

THE UNIVERSITY OF ROCHESTER

AREA PAPER
for the Degree Master of Science

Binaural Audio as an Accessible User Interface

Author:
Androwis ABUMOUSSA

Supervisor:
Dr. Jeffrey BIGHAM

April 23, 2013

Contents

1	Acknowledgements	iv
2	Abstract	v
3	Introduction	1
3.1	Motivation	1
3.2	Binaural Audio	2
3.3	Current Solutions	3
3.3.1	Transaural Audio	3
3.4	Research Domain	4
3.4.1	Audio Used for Analytics	4
3.4.2	Audio Used for Interfaces	4
3.5	Our Approach	4
4	Background	5
4.1	Audio’s Role in Accessibility	5
4.2	Current Audio Based Accessibility Solutions	5
5	Related Work	6
5.1	Interface Design	6
5.2	Interfaces	6
5.3	Engines	6
5.4	Analytics	6
5.5	6
5.5.1	Design Principles	6
5.5.2	Design Patterns	6
5.5.3	Design Methodologies	6
5.6	Artificial Intelligence	6

6	Our Approach	7
7	Future Work	8
7.1	Analytics Domain	8
7.2	Human Computer Interaction	8
8	Conclusions	9

Acknowledgements

Abstract

Introduction

Our society is driven by information. Fields of research explore how to communicate constantly changing information to interested parties at the appropriate time. With the influx of mobile devices that provide an always on channel, research has explored the effect these disruptions have on a multitasking computing environment. The goal of much of this research has been to study how relevant and correct information can be efficiently delivered to a user in a manner that does not distract from their current tasks [5].

Armed with devices that are constantly connected, our society has access to all sorts of information, everything from weather forecasts and traffic conditions, to neighboring attractions, restaurant schedules, store specials, even to the location and discoveries of our friends.

The essence of any event or update to user data is condensed into a notification. Current research has explored how notification systems attempt to deliver current, important information to computer screens. Many works have also explored the costs, benefits, and optimal displays of these notifications from psychological perspectives that overlap with our ability to handle interruptions and distractions [2, 5].

Given the stochastic nature of notifications, probabilistic reliability can be used to create intelligent messaging systems that gain a user's trust, rather than lose it [3]. Empirical evidence shows that users will ignore all notifications that are not highly valid when performing demanding visual tasks [4].

The question of a user's trust in a system is therefore important when designing notification interfaces, with the goal being mitigating all negative first impressions. Empirical evidence shows that users carry a historical bias when dealing with actionable notification [3].

The study of information design and options that are suitable for three often conflicting design objectives of notifications - interruption to primary tasks, reactions to specific notifications, and comprehension of information over time - are necessary.

3.1 Motivation

Imagine driving to work on a typical morning. Eric Horvitz dream of an intelligent interface has been realized, so your navigation system is checking

current road conditions relative to the location of the GPS inside of the car. Your smartphone has resumed polling your work email address, your calendar has been updated by a colleague, and your family is messaging you reminding you of a prior engagement.

system that knew exactly how much or how little to say. The system is able to provide just the right amount of context for each task you're performing.

3.2 Binaural Audio

Human auditory localization has been studied extensively [1, 9]. Humans are especially adept at localizing sounds in three dimensions. Consider a sound source to the left of a listener. Sounds from the source arrive at the left ear first, and a short time after reach the right ear. The amplitude of the left ear sound will be attenuated due to head shadowing.

The predominant auditory cues for determining whether a sound is coming from the left or right directions are the interaural intensity differences and the interaural time differences. Humans are also adept at identifying sound position that are in front of or behind them, along with estimating the sound source's elevation. This is possible because the incident sound waves interact with the torso, head, external ear (pinna) prior to arriving at the inner ear.

The directional dependent filtering to each of a subject's ears can be expressed as a frequency response, called a head related transfer function (HRTF), and thus a pair of HRTFs describe how sound from one location reaches the two ears. HRTFs are usually measured using human subjects or dummy-head microphones which consist of response pairs for the left and right ears corresponding to a large number of source positions surrounding the head.

There are two regions of interest when considering source locations of sound. When the sound is close to the head, the spherical curvature of the incident sound waves cause the HRTFs to change qualitatively as a function of distance, but at moderate distances, the incident waves can be considered planar. At extreme distances, humans are only capable to process auditory cues that only depend on the sound sources volume.

Given that humans are capable of processing sound signals to place the audio source in space, a binaural spatializer can then be implemented to simulate the auditory experience of one or more sound sources arbitrarily



Figure 3.1: Binaural microphone with exact dimension and density as human head

located around a listener. The general idea is reproduce the acoustical signals at the two ears that would occur normally in a natural listening situation.

This process is accomplished by convolving each source signal with a pair of HRTFs corresponding to the sound sources intended location. The resulting signal is presented to the user through headphones.

3.3 Current Solutions

3.3.1 Transaural Audio

Transaural audio is a method used to deliver signals to the ears of a listener using stereo loudspeakers. The idea is to filter the binaural signal such that the subject can process the subsequent stereo representation as a binaural signal. This technique was first put into practice by Shroeder and Atal [6,7]. Song et al. demonstrated how computer vision techniques could be applied to the field of Transaural Audio. By tracking a subject's head, the computer can recalculate the HRTFs necessary for the sound, as well as reposition the sound sources relative to the subject to provide an adaptive 3D audio system [8].

3.4 Research Domain

3.4.1 Audio Used for Analytics

3.4.2 Audio Used for Interfaces

3.5 Our Approach

Background

4.1 Audio's Role in Accessibility

4.2 Current Audio Based Accessibility Solutions

Related Work

5.1 Interface Design

5.2 Interfaces

5.3 Engines

5.4 Analytics

5.5

5.5.1 Design Principles

5.5.2 Design Patterns

5.5.3 Design Methodologies

5.6 Artificial Intelligence

Our Approach

Future Work

7.1 Analytics Domain

7.2 Human Computer Interaction

Conclusions

Bibliography

- [1] Jens Blauert. *Spatial hearing: the psychophysics of human sound localization*. The MIT press, 1997.
- [2] Edward Cutrell, Mary Czerwinski, and Eric Horvitz. Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. 2001.
- [3] Scott LeeTiernan, Edward Cutrell, Mary Czerwinski, and Hunter Hoffman. Effective notification systems depend on user trust. In *Proceedings of Human-Computer Interaction-Interact*, 2001.
- [4] Masha Maltz and Joachim Meyer. Cue utilization in a visually demanding task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 44, pages 283–283. SAGE Publications, 2000.
- [5] D.Scott McCrickard, Mary Czerwinski, and Lyn Bartram. Introduction: design and evaluation of notification user interfaces. *International Journal of Human-Computer Studies*, 58(5):509 – 514, 2003. `Notification User Interfaces`.
- [6] Manfred R Schroeder. Digital simulation of sound transmission in reverberant spaces. *The Journal of the Acoustical Society of America*, 47:424, 1970.
- [7] MR Schroeder and BS Atal. Computer simulation of sound transmission in rooms. *Proceedings of the IEEE*, 51(3):536–537, 1963.
- [8] Myung-Suk Song, Cha Zhang, Dinei Florencio, and Hong-Goo Kang. Personal 3d audio system with loudspeakers. In *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, pages 1600–1605. IEEE, 2010.

- [9] William A Yost and George Gourevitch. *Directional hearing*. springer-Verlag New York, 1987.