# Soundscape Streaming and Visualization for HCI Design

Hill Hiroki Kobayashi
Center for Spatial Information Science, The University of Tokyo, Chiba, Japan.

**Summary**

Currently, human–computer interaction (HCI) is primarily focused on human-centric interactions; however, people experience many nonhuman-centric interactions during the course of a day. Interactions with nature, such as experiencing the sounds of birds or trickling water, can imprint the beauty of nature in our memories. In this context, this paper presents a system for HCI design. This novel system can stream and visualize the presence of wildlife in undeveloped natural locations through HCI. The application results indicate that users take the sound as a cue that triggers nonlinguistic believability in the form of the mythological metaphor of wildlife. This paper describes a basic system of soundscape streaming and visualization for HCI design.

**PACS no. xx.xx.Nn, xx.xx.Nn**

## 1. Introduction

Human beings can imagine the presence of animals in the forest through their voices because our brain has the ability to record and store the images as a past experience. Since each person's history and experience are different, the imagined presence of the animals differs in relation to the sizes, shapes, and types of animals. A contemporary approach to observing the difference in the animal presences among users is to place a remote microphone and camera in a forest and to record the animal presence. By listening to and seeing the captured streams and sharing the experience, we can confirm the actual animal through HCI without being there.

This study has attempted to understand the processes of nonhuman-centric interaction between users and remote uninhabited environments through the use of information technologies and to reveal new knowledge regarding such interactivity. In doing so, this study hopes to discover the cognitive processes of our imagination mechanism. Such a discovery would help us design an interactive system that leverages the boundary of the real and virtual worlds by engaging the cognition to perform a nonhuman-centric interaction through HCI.

This paper will discuss the design, development, and evaluation of a system to tackle the previously mentioned mythological phenomenon in a contemporary real-world setting. This study is not intended to propose a solution to any one single technological or ecological problem; however, it proposes a new viewpoint of multidisciplinary HCI design and interfaces. The structure of this paper is as follows. Section 2 details related studies and previous studies. Section 3 describes the proposed bioacoustic streaming system. Section 4 describes the proposed bioacoustic visualization system and offers conclusions.

## 2. Related Studies and Previous Studies

This study is based on the concept of human–computer–biosphere interaction (HCBI) [7]. HCBI is an extension of human–computer interaction (HCI) [3] and human–computer–pet interaction (HCPI) [9]. Computer-supported cooperative work (CSCW) is based on computer-interaction paradigms to support specific activities. For example, we can exchange ideas, thoughts, theories, and messages by encoding them into transferable words, communicating them through computer systems, and then decoding them. However, in our daily lives, we implicitly exchange and share a great deal of additional nonverbal information to maintain our social relationships, such as messages acknowledging the presence and mood of others [4]. The consideration of implicit (background) information opens up new possibilities for interaction through nonlinguistic, wearable computing devices and

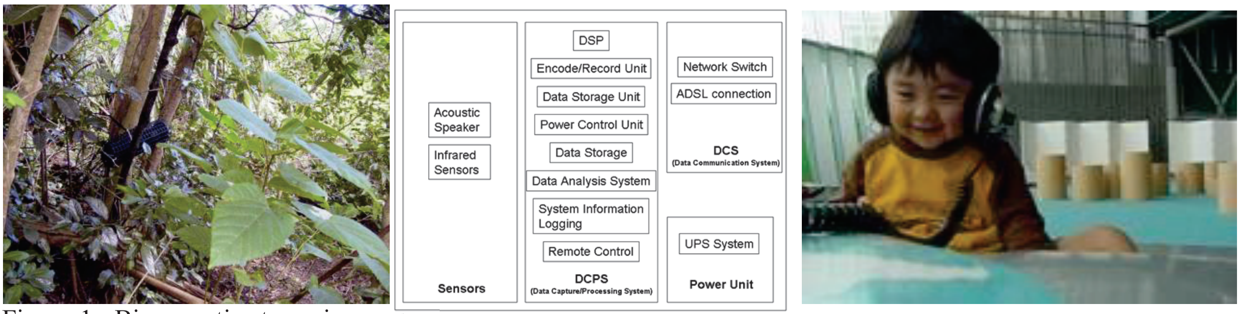


Figure 1. Bioacoustic streaming system

nonverbal remote coomuniation among different species. Werable computing deviecs enable us to extend our spatial interactions and develop human–human communication beyond physical distance through [2].

Development in two stages has been ongoing since the year 1997. The first development stage was to build a networked bioacoustic streaming and recording system in an uninhabited mountain forest in the northern part of Iriomote Island (24° 20' N, 123° 55' E) in the southern Ryukyu Islands, Japan. The real-time streaming system has been upgraded several times over the years to improve its long-term stability under unmanned operating conditions. The second development stage was to build a networked bioacoustic streaming and visualization system in an uninhabited mountain forest at Chichibu University Forest, the University of Tokyo, in Japan (35° 56' 17.28" N, 138° 48' 11.04" E, 150 km from Tokyo).

## 3. Bioacoustic Streaming System

Natural communities contain a spectrum of lifeforms that interact with each other. Many scientists agree with the judgment that the essence of ecology is the study of interactions among species in their native habitat [1]. Using ubiquitous technologies, it is generally considered that typical environmental monitoring near urban areas can be performed effectively. However, in those remote areas, such as the home range of wildlife, the availability of electric power and information infrastructures for monitoring wild animals is either limited or non-existent. This is primarily because the profit generated by infrastructure-based services is usually low in areas such as the sanctuary forest shown in Figure 1 (left), where the number of users is small [7]. Furthermore, in areas where no technological infrastructure is present, observation over a long duration (e.g., over several years) is difficult. Thus, it was necessary to develop methods that make the system work while using the fewest possible resources.

Based on this requirement and the construction and operating experiences of the first developments, a new observation tool for a monitoring system was created. The system enables users to listen to live stream audio in real time. The system diagram is shown in Figure 1 (middle) [6]. A weatherproof microphone placed within the ecosystem collects environmental sounds 24 hours a day, 365 days a year. This part of the audio digitization system (ADS) converts the analog audio signal from the microphone to a digital audio signal and transfers the signal to the recoding system with extremely low noise. The weatherproof microphone setup consists of a non-directional microphone wrapped in a sheet of thick waterproof sponge and a hard plastic mesh. The microphone and cable joint are covered with waterproof putty and tape to protect them from moisture. Microphones were attached in pairs to trees with elastic bands, as shown in Figure 1 (left). Digitized signals are then sent to the DCPS. The digital signal cable can be extended up to 10 km without incurring digital distortion. Next, the digital audio signal is connected to the audio processing system. This audio signal is digitally processed to enhance its quality using remotely controllable real-time audio processing software. The processed audio signal is sent to the encoding/recording system, encoded into an MP3 live stream, and recorded as WAVE sound format files. The MP3 live stream is sent to a stream server at the data archive system directly through the Internet and is played on various MP3-based audio software formats at different locations around the world, as shown in Figure 1 (right) [7].

This project received a Jury Recommended award at the Japan Media Arts Festival in 2002. Since 1997, real-time environmental sounds from Iriomote Island's subtropical forests have been monitored by networked microphones and transmitted via the website as "Live Sound from Iriomote Island" 24 hours a day, 365 days a year [7]. The technology behind this networked bioacoustic streaming and recording system is also

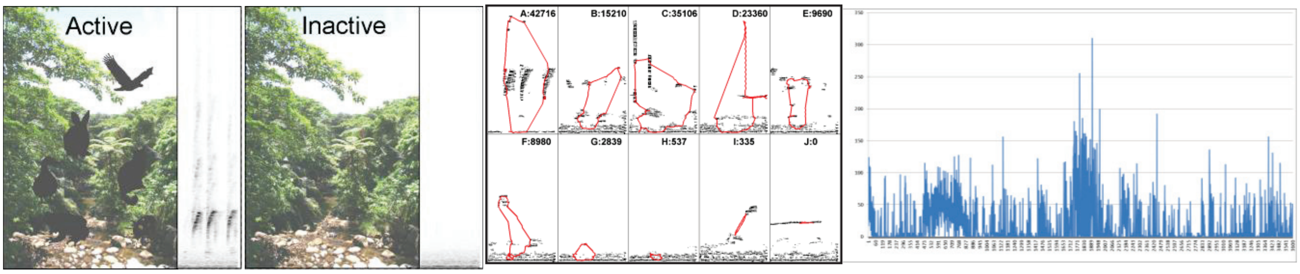


Figure 2. Bioacoustic visualization system

used for "AQUASCAPE: the stethoscope of the Earth's water", through which Internet users able to listen to sounds in real time, for example the moving water of a pond in Tokyo, living creatures in a Japanese garden in Kyoto, and a street in Mumbai City, India [7].

## 4. Bioacoustic Visualization System

To visualize the bioacoustic activity contained in the remote forest soundscape as a HCI design in real time, we proposed a bio-activity index (BAI) to convert this activity into numerical data using the Internet [8]. BAI uses active contour models to quantify the bioacoustic activity of the calculated area into the shape of the visualized bioacoustic patterns of the live soundscape data.

Examples of bioacoustically active and inactive scenes with spectrum data are shown in Figure 2 (left) [7]. This figure reflects actual records obtained from the developed system and sonographic data. As can be seen in the figure, in the active condition shot, several sets of vertically long but horizontally short tracks are present. The points shown are radically broad spectrum prints recorded over a short period of time. In contrast, in the inactive scene, ash-colored areas show sounds present in the low-frequency regions (wind and trees), compared with silence in the high frequency regions. Through this comparision of the two conditions, it can be concluded that it is possible to monitor the presence of wildlife in subtropical forests, where species exchange information on their presence bioacoustically, from animal voiceprints. However, the figure does not identify the species of animals calling, and there are always undiscovered species in any remote area richly endowed with nature.

To extract bioacoustic prints from streamed sounds, a bio-activity index (BAI) was first used to convert the bioacoustic activity into numbers. A BAI applies active contour models [5] to quantify activity from the calculated shape of visualized bioacoustic patterns in the live soundscape data obtained from the forest.

The active contour model algorithm forms contours that highlight features of interest within an image **[5]** The model is a controlled continuity spline under the influence of image forces and external constraint forces. The internal spline forces serve to impose a piecewise smoothness constraint. The image forces push the snake in the model algorithmtoward salient image features like lines, edges, and subjective contours. The external constraint forces are responsible for placing the snake near the desired local minimum. Representing the position of a snake parametrically, $v(s) = v(x(s), y(s))$, we can write its energy functional as

$$E^*_{snake} = \int_0^1 E_{snake}(v(s))ds$$
$$= \int_0^1 E_{\text{int}}(v(s)) + E_{image}(v(s)) + E_{con}(v(s))ds \qquad (2)$$

where $\boldsymbol{E_{int}}$ represents the internal energy of the spline due to bending, $\boldsymbol{E_{image}}$ gives rise to the image forces, and $\boldsymbol{E_{con}}$ gives rise to the external constraint forces and provides examples of $\boldsymbol{E_{con}}$ for interactive interpretation. Given an approximation of the boundary of an object in an image, an active contour model can be used to find the actual boundary. Again, this development was designed to detect wild animals at mid-range distances, and thus allows the system to detect animals as they approach.

To evaluate the BAI method, we set up a system to receive real-time audio transmissions from the forest and to calculate and evaluate the BAI using active contour models. The remote system shown in Figure 1 was installed on the island (and in the Chichibu University Forest) to capture and transmit live sounds from the two forests in CD-quality format over the Internet. The local system, OpenCV on Linux Platform installed in the University of Tokyo, was set up to receive and calculate the BAI of the live sounds every second.

Figure 3. Designed Interfaces.

As shown in Figure 2 (middle), the BAI could be successfully measured. It was found that bioacoustical information usually occupies several bands of a 20–10000 Hz frequency spectrum. Other environmental sounds are usually detected below 1000 Hz. Ten spectrum image results are shown in Figure 2 (middle). Darker lines indicate a greater amplitude in a particular range. The red contour on each result is the pattern detected using the active contour model. The BAI is the calculated area of the pattern. A large area indicates that biological activity in the forest is high. In the case of a quiet forest, smaller areas indicate that biological activity is absent.

Results A, D, and F, which show high BAI, indicate that bioacoustic activity has definitely been detected by the active contour model and its area is calculated in pixels. However, results B, C, E, and I in Figure 2 (middle) show that portions of the pattern were not detected. This indicates that software improvements are necessary. Results G and H, which show low BAI levels, are typical inactive states in bioacoustic environments. The contours are placed at the bottom part of the spectrum, where other environmental sounds are displayed. Result J is a methodological problem for this application when using the active contour model. Despite the pattern of horizontally long and vertically thin results, a call from a wild animal with a narrow band spectrum could not be properly included in the calculation. Figure 2 (right) indicates the result of the continuous process of the data on August 1, 2013.

## 5. Discussion

Wearable Forest [8] in Figure 3 (left) is a garment that bioacoustically interacts with distant wildlife in a remote forest through a networked remote-controlled speaker and microphone. It expresses the unique bioacoustic beauty of nature and allows users to interact with a forest in real time through a network to acoustically and visually experience a distant forest soundscape, thus merging humans and nature without great environmental impact. This novel interactive sound system can create a sense of unity between users and a remote soundscape, enabling users to feel a sense of belonging to nature even in the midst of a city.

## 6. Conclusions

This study aims to increase nature awareness in the midst of a modern city through HCI. This novel system can stream and visualize the presence of wildlife in undeveloped natural locations through HCI. The application results indicate that users take the sound as a cue that triggers non-linguistic believability in the form of the mythological metaphor of the wildlife. This paper describes the basic system of soundscape streaming and visualization for HCI design.

## Acknowledgments

This research is supported by MIC: Strategic Information and Communications R&D Promotion Programme (SCOPE) (No. 142103015).

## References

[1] M. Begon, *et al.*: *Ecology: Individuals, Populations, and Communities*, 3rd ed., Blackwell Science, Oxford, Cambridge, MA., 1996.

[2] R. Bray, *et al.*: *Botanicalls*, 2009. Available: http://www.botanicalls.com/

[3] B. Hewett, *et al.*: *Definition of Human Computer Interaction*, 1992. Available: http://old.sigchi.org/cdg/cdg2.html

[4] Y. Itoh, *et al.*: 'TSUNAGARI' communication: fostering a feeling of connection between family members. In: *CHI '02 Extended Abstracts on Human Factors in Computing Systems*, ACM 2002, 810–811.

[5] M. Kass, *et al.*: Snakes: Active contour models. *International Journal of Computer Vision* 1, 4 1988/01/01 (1988), 321–331.

[6] H. Kobayashi, *et al.*: Tele echo tube: beyond cultural and imaginable boundaries. In: *Proceedings of the 21st ACM International Conference on Multimedia*, ACM 2013, 173–182.

[7] H. Kobayashi, *et al.*: Human computer biosphere interaction: towards a sustainable society. In: *CHI '09 Extended Abstracts on Human Factors in Computing Systems*, ACM 2009, 2509–2518.

[8] H. Kobayashi, *et al.*: Wearable forest – feeling of belonging to nature. In: *Proceedings of the 16th ACM International Conference on Multimedia*, ACM 2008, 1133–1134.

[9] P. Lee, *et al.*: A mobile pet wearable computer and mixed reality system for human poultry interaction through the internet. *Personal and Ubiquitous Computing* 10, 5 (2006), 301–317.