Recognizing Shapes and Gestures Using Sound as Feedback

Javier Sanchez
CCRMA - Center for Computer Research in Music and Acoustics
The Knoll, 660 Lomita Dr.
Stanford, CA 94305 USA
jsanchez@ccrma.stanford.edu

CHI 2010, April 10–15, 2010, Atlanta, Georgia, USA.
ACM 978-1-60558-930-5/10/04.

Abstract
The main goal of this research work is to show the possibility of using sound feedback techniques to recognize shapes and gestures. The system is based on the idea of relating spatial representations to sound. The shapes are predefined and the user has no access to any visual information. The user interacts with the system using a universal pointer device, as a mouse or a pen tablet, or the touch screen of a mobile device. While exploring the space using the pointer device, sound is generated, which pitch and intensity vary according to a strategy. Sounds are related to spatial representation, so the user has a sound perception of shapes and gestures. They can be easily followed with the pointer device, using the sound as only reference.

Keywords
Auditory (non-speech) feedback, auditory display, gestures, non-visual visualization, parametric curves, proprioception, sonification

ACM Classification Keywords
H.5.2. Information interfaces and presentation (e.g., HCI): User Interfaces: Auditory (non-speech) feedback. H.5.5. Information interfaces and presentation (e.g., HCI): Sound and Music Computing.

General Terms
Design, Experimentation, Performance.

Introduction

The aim of this research project is to use sound as feedback with the aim of recognizing shapes and gestures. The proposed system has been designed with the idea of relating spatial representations to sound, which is as a way of sonification.

Sonification can be defined as the use of nonspeech audio to communicate information [6]. Basically, the idea is to relate some parameters of the shapes and gestures we want to communicate, with some sound parameters as pitch, amplitude, timbre or tempo between others. By nature, sonification is an interdisciplinary field, which integrates concepts from human perception, acoustics, design, arts, and engineering.

The best-known example of sonification is the Geiger counter, invented by Hans Geiger in the early 1900's. This device generates a "beep" in response to non-visible radiation levels, alerting the user of the degree of danger. Frequency and intensity vary according to the existing radiation level, guiding the user.

Another example of sonification is given by the Pulse-oximeter, which was introduced as medical equipment in the mid-1980's. This device uses a similar concept that the Geiger counter. It outputs a tone, which varies in frequency depending on the level of oxygen in the patient blood.

Other known example of sonification is the Acoustic Parking System (APS) used for parking assistance in some cars. It uses sensors to measure the distance to nearby objects, emitting an intermittent warning tone inside the vehicle to indicate the driver how far the car is from an obstacle.

Sonification has been used to develop navigation systems for visually impaired people [8] allowing them to travel through familiar and unfamiliar environments without the assistance of guides.

Other works [2,11] are focused on creating multimodal interfaces to help blind and impaired people to explore and navigate on the web. The design of auditory user interfaces to create non-visual representations of graphical user interfaces has been also an important research activity [1,9].

Some systems have been developed to present geographic information to blind people [5,7,10]. It allows the user to explore spatial information.

In some works the aural feedback is added to an existing haptic force feedback interface to create a multimodal rendering system [3,4].

Although our system would be used to assist visually impaired people in the recognition of shapes and gestures, we do not want to limit its scope to this field of application.

System description

The interaction with the system is made by means of any universal pointer device, as a mouse, a pen tablet, a pen display or a touch screen of a mobile device.

This was the most important design specification of the system, which allows the user to use it with any existing device. Figure 1a shows how the user interacts with the system.

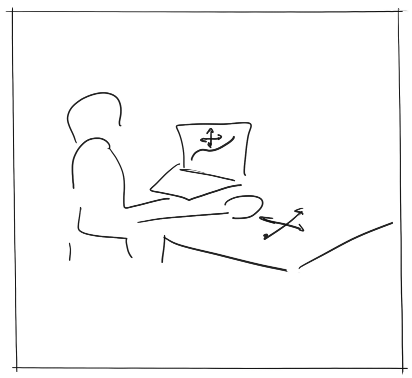


figure 1. Using a universal device facilitates the interaction.

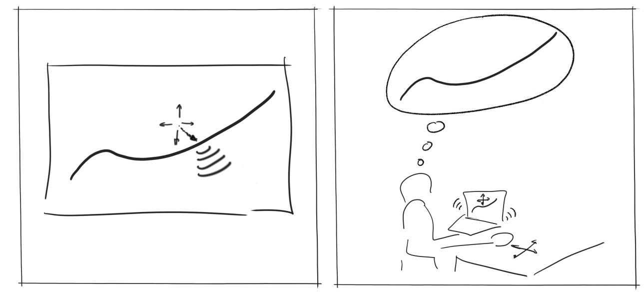
Some shapes have been defined previously and imported into the system, so the user is ready to identify them. It must be clearly stated that the user has not access to any visual information.

In order to identify the shapes, the user should explore the space by moving the pointing device. The movement of the pointer is directly associated with the movement of a virtual point in the imported shape.

As the user approaches to the shape, a sound is generated (see figure 2a), which pitch, timbre and intensity can vary according to a specific spatial to sound mapping strategy.

The system is based on the proprioception sense, which provides a relation between the gesture made by the user and the spatial representation of this gesture. Using the sound as feedback, the curve can be easily followed by the user.

Thanks to the proprioception sense, the hand gesture made while following the sound is transformed into a spatial representation of the shape (see figure 2b).



(a) (b)

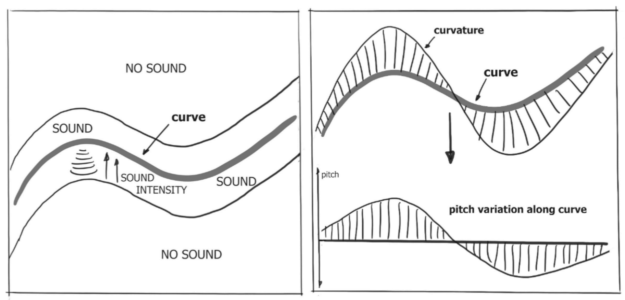
figure 2. The user has not access to any visual information, using only the sound as feedback. (a) a sound is generated when the user approaches to the shape. (b) the gesture made while following the sound is transformed into a spatial representation of the shape.

In order to provide a relation between the gesture made and the sound feedback, a perfect synchronization of perceived audio events with expected tactile sensations is needed. The sound intensity increases as the distance to the curve decreases, as shown in figure 3a.

In addition to this, some curve parameters, as position, slope or curvature are used to enrich the sound information given to the user.

A pitch variation is telling the user about the curvature of the shape at each point. A strategy consists on varying the sound pitch along the curve shape, according to the curvature at each point of the curve. As example, a straight line will generate a constant pitch. The curve represented in figure 3b has a variable curvature, so the user will have different pitch perceptions while he is moving along the curve.

Other strategies use the slope of the curve at each point to generate different sound pitches. Sound duration, rhythm, timbre or panoramization are other parameters that can be also related to the spatial movement of the user.



(a) (b)

figure 3. Sound to spatial relationship. (a) The sound intensity increases as the distance to the curve shape decreases. (b) Curve properties as slope or curvature are associated with sound parameters to enrich the sound feedback.

The curve shapes are represented by means of parametric curves, which are a standard in 2D drawing representation. Since Drawing Exchange Format (DXF) is used to store the graphic information, it is very easy to generate curve shapes using any commercial CAD application and import them into our system.

Other system Details

The analysis of the user motion, the curve representation and the output sound has been computed using MAX/MSP, a visual programming environment specifically designed to simplify the creation of acoustic and visual applications. Using the Jitter module, the graphic output has been also represented in real time (see figure 4)

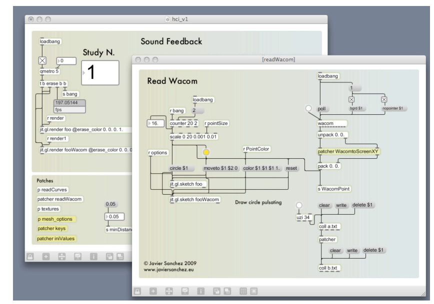
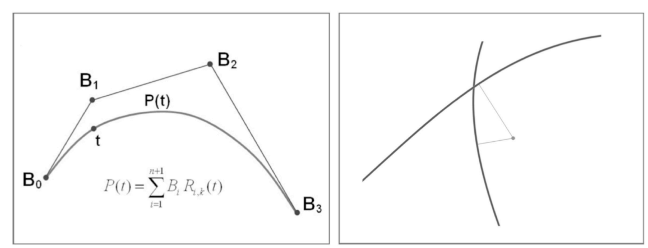


figure 4. MAX/MSP is an excellent programming environment to test a prototype system, adjust sound parameters or communicate with any universal device.

The curve shapes are represented by means of parametric curves, which are a standard in 2D drawing representation (see figure 5a). Since Drawing Exchange Format (DXF) is used to store the graphic information, it is very easy to generate curve shapes using any commercial CAD application and import them into our system.



(a) (b)

figure 5. Parametric curves are used to define shapes. (a) Parametric curve definition (b) multiple curves can be defined.

Multiple curve shapes can be defined into the same scenario using different sound pitches and timbres for each curve (see figure 5b). Distances to the curves are evaluated as the user interacts with the model.

Conclusions

This paper has presented a novel method that consists in the use of sound to recognize shapes, capturing the gesture of the user.

A universal pointer device is used to control the system, facilitating the human computer interaction. Parametric curves are used, as they are a standard in 2D drawing representation. Some of the curve parameters, as slope, curvature or position, are related to the sound output, helping the user to follow the curve. Multiple curve shapes can be defined in the same scenario using different timbres for each curve.

Current work is related with the use computer vision techniques to track the hand movement of the user. By means of this, the user can interact directly using the webcam of the computer.

It is also being evaluated the possibility of using the system as an extension (add-on) of some existing computer application.

Other applications are also been studied in which the sound can be related to a gesture to assist the user in common tasks.

The overall low cost of the system and its easy implementation is also an important point in favor.

References

[1] Buxton, W., Using Our Ears: An Introduction to the Use of Nonspeech Audio Cues. *Extracting Meaning from Complex Data: Processing, Display, Interaction*, edited by E.J. Farrel, Vol. 1259, SPIE 1990, p. 124-127.

[2] Donker H, Klante P, Gorny P. The design of auditory user interfaces for blind users. In Proc. of the second Nordic conference on HCI, pp. 149-156 (2002)

[3] Grabowski,N.A., Barner, K.E. Data visualization methods for the blind using force feedback and sonification. *SPIE Conference on Telemanipulator and Telepresence Technologies*, 1998

[4] IFeelPixel: Haptics & Sonification http://www.ifeelpixel.com/faq/#whatitwill

[5] Kamel, H. and J. Landay. Sketching images eyes-free: a grid-based dynamic drawing tool for the blind. *Proc. of ACM SIGCAPH Conference on Assistive Technologies* (ASSETS). pp. 33-40. (2002)

[6] Kramer G., Walker B., Bonebright T., Cook P., Flowers J; Miner N., Neuhoff, J: Sonification Report: Status of the Field and Research Agenda. International Community for Auditory Display, www.icad.org (1997)

[7] Krueger, M. KnowWare™: Virtual Reality Maps for Blind People. *SBIR Phase I Final Report, NIH Grant #1 R43 EY11075-01*, (1996)

[8] Loomis, J.M., Reginald G. G., Roberta L. K. Navigation System for the Blind: Auditory Display Modes and Guidance. *Presence*,V.7,N.2,193–203(1998)

[9] Mynatt, E. & G. Weber. Nonvisual Presentation of Graphical User Interfaces: Contrasting Two Approaches. *In Proc. of the Computer*, CHI 94. (1994)

[10] Parente, P. and G. Bishop. BATS: The Blind Audio Tactile Mapping System. *ACMSE.* (2003)

[11] Yu, W., Kuber R., Murphy E., Strain P., McAllister G. A Novel Multimodal Interface for Improving Visually Impaired People's Web Accessibility. *Virtual Reality*, Vol 9: 133-148 (2006)