

ROOM#81 - Agent-Based Instrument for Experiencing Architectural and Vocal Cues

Nicolas d'Alessandro
MAGIC Lab, University of
British Columbia
Vancouver, BC, Canada
nda@magic.ubc.ca

Roberto Calderon
MAGIC Lab, University of
British Columbia
Vancouver, BC, Canada
rvca@interchange.ubc.ca

Stefanie Müller
Hasso Plattner Institute
Potsdam, Germany
stefanie.mueller@student.
hpi.uni-potsdam.de

ABSTRACT

ROOM#81 is a digital art installation which explores how visitors can interact with architectural and vocal cues to intimately collaborate. The main space is split into two distinct areas separated by a *soft wall*, i.e. a large piece of fabric tensed vertically. Movement within these spaces and interaction with the soft wall is captured by various kinds of sensors. People's activity is constantly used by an agent in order to predict their actions. Machine learning is then achieved by such agent to incrementally modify the nature of light in the room and some laryngeal aspects of synthesized vocal spasms. The combination of people closely collaborating together, light changes and vocal responses creates an intimate experience of touch, space and sound.

Keywords

Installation, instrument, architecture, interactive fabric, motion, light, voice synthesis, agent, collaboration.

1. CONCEPTS AND OUTLINE

Human social interaction is versatile and pervasive in human life. Yet, when creating machines for interaction we often forget the subtlety of unconscious cues and focus on conscious models. Looking at intimate inter-personal relations between humans, we can state that these unconscious parts – such as small gestures or digressing eyes – lay out the foundation for emotional commitment. We propose that human-computer interaction needs to be defined through both conscious and unconscious interactions that rely on meaningful feedback systems.

Although art has made use of human-computer interaction through *new media installations*, it generally focuses on conscious and direct interaction paradigms, like the feedback loop, to create a simplistic illusion of control. We believe that there is a need to explore the integration of such uncontrollable and, at times, ungraspable nature of unconscious interactions between humans and their machines. As a result, the dialogue between humans and their technologies remains personal and intimate.

In ROOM#81 we examine interactive places that explore interaction through subtle contexts. Visitors are welcomed in a room where architectural and vocal cues are the main components that structure the nature of such space. A lar-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

NIME'11, 30 May–1 June 2011, Oslo, Norway.
Copyright remains with the author(s).

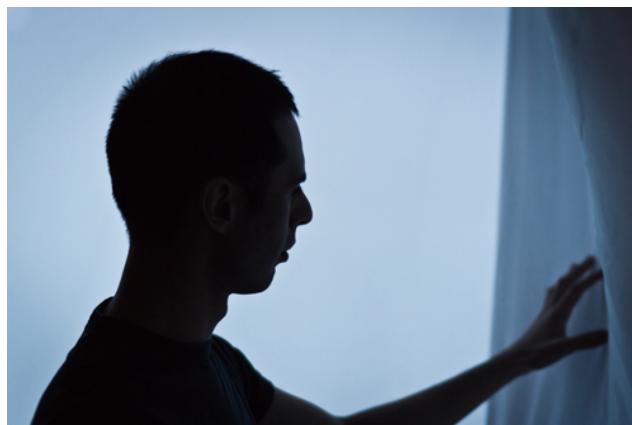


Figure 1: Movements in the installation space and collaborative interaction with the fabric influence the lighting properties and the voice quality of an artificial vocal character, screaming in the room.

ge piece of fabric is hung up in the middle of the room to create a *soft wall* separating the space in two areas. Visitors, who have never seen each other before, can access the installation from both sides of the fabric and have an interaction between themselves by pulling and pushing the fabric. Their movements in space, along with their haptic interaction with the fabric affect the sentient nature of the room, which responds with changes in light and voice modulations. Visitors experience an invisible, yet personal, vocal character that screams in agony, pleasure, or concern somewhere in the room they intimately share.

The soft wall: a mediation tool for intimacy

We believe that these three simple cues – the movement of a foreign person towards you through a piece of fabric, the changes in light quality and the changes in the tension of a voice – open up a large space for aesthetic interpretation. Based on intimate and sensual displacements of the fabric, one begins to wonder the nature and story of the person behind it. How does this person look like? It is a man or a woman? What is his/her personality? Where does he or she come from? These questions make the *soft wall* both an invitation to play and a collaborative effort to affect the nature of the room. Figure 2 shows a close-up on interlaced hand gestures of people that have never seen each other.

The vocal character: a subjective response

The large spectrum of vocal solicitations adds a second layer to our exploration. We can easily imagine some visitors being amazed by the sharp and quick spasms of the voice synthesizer. Yet, we also think of visitors who might regard



Figure 2: Confident interaction from both sides of the fabric with people who do not know each other, making the fabric a mediation tool for intimacy.

such sounds as painful screams, or clear sexual references. The same is true for slow and languorous sounds or whispering noises. Because people react differently to the same kinds of vocal stimuli, our array of stimuli becomes infinitely large. ROOM#81 positions itself at the subjective level of human interaction, opening a wide space for interpretation.

The agent: analog instrument and social control

With this collaborative instrument we also offer a greater social sense of control. Visitor's interaction is never directly mapped onto the vocal or light spaces, but has a non-obvious, adaptive impact on the vocal and light stimuli. Gestural inputs from sensors are used in an ongoing machine learning process that constantly changes the behaviour of an agent. As the agent forms a model of its world and acts upon it, its "thoughts" are mapped on the vocal and light spaces. With this in mind, a visitor can only control the other person's reaction to his own usage of the fabric. As such, we conceive ROOM#81 as an analog instrument of a tripartite nature, that is, played by two humans that intimately discover each other through a *soft wall* and an architectural/vocal agent.

2. RELATED WORK

ROOM#81 is related to interactive art for connecting people through sound haptics, voice-related digital musical instruments and interactive light/sound installations.

Connecting people through sound

Contacts [6] is an interactive sound installation for two or more people that consists of a small ball on a stand. When the visitor places his or her hand on the ball, his body becomes an interactive sound space that is sensitive to other people's touch. Shaking hands, caressing and kissing create different sonorous sounds. If the visitor remains alone, nothing happens. Following this, the visitor is encouraged to explore the intimate space of touch with a second person. To detect touch, the installation makes use of the small electrical tension every human carries on his skin surface.

Akousmaflore [5] is based on the same concept as Contacts. In Akousmaflore the visitor strokes musical plants that are arranged in a small garden. Each plant reacts in a different way to contact and warmth based on its individual leaf structure. Visitors can interact with each other by using the plants as different instruments in their musical arrangements.

Both installations aim at creating a complex and subtle interaction between gestures and sound. However, they often implement random factors on their process of sound

mapping through various modifications on the data. Furthermore, both have used digital audio samples in order to create harmonic feedback.

Voice-related digital musical instruments

The use of interactive voice synthesis for both performative and installation purposes has not been studied further than sample playback. We can highlight HANDSKETCH [2] and DiVA [3] as two more advanced projects.

The HANDSKETCH [2] is a digital instrument made for the bi-manual control of voice quality dimensions: pitch, intensity and glottal flow parameters. The instrument aims at exploring the expressivity of voice by simulating laryngeal behaviours with realtime gestural control.

DiVA [3] allows for direct manipulation of phonetical and prosodical spaces by using hand gestures. Hand gestures are intermediately converted to articulator (e.g., tongue, jaw, lip, vocal chords) parameters of a 3D vocal tract model.

Interactive light and sound installations

Shortcut [1] is an artwork that responds to the speed, rhythm and number of people in a passageway by building up a pattern of light that reflects the recent movements. A similar concept is deployed by Dune [4] which maps a sound and light space onto a visitor's movements through space.

3. INSTALLATION SETUP

In this section, we explain how the installation is laid out. We also give further details on the inputs (webcam, stretch sensor and light sensors) and outputs of the system (two-channel audio system, beamer lighting).

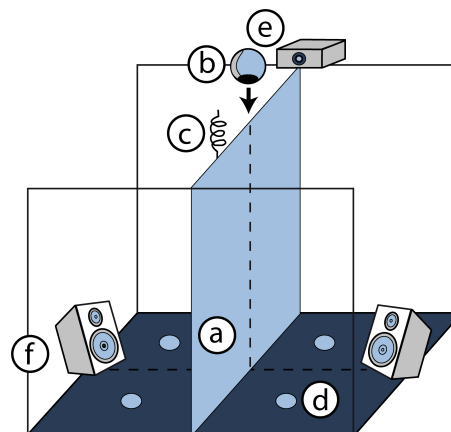


Figure 3: Setup of ROOM#81: (a) large piece of fabric, (b) webcam, (c) stretch sensor, (d) four light sensors, (e) beamer, (f) two loudspeakers.

Spacial arrangement

ROOM#81 is an installation that is setup on a rectangular floor space of 4x4 meter. At least one side of this space requires a wall from the building. Colour and material of this wall can be of any kind, but light colours are preferred. In Figure 3, this required pre-existing wall is the vertical face of the cube that is closest to the viewpoint.

Collaborative piece of fabric

Perpendicular to this wall, we place a piece of fabric (a). The fabric divides the room into two parts of equal size. The width of the fabric is at least 4 meter to give visitors a surface large enough to discover different kinds of interaction. The height of the fabric is at least 2 meter to avoid

that visitors look over it to the other side of the installation. In our current prototype of ROOM#81, we hang up the fabric between two solid tripod stands (people push and pull the fabric). Visitors can access the fabric from both sides ideally by two separate entrances. A more open configuration is possible, but the notion of two distinct area should be maintained.

Webcam, stretch sensor, light sensors

A webcam (b) is placed 1 meter above the fabric to capture the actions of visitors on both sides of the installation. In our setup, the webcam is attached to the ceiling of the room. However, if the ceiling is too high or not appropriate for positioning a webcam, it can also be placed on the side of the fabric to track lateral movements. Figure 4 shows results of the webcam-based motion tracking.

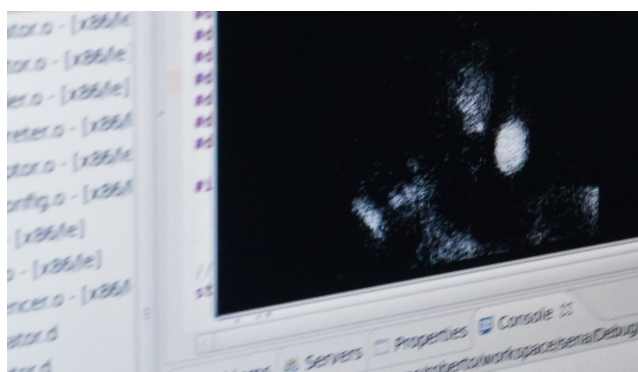


Figure 4: Results of the webcam-based motion tracking inside the installation space (left side).

A stretch sensor (c) is attached on top of the fabric. It measures the degree of tension in the fabric. The stretch sensor enables us to reason if visitors applied soft or hard manipulation when pulling or pushing the piece. In our current prototype, we horizontally hang up a metal bar between the two tripod stands to fix the stretch sensor.

Four light sensors (d) are placed on the ground, one in each quarter of the installation space. The light sensors capture additional information on visitors actions in the installation. People's movements relatively to the light source create various shades that are captured by these sensors. The captured data provides a rough sense of where people are located in the room.

Ambient light and vocal sound diffusion

We place a video projector (e) at the top corner of the fabric that is not connected to the wall. We use the video projector to vary the ambient light in the installation by going over a range of single colours that are projected full-screen on the wall. By projecting different colours on each half of the wall, we are able to create different moods on each of the two sides of the installation.

One loudspeaker (f) on each side of the installation diffuses the vocal sounds, avoiding the creation of an immersive sound field. This localizes the voice on various sources across space. Consequently, the virtual vocal character moves from one area to the other depending on visitors' behaviour.

Second wall: closing the space

Optionally, a second wall can be placed on the opposite of the pre-existing wall. In Figure 3, this second wall is the vertical face of the cube that is furthest to your viewpoint. Using a second wall helps to close the space which gives visi-

tors a sense of a semi-private surrounding. Visitors interact in a less constrained manner with such an arrangement.

4. THE AGENT BEHAVIOUR

In this section, we explain how we make use of a constantly learning agent impacting on architectural and vocal cues, in order to encourage the visitor to reflect upon his behaviour. We describe how this machine learning process works by using a Bayesian network. Afterwards we give more details on the voice synthesis algorithm for the production of spasms.

Agent-based interaction

There is an interesting space to explore between the instrument and the installation. For an instrument we expect a predictable behaviour that allows for practice. Opposed to this is the installation which introduces unpredictable aspects from the visitor's point of view. Indeed the visitor embodies a part of the system, but is also partially embodied by the system. In this context, the use of an agent as another contributor to the experience is particularly suitable. Participants can enjoy and refine their use of a predictable part of the installation – the two-sided fabric – but their movements are used to train an agent that takes part in the experience by influencing light and sound.

Self-learning Bayesian network

The inputs of the system are the image captured by the webcam, the stretch sensor and the light sensors. Such data is fed into a Bayesian network that aims at predicting human behaviour in the installation. As visitors interact with the installation, data is created and the installation becomes more accurate. An agent then uses the network in a statistical manner to predict visitors' behaviours or promote them. Our installation makes use of no predefined mapping, but uses adaptive machine learning to create the visual and audible cues. Therefore, our instrument is self-determined and self-learned based on the visitor's interaction.

Visitors contribute to the ongoing learning process of the agent which allows for complex scenarios. For example, if a visitor pulls hard over a long time, the room will not necessarily stop screaming after the visitor disrupted his interaction. Based on the learning algorithm, ROOM#81 will behave differently for each visit, sometimes extrapolating the solicitation, sometimes provoking the change.

Voice synthesis algorithm

The agent shares his beliefs (statistical probabilities of certain actions) with an interactive voice synthesizer. The voice synthesis algorithm is based on the RAMCESS synthesis engine, the same as used in the HANDSKETCH digital instrument. RAMCESS is a concatenative synthesis (using FTM for Max/MSP) with realtime frame selection and sound transformation. This algorithm produces primitive vocal spasms – like a big open /a/ – with realtime control on the pitch, intensity, vocal fold tenseness and breathiness.

The synthesis parameters are not controlled directly but are mapped to a vocal space. The agent changes the way the system cycles through this vocal space. There are three dimensions in how this cycle changes based on agent probabilities: the speed of the cycle which determines the abruptness of the spasm, the overall pitch zone of the spasm which can be low or high, and the overall pitch range of the spasm which can be flat or abrupt. The sound is also spatialized between two loudspeakers, in order to set the voice where there is the least chance that something new happens, as a way of triggering a too predictable visitor.

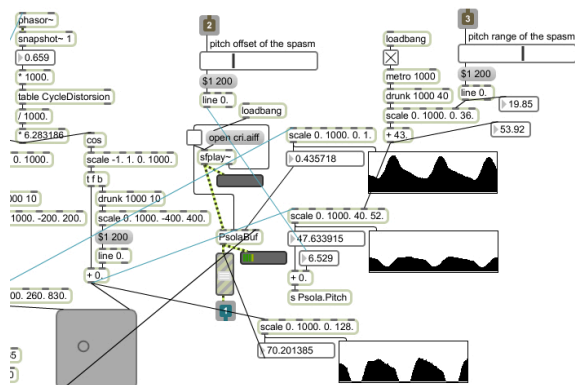


Figure 5: Voice synthesis mapping in Max/MSP. Trajectories in the vocal space can be changed: spasm, speech pitch offset and pitch range.

5. EQUIPMENT

ROOM#81 requires the following equipment: 1 piece of fabric: size = 4 x 2 meter; 2 tripods: height = 2 meter; 1 metal bar: length = 4 meter; 1 webcam; 1 beamer; 1 stretch sensor; 4 light sensors; 2 loudspeakers; elements to put vertically and create a second wall, e.g. poster grids work (optional). We can bring all of the equipment except for the tripods, the loudspeakers, the metal bar and the optional second wall. We would like to ask conference organizers to provide us with these items. There are no more technical requirements except for a free white wall. ROOM#81 can be set up as both a foyer location or a room-based installation.

6. VIDEO DEMONSTRATION

A video of the installation has been taped during prototyping at the University of British Columbia (Vancouver): <http://www.nicolasdalessandro.net/room81>

7. ACKNOWLEDGEMENTS

The authors would like to thank the Media and Graphics Interdisciplinary Centre (MAGIC) of the University of British Columbia, and its Director: Prof. Dr. Sidney Fels. First because this lab has created the opportunity for us to meet and discuss about connecting our respective matters of interest. Secondly because we could use MAGIC facilities to mount the prototype and develop the software.

8. BIOGRAPHIES

Nicolas d'Alessandro

Nicolas d'Alessandro is a researcher and musician who has been exploring the interactive side of artificial voice production for the last eight years. He built several digital instruments for performing synthetic speech and singing, such as the HANDSKETCH, and played them on stage. He holds a PhD in Applied Sciences from the University of Mons (Belgium) and is now Research Associate at the University of British Columbia (Canada), where he supervises the DiVA project and co-directs the UBC Laptop Orchestra.

Roberto Calderon

Roberto Calderon is an architect and artist interested in the human perception and interaction with ubiquitous technology and interactive environments. His work deals with public displays, interactive architecture, wearable and mobile devices. He is interested in the concept of agent based

architecture able to form intimate relationships with its inhabitants. He is currently pursuing his PhD at the Media and Graphics Interdisciplinary Centre at the University of British Columbia.

Stefanie Müller

Stefanie Müller is a computer scientist and author interested in transferring the story behind everyday experiences into interactive artwork. Thereby, she is drawing on her experiences as a writer of modern poetry for which she received several scholarships. Stefanie is working as an anthologist and recently published two books in collaboration with the canadian photographer Darren Holmes.

9. REFERENCES

- [1] J. Bruges. Shortcut. <http://www.jasonbruges.com/projects/uk-projects/shortcut>, 2010.
- [2] N. d'Alessandro and T. Dutoit. HandSketch Bi-Manual Controller: Investigation on Expressive Control Issues of an Augmented Tablet. In *New Interfaces for Musical Expression*, pages 78–81, 2007.
- [3] S. Fels, R. Pritchard, and A. Lenters. Fortouch: a wearable digital ventriloquized actor. In *New Interfaces for Musical Expression*, pages 274–275, 2009.
- [4] D. Roosegaarde. Dune. <http://www.studio Roosegaarde.net/project/Dune>, 2010.
- [5] Scenocosme. Akousmaflöre. http://www.scenocosme.com/akousmaflöre_en.htm, 2009.
- [6] Scenocosme. Contacts. http://www.scenocosme.com/contacts_installation_en.htm, 2009.