# Auditory Illusions for Audio Feedback

Michel Beaudouin-Lafon and Stéphane Conversy Laboratoire de Recherche en Informatique, URA CNRS 410 LRI - Bâtiment 490 - Université de Paris-Sud 91 405 ORSAY Cedex - FRANCE e-mail: mbl@lri.fr - www: http://www-ihm.lri.fr/~mbl

## ABSTRACT

Sheppard-Risset tones are sounds that seem to go up (or down) indefinitely. We have designed an "elevator" sound based on this auditory illusion and have implemented it in the ENO audio system. The sound is synthesized in realtime and can be controlled in real-time through high-level parameters. We have used this sound for audio feedback when scrolling and for monitoring the progress of long system operations.

**KEYWORDS:** Non-speech audio, auditory icons, auditory illusions, feedback, notification.

## INTRODUCTION

As our screens become more and more crowded, we need to develop alternative ways to convey information to users. We are investigating the use of non-speech audio cues to provide feedback about user actions, notification of system's state changes, and awareness of other users' actions [1].

Our approach structures the sound space using a set of organizing concepts useful to application designers. This approach builds on Gaver's auditory icons [4] and, to a lesser extent, on Blattner's earcons [2]. It contrasts with most uses of sound in today's user interfaces, which play pre-recorded audio clips. In order to make it easier for applications to include audio cues, we have developed the ENO audio system [1]. ENO is based on the concept of sound sources that produce sound when subjected to an interaction. For example, a machine source produces sound when it is started, and an object source produces sound when it is hit or scraped. The sounds are generated in realtime and can be controlled in real-time by high-level parameters such as machine speed or object material.

We believe that auditory illusions have an interesting potential for user interaction. Like visual illusions, they exploit imperfections in our perceptual systems. For example, a common visual illusion in computer graphics consists of generating an animation by displaying a series of fixed images. In this paper we introduce the use of an auditory illusion known as Sheppard-Risset tones to convey

ACM CHI'96, Human Factors in Computing Systems, Vancouver, Canada, April 14-18, 1996 Conference Companion (Short Papers) ACM, New York, 1996. the idea of motion in the user interface. We first introduce Sheppard-Risset tones and then describe how they can be used for feedback when interacting with a scrollbar and for monitoring the progress of long system operations.

## SHEPPARD-RISSET TONES

Sheppard-Risset tones [5] are sounds that go up (or down) indefinitely. They are the auditory equivalent to M.C. Escher's famous etching "Ascending and Descending" which depicts an endless staircase. Sheppard-Risset tones are made of a set of partials whose amplitude is bounded by a bell-shaped curve (figure 1). Over time, the frequency of each partial is shifted upward, while its amplitude is adjusted to fit the bell curve. New partials enter the lower end of the bell curve while partials reaching the other end of the curve disappear. As a result, the overall frequency of the sound is constant, but is perceived as going up continuously. Descending sounds are created similarly by shifting the partials downward.

Sheppard-Risset tones are controlled by 4 parameters :

- Base frequency: the center frequency of the bell curve;
- Bandwidth: the ratio between the highest and lowest frequencies of the envelope curve;
- Speed: the relative shift in frequency per unit of time. The sound is steady when the speed is zero, it moves upward when the speed is positive and downward when the speed is negative;
- Density: the ratio between the frequencies of successive partials. For a given bandwidth, the density determines the number of partials.

We determined empirically the bounds of the parameters necessary to maintain the auditory illusion. Six partials and



Figure 1: a semi-logarithmic plot of the spectrum of a Sheppard-Risset tone. Over time, the partials' frequencies are shifted upwards.

a density around 1.2 give satisfying results. More partials actually make the sound richer and more intrusive. With these values, the optimal bandwidth is defined by a ratio of 3 (1 1/2 octave). We also found that the bell curve can be approximated with a triangle envelope with no perceivable loss in sound quality. This allows for a more efficient synthesis algorithm.

#### AUDIO FEEDBACK OF MOTION

We have designed an "elevator" sound based on Sheppard-Risset tones and have implemented it in ENO<sup>\*</sup>. We call it an elevator because it conveys the concept of vertical motion and the tone of the sound is not unlike an actual elevator. The source has two parameters: the size, bound to the base frequency (bigger lifts sound lower), and the speed. The bandwidth and density are fixed. We have used this "elevator" sound for audio feedback from scrollbars and progress bars.

Scrollbars are used in graphical user interfaces to allow users to view large documents through small windows. A typical scrollbar has two arrows for incremental moves, a thumb (sometimes called an elevator) for direct access and two paging areas, between the thumb and the arrows, for incremental moves by pages.

The main problem with scrollbars stems from the constant shift in attention between two points: the document being scrolled, and the scrollbar itself. One specific problem, known as kangarooing [3], occurs when paging: when the thumb is close to the position being clicked in the paging area, successive clicks page up and down because the thumb jumps around the clicked position. In general, it takes a while for users to understand what is happening, especially since they are not watching the scrollbar. Another problem occurs when the cursor is inadvertently moved out of the scrollbar: users do not notice it and get confused when the interface does not respond as expected.

The elevator sound solves these problems: whenever the cursor is in the scrollbar, a faint sound is played. Its speed is zero in the thumb, slow in the arrows and faster in the paging areas. This way, users know where they are without looking. When clicking or dragging, the sound gets a bit louder. If kangarooing occurs, the user is immediately notified by the fact that the sound changes direction.

This design has a drawback: it does not provide audio feedback of the relative position in the document. We have tried to change the base frequency of the sound to reflect the position in the document, but this tends to conflict with the motion effect. In addition, since most listeners cannot recognize absolute pitch, it is not a good idea to use frequency to convey an absolute value. This is still an open problem.

The second use of the elevator sound is for audio progress bars. Progress bars act as a notification mechanism when the system is engaged in a long operation, such as copying a large file, and the progress of the operation (percent-done) is known. The problem with visual progress bars is that they need to be looked at to know the progress of the operation. Since the operation is long, the user is likely to engage in other tasks.

We have used the elevator sound to complement the visual display of a progress bar: the speed conveyed by the sound reflects the speed of the progression. Therefore, the sound conveys the instantaneous speed at which the operation progresses rather than the current "percent-done". This way, it is very easy to know when the operation stops progressing: the sound stops moving. This technique also works when the percent-done is not known but the rate of progression can be monitored. For example, when downloading a document, the Netscape World-Wide Web browser displays the current throughput of the connection, even if the size of the document is not known. The throughput can be mapped to the speed of the lift sound, giving an accurate and non-intrusive way to monitor the progress of the operation.

## CONCLUSION

Most objects we use in the real world produce sound when being manipulated. However faint, we perceive them and they contribute to our representation of the world. Computer objects should also produce sound in order to be easier to apprehend and manipulate. These sounds need not be "natural", as long as they convey useful information. Among others, auditory illusions can be used to add auditory feedback to interactive applications. We intend to build on this work by developing other types of sounds and integrating them into ENO.

## ACKNOWLEDGMENTS

Thanks to Wendy Mackay and Stéphane Chatty for comments on an earlier draft of this paper.

#### REFERENCES

- [1] Beaudouin-Lafon, M. and Gaver, W., "ENO: Synthesizing Structured Sound Spaces", *Proc. Seventh Annual Symposium on User Interface Software and Technology (UIST'94)*, pp 49-57, ACM, Nov. 1994.
- Blattner, M., Sumikarwa, D. and Greenberg, R.,, "Earcons and Icons: Their Structure and Common Design Principles", *Human-Computer Interaction*, 4(1), pp 11-44, 1989.
- [3] Brewster, S., Providing a Structured Method for Integrating Non-Speech Audio Into Human-Computer Interfaces, PhD dissertation, University of York, U.K., 1994.
- [4] Gaver, W., "Auditory icons: Using sound in computer interfaces", *Human-Computer Interaction*, 2, pp 167 -177, 1986.
- [5] Risset, J-C., "Paradoxes de Hauteur", IRCAM Report no. 10, IRCAM, 1977.

<sup>\*</sup> ENO is available at URL http://www-ihm.lri.fr/eno and/or by e-mail to mbl@lri.fr.