Rokeby invites the user to cooperate with this system to 'bring forth' a sonic environment. A schematic of the basic operational loop is shown in Figure 7.3 (right).

Whereas Videoplace manipulates the visitor by evoking a very precise attribution of cause and effect, Rokeby plays with a more complex feedback between movement

Figure 7.3: David Rokeby in Very Nervous System in the street in Potsdam (left) and a schematic of the basic operational loop (right).

and system response. He aims not to invert control, but to develop a relationship of "encounter and involvement". Physical movements over several cubic metres are tracked and intricately analysed with computer vision techniques. The synthesis component is heavily composed, using banks of 'instruments' with preset tendencies. For example, one instrument might be voiced on a snare drum, tend to play on off beats and double its rhythm if you move faster; another might synthesise a brass section and rise in register if you lift up your right arm. Various instantations of the piece have been exhibited, primarily in gallery environments and outdoors in public spaces (see Figure 7.3 (left)), but the systems have also been used in live performance.

Whilst Rokeby had to work hard in early VNS predecessors to achieve the real time motion capture necessary to support the transformation of thin air into a persuasive interface, this can now be easily achieved on a laptop. The use of 'invisible' interfaces in interactive arts is now very common: their immediacy and intuitiveness underlying their attraction as both installation and performance based interfaces.

Across the interactive arts, the use of 'invisible interfaces', whether focused for musical expression, or extended for installation environments creates a very intuitive and flexible mode of interaction which encourages acceptance of uncanny or complex man-machine dialogues. In both Videoplace and VNS, the artists play with the user's concept of control using carefully designed instructions which mediate between sensor input, and the graphical or sonic effectors. These rules govern the appearance or movement of the Critter, or the controls of a particular instrument specification in a VNS performance.

7.1.2 Designing Feedback

As in all digital art, the implementation and materialisation of formal processes and mechanisms is as important than the processes themselves. Just as the mapping from mathematical procedure to sound parameters defines the musical success of the formal structures, so the visual or sonic environments in which the interactive features of the composed environments are made available to perception, define the interactive experience.

Two considerations in particular drove the design of Ashby's Grandmother's Footsteps. Firstly the use of explicit rules mediating the sensory inputs (video camera/sensors etc) was replaced with the homeostat which acts as a self-modulated control system. Secondly thought was given to the importance of how the sonic feedback was delivered. As mentioned above, with responsive environments of this type, where the user 'brings forth' the content, it is vital that they are immediately and sustainably engaged. At the lowest technical level, this means ensuring that the interfaces are accurate and fast, and that they evoke meaningful perceptual experiences (even if this is confusion and frustration as in many cases). In Rokeby's system this is assured by his well developed VNS software. Arguably in his system, user's interest is sustained in part by the complexity of mappings from movement to sound. As he describes it:

> "The feedback is not simply 'negative' or 'positive', inhibitory or reinforcing; the loop is subject to constant transformation as the elements, human and computer, change in response to each other. The two interpenetrate, until the notion of control is lost and the relationship becomes encounter and involvement." - Rokeby (1990)

In VNS, addition, the sonic feedback which users get is heavily composed, so we get fragments of pentatonic panpipes, flurrying shakuhachi and bubbling brooks giving the audiences the sensation that they are composing musical works with their bodies.

In Ashby's Grandmother's Footsteps, some of the feedback material is composed to an extent (generatively), but more central is the exploration of the effect of including feedbacks on a number of timescales. Inspired by Ashby's notion on ultra-stability, in which two levels of feedback subserve an organism's ability to adapt to ongoing environmental interactions, the hope was that providing feedbacks at multiple timescales would engage the visitor more deeply in the system, creating a richer interactive experience.

Such considerations are central to installations in responsive environments, but are also of course of relevance to all forms of interaction. Several researchers have suggested that exploration of mappings in interactive sonic installations is crucial to developing understandings of interactivity which will push performance software forward.

> "I propose that the public exhibition of interactive, responsive sound installations and environments is a good platform for the investigation of mappings that may be inherent to the process of interaction. Of course, the interface design dictates the nature and the scope of all interaction to some extent, but public exhibition exposes the work to an untrained and inquisitive audience, who are prepared to invest time in the development of a relationship with the interactive system. They have no prior knowledge of the rules of engagement, and therefore set out to develop a cognitive map of possible relationships with the system, a map that deepens over time." - Paine (2002), p.298

7.2 Design and realisation

One of the aims of the installation was to play with the user's feeling of control in the space. Ultimately, they have to play the grandmother's footsteps game on the machine's terms and try and cheat it with stealth-like movement. In addition to this simple game playing, the piece contains elements of interactive and generative sound. 'Grandmother' is concerned only with whether or not they move too quickly. Others aspects of their movement through the space evoke and control several other layers of sound. This was designed to encourage exploration of small movements in the space and give the user a complementary sense of control.

- Track 21 gives an example of the output of the system

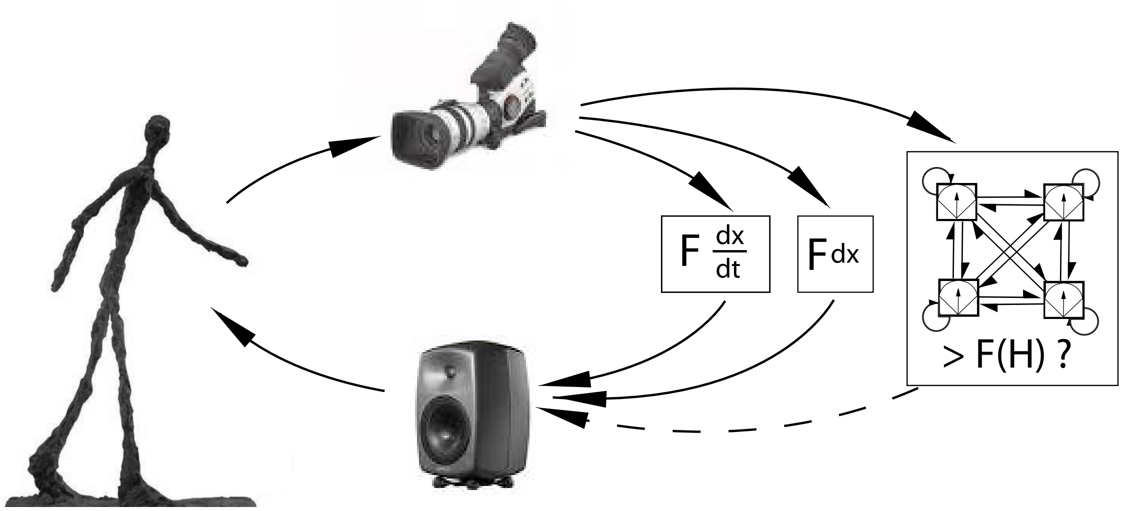

Figure 7.4: Schematic showing basic feedbacks between user, sensors, mediating devices and sound.

7.2.1 Overview

Within the space, the visitor's movements are implicated within three principle feedback loops. Each is associated with a different measure of movement and contributes to different aspects of the sound scape. The principle 'grandmother' control triggers a loud aversive metal-on-metal screeeach if the user moves too suddenly. The actual threshold is determined by the current state of the homeostat (shown as [$> F(H)?$] in Figure 7.4). A low level continuous feedback is given by directly sonifying the rate of change of movement (shown as $F(\frac{dx}{dt})$). This acts to augment the standard sensory-motor loop which engages us in the world, and aims to create a very personal and immediate sense of artificial reality by heightening awareness of movement. Finally the movement of visitors down the corridor (labelled $F(dx)$) triggers progressive changes in the harmonic drones, as well as increasing levels of some of the incidental sounds; this was designed to give a basic sense of progression in the sound world, reflecting the visitor's progression in physical space. In contrast to these sets of mappings between aspects of movement and sound, if the user stops completely still for a certain period, a contrasting set of sounds is introduced.

All three movement measures are derived from a motion detection algorithm operating on a live video feed. The DVCam was situated at the end of the corridor and covered its length. Details are given below in Section 7.2.3. All sounds were delivered via wireless headphones. Originally the installation was conceived to be set in the dark and to deliver sounds via speakers. The use of headphones rather than speakers changed the impact of the sound: the piece is very rich in bass frequencies, using sounds that have quite a strong impact physically, and adding to the sense of 'artificial reality'. On the other hand, the use of headphones makes the more delicate sound of each movement much closer and more immediate. Having only one pair of headphones also provided a much appreciated method of preventing more than one person entering the space at a time.

7.2.2 Composition and Implementation

As with all interactive generative pieces, composition must be approached as the design of sets of possibilities which a user can wander through: fields of sound, if you like which may or may not be heard in any way that you as the composer have ever experienced. In this setting, the qualities of the sounds themselves were also vital in evoking the desired combination of submission and exploration. The mappings from movement to sound are of course also key in creating a sense of contingency which is convincing enough to lure them into the game.

Sound design was approached with the aim of supporting a sense of 'artificial reality'. All material used in the piece were sampled from natural materials, and manipulated through granular synthesis, time stretching and reversing. All sounds are manipulated or processed in a Max/MSP patch which was controlled via OSC from another laptop running Jitter. There are three layers of sound which are directly influenced by the user: a sustained drone which is present as the user enters and differentiates harmonically as they travel through the space; a delicate crunching which is triggered with their every movement and the loud screech from Grandmother when they move too quickly. Other lines are more indirectly influenced as described below.

Grandmother Control in the Homeostat

In seeking a device to play the role of Grandmother in this installation, Ashby's homeostat offers an interesting mechanism for detecting change: it is richer than simply implementing a fixed cut-off point as it responds to minimal changes, but triggers a sudden change above a certain threshold, a threshold which is self-determined. This internal control over sensory-motor mappings gives the system some degree of dynamical independence.

The model homeostat used in Ashby's Grandmother's Footsteps is almost identical to the basic simulation described in Chapter 5, except that rather than issuing perturbations via a button press or feed from audio analysis, the simulated units are joined to the real world, being influenced by the user's movements via video analysis. The network in Ashby's original description represents the couplings between an organism and its environment. Each unit, or collection of units, in the mechanical device then, can be arbitrarily conceived as an organism, or its environment (which could be another organism). The full homeostat can be seen as a formal implementation of the concept of a 'responsive environment'.

The homeostat receives input into one unit via a motion-detection algorithm working on the camera feed. As this is invariably connected to the rest of the network the user's movements impact on the state of the homeostat as a whole. When stable, the homeostat can adjust to small fluctuating values, but if the input is large enough to cause any of the units to exceed their critical values, the network reconfigures as described previously. This triggers a loud screeching sound, but also of course means that as the weights have almost certainly changed the effective sensitivity is altered. The critical threshold for movement is therefore determined by the current internal state of the homeostat.

The sound delivered is a loud metal-on-metal screeeach: a recording of two pieces of resonant metal being scraped against each other[2]. This is simply triggered as a sample in Max/MSP, but modulated incidently each time it is played in an attempt to preserve its aversive properties.

Reconfiguration of the network also triggers updates in the processes controlling the drone harmony (described below). The output of one of the units of the principle homeostat is also used to control the playback speed of an indian bell sample. This is modulated over a very large range producing very different sounds according to the current state of the network. For some states it is inaudible - for example when the outputs oscillate at high speed. At times it appears as a deep resonant bell, at others a delicate tinkle. Other states trigger a slow modulation creating a flanging effect on its natural harmonics.

[2]This is a sample taken with kind permission from Arve Henriksen's Planting Trees, Creating Beauty. From the album Sakuteiki (2001) (Rune Grammofon)

Sonified Movements

The sensitivity of the video analysis was set such that the tiniest movement - even of a finger being bent - can be detected. This was crucial in this setting. In many applications motion detection is applied to fairly large movements, aiming to capture gross physical gestures of expression. In this situation, large or sudden movements are banned. In order to provide adequate room for exploration, the space of tiny movements needed to be augmented.

This was achieved by using the motion detection output to trigger and control the playback speed of a sample of some empty snail shells being rolled around in a tea-towel. This source of the sound isn't making any artistic statement, its just a good noise that is both very delicate and also incredibly rich harmonically: at low playback speeds it gives reverberating clunks, at close to original pitch it sounds like someone walking on eggshells, and at high speeds gives an electro schraaaunch. Two different samples are looped at playback speeds that differ by a factor of ten to give a richer sound. Although this is looping, it is rarely recognised as such due to the continuous changes in playback speed.

Harmonic Progressions

Underneath the intricate snail-shell movements and metalic screeches is a bed of drones which differentiate, aiming to draw the visitor along the corridor. These shift harmonically and increase in volume as the visitor moves through the space. A slow constant pounding is drawn through the whole space, increasing from an almost imperceptible breath to a fairly unnerving thud toward the end.

Originally the intention was to use infrared or ultrasonic sensors to give information regarding the person's position along the corridor. This would have given the possibility of creating more sophisticated compositions in the space. However determination with wishing to work with low level components and lack of time to tune what turned out to be a rather erratic device meant that an alternative method of estimating their location had to be devised. After testing several possibilities, it was found that a sufficiently accurate estimate could be made by taking a cumulative reading of movements. [3]

The sound itself is a sample of a bowed cymbal, processed with a purpose-built granulator to produce a fluid, pitch-able and continuous sound which retains the characteristic metallic harmonics of the source. Before someone enters the space, this is a monophonic drone. As the person moves through the space (i.e. every X times movement is detected) this differentiates to produce microtonal harmonies which change and build as they move down the corridor. These were implemented in a similar way to that described in sound installation AdSyMII, i.e. outputs are mapped pitch deviations to create microtonal harmonies. The principal Grandmother network needed to be run at a highspeed in order to retain the required temporal sensitivity, so a second network was implemented which was updated intermittently according to how fast the person moved.

Finally, a little 'hidden' generative sequence was designed which only appears if the visitor stands completely still for more than a fixed number of seconds (typically around 30). This is a slightly lighter series of melodic chimes, similarly produced by granulating a bowed cymbal sample. The period for this happening is set by hand, but the pitches of the chimes are controlled by the same homeostat which sets the pitches of the drones.

---

[3] Even if someone has been in the space a long time, but has not moved very much, it is generally fair to assume that they are somewhere near the start of the corridor. The only time that this is upset is when someone stands stationary part way down the corridor and moves theirs arms or head about a lot. In this case a false measurement is given resulting in an increase in volume of many parts. But at least they were enjoying themselves!

Figure 7.5: A screen shot from Jitter showing the visual effect of the frame differencing algorithm which is used for motion detection.

7.2.3 Technical Details

The installation was run inside Max/MSP/Jitter on two G4 PowerBooks linked via Open Sound Control (OSC) on a local area network. A diagram of the set up is given in Appendix B, Figure B.1. OSC is a communication protocol which enables high speed data sharing between networked comptuers. It is comparable to MIDI, but has much lower lag, allows the specification of data types and formats and can be sent via UDP or TCP/IP.

Video Analysis

Video analysis was performed inside Jitter. Jitter is an extension of Max/MSP which supports realtime manipulation of video, 3D graphics and other data sets within a unified processing architecture. This makes it relatively straight forward to grab and analyse a live feed from a digital camera connected via firewire. Motion detection is carried out with a simple frame-differencing algorithm. A video feed from a digital video camera in manual mode is grabbed at a resolution of 320 * 240 at 30 frames per second. To ensure no movement outside the installation area was detected, an adjustable mask was made, screening areas outside the corridor.

The frame differencing algorithm first calculates the difference in every pixel value between successive frames. The visual result of this is shown in Figure 7.5: nothing moves in the background areas, so the difference between frames for these pixels is [0 0 0 0] i.e. black; movement is greatest at the edges of the figure, or a limb which is moved suddenly (as visible on the left arm in the middle image) producing values approaching [1.0 1.0 1.0] i.e. white. A global measure for the whole frame is then calculated by taking the average difference and normalising. Although very simple, this was effective and could be tuned to suit environments with different lighting levels

Sensors and Switches

One of the installation issues was to ensure that the system was reset after each use, and to be able to ascertain when someone had entered the installation area. There are many ways this could be achieved. Here a combination of simple sensors and an engineering of people's movements with the physical space proved successful. The layout of the installation space made it possible to place the headphones just outside of the area under surveillance, this meaning that you could be fairly certain that they would engage with the piece before the sensor had detected their presence.

A modified passive infrared (PIR) sensor was used to detect when someone has entered the space. PIRs work by detecting changes in infrared radiation which is given off by all objects above absolute zero. These are used commonly in domestic burglar alarms. This was focused using the very low-tech but proven technique of sticking parallel strips of black electrical tape on the front of it, reducing its area of visibility to a narrow beam across the entrance of the space.

At the other end, a push-button switch was placed in a position prominent enough for most people to realise its function. If people did not use it, a fall-back procedure was implemented which reset all software for the next person.

7.3 Reactions and Discussion

Audience activity suggested that the piece achieved its basic function as a game of control: despite no instruction or explanation all users (bar a couple of ardent noise-core fans) quickly understood the game and submitted to the stealth-like movements necessary to play. In many cases visitors returned a second time. In the first run they seemed to quickly grasp the nature of the game and experiment with the range of allowed movement. On the second run they would enter very tentatively and try and traverse the whole corridor without triggering the screaach. The basic game is very simple, but there seemed to be a basic sense of 'achievement' with successful completion - and corresponding irritation by those unable to move sufficiently slowly !

It could be argued that simply playing a highly aversive sound is a fairly cheap way of controlling people's movement, but observing visitor's movements and speaking to them afterwards suggests that the experience was a little richer than just a game of control.

The ferocity of the first screeach inevitably made people stop in their tracks, so opening the space of minimal movements in which the snail-shells operated. Without these continuous feedbacks, people would perhaps would have given up and just stepped back out of the space. The snail-shell feedback was surprisingly effective in not only providing some form of entertainment, but apparently actively slowing down people's movements. In tests carried out in the lab, it was found that you could almost directly control the average speed of movement by adjusting the base-rate of the playback speed. The auditory feedback seemed to act as a positive feedback loop, with slow sounds making slow movements slower, and fast sounds making fast movements faster. The sound seemed to not only heighten awareness of movement, but create an illusion. Many visitors said that it felt like they were walking through gravel, or slurry (depending on the base-rate of playback speed). These two levels of feedback then were mutually supportive in that the screeach set an initial precedent, forcing people onto a slower pattern of movement, and the snail-shell reinforced their slow movements, keeping them away from the zone of detection.

One of the major issues was in setting the gain on the motion detection reading which was fed into the homeostat. Using the frame differencing method for motion detection was perhaps too simplistic as it does not take into account the distance of the person from the camera. As people approach the camera, becoming larger in the frame, it becomes effectively more sensitive. In practice, this wasn't too much of a problem as people became quite stealthy by the end of the corridor. If installed again it would be fairly straightforward to scale the output according to the percentage of frame filled by a moving object.

Whilst the homeostatic network was conceptually attractive, in practice a similar, if not more succesful effect could have perhaps been achieved with a simple switch. Self-modification is an conceptually attractive characteristic, but in practice, in this implementation, it just meant that users were given inconsistent feedback due to the changing threshold.

The balance of positive and negative feedback loops on a variety of timescales seemed to bolster engagement with the system. The organic, and slightly threatening nature of the sounds also supported the sense of an artificial reality which by all accounts was created by the experience. Whilst the use of such distributed interfaces and feedbacks is not perhaps suitable for on-stage performances, such environments provide fertile ground for exploring modes of interaction and engagement.