Nonspeech Audio-Based Interfaces

Shigueo Nomura
Kyoto University, Japan

Takayuki Shiose
Kyoto University, Japan

Hiroshi Kawakami
Kyoto University, Japan

Osamu Katai
Kyoto University, Japan

INTRODUCTION

Visual and auditory imagery combination offers a way of presenting and communicating complex events that emulate the richness of daily experience (Kendall, 1991). It is notable that sound events arise from the transfer of energy to a sound object in everyday life. Even in childhood, we learn to take the following attitudes about the sound events:

- Recognize the occurrence of sound events and relate them to physical events.
- Classify and identify heterogeneous sound events through a lifetime of experience.

Important distinctions in the data can be communicated by exploiting simple categorical distinctions of sound events. Taste, smell, heat, and touch are not suitable channels for data presentation because our perception of them is not quantitative. However, the auditory system constitutes a useful channel for data presentation (Yeung, 1980).

Furthermore, sounds play an important role in the study of complex phenomena through the use of auditory data representation according to Buxton (1990) and Kendall (1991). It is known that our ears and brains can extract information from nonspeech audio that cannot be, or is not visually displayed (Buxton, 1990).

BACKGROUND

Nonspeech Audio

According to Wall and Brewster (2006), nonspeech audio can be delivered in a shorter time than synthetic speech. Synthetic speech audio can be laborious and time consuming to listen to and compare many values through speech alone. So, nonspeech audio can be a better means at providing an overview of the data.

Researchers, such as Bronstad, Lewis, and Slatin (2003) have investigated whether the use of nonspeech audio cues can reduce cognitive workload to users performing very complex tasks that they would otherwise find impossible.

Nonspeech audio researchers have investigated sounds more complex than ubiquitous interrupting beeps to provide information about spatial structure to computer users. In this way, the vOICe Learning Edition (Jones, 2004) is an actual example of interface that translates arbitrary video images from an ordinary camera into nonspeech sounds. However, the artificial sounds adopted by the vOICe have no analogs in everyday listening. So, this kind of interfaces has required extensive trials from users before their effective use.

Everyday Listening

According to Buxton (1990) and Gaver (1988), everyday listening is the experience of listening to events rather than sounds. This experience is different than that enjoyed by traditional psychoacoustics. It consists of hearing which things are important to avoid and which might offer possibilities for action. Instead of perceiving attributes, such as frequency, spectral content, amplitude of sounds, everyday listening is concerned with the attributes of events in the world. Examples of events are the speed of an approaching automobile, the force of a slammed door, and the direction of a walking person (Gaver, 1988). Everyday tasks, such as driving and crossing the street, are examples due to everyday listening. Listening to airplanes, water, birds, and footsteps are other examples of everyday listening. Also, Gaver (1993) has investigated several studies of everyday listening based on the ecological perspective.

Surprisingly, everyday listening skills have been virtually ignored in computer-based interfaces. Listening to events is not well understood by traditional audition approaches.

Studies suggest that a comprehensive account of everyday listening has yet to emerge. We have investigated a first study on comprehensive account of everyday listening by using nonspeech sounds as events (Nomura, Utsunomiya, Tsuchinaga, Shiose, Kawakami, Katai, & Yamanaka, 2007a).

We have concentrated on such everyday listening and ecological approach to perception in our previous works (Nomura, Shiose, Kawakami, Katai, & Yamanaka, 2004; Nomura, Yamanaka, Katai, Kawakami, & Shiose, 2004; Shiose, Ito, & Mamada, 2004). One of the works concerned with perceiving an approaching automobile is useful in understanding the scope of an ecological approach to perception (Shiose et al., 2004). The auditory system captures the experience of everyday listening by the idea that a given sound provides information about the interaction of materials at a location in an environment. Also, alterations in loudness caused by alterations on distance from a source may also provide information about time-to-contact in an analogous fashion to alterations in visual texture (Shaw, McGowan, & Turvey, 1991).

Echolocation

Spallanzani discovered that blinded bats could fly, avoid obstacles, land on walls and ceiling, and survive in nature as well as sighted bats. However, this discovery remained unanswered as how bats possess a sixth sense for orientation and navigation (Raghuram & Marimuthu, 2005). Griffin answered the question in 1938 and called this sixth sense "echolocation" (Griffin, 1958).

Some bat-like sonar systems have been developed using echolocation (Barshan & Kuc, 1992; Bitjoka & Takougang, 2007; Waters & Vollrath, 2003).

On the other hand, Daniel Kish (Roberts, 2006), as an example of human echolocation, lost his sight as an infant and taught himself to "see" with sonar by clicking his tongue. He learned to see without sight.

In our recent work (Nomura, Chiba, Honda, Shirakawa, Shiose, Katai, Kawakami, & Yamanaka, 2008), we looked at the possibility to emerge a comprehensive account of human echolocation that enables spatial structure perception of the environment using acoustics.

Virtual 3-D Acoustic Environment

Virtual 3-D acoustic environments created by computers are an emerging technology that may be used to teach blind children in actual acoustic environment (Inman, Loge, & Cram, 2000).

Advantages of using virtual 3-D acoustic training environments are to provide learners with guided and unguided practice controlling audio parameters by software. The parameters can be adjusted to suit the specific needs of a learner's auditory experience (Inman, Loge, & Cram, 2001).

The virtual 3-D acoustic environment is modeled on the binaural human hearing system (Inman, Loge, & Cram, 2001) and described by a complex response function (head-related transfer function - HRTF) (Mφller, Sφrensen, Hammershφi, & Jensen, 1995). HRTFs contain all the information about the sound source's location (its direction and distance from the listener), and can be used to generate binaural cues (interaural time differences - ITDs; interaural intensity differences - IIDs). Consequently, measured and implemented HRTFs can generate virtual 3-D acoustic environments.

In our previous work (Nomura, Utsunomiya, Tsuchinaga, Shiose, Kawakami, Katai, & Yamanaka, 2007b), we have investigated an approach to enhance spatial conceptualization performances of subjects using the virtual 3-D acoustic environment system.

WORK'S PURPOSE

We propose novel nonspeech audio-based user interfaces to provide visually impaired and elderly people with the opportunity to perceive and conceptualize spatial structure through echolocation. We hope that the users have opportunities to freely access the available facilities and friendly interaction with these targets (spatial structures) without depending on such expert systems as pattern recognizer, spoken language converter, or voice synthesizer.

Also, the purpose is not only or simply based on converting inaudible ultrasound echoes into audible sounds like the sonar system modeled by Bitjoka and Takougang (2007).

In this way, we believe that the user may take advantage of the similar skills employed by bats in everyday listening through echolocation. For example, the association of reverberations with empty environment is a kind of such skill. When a room is reverberant, it means that it is spacious, considering all other things equal. The idea is that the users can utilize their experience and familiarization with everyday listening to friendly conceptualizing spatial structure information without hard cross-modal training. The strategy is based on skill transfer process (Shiose, Sawaragi, Nakajima, & Ishihara, 2004) to embody the cues for spatial structure conceptualization process by avoiding heavy computational load, due to the previously mentioned expert systems (like the vOICe).

We suppose that the eventual interface users can embody the skills employed in everyday tasks by navigating and training in a virtual 3-D acoustic environment. Then, we experimentally evaluate performances of subjects on tasks to perceiving spatial structures represented by various aural surfaces in the 3-D acoustic environment.

8 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/chapter/nonspeech-audio-based-interfaces/13992

Related Content

Electronic Government Strategies and Research in the U.S.

Charles C. Hinnantand Steve Sawyer (2005). *Encyclopedia of Information Science and Technology, First Edition (pp. 1012-1017).*

www.irma-international.org/chapter/electronic-government-strategies-research/14378

Developing a Multi-Agency E-Participation Strategy for Disadvantaged City Communities: A Case Study

John N. Walshand Fergal McGrath (2016). *Handbook of Research on Innovations in Information Retrieval, Analysis, and Management (pp. 358-376).*

www.irma-international.org/chapter/developing-a-multi-agency-e-participation-strategy-for-disadvantaged-city-communities/137485

Developing a Hypertext GUIDE Program for Teaching the Simple Tasks of Maintaining and Troubleshooting the Educational Equipments

Kamel Hussein Rahoumaand Peter Zinterhof (2002). *Annals of Cases on Information Technology: Volume 4 (pp. 58-72).*

www.irma-international.org/article/developing-hypertext-guide-program-teaching/44498

A System Dynamics Approach to Changing Perceptions about Thermal Water Quality Trading Markets

Asmeret Bier (2010). *International Journal of Information Systems and Social Change (pp. 1-12).*

www.irma-international.org/article/system-dynamics-approach-changing-perceptions/45873

Web-Based 3D Real Time Experimentation

C.C. Ko, Ben M. Chenand C.D. Cheng (2009). *Encyclopedia of Information Science and Technology, Second Edition (pp. 4088-4092).*

www.irma-international.org/chapter/web-based-real-time-experimentation/14190