Designing Sound Representations for Responsive Environments

Takuya Yamauchi

yamauchi@yinteraction-design.com

ABSTRACT

In this paper, we demonstrate a responsive sound installation consisting of computer-linked thermometers and cameras installed in both interior and exterior locations that detect the states of these spaces based on image and temperature data. The system simultaneously produces and modifies sound pictograms in different spaces in order to convey information on motion or temperature changes. We also propose a sound design method that represents sounds made by physical objects using the above-described installation and sound design method.

1. INTRODUCTION

Herein, we demonstrate a method for creating a sound installation that uses sensors and computers to present computer-generated sound and novel pictogram imagery using information obtained from cameras and thermometers. The computer-generated pictogram sound and image data are presented in a different location from the collection point, and the audience can recognize the changing states by listening to the sound and/or watching the imagery.

The system uses thermometers and cameras that are installed in both interior and exterior spaces to detect the state of those spaces based on the obtained image and temperature data. When the state of the captured images or the temperature changes, such as when a moving object is detected, a signal tone and a pictogram are created and assigned to describe the state based on a temperature and image data-mapping table.

The designed sounds change according to the detected states, and the proposed installation simultaneously produces sounds and pictograms in different spaces in order to convey information related to motion or temperature states. For example, if the temperatures of the interior and exterior spaces are different, different sound pitches are assigned to each space. Thus, individuals walking from outside to inside, or vice versa, will experience a change in sound.

If the installation were demonstrated in both interior and exterior spaces, the temperature in the spaces would also change with the season. For example, if the installation were demonstrated in winter, the inside space would be hot, and outside would be cool, or vice versa in summer.

If the installation is considered from a long-term point of view, individuals can experience sounds of varying pitch, which change with temperature, as the seasons change. Consequently, the assigned sound provides information on both motion concerning a moment and temperature regarding a long period of time.

2. BACKGROUND

2.1. Sound Pictogram

When a subject hears various sample sounds, his or her auditory reaction time will differ for each sound sample. Our proposed auditory system quickly and easily recognizes signal tones, which are characterized by a noticeable sound attack, as compared with environmental sounds. Similarly, sounds that include only a few frequencies are more easily detected than complex sounds that combine several frequencies.

The above-mentioned characteristics are used in various sound design methods in order to provide interaction tones for home appliances, and notification and warning sounds in various environments, as well as navigation information for people with visual impairments. Characteristics such as pitch, tempo, and amplitude can be used to classify whether a sound sample can be easily recognized, and it is important for sound designers to consider these characteristics.

2.2. Sound Design for Interaction

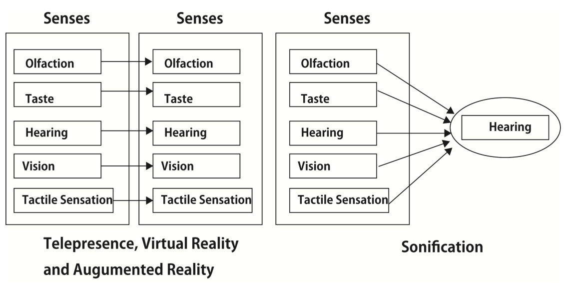

Figure 1 Senses and interaction

Auditory icons are icons that designate sound events on computers, and a number of sound icon design methods, such as SonicFinder, have been proposed for use in mapping sounds to computer events via a graphical user interface. [1] Mapping methods for use between physical objects and sounds have also been presented, and design methods have been applied to a tangible user interface between physical objects and computers. [2][3][4]

An auditory icon assumes that there are several layers that can be used to differentiate sound types. If a sound type is easily discerned, it may be an environmental sound that is generated by some physical phenomenon in the everyday environment. Sounds can evoke concrete images, and we can visualize physical objects based on sounds related to those objects. In contrast, in the case of simple sounds that consist of only a few frequencies, it may be difficult to imagine where the sounds are generated because of their abstract nature.

Computerized sound engines can generate a number of waveforms in order to produce artificial sounds that do not exist in the everyday environment because of their abstract nature. Both artificial sounds and environmental sounds cover a wide range of frequencies. The abovementioned layers range from concrete to abstract, and auditory icons are based on the concept of layers. Responsive systems analyze signals input from the environment and provide feedback by means of actuators. There are also cases in which feedback is returned as a system.[5] Figure 1 shows the relationship between our senses and demonstrates how the system transforms input into output. In cases involving virtual reality, telepresence, and augmented reality, input data that are detected by sensors are equivalent to output senses.

In other words, the system transforms visual input data to same-sense (visual) output data and auditory data to same-sense (auditory) data. This means that equivalent senses are mapped by the system. On the other hand, there are cases of non-like sense mapping, in which, for example, visual sense data are mapped to auditory sense data. Sonification and auditory display are examples of non-like sense mapping. Such approaches attempt to change visual input data into auditory output data in order to represent the detected data using a sound element, such as pitch or tempo. In the present paper, we focus on auditory representation.

2.3. Sonification timescale

One sound design for navigation systems or home appliances involves sound mapping of an action. In order to represent the target action, a mapped sound corresponding to that action is generated as soon as the action occurs. One such example would be an interaction sound that allows a user to determine whether a button has been correctly pushed based on the interaction sound.

We can begin by assuming that the response system is connected to a camera and that the system enables real-time image analysis. When the camera captures moving objects, the system analyzes their trajectories via image processing and the sum of the vectors of the trajectories is calculated as the state in the image. If we want to express a state via sonification, there are a number of sound mapping methods available.

The interaction of moving objects and sounds must be responsive, and a sound should be generated and expressed as soon as a moving object is detected. This method enables sounds to be recognized effortlessly so that listeners can easily imagine what happened in the space. In other words, the sound is a responsive sign pictogram that expresses the state of the detected space captured by the camera. Since sound feedback from an action is very quick, the response time is very short for the responsive sounds that are used in these methods.

Next, we will consider sonification-related data obtained from the natural environment. Examples of data detected in the natural environment include temperature and humidity over the course a month. The variability of data collected from the natural environment over a short period of time is lower than that of data collected from interactions between moving objects because such phenomena from the natural environment change very slowly.

Data such as solar wind and cosmic background radiation are measured over long periods in order to reduce data variability. If sonification was designed for these phenomena, sound mappings would be required in order to express changes in phenomena over extended periods of time. Changes in micro-world or space environments would need extended detection periods, and any sonification produced using that data would be difficult to comprehend intuitively. Therefore, it is clear that sound design for sonification should consider both short-period and long-period phenomena.

(1) Sound mapping for short periods of time

Sounds must quickly return feedback in order to report action and motion, and the pitch or tempo in the sound should be assigned to actions and motions.

(2) Sound mapping for a long period of time

Changes associated with the natural environment occur continually and gradually over long periods of time, and sound assignments must report the state of the phenomena. Thus, a new sound method is required.

We investigated the use of thermometers in sound mapping for both the interactions of moving objects and the natural environment, and decided to focus on a design method for interior and exterior spaces.

2.4. Temperature sonification

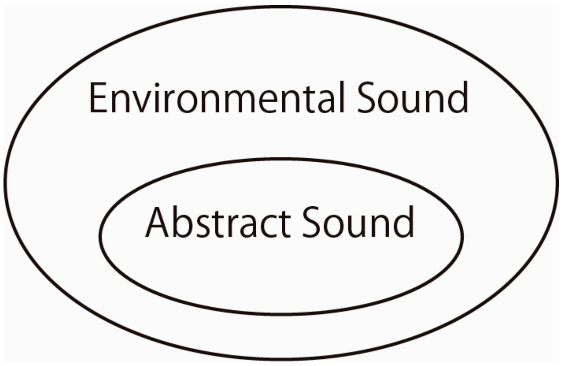

Figure 2 Conceptual diagram of environmental sound and abstract sound

In the following sub-section, we describe our attempt to implement a sound installation system that captures temperature and image data. One purpose of the installation is to create sound pictograms based on images taken by a camera and detecting temperature via a thermometer, from which the system generates a sound that integrates the image and temperature data. In order to present the state and temperature via sound, a new method of designing sound pictograms is required.

The results of psychological experiments using the semantic differential (SD) method indicate that the timbre of the sound of hitting metal provide a cold impression, whereas low tempo and pitch provide a hot impression. High pitch and fast tempo also provide a cold impression.[7][8] We designed the temperature-related sound based on these findings.

Next, we considered how to represent sounds related to sensor data and how sound should be used to present thermometer temperature data. In our installation, a sound pictogram and environmental sounds generated by granular synthesis are considered. Figure 2 shows a conceptual diagram of environmental and abstract sounds. When creating sound pictograms, we assumed that all sounds were composed of physical pitch, rhythm, and amplitude. The abstract sound in Fig. 2 consists of artificial sounds, such as pure tones.

Environmental sounds are composed of numerous abstract sounds. As in the case of environmental sounds, increasing the number of abstract sounds in an artificial design increases the complexity of the resulting sound. Much like an environmental sound, concreteness depends on the number of abstract sounds that are added. We designed environmental sounds and sign pictograms based on the conceptual diagram.

2.5. Environmental and Signal Tone

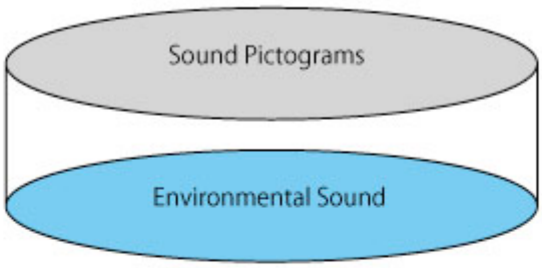

Figure 3 Sound design concept

Our sound design model consists of sound pictograms and environmental sound. Figure 3 shows the concepts behind this sound installation model. As can be seen in the figure, the concept involves the use of abstract sound in the form of sound pictograms and concrete sound as environmental sound.

As described in the previous section, both the abstract sounds used as the sound pictogram and environmental sounds have psychological and physical characteristics that compliment each other. We designed our sound installation based on category of the concept, and the sound pictogram and the environmental sound are used for each purpose.

One of the reasons environmental sound is used is to provide background sounds. In our model, environmental sound was generated by granular synthesis based on 10 sound types. Another purpose of the sound pictogram is to provide notifications and intentions. The sound pictograms represent temperature and provide information on states outside and inside a room, thereby informing listeners in the room by means of the designed sound.

2.6. Background Sound

Figure 4 shows a waveform chart and sound spectrographs based on 10 environmental sound samples that were generated via granular synthesis. The granular synthesis technique is used for combining sound grains, which are defined as pieces of sound in the sound samples, via numerical calculation. The resulting combined sounds can be flexibly changed by modifying the frequencies and amplitude of the sound grains. The spectrographs include periodical waveforms that are composed of numerous frequencies with flat peak levels. These 10 sound samples were used for synthesis by Max/MSP, which is a visual programming language.

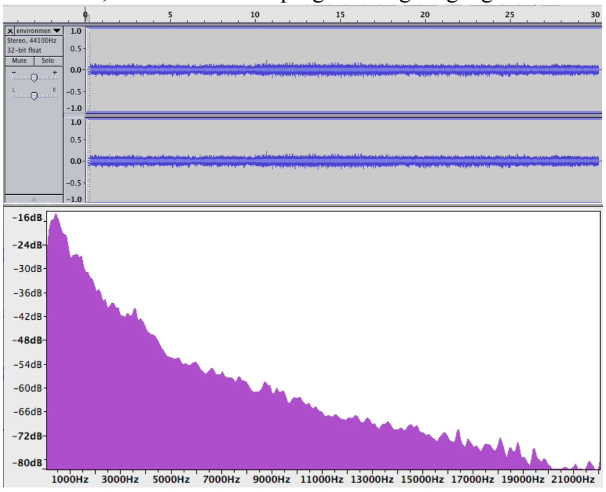

Figure 4 Waveform and spectrograph of environmental sound

2.7. Sound Pictogram

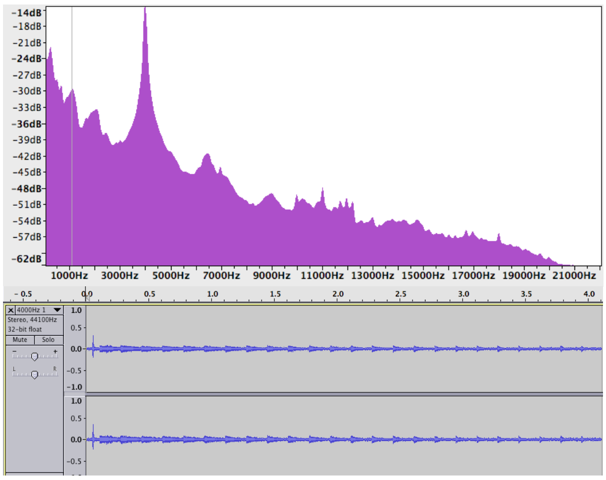

Figure 5 Waveform and spectrograph of a sound pictogram

The timbre, volume, and frequency of sounds are indicated via sine waves or noise in the sound pictograms of artificial sounds in order to indicate changes in state or motion.[6] For example, timbre changes according to temperature. If the timbre includes metallic factor in SD scale collections produced by factor analysis, then the sound gives listeners the impression of cold. High-frequency sound also gives listeners the impression of cold, whereas low-frequency sound gives listeners the impression of low activity.

In our sound installation, SuperCollider was used to create a signal tone generated by a ringing filter in order to adjust the sound timbre. Figure 5 shows the waveform and spectrum of the signal tone. The signal tone is a periodic sound, and includes only a few frequencies. The spectrum has a frequency peak of 4,000 Hz and also has only a few frequencies. The following items were considered in designing the signal tone:

(1) To generate an easily heard pure sound to report the state that occurs.

The signal tone, which is categorized as an abstract sound, is a periodical waveform that contains just a few frequencies. State changes such as detected motion are reported by the signal tone. We next considered how sound should be designed in order to indicate the reported state.

(2) To report temperature by the signal tone.

In our demonstration, assignment of sound for a continuously changing parameter such as temperature was considered. The temperature in a room changes continuously over long periods of time but does not change drastically over short periods of time. We also considered how assignment should be made.

The proposed system assigned a sound depending on the temperature as measured by the thermometer, and sounds were generated when moving objects were detected. The table shows the mapping relationship between temperature and pitch of the generated sound. In the case of a low temperature in the detection space, moving objects in the space are indicated by low-pitch sounds, whereas in the case of a high temperature, movements are indicated by high-pitch sounds.

Temperature	Pitch
Low	Low
High	High

Table 1 Mapping relationship between temperature and pitch

3. CONCEPT

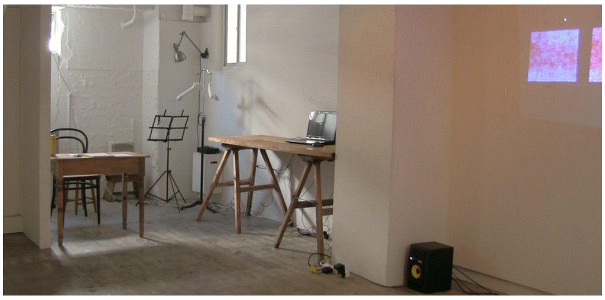

Figure 6 Sound installation

Next, we demonstrated an interactive artwork using sensors and sounds. Figure 6 shows the space used for the interactive artwork. The purpose of the interactive artwork is to represent various data types, such as temperature or motion, using sound and images. In particular, the present study focuses on the simple representation of a large amount of sensor data.

It is expected that individuals will be able to recognize not only states of motion or noise based on auditory and visual pictograms, but also continuous changes in those states, such as multiple individuals simultaneously walking from an interior to an exterior space. When numerous people walk through the space, the system also detects their trajectories. A signal tone for the people is then reported in the room, and we can recognize the new state from the resulting noise.

Since representation sounds change according to temperature and motion, seasonal changes can be recognized based on pictogram sound in the long-term point of view. Here, the signal tone is reported when the system detects state changes, such as motion, via the camera, and the pitch of the signal tone also changes according to temperature changes. Changes to the signal tone created by temperature changes can be easily recognized if the system is active for long periods of time.

A responsive system for an interactive artwork presents not only the motion or action interactions, but also long-period changes, such as seasonal changes.

4. Implementation

4.1. Visual Pictogram

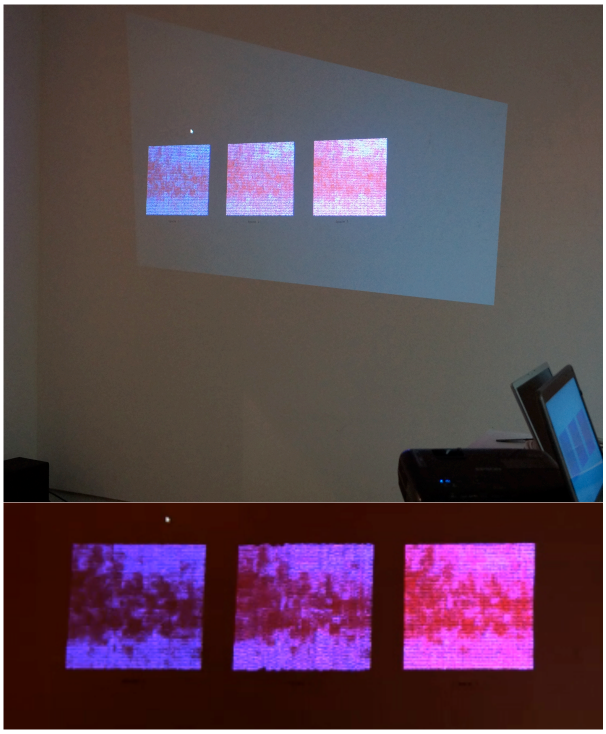

Figure 7 Visual pictogram in the sound installation

The interactive artwork utilizes simple icons indicating temperature and motion. One purpose of the display is to indicate complex state and temperature information via simple visual and auditory icons. Another purpose is to represent these icons as real, existing objects, such as furniture or shadows. The icon color provides an indication of temperature. High temperatures are shown in red, and low temperatures are shown in blue (Fig. 7).

4.2. System

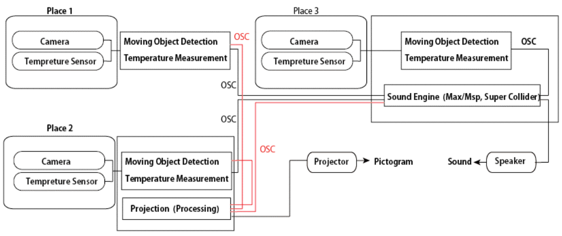

Figure 8 System Architecture

The proposed sound installation system is composed of three cameras and three thermometers connected using network protocols. The OSC protocol sends data to Max/MSP and a number of projectors in real time. Figure 8 shows the system and network used in our installation. After captured image and temperature data are sent to these engines, pictograms are projected as information on the states that exist in these spaces.

4.3. Camera Detection

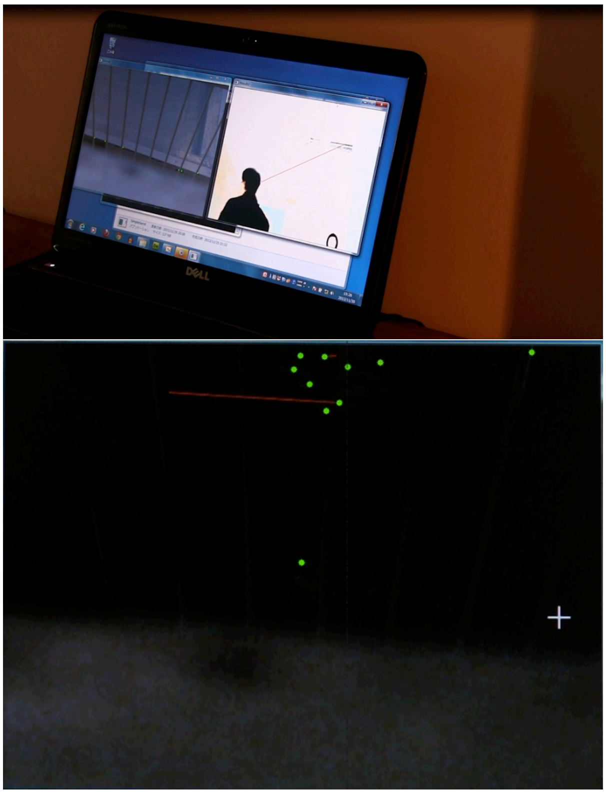

Figure 9 Camera detection system

Three cameras for detecting moving objects were included in our installation (Fig. 9), and the moving objects were analyzed by an optical flow algorithm of open source computer vision (OpenCV). The algorithm used for calculating spatial vectors of luminance estimates the flow related to the vectors in order to detect moving objects.

The system analyzed the states of interior and exterior spaces based on captured images, regardless of whether the images were captured in an active space. For example, for a case in which there are moving objects in a cold outside space, the system first detects the moving objects and the temperature information. Next, computer-generated images and sounds are created based on the temperature information and the audience can hear temperature sounds based on the activation in the cold space.

4.4. Temperature Detection

Figure 10 Thermometer in the sound installation

A thermometer with an embedded microcontroller is used to generate the sounds and operate the projectors (Fig. 10). Three thermometers were positioned in interior and exterior locations, and the temperature levels were observed in real time. Observed data were then analyzed and sent to visual server via processing and sound engine. The pictogram that results from the processing reflects ambient light based on the temperature in the computer graphics, and thus represents the temperature state.

5. Conclusion

Our proposed system, which can measure temperature in interior and exterior spaces and can detect moving objects by camera, presents the state of these spaces by mapping sounds and pictograms. The pitch of a generated sound changes according to temperature and the activity of the moving object. Based on the detected image, if the temperature of a space is determined to be high, a high-pitch sound is generated. Individuals can then simultaneously recognize activities occurring in both interior and exterior spaces based on generated sounds.

Sound pictograms were designed by combining environmental and abstract sounds, which each have a role in representing the state in the spaces and ambient sounds. If individuals can feel or recognize information about the location states and temperature levels by hearing the mixed sound, they can also understand auditory, visual, and tactile sensations by simultaneously hearing those sounds. Sonification is an augmented reality that is realized by sound design.

The method used by the designer to map the sounds is important, and sensor data must be assigned to valid sounds. These assignments depend on the designer's receptivity, and it is important to remember that the impressions of individuals will also change based on the sound design. For instance, a heartbeat can be accurately mapped to a pulse sound in electrocardiography, which then provides information on the cardiac state by means of an electrocardiogram and auditory information. The designed pulse sound can provide accurate information for predicting imminent cardiac events. When designing the sound to be used, it is important to map the sound to the actual object.

6. ACKNOWLEDGMENT

We would like to thank the members of Cyber Sound Project for their helpful assistance.

7. REFERENCES

[1]"The SonicFinder: an interface that uses auditory icons", William W. Gaver

,Journal Human-Computer Interaction archive Volume 4 Issue 1, March 1989 Pages 67-94

[2] Ref Till Bovermann, Rene Tünnermann, Thomas Hermann (2010). Auditory Augmentation. Int. J. on Ambient Computing and Intelligence (IJACI), 2 (2) , p. 27 – 41

[3] "Sound Jewelry" Takuya Yamauchi and Toru Iwatake Leonardo Music Journal No 18 MIT Press

[4] “the transfinite” Ryoji Ikeda, audiovisual installation 2011

[5] Takuya Yamauchi and Toru Iwatake "Design of a process for interactive product in ubiquitous space " International Journal on Interactive Design and Manufacturing (IJIDeM) Volume 2, Number 2 / 2008,5 Springer Verlag

[6] O.Kitamura, S. Namba and A.Matsumoto “Factor analytical study of tone color”,Proc. 6th ICA,A-5-II”

[7] Grey, J.M. “Multidimensional Perceptual Scaling of Musical Timbres”, Journal of the Acoustical Society of America, 61(5): 1270-1277

[8]Wessel, D. L., “Timbre Space as a Musical Control Structure”, Computer Music Journal, 3(2): 45-52