Augmenting Human Hearing Through Interactive Auditory Mediated Reality

Evgeny Stemasov, Gabriel Haas, Michael Rietzler, Enrico Rukzio
Institute of Media Informatics
Ulm University, Ulm, Germany
<*firstname*>.<*lastname*>@uni-ulm.de

ABSTRACT
To filter and shut out an increasingly loud environment, many resort to the use of personal audio technology. They drown out unwanted sounds, by wearing headphones. This uniform interaction with all surrounding sounds can have a negative impact on social relations and situational awareness. Leveraging mediation through smarter headphones, users gain more agency over their sense of hearing: For instance by being able to selectively alter the volume and other features of specific sounds, without losing the ability to add media. In this work, we propose the vision of interactive *auditory mediated reality* (AMR). To understand users' attitude and requirements, we conducted a week-long event sampling study ($n = 12$), where users recorded and rated sources ($n = 225$) which they would like to mute, amplify or turn down. The results indicate that besides muting, a distinct, "quiet-but-audible" volume exists. It caters to two requirements at the same time: aesthetics/comfort and information acquisition.

CCS Concepts
•**Human-centered computing** → **Human computer interaction (HCI);**

Author Keywords
Auditory Mediated Reality; Hearables; Augmented Hearing

INTRODUCTION
Today's urban environments are populated by rich and dynamic soundscapes (i.e. acoustic environments [16]). Parts of these soundscapes are aesthetically pleasing to the population. Noise pollution however, is composed of unpleasant, loud or irrelevant sounds. Unlike vision, hearing is not directed and can not be averted easily. These aspects, in conjunction with the rise of ubiquitous personal audio devices, are reflected in the increasing use of headphones and similar technology. Personal audio is an interaction with the entire soundscape in a uniform fashion – mainly filtering and drowning out all sounds alike. Putting on headphones therefore drowns out every single source, independently of aesthetic and relevance. While this allows users to curate what they hear [3], personal audio impacts situational awareness [9] and has distinct health risks [4].

Apart from being merely a tool for media consumption, personal audio is an attempt to re-gain agency over one's own hearing: users attempt to filter distracting, unpleasant or irrelevant sounds by drowning them out [3]. With devices like *Hearables*, auditory perception can be mediated through a digital device. This is already implemented by some products [2, 5] and was subject of previous research [6, 11, 13]. Mediation, in turn, opens up a large design space for addition and alteration [12, 18]. Ultimately, users would re-gain full agency and control over their sense of hearing.

While this is a compelling vision, practical requirements for such devices have not been determined yet. An evident aspect is the alteration of volume or loudness of certain sound sources. In an initial step towards understanding the requirements for AMR, we conducted a week-long event sampling study with 12 participants, followed up by semi-structured interviews. They recorded and rated sources they would like to increase or decrease in volume. Based on the collected data, we were able to gain insights into relevant aspects of interactive AMR. There is a level between "muted" and "unaltered", which retains awareness and completeness of the soundscape, but caters to users' desire for silence and acoustical comfort. Furthermore,

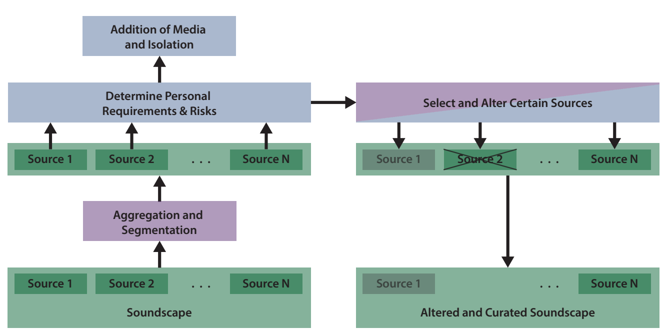

Figure 1. Schematic flow of AMR: Instead of interacting with the soundscape as a whole, users are able to curate their hearing on a *per-source-basis*, while retaining functionality of personal audio *and* their vital sense of hearing.

UIST '18 Adjunct October 14–17, 2018, Berlin, Germany

ACM ISBN 978-1-4503-5949-8/18/10.

DOI: https://doi.org/10.1145/3266037.3266104

users identify sound sources predominantly as objects, rarely as textures and never by their acoustical properties. If given the chance to change the volume of any sound source, users mostly reduce or mute and rarely increase the level. Furthermore, systems have to follow certain guidelines to ensure soundscape understanding, allow for brief interaction, and respect user's need for information via the auditory channel.

CONCEPT AND VISION

Mediated reality is the foundation for altered, modulated perception of the world. For auditory perception, it is already present in hearing aids, where it restores abilities [7]. However, it may also be used to add *new* abilities (e.g. the ability to change sound volume on a per-source basis). Deliberate, augmentative *alteration* or *modulation* was discussed far less often [18, 19] than *additive* augmentation [1, 15, 17].

Interactive AMR can provide people with agency on a more refined level. Each user is able to alter the volume and all other features of sound sources based on their personal preference. This preference consists of an individually weighted set of requirements: comfort, information acquisition, desire to isolate. Parts of this alteration may be automated, while others have to remain under the users' control.

A schematic process of altering one's own personal soundscape can be seen in figure 1. A system provides the users with a segmented soundscape, essentially defining interaction targets. Users can then change the volume of each source (increase, decrease, mute), after determining their own personal need for information and comfort. The ability to add media, as provided by personal audio devices, is kept.

PRELIMINARY STUDY

To gain better insights into user requirements in real-world environments, we created an app which allowed users to record and rate sound sources which they would like to make louder or quieter. The methodology is comparable to previous app-based research on noise-pollution [10, 14]. After receiving an introduction to the concept and the application, the participants received a week to collect data. Users were asked to record a brief sample of the sound they wanted to alter and which level they want it to be. The "desired level" ranged from 0 ("inaudible") to 10 ("as loud as possible"). Additionally, they were asked to label it and rate it on 7-point scales, adapted from [8]: *Naturalness*, *Pleasantness*, *Regularity*, *Proximity* and *Directionality*. Unlike interviews, the app-based study yielded more organic results, as users actively surveyed their context with the right priming.

The 12 participants recorded 225 sound sources. For a classification of the sources, we referred to the scheme proposed by Schafer [16]. The main category of altered sources, were mechanical sounds (56%), which were mostly reduced (18.7%) or muted (33.8%). Human sounds made up 26.7% – with 10.2% muted, 12.4% decreased and 4% increased. Natural (9.3%, with 42% increases), societal (5.8%)

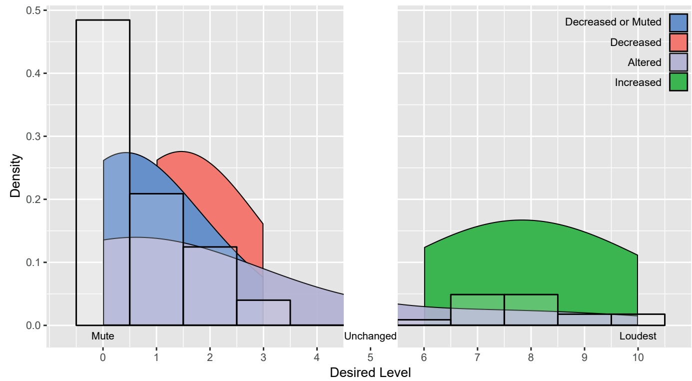

Figure 2. Histogram and distributions of the desired levels which participants indicated during the study. The scale ranged from "inaudible" to "as loud as possible".

and indicator (1.8%) sounds accounted for the remaining classes.

The following insights were gathered based on the collected data: 1) Muting sound sources as an action is not sufficient – it caters to comfort only, ignoring the informativity of a soundscape. A level which is "quiet-but-audible" likely suffices (see figure 2). 2) Users associate sound sources primarily with objects, not with physical properties and not with aggregations like "background". 3) "Pleasantness" as a subjective rating, can serve as a predictor for most cases. However "relevance" has to be taken into account, if destructive alterations like muting are used. 4) The desired level of sources may follow dynamically changing requirements, making interactivity a necessary component.

DESIGNING FOR AMR

To gain more specific requirements, we conducted interviews about the general vision of AMR directly after the study. After a thematic analysis, the following categories were developed:

- **Clear Metaphors**: When selectively altering level, objects are helpful for selection. Abstract concepts like "background" and physical concepts like frequency are less useful.
- **Transparency of Alterations**: Whenever a source is altered, the effect has to be clear to the user. This allows them to weigh advantages and risks appropriately.
- **Leveraging Human Abilities**: Humans are already able to segment a soundscape into sources. Leveraging selective attention to define filters would make AMR open to users.
- **Mediating Risks**: While interactive AMR may provide more comfort, filtering relevant sounds is a source of danger. Muting has to be treated carefully, especially if auditory comfort can be reached without total removal of sounds.
- **Interactivity**: Participants were divided concerning the modality of the interaction. However, they all highlighted that it is necessary and has to be brief and unobtrusive.

CONCLUSION & FUTURE WORK

We proposed the concept of interactive AMR for per-source alteration of volume and presented insights from a preliminary study. Paired with suitable interaction techniques, it may hand back agency over the sense of hearing to users. Current personal audio technology forces users to adapt their usage or accept consequences of it, as it implies a uniform, almost binary, alteration of the soundscape. Interactive AMR is able to bridge the gap between unobstructed hearing and mobile media consumption, without sacrificing the sense of hearing. With this augmentation, users could balance informativeness of the soundscape with their own dynamic requirements for acoustical comfort. This is reflected in the study results, where ”muting” sounds (equivalent to headphones with loud media), is the most common, but by far not the only type of alteration.

This work focused on one dimension of sound: volume. While it is the most ”graspable” one to users, aspects like spatial position, frequency and various others are relevant for future research. Additionally, specific interactions and visualisations can be implemented and evaluated with the proposed guidelines in mind.

REFERENCES

1. Benjamin B. Bederson. 1995. Audio Augmented Reality: A Prototype Automated Tour Guide. In *Conference Companion on Human Factors in Computing Systems (CHI '95)*. ACM, New York, NY, USA, 210–211. DOI: http://dx.doi.org/10.1145/223355.223526
2. BRAGI GmbH. 2018. Wireless Earphones - The Dash Pro. (2018). https://www.bragi.com/thedashpro/ (Accessed 16.08.2018).
3. Michael Bull. 2008. Sounding Out The City: Personal Stereos and The Management of Everyday Life; Music in Everyday Life. In *Sounding Out the City: Personal Stereos and the Management of Everyday Life (Materializing Culture)*. Berg Publishers, Oxford, UK.
4. Peter J. Catalano and Stephen M. Levin. 1985. Noise-Induced Hearing Loss and Portable Radios with Headphones. *International Journal of Pediatric Otorhinolaryngology* 9, 1 (June 1985), 59–67. DOI: http://dx.doi.org/10.1016/S0165-5876(85)80004-5
5. Doppler Labs Inc. 2018. Here One - Wireless Smart Earbuds. (2018). https://hereplus.me/products/here-one/ (Accessed 15.08.2018).
6. Aki Härmä, Julia Jakka, Miikka Tikander, Matti Karjalainen, Tapio Lokki, Jarmo Hiipakka, and Gaëtan Lorho. 2004. Augmented Reality Audio for Mobile and Wearable Appliances. *Journal of the Audio Engineering Society* 52, 6 (2004), 618–639. http://www.aes.org/e-lib/browse.cfm?elib=13010
7. Jochen Huber, Roy Shilkrot, Pattie Maes, and Suranga Nanayakkara. 2017. *Assistive Augmentation*. Springer, Singapore, New York, NY. https://doi.org/10.1007/978-981-10-6404-3
8. Jian Kang and Mei Zhang. 2010. Semantic Differential Analysis of the Soundscape in Urban Open Public Spaces. *Building and environment* 45, 1 (2010), 150–157. http://www.sciencedirect.com/science/article/pii/S0360132309001309
9. Richard Lichenstein, Daniel Clarence Smith, Jordan Lynne Ambrose, and Laurel Anne Moody. 2012. Headphone Use and Pedestrian Injury and Death in the United States: 2004–2011. *Injury prevention* (2012), 18:287–290. https://doi.org/10.1136/injuryprev-2011-040161
10. Nicolas Maisonneuve, Matthias Stevens, Maria E. Niessen, and Luc Steels. 2009. NoiseTube: Measuring and Mapping Noise Pollution with Mobile Phones. In *Information Technologies in Environmental Engineering*. Springer, Berlin, Heidelberg, Berlin, Heidelberg, 215–228. DOI: http://dx.doi.org/10.1007/978-3-540-88351-7_16
11. Aadil Mamuji, Roel Vertegaal, Changuk Sohn, and Daniel Cheng. 2005. Attentive Headphones: Augmenting Conversational Attention with a Real World TiVo. In *Extended Abstracts of CHI*, Vol. 5. ACM, Portland, Oregon, USA, 2223–.
12. Steve Mann. 1999. Mediated Reality. *Linux Journal* 1999, 59es (March 1999), 5. http://dl.acm.org/citation.cfm?id=327697.327702
13. Florian Mueller and Matthew Karau. 2002. Transparent Hearing. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems (CHI EA '02)*. ACM, New York, NY, USA, 730–731. DOI: http://dx.doi.org/10.1145/506443.506569
14. Rajib Rana, Chun Tung Chou, Nirupama Bulusu, Salil Kanhere, and Wen Hu. 2013. Ear-Phone: A Context-Aware Noise Mapping Using Smart Phones. *arXiv:1310.4270 [cs]* (Oct. 2013). http://arxiv.org/abs/1310.4270
15. Spencer Russell, Gershon Dublon, and Joseph A. Paradiso. 2016. HearThere: Networked Sensory Prosthetics Through Auditory Augmented Reality. In *Proceedings of the 7th Augmented Human International Conference 2016 (AH '16)*. ACM, New York, NY, USA, 20:1–20:8. DOI: http://dx.doi.org/10.1145/2875194.2875247
16. R. Murray Schafer. 1993. *The Soundscape: Our Sonic Environment and the Tuning of the World*. Destiny Books, Rochester, Vermont.
17. Christoph Stahl. 2007. The Roaring Navigator: A Group Guide for the Zoo with Shared Auditory Landmark Display. In *Proceedings of the 9th International Conference on Human Computer Interaction with Mobile Devices and Services (MobileHCI '07)*. ACM, New York, NY, USA, 383–386. DOI: http://dx.doi.org/10.1145/1377999.1378042
18. Yuichiro Takeuchi. 2010. Weightless Walls and the Future Office. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10)*. ACM, New York, NY, USA, 619–628. DOI: http://dx.doi.org/10.1145/1753326.1753419
19. Selman Yücetürk, Mohammad Obaid, and Asım Evren Yantaç. 2016. Probing Human-Soundscape Interaction Using Observational User Experience Methods. In *Proceedings of the 9th Nordic Conference on Human-Computer Interaction (NordiCHI '16)*. ACM, New York, NY, USA, 33:1–33:4. DOI: http://dx.doi.org/10.1145/2971485.2971495