HETEROGENESIS: COLLECTIVELY EMERGENT AUTONOMY

Carlos Castellanos, Diane Gromala, Philippe Pasquier

Simon Fraser University
Surrey, BC CANADA
{carlos_castellanos, dgromala, pasquier}@sfu.ca

ABSTRACT

Heterogenesis is an interactive sound and tactile installation consisting of a group of autonomous artificial agents that collectively generate and evolve a soundscape in response to one another and to the presence of human participants. Taking the form of long pillar-like sculptures and forming a rudimentary artificial neural network, the agents respond to their environment in three ways: (1) by altering their generated sounds, (2) by communicating with one another so as to "warn" of human presence nearby, and (3) by producing an inaudible acoustic pressure field that "pushes" participants as they approach. *Heterogenesis* represents an attempt to manifest complex emergent behavior through a rich set of interactions with autonomous systems — both human and machine.

Keywords**—** Artificial Intelligence, Artificial Life, Art, Interactive Art, Sound, Cybernetics

1. INTRODUCTION: CONTEXT

When dealing with aesthetics of "intelligent" machines, it becomes necessary to shift from purely technical concerns to those of process, flow and dynamic interactions. To this end we describe *Heterogenesis*, an interactive sound and tactile installation consisting of a group of autonomous artificial agents that collectively generate and evolve a soundscape in response to one another and to the presence of human participants. This piece, currently in the early stages of development and testing, builds upon the work of Gordon Pask, Francisco Varela and others [9, 10, 12, 13, 14, 15] who investigate the idea of "collectively emergent autonomy." This refers to how humans and physically situated autonomous technological systems can co-construct and co-evolve with their environment through their interactions [1, 2]. The term "heterogenesis" refers to diversity or to "make different." Through this piece we wish to investigate the experience of how humans and machines can enact such difference. It is hoped that these spatial-tactile interactions will deliver an aesthetic experience which motivates a sense of being embedded in an increasingly technological environment.

A relatively recent development, the notion of intelligent systems co-evolving with their environments is founded upon links between artificial intelligence (AI) and the enactive paradigm in cognitive science [2, 15]. This paradigm of "embodied cognition" describes the processes whereby the nervous system links with the sensory and motor capabilities of an organism to connect that organism to its physical environment. Rooted in the concept of autopoiesis [4, 5, 12] — which Varela describes as "a characterization of the mechanisms which endow living systems with the property of being autonomous" [12:14] — enaction explains how cognition emerges when system and environment, through networked architectures, trigger and select in each other structural changes. This dimension of *structural coupling* helps elucidate how living systems (real or artificial) exhibit emergent properties such as autonomy via cooperative self-organizing actions with other living systems and with their surrounding environment.

Enaction's overall holistic approach — coupled with its philosophical roots in embodied phenomenology [6] — provides a contrast to the traditional reductionist approaches of science and engineering and gives it a natural resonance with artistic approaches. Thus when applied in an AI or artificial life (A-life) context, the environment — like any other agent — becomes a first-class actor and human sensorimotor interaction becomes an important co-evolutionary component. Autonomy is constructed through these co-evolutionary interactions. The insight explored here is that the integration of this human interaction — recurring longitudinally within the context of heightened aesthetic experiences that are characteristic of the arts — may lead to changes in perception and awareness.

2. PRECEDENTS

The use of AI and A-life techniques have yielded a rich and diverse set of artworks since their earliest iterations in the 1950s. The artworks most relevant to the discussion here however are those where the system in question is situated in the real world, occupying and operating in physical space, as opposed to those that feature elegant computer-based simulations that exist primarily in the abstract world of software. One of the earliest of these types of interactive artworks were those of cyberneticist Gordon Pask. Probably best known for his development of Conversation Theory [7, 8], Pask's career also featured the construction of interactive artworks, some of which portended his later work in cybernetics and constructivist learning models. According to Pask "[m]an is prone to seek novelty in his environment

978-1-4244-7493-6/10/$26.00 ©2010 Crown

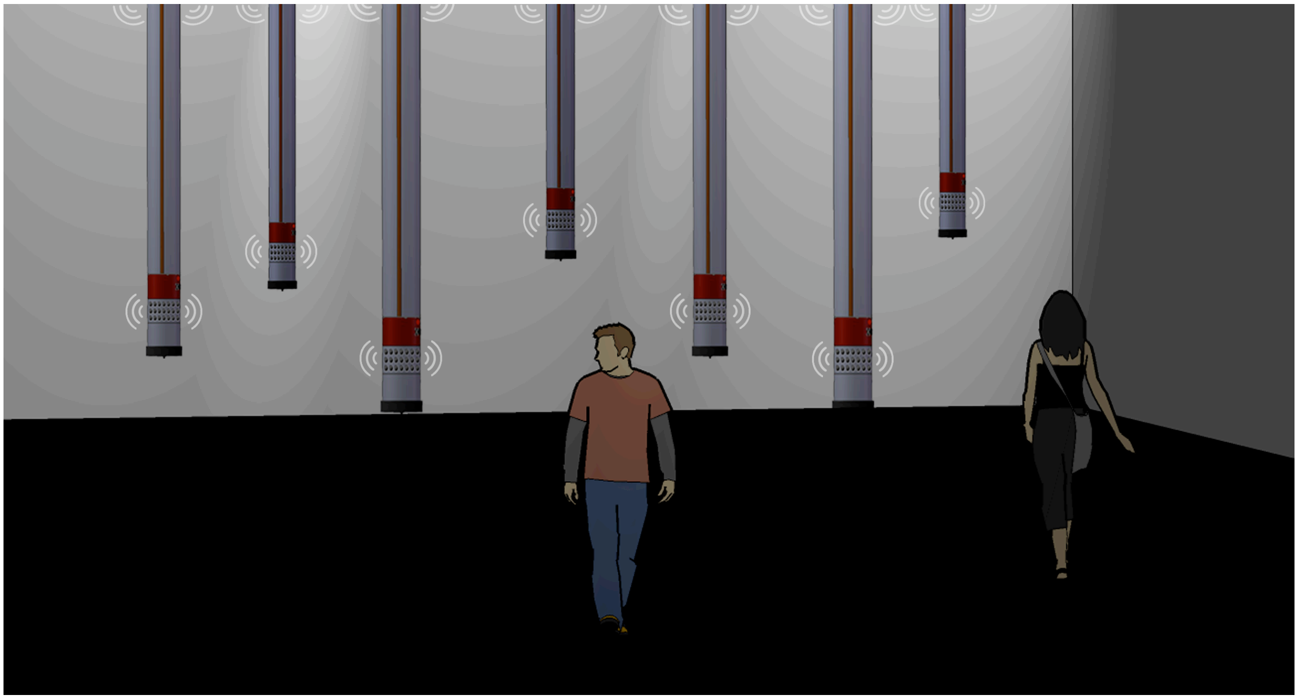

Figure 1. Concept rendering of *Heterogenesis* installation.

and, having found a novel situation, to learn how to control it" [9:76]. Perhaps nothing demonstrates this better than *Musicolour* (1953). One of the earliest artworks that had what could be considered a legitimate learning mechanism, *Musicolour* was comprised of a device that picked up the sound of a live musical performance with a microphone. The signal was then passed to a group of filters which were used to control banks of multicolored lights. The unique aspect of the *Musicolour* system was its capacity to become "bored" if the musical patterns became too repetitive [9:80]. The system would initiate change on its own, encouraging the performer to do the same. This was accomplished via circuity — analogous to biological neurons — that enabled certain lights to be activated only if a relevant filter surpassed a given threshold.

Similarly, American sculptor James Seawright has constructed systems that respond to environmental inputs and produce what he has called "a kind of patterned personality" [16:17]. In works such as *Watcher* (1965) and *Searcher* (1966), Seawright constructed systems that respond to light and participant presence in unpredictable ways, giving the appearance of intelligence or sentience. Seawright described the approach to constructing artworks of this type as "...setting up a generalized set of options and throwing them, so to speak, at an audience" [11:89].

More recently, Simon Penny has experimented with systems that generate conditions for emergence (defined here as unanticipated behavior that cannot be predicted from the operation of the system's individual parts), complexity and at least the appearance of intelligence or sentience. Of particular note here is Penny's *Sympathetic Sentience* (1995-96) [16:128-130]. Realized in collaboration with Jamieson Schulte, this piece consists of a group of small wall-mounted electronic devices that generate a series of rhythmic chirps. The units communicate by sending their rhythmic patterns to the next unit in the chain via an infrared signal, with the data stream looping through the entire group. Each unit combines its own rhythms with the ones it receives from other units. Participants can interrupt this chain of communication by moving through the space and blocking the infrared beams, thus altering the rhythmic patterns.

Penny's piece is a good example of a system that utilizes quite simple techniques for generating emergent complexity. More importantly we can discern in all of these pieces a view of interacting emergent systems (of which humans are just one variety) whose constitutive autonomy is seen as arising from situated, contingent and perhaps most importantly (and a bit counterintuitively) collective networked interactions with their surrounding environment. These pieces emphasize the ontological nature of autonomous systems. Their capacity to simply *be,* "to assert their existence" and — through their interactions with their environment — "shape a world into significance" [14:xi].

This is the approach behind *Heterogenesis.* It is a system that cannot be fully formed and cannot fully realize itself except through interaction with its environment. In a sense the piece *is* the interactions with its physical environment.

3. PROJECT DESCRIPTION

3.1. Overview

Each agent in the *Heterogenesis* system takes the form of long, hanging, pillar-like sculptures embedded with electronics that enable them to respond to the sonic environment, interact with human participants and communicate with other nearby agents. Sonic complexity emerges through a network of interactions and communication among humans and machines. A typical participant experience would be of a "multichannel" sound environment where if successful, would be characterized by simultaneously observed and induced complexity and emergence. A concept rendering of the piece is shown in Figure 1.

Each agent selects from a predetermined range of pitches (within a certain range, say 3 octaves) and velocities in response to the sonic and rhythmic character of the environment. After randomly selecting a tone to play at start-up, the agent begins listening to the environment. If it becomes too repetitive or predictable the agent may respond by outputting a progressively different set of sound patterns; or it may cease responding altogether until it hears something "interesting." Thus the environment as a whole exhibits a sort of metastability, appearing more or less stable for a time but with the possibility of cresting to a sonic display of "restlessness." The stability of the environment may also be disrupted when humans enter and begin making their own sounds (actively or passively) or approaching the pillars themselves. Both of these actions will cause further changes in the agents' sonic output, slightly perturbing their generation patterns. Furthermore, if a participant gets too close to a pillar, it may physically "push back" by generating a non-audible (but felt) acoustic pressure field which acts as a form of non-contact tactile feedback. As the participant walks away, the agent may emit a "warning" signal to the pillar nearest to the anticipated direction that the participant is heading toward. This signal may cause the receiving pillar to respond by further altering its sonic output. All of these scenarios will of course alter the overall sonic environment which all of the agents (and humans) are embedded in.

3.2. Pillar Construction

Each pillar will be made from a durable plastic (such as plexiglass acrylic) and hangs from the ceiling, approximately 1 m from the floor. It measures approximately 8 cm in diameter with a hollow center approximately 6 cm in diameter for running cables and other electronics. Figure 2 shows a pillar with indications of the location and placement of components. All electronics and hardware contained in the pillar itself, thus making it a self-contained system.

The system hardware is as follows:

- 3 equidistant speakers situated near the top of the pillar, angled downward. These speakers all emit the same mono sound.
- 3 equidistant metal strips running down the pillars. These strips are antennae of sorts, acting as capacitance sensors which allow the system to measure the distance and position of participants (relative to the pillar) as they approach.
- 3 equidistant infrared (IR) receivers situated approximately 35 cm from the bottom end of the pillar. These are used to receive the warning signals of nearby pillars.
- 3 equidistant IR emitters situated just above the IR receivers. These are used to send warnings to nearby pillars.
- An array of airborne tactile feedback actuators, situated just below the IR receivers, running around and down the pillar (ending just above the end). This technology is explained in greater detail below.
- 1 omni-directional microphone which sits at the bottom end of the pillar and picks up sounds from the environment.
- A microcontroller for measuring and processing sensor inputs.
- A rudimentary analog synthesizer constructed out of simple resistors/capacitor oscillator circuits; controlled from a microcontroller and digital potentiometers (which provide the varying voltage).

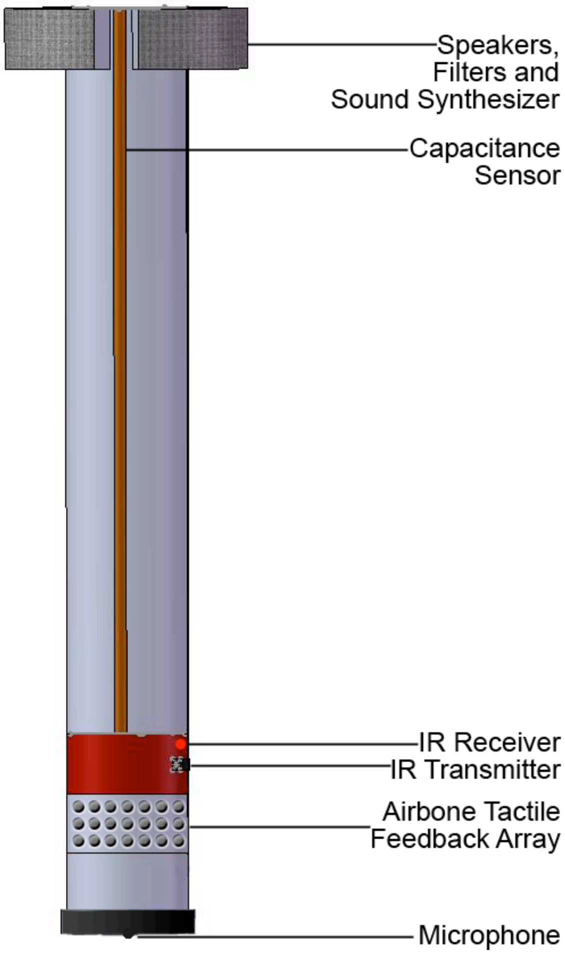

Figure 2. Pillar concept rendering showing placement of components.

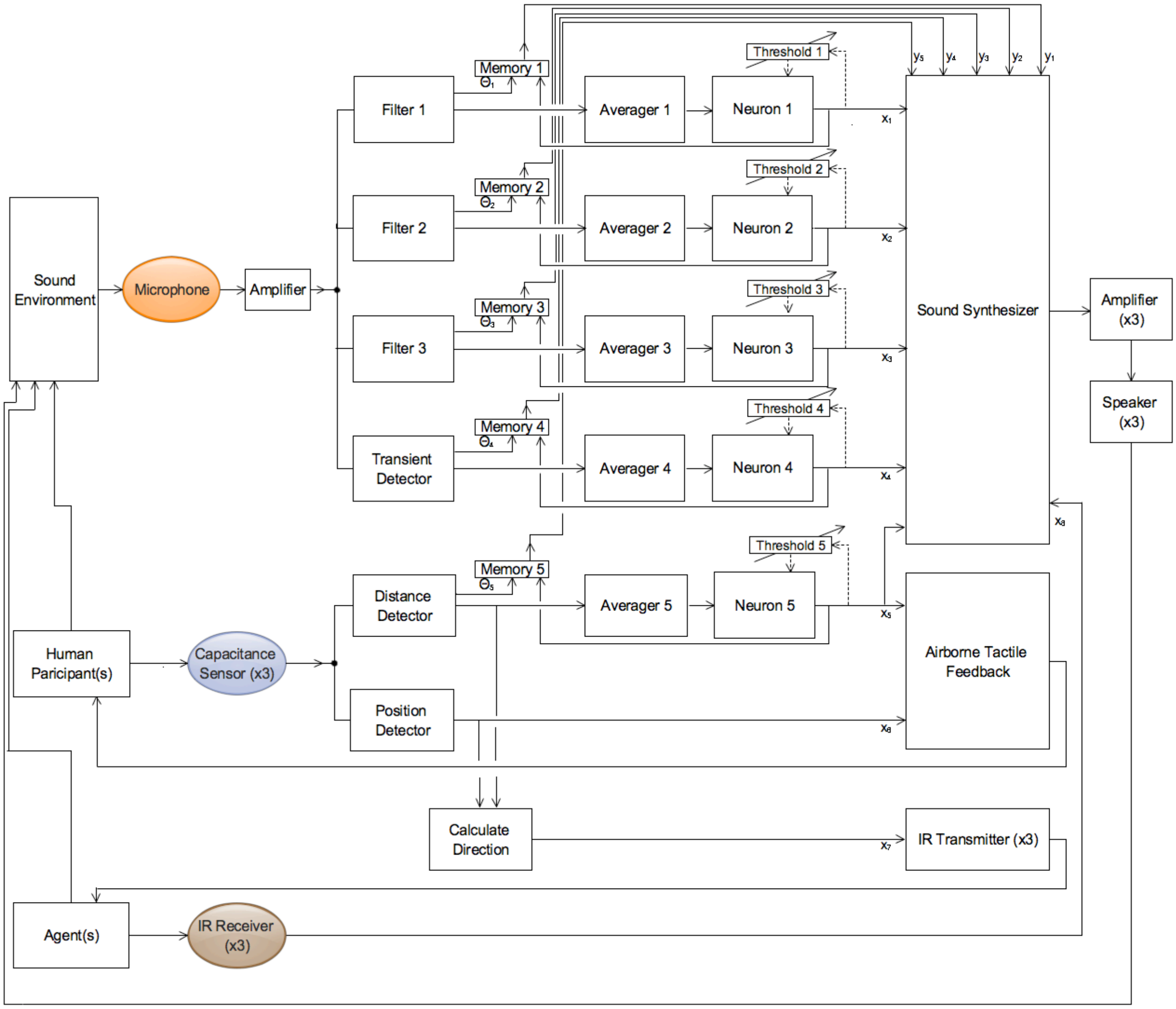

Figure 3. System Diagram.

3.3. System Design and Behavior

Each agent is designed to be a self-contained system that combines both analog and digital electronics. Agent behavior is rooted in the ability to learn from what is detected and from interactions with participants. Collectively, the agents form a rudimentary artificial neural network which allows them to intelligently adapt and respond to their environment. This configuration values diversity and novelty over consistency and stability. What is produced is an open-ended situated interaction among system, environment and human.

Figure 3 is the system diagram which describes the overall design and behavior of each pillar. Sound is first picked up by the microphone, amplified and then sent to four processors. The first three processors (Filters 1, 2 and 3) are bandpass filters while the fourth is a transient detector. The outputs of these processors are then continuously averaged over a given time span (say 5 seconds) using an Averager. In the case of the bandpass filters this will produce an average amplitude level within the given frequency ranges. In the case of the transient detector, an average velocity level is calculated[1]. The averagers then output a value that changes with the presence of signal at their inputs. For example if the agent "hears" sound that falls within the frequency range of Filter 1, then that signal going into the Neuron will be of higher value. The "neuron" is a device that emits an impulse if its input exceeds some threshold value. This impulse is designated x_i=1, where i=1, 2, 3, or 4 and indicates the signal going into the Sound Synthesizer. However the neuron's threshold changes based on how often it fires, becoming less sensitive

---

[1] In the case of the transient detector, unless the signal received is above a predetermined threshold, a transient is not recognized and thus the velocity would in effect be zero in that instance.

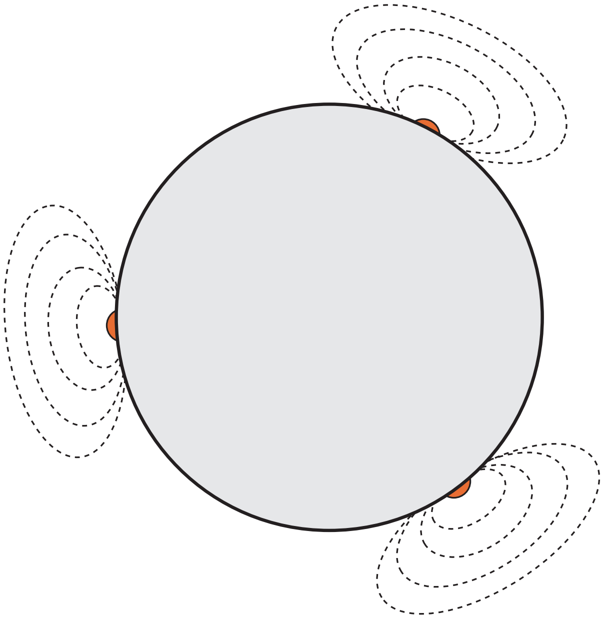

Figure 4. Top view of pillar showing the three capacitance sensors and their sensing characteristics.

(its threshold increases) the more it fires and more sensitive (its threshold decreases) otherwise.

Once the neuron fires, the signal is fed into the synthesizer which recognizes it simply as a command to play some note, or in the case of Neuron 4 (the transient detector) to select some velocity or attack. It will continue to play (e.g. hold down the note) for a predetermined amount of time (say 1 second) and then cease playing. The synthesizer has three oscillators for generating sound, each of which correspond to the three band-pass filters. The neuron's signal and the output of the filter (designated signal Θ_i) are also fed into a Memory module. This is essentially a timing buffer that compares the difference between the last time the neuron fired and the time since the filter last received input. This comparison is also weighted with respect to the amplitude of signal Θ_i. From this comparison the module sends a signal (designated y_i) which causes the synthesizer to either switch to another tone (or velocity in the case of the transient detector) or play the same one.

A participant's distance and relative position to a pillar is determined by 3 capacitance sensors (see Figure 4). When voltage is applied to the sensors, an electrostatic field is created. This field reacts to changes in capacitance caused by the presence of an object (such as a human body)[2]. Distance can be calculated by measuring changes in the electrostatic field. This is achieved by measuring how much time it takes for the pin on the microcontroller that the capacitance sensor (which is simply a capacitor) is connected to to go from low to high. Thus the distance from ground (a human body) to the sensor will be the primary variable contributing to the capacitance value. A lower value indicates the capacitor is charging quicker, which corresponds to a participant's nearer proximity to the pillar. Relative position is estimated by computing the relative distance of a participant from each of the three sensors. Then, just like the filters and transient detector, the signal from the distance detector is averaged and sent to the neuron which only fires if the incoming signal is above a given threshold. This signal corresponds to the participant's proximity to the pillar, where closer proximity makes the neuron more likely to fire.

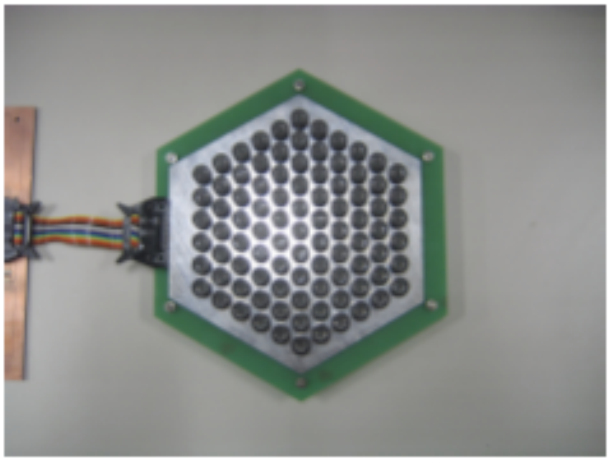

Figure 5. Example of an airborne (non-contact) tactile feedback array showing the ultrasonic actuators arranged in a hexagonal pattern. Image taken from Iwamoto, et al [3]. A demonstration video of this array is available at http://www.youtube.com/watch?v=hSf2-jm0SsQ

Once this neuron fires, the signal is fed into the synthesizer which recognizes it as a command to increase or decrease its pitch output by an amount proportional to the participant's distance from the pillar. Just like the others however, this neuron's threshold changes based upon how often it fires. Thus if a participant or group of participants hover near a pillar for too long it will become "accustomed" to the human presence and cease altering the pitch. Furthermore, if a participant begins to walk away, the agent may direct a "warning" signal, via its IR emitter, to a nearby pillar in the direction that it estimates the participant is heading. To determine which of its IR emitters to send the signal from, the agent calculates the participant's direction based on changes in both her distance and position. The emitter will then fire after a predetermined distance threshold. The receiving pillar will then emit a loud high pitched tone. This pitch is selected randomly (within some predetermined range) from the synthesizer's upper register.

3.4. Airborne Tactile Feedback

In addition to the changes in pitch, each pillar also emits an acoustic pressure field which acts as a form of "airborne" (non-contact) tactile feedback in the direction of

---

[2] Although capacitive fields are normally omnidirectional, their shape can be altered by the use of a nearby grounded metal source. Our approach is to shape the field so that it projects away from the pillar (and toward participants).

participants as they approach. This is achieved by an array of small ultrasound transducers (typically about 1-3 cm in diameter). This technique of using acoustical radiation as a form of tactile feedback is based on work done by Takayuki Iwamoto, et al [3]. Figure 5 shows an example of such an array. This array radiates an ultrasonic pressure field that while inaudible to humans, is powerful enough to be felt as a series of vibrations against the body. The array generates the field when it receives signal x_5=1, just like the synthesizer. If participants hover near the pillar for too long, the tactile feedback will stop.

4. IMPLEMENTATION DETAILS

Though currently in the early stages of development, most of the necessary hardware and software components have been identified. As previously mentioned, the project combines both analog and digital electronics with all of the components necessary for operation being housed within the pillar itself. Thus no standard desktop or laptop computers are necessary except for the programming of the hardware.

Technical components are as follows:

- 1 Arduino microcontroller per pillar (a relatively powerful Arduino with plenty of I/O pins like the Mega should suffice).
- 1 small omni-directional microphone per pillar.
- Approximately 50-75 ultrasound transducers (such as the T4010A1 from Nippon) per pillar.
- 3 small powered speakers (such as computer speakers) per pillar (it may be necessary to build custom speaker enclosures).
- Durable plastic material (such as plexiglass acrylic) for the construction of the pillars.
- 1 digital potentiometer (such as the AD5206 from Analog Devices), 1 Schmitt Trigger hex inverter (such as the 74C14) and various resistors and capacitors for building 1 sound synthesizer per pillar.
- 3 IR emitters (such as standard IR LED) and receivers (such as the PNA4601M) per pillar.
- Aluminum or copper sheets for the construction of the capacitance sensors.
- Other miscellaneous items such as wires, wood, paint, glue, etc.

5. CONCLUSIONS AND FUTURE WORK

Early investigations regarding the arrangement and testing of components, materials and learning and interaction strategies is ongoing. Thus the system just described is best viewed as a “first draft.” The overall goal however is focused not on technical capacities or behavior but on interactions with the physical environment. The tension between observing this continuously iterating “community” of sonic devices while simultaneously being part of it — thus influencing the very dynamics that one is observing — has resonances with the fields of cybernetics and autopoietic theory discussed earlier and may offer fertile conceptual ground for further artistic investigations. More broadly however, our stance is that phenomena such as autonomy and intelligence — especially with regard to artificial agents — are not innate properties but rather are contingent and relative to the process and contexts from which they arise. Whether constructed by artist or scientists, agents are (incomplete) mirrors of their creators; representations that may help us develop not only a sense of being embedded in an increasingly technological environment, but of being in a world full of dynamic, fluctuating and continuously emerging systems.

6. REFERENCES

[1] Castellanos, C and Gromala D, “Symbiogenic Experiences in the Interactive Arts,” in *Proceedings of the Digital Arts and Culture Conference.* Irvine, CA: UC Irvine, 2009.

[2] De Loor, P., Manac'H, K., and Tisseau, J., “Enaction-Based Artificial Intelligence: Toward Co-evolution with Humans in the Loop,” Minds and Machines vol. 19, no. 3, pp. 319-343, 2009.

[3] Iwamoto, T., Tatezono, M., and Shinoda, H., “Non-contact Method for Producing Tactile Sensation Using Airborne Ultrasound,” in *Proceedings of the 6th international conference on Haptics: Perception, Devices and Scenarios*. Madrid, Spain: Springer-Verlag, pp. 504-513, 2008.

[4] Maturana, H.R. and Varela, F.J., *Autopoiesis and Cognition: The Realization of the Livin*g. D. Reidel, Dordrecht, Holland; Boston, 1980.

[5] Maturana, H.R., and Varela, F.J., *The Tree of Knowledge: The Biological Roots of Human Understanding*. Shambhala, Boston, 1992.

[6] Merleau-Ponty, M., *Phenomenology of Perception*. Routledge, London/New York, 2002.

[7] Pask, G., *Conversation, Cognition and Learning: A Cybernetic Theory and Methodology*. Elsevier, Oxford, 1975.

[8] Pask, G., *Conversation Theory: Applications in Education and Epistemology.* Elsevier, Oxford, 1976.

[9] Pask, G., “A comment, a case history and a plan,” in Reichardt, J. ed., *Cybernetics, Art, and Ideas*. Studio Vista, London, pp. 76-99, 1971.

[10] Pickering, A., “Ontological Theatre: Gordon Pask, Cybernetics, and the Arts,” Cybernetics & Human Knowing vol. 14, pp. 43-57, 2007.

[11] Seawright, J., “Phenomenal Art: Form, Idea, and Technique,” in *On the Future of Art*. Viking Press, New York, pp. 77-93, 1970.

[12] Varela, F.J., “Autonomy and Autopoiesis,” in Roth, G. and Shwelger, H., eds., *Self-Organizing Systems: An Interdisciplinary Approach.* Campus Verlag, pp. 14-23, 1981.

[13] Varela, F.J., *Principles of Biological Autonomy*. North Holland, New York, 1979.

[14] Varela, F.J. and Bourgine P., “Towards a Practice of Autonomous Systems” In Varela F.J. and Bourgine P., eds., *Toward a Practice of Autonomous Systems*. MIT Press, Cambridge, MA, pp. xi-xvii, 1992.

[15] Varela, F.J., Thompson, E.T., and Rosch, E. *The Embodied Mind: Cognitive Science and Human Experience*. MIT Press, Cambridge, MA, 1992.

[16] Whitelaw, M., *Metacreation: Art and Artificial Life*. MIT Press, Cambridge, MA, 2004.