Augmented Reality: Audio Applications

Kim Tarvainen
Aalto University
Master's Programme CCIS / AAT

kim.tarvainen@aalto.fi

Abstract

There are many situations in life, where seeing is not enough. One can only get a limited view of things at once. The use of augmented reality audio can give sonic cues of things that would otherwise be left unnoticed or would be noticed too late. For visually impaired individuals this technique can be especially important allowing more autonomous way of life. In this paper several applications of augmented reality audio previously presented in scientific articles are briefly introduced. It can be observed, that augmented reality audio has proven useful in many areas of life, such as navigation, guiding, games and arts.

1 Introduction

When using devices based on augmented reality (AR), events outside the display area are not necessarily registered by the user accurately or at all. For example, a quiet electric car approaching from behind is easily left unnoticed [7]. These issues can be solved with augmented reality audio (ARA). ARA can provide sonic cues of things that can't be seen either due to the things being outside the field-of-view or the user's disability. The disability can mean that the user is visually impaired, temporarily blinded by the sunlight or something else. ARA can be location- or orientation-dependent, which is the case in, for example, navigation applications [1].

In this paper current audio applications of augmented reality are represented. First, the term wearable augmented reality audio is introduced. Next, current ARA applications are represented. Then, the next chapter concentrates on the concept of Augmented Reality Audio Network. Finally, the paper is concluded with some speculation about the future of ARA.

2 Wearable Augmented Reality Audio

Closely related concepts to ARA are wearable augmented reality audio (WARA) and mobile augmented reality audio (MARA) [1]. WARA includes all the techniques that add virtual audio on top of the real acoustic environment. MARA is a special case of WARA, where particular attention is paid on the situations, where the user is moving.

WARA applications can have either localized or freely floating virtual events [1]. Localized acoustic events allow audio content to be attached to objects or places visible in the real environment. Freely floating events aren't connected to any particular objects in the surroundings. They only use the head of the user as an anchor point. They can provide different information services, such as news or entertainment, for example, in the form of music.

Many ARA applications need GPS information to get the location of the user [1]. Especially those applications that have localized acoustic events make use of the user orientation. For these cases, functioning head-tracking has to be applied.

3 Current Applications

In this section several ARA applications will be discussed. First, a platform for location-based games is introduced. Then, a navigation system made especially for visually impaired individuals is introduced. Finally, some other applications are briefly covered.

3.1 Games

Great use for MARA is found in location-based games. Traditionally, interaction with mobile devices has been based on visual means. But there's a great potential in the auditory domain. In addition to being of essential help for those who are visually impaired (this will be discussed in the next section), audio can be used to provide other content.

A system has been developed, which allows both recording and reproduction of binaural audio recorded at a certain point in time, place and orientation depending on the current location of the user [4]. This 3-D audio content can be shared with several users creating a great platform for multi-player games. These kinds of pervasive games are often played outside in urban environments.

As illustrated in Figure 1, the system is implemented so that in addition to wearing a headset, the user carries a computer, MARA mixer and a GPS device in a backpack [4]. In this particular case, a Nintendo Wii Remote is used for interaction.

Figure 1: The user of the MARA outdoor platform [4]

A structural view of the system can be seen in Figure 2 [4]. There are microphones attached to the earphones of the headset for capturing the acoustic environment binaurally. The MARA mixer is used to equalize the captured sound for proper reproduction. HRTFs are used to place virtual sound sources around the user. The GPS device is used to track the position of the user for accurate audio panning. Also orientation needs to be tracked for the same purpose. This is done with yet another device. The position and orientation data of the audio is saved as meta data in audio files.

There was an application introduced in the paper called Audiomemo [4]. It can be used to click on a certain point belonging to the path crossed on a map to hear binaural audio recorded at that location.

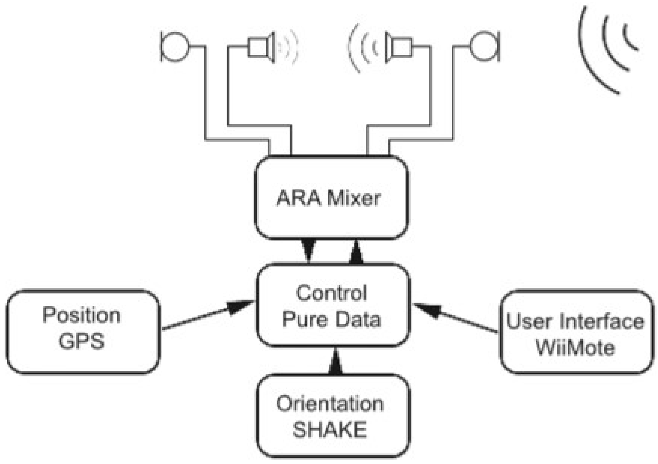

Figure 2: Structural layout of the presented MARA platform [4]

3.2 NAVIG

ARA can be especially useful for visually impaired individuals. Navigation, which is a huge challenge for these kinds of individuals, can be divided to two main components, mobility and orientation [2]. The term Micro-Navigation, a.k.a. Mobility, is used to describe the sensing of the immediate environment with the aim to get to a point in the vicinity. This includes possible obstacles, that might affect the chosen path. The other component is Macro-Navigation, a.k.a. Orientation. It comprises of having the proper orientation, selecting an appropriate path and being aware of when the destination has been reached. In the case of the visually impaired, the visual cues needed to sense the environment for successful execution of the mentioned processes are missing. This is where ARA comes in.

NAVIG is a system that greatly helps visually impaired people in autonomous navigation [2]. It uses binaural rendering to provide both spatialized text-to-speech and sonifications to guide the user. The system combines information from satellites, images etc. for high precision geolocalization. Some geolocated visual landmarks are shown in Figure 3. A fast image recognition platform and spatial audio rendering are used to make detailed trajectories. With those, the user will get macro- or micro-navigational destinations. NAVIG is meant to be used along with more traditional aids, such as a cane or a guide dog.

3.3 Other applications

Auditory Sticker is a communication application, that can be used to leave audio messages in particular locations [1]. When another person using a WARA headset gets to the location where the note was left, that person will be able to hear the message.

Figure 3: Examples of geolocated visual landmarks used for user-positioning (shop sign, facade, road sign, mailbox) [2]

Auditory Calendar is an application based on the use of freely floating audio events [1]. Calendar entries can be heard from different directions depending on the time stamp of each event.

Acu-Notch is a plugin, that creates an acoustic hole into amplitude-panned stereo signal [1]. This is useful when listening to music and something important requires user's attention. The direction of the important sound source is registered and the amplitude of the music is attenuated at that azimuth angle. This makes it possible for the user to hear the important sound clearly without having the need to pause or lower the total volume of the music. Furthermore, sounds that wouldn't normally be heard over the music, will not be left unnoticed.

WARA technology can also be integrated in voice-over-ip (VoIP) applications [1]. The operation principle of a binaural telephone is presented in Figure 4. Using a binaural telephone, the audio environment in a negotiation room can be transmitted to a user wearing a headset as if the user was sitting in the same room with the others in the far end.

Augmented sound events have many possible uses [1]. They can provide personalized advertisements, announcements, warnings, etc.

One special kind of an application is a user-adaptive audio-augmented museum guide [9]. The user is presented with a customized ARA experience based on the user's preferences, interests, current location and orientation. The interface is based on user's motion. The implementation consists of a wireless headphone with two attached antennas for localization. In the background, there are several computer

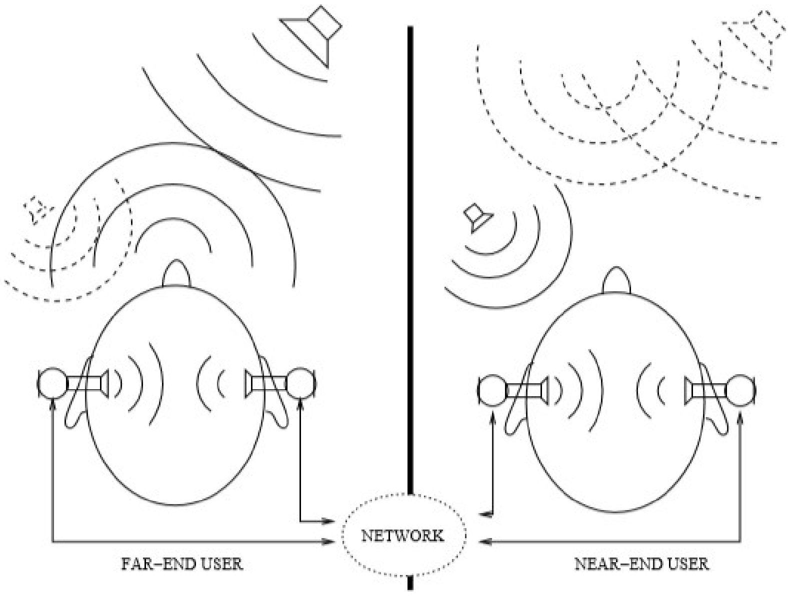

Figure 4: Binaural telephone. The far-end user's voice should be rendered outside the near-end user's head [1]

systems. One of those is responsible for rendering the transmitted audio signal into 3D surround sound. Another one is used for tracking purposes.

When used with visual AR, ARA can give important sonic cues about events that are outside the narrow field of view of a display [6]. One particular system is illustrated in Figure 5. In this 3D audio augmented reality (3DAAR) system, ARA is used to provide alerts and informational cues. The device is targeted at mobile security applications, where it helps mobile personnel to spot suspected or detected threats. It also helps in navigation.

An application integrating ARA with ecological simulation, electronic music composition and human-computer interaction has been introduced [8]. This invention set on the creative side makes it possible to create an interactive augmented reality music composition. Animal behaviors in an artificial ecosystem are sonified and spatialized and then reproduced to the listener via headphones. The output of this binaural signal depends on the listener's position, which is tracked with a Kinect sensor.

Augmented reality has found new use in orchestral conducting [5]. Conductors leading an orchestra often use a click track or a soundtrack fed through headphones. They are faced with the challenge, among other ones, of the headphones blocking the real acoustic environment. Earlier, a solution has been to take another element off from the ear, which isn't ideal. The new ARA system uses binaural rendering and head tracking to reproduce the real acoustic sound of the orchestra in addition to the other sound being played.

Figure 5: 3DAAR with head tracker, GPS and 802.11b [6]

4 Augmented Reality Audio Network

Augmented Reality Audio Network (ARAN) is an extension, which brings yet new kinds of possibilities compared to the traditional ARA framework [3]. In the traditional techniques, the real acoustic component is the one the listener would hear without any headphones. However, in ARAN, this component can be spatially expanded to include soundscapes from afar in the listener's near-field acoustic space. It is based on connecting several spatially-distributed soundscape nodes into a single acoustic field. This technique is useful for applications such as live music situations and interactive installations.

5 Conclusions

This paper took a view on the current augmented reality applications involving audio on the market and those being researched. The concept of wearable augmented reality was introduced. It usually features a headset, which is acoustically transparent and able to add virtual sound sources. There was also talk about Augmented Reality Audio Network, which expands the original acoustic space of the user with long-distance soundscapes. Applications include live music situations.

It seems that new innovations are especially made in the mobile augmented reality area. This is supported by the increased usage of different mobile devices and applications in them that could benefit from augmented reality. This has earlier happened with navigation, where spoken guidance is often provided. The use of augmented reality audio in this context is particularly useful when driving a car. It's easier to concentrate on driving when the driver is not required to look at the map all the time. Visually impaired individuals can live more autonomously with the help of smart technologies using augmented reality audio. Countless exciting possibilities lie in gaming, music and arts.

References

[1] Karjalainen, M., Lokki, T., Nironen, H., Harma, A., Savioja, L., and Vesa, S. Application scenarios of wearable and mobile augmented reality audio. In *Audio Engineering Society Convention 116* (May 2004).

[2] Katz, B. F. G., Kammoun, S., Parseihian, G., Gutierrez, O., Brilhault, A., Auvray, M., Truillet, P., Denis, M., Thorpe, S., and Jouffrais, C. Navig: augmented reality guidance system for the visually impaired. *Virtual Reality 16*, 4 (Nov 2012), 253–269.

[3] Moustakas, N., Floros, A., and Kapralos, B. An augmented reality audio live network for live electroacoustic music concerts. In *Audio Engineering Society Conference: 2016 AES International Conference on Audio for Virtual and Augmented Reality* (Sep 2016).

[4] Peltola, M., Lokki, T., and Savioja, L. Augmented reality audio for location-based games. In *Audio Engineering Society Conference: 35th International Conference: Audio for Games* (Feb 2009).

[5] Soudoplatoff, D., and Pras, A. Augmented reality to improve orchestra conductors' headphone monitoring. In *Audio Engineering Society Convention 142* (May 2017).

[6] Sundareswaran, V., Wang, K., Chen, S., Behringer, R., McGee, J., Tam, C., and Zahorik, P. 3d audio augmented reality: implementation and experiments. In *The Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003. Proceedings.* (Oct 2003), pp. 296–297.

[7] Valimaki, V., Franck, A., Ramo, J., Gamper, H., and Savioja, L. Assisted listening using a headset: Enhancing audio perception in real, augmented, and virtual environments. *IEEE Signal Processing Magazine 32*, 2 (March 2015), 92–99.

[8] Özcan, Z., and Çamci, A. An augmented reality music composition based on the sonification of animal behavior. In *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality* (Aug 2018).

[9] Zimmermann, A., and Lorenz, A. Listen: a user-adaptive audio-augmented museum guide. *User Modeling and User-Adapted Interaction 18*, 5 (Nov 2008), 389–416.