HAL open science

# Multipresence-Enabled Mobile Spatial Audio Interfaces

Michael Cohen, Owen Noel Newton Fernando, Adrian David Cheok

## To cite this version:

Michael Cohen, Owen Noel Newton Fernando, Adrian David Cheok. Multipresence-Enabled Mobile Spatial Audio Interfaces. 9th International Conference on Entertainment Computing (ICEC), Sep 2010, Seoul, South Korea. pp.434-436, 10.1007/978-3-642-15399-0_52 . hal-01055610

**HAL Id: hal-01055610**

**https://inria.hal.science/hal-01055610v1**

Submitted on 13 Aug 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

# Multipresence-Enabled Mobile Spatial Audio Interfaces

Michael Cohen[1], Owen Noel Newton Fernando[2], and Adrian David Cheok[2]

[1] Spatial Media Group, University of Aizu, Aizu-Wakamatsu, Japan 965-8580
mcohen@u-aizu.ac.jp
[2] Keio-NUS CUTE Center, National University of Singapore, Singapore 119613
{newtonfernando,adriancheok}@mixedrealitylab.org

**Abstract.** Mobile telephony offers an interesting platform for building multipresence-enabled applications that utilize the phone as a social or commercial assistant. The main objective of this research is to develop multipresence-enabled audio windowing systems for visualization, attention, and privacy awareness of narrowcasting (selection) functions in collaborative virtual environments (CVEs) for mobile devices such as 3rd- and 4th-generation mobile phones. Mobile audio windowing system enhances auditory information on mobile phones and encourages modernization of office- and mobile-based conferencing.

## 1 Introduction

Vision and audition are the two main human senses for obtaining information about the outside world, and rich interfaces need both modes (at least!). Visual windowing systems allow multiple and multiwindow applications to share display resources; audio windowing systems, in analogy to graphical windows, can bring order to a cacophony of multiple simultaneous sound sources. Audio windowing can be thought of as a frontend, or articulated user interface, to a system with a spatial sound backend. Audio windowing systems are especially appropriate on small screen devices where there limited space for information and when control changes are not needed.

Researchers are working on interactive spatial audio to support synchronous groupware like teleconferences, either desktop applications or mobile device applications. For instance, Herder and Yamazaki [2] describe a networked virtual environment with 3D audio to support a chatspace while Aoki et al. [3] present a mobile audio space intended for use by gelled social groups. To the best of our knowledge, there is no existing presence awareness systems that can handle multiply present sources and sinks. The narrowcasting attributes [1] presented here suggest an elegant solution for such multipresence environments.

The suite of inclusion and exclusion narrowcasting commands for sources and sinks are like analogs of burning and dodging (shading) in photographic processing. The duality between source and sink operations is tight, and the semantics are identical: an object is inclusively enabled by default unless, a) it

explicitly excluded (with `mute` or `deafen`), or, b) peers (either `self`-designated or non `self`-designation of the selected object) are explicitly included (with `select` [`solo`] or `attend`) when the respective avatar is not. We have developed a mobile phone interface which allows users to multicast their position to multiple receivers and control crowded soundscapes using these narrowcasting operations. Programmed with J2ME, our application runs on NTT DoCoMo iappli mobile phones.

## 2 "ı·Con": Mobile Device Multispace Interface

Featuring selectable icons with one rotational and two translational degrees of freedom, the "ı·Con" 2.5D dynamic map interface is used to control position, sensitivity, and audibility of avatars in a groupware session. Its isosceles triangle icons are representations of symbolic heads in an orthographic projection, including narrowcasting attributes.

We have developed symbolic representations of narrowcasting operations (as shown in Figure 1) for the mobile interface. In our mobile application, narrowcasting attributes graphical displays are triply encoded— by position (before the "mouth" for `mute` and `select`, straddling the "ears" for `deafen` and `attend`), symbol ('+' for assert '-' for inhibit), and color (green for assert red, yellow, and orange for inhibit). The attributes are not mutually exclusive, and the encoding dimensions are orthogonal. For instance, a sink might be first `attend`ed, perhaps as a member of some non-singleton subset of a spaces sinks, and then later `deafen`ed, so that both attributes are simultaneously applied.

The `select` and `attend` attributes are denoted by characteristic features. If any avatar has been `selected`, non-`select`ed avatars are implicitly `mute`d if their self-designation state is the same as the selected avatar. For example, if a self-designated avatar is `select`ed, only other self-designated avatars are implicitly `mute`d. In the same manner, if any `attend`ed avatars exist in a given space, non-`attend`ed avatars are implicitly `deafen`ed if they are in the same class (`self` or non-`self`) as an `attend`ed avatar. Orange color is used to represent these implicit effects, implicit `mute` represented by a '-' before the mouth and implicit `deafen` by '-'s straddling the ears.

A unique feature of our system is the ability of a human pilot to delegate multiple avatars simultaneously, increasing quantity of presence; such multipresence enables one to overcome some fundamental constraints of human condition. Our interface will allow each human user to have sources and sinks present in multiple conversations at once. The ı·Con interface will allow users to multicast their voice to multiple receivers and control crowded soundscapes using narrowcasting operations. Users will be able to receive multiple audio streams and transmit their voices to multiple recipients, controlled according to custom preferences. Multiple sources are useful, for instance, in directing one's remarks to specific groups, decreasing the granularity of audibility control. Multiple sinks are useful in situations in which a common environment implies social inhibitions to rearranging shared sources like musical voices or conferees, as well as individ-

ual sessions in which spatial arrangement of sources, like the configuration of a concert orchestra, has mnemonic value.

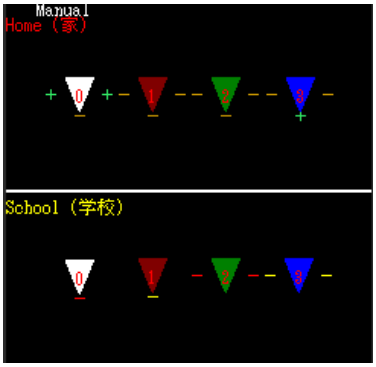


**Fig. 1.** Narrowcasting attributes graphical displays: In the Home space, avatar #0 is `attend`ed, so its complement (comprising all the other avatars) is `deafen`ed; and avatar #3 is `select`ed, so its complement is implicitly `mute`d. In the School space, avatars #0 & #2 are respectively `mute`d and `deafen`ed by self while avatars #1 & #3 are respectively `mute`d and `deafen`ed by others.

## 3 Conclusion

The basic goal of this research is to develop rich idioms for selective attention, presence, privacy, and narrowcasting for presence awareness applications for a nomadic device like a mobile phone. A multipresence feature using these idioms encourages users to install avatar representatives of themselves in several places and spaces at once. Spatial sound enhances the conferencing capabilities in collaborative virtual environments. Multipresence-enabled systems will enhance office communication and make it easier to coordinate work groups. The system also enhance human communication, allowing users to interact with friends, family, and colleagues "anytime anywhere." We expect that commercial development of this research will involve partnerships with network providers, who might license such technology to offer to their subscribers as an added value service.

## References

1. Cohen, M.: Exclude and include for audio sources and sinks: Analogs of mute and solo are deafen and attend. Presence: Teleoperators and Virtual Environments. 9(1), 84–96, ISSN 1054-7460 (2000)

2. Yamazaki, Y., Herder, J.: Exploring spatial audio conferencing functionality in multiuser virtual environments. In: 3rd International Conference on Collaborative Virtual Environments, pp. 207–208. IEEE Press, San Francisco, California, USA, ISBN 1-58113-303-0 (2001)
3. Aoki, P.M., Romaine, M.H, Szymanski, J.D., Wilson, T.D., Woodruff, A.: The mad hatters cocktail party: a social mobile audio space supporting multiple simultaneous conversations. In: Proc. of the Conference on Human Factors in Computing Systems, pp. 425–432. Ft. Lauderdale,Florida, USA, IEEE Press, ACM Press (2003)