AN AUGMENTED REALITY AUDIO HEADSET

Miikka Tikander, Matti Karjalainen, and Ville Riikonen

Department of Signal Processing and Acoustics,
Helsinki University of Technology, TKK, Espoo, Finland
firstname.lastname@tkk.fi

ABSTRACT

Augmented reality audio (ARA) means mixing the natural sound environment with artificially created sound scenes. This requires that the perception of natural environment has to be preserved as well as possible, unless some modification to it is desired. A basic ARA headset consists of binaural microphones, an amplifier/mixer, and earphones feeding sound to the ear canals. All these components more or less change the perceived sound scene. In this paper we describe an ARA headset, equalization of its response, and particularly the results of a usability study. The usability was tested by subjects wearing the headset for relatively long periods in different environments of their everyday-life conditions. The goal was to find out what works well and what are the problems in lengthened use. It was found that acoustically the headset worked fine in most occasions when equalized individually or generically (averaged over several subjects). The main problems of usage were related to handling inconveniences and special environments.

1. INTRODUCTION

In augmented reality audio (ARA), natural surrounding sound environment is enriched with virtual sounds, played to the user with a special ARA headset, which consists of a pair of headphones with integrated microphones. In normal usage, the microphone signals are transmitted to the earphones, exposing users to natural surrounding sounds. To separate the situation from listening without a headset, this is called pseudo-acoustics [1]. Ideally, the ARA headset should be acoustically transparent, with no difference between pseudo-acoustics and listening to environment without a headset.

For creating a realistic augmented reality audio environment, different kinds of external hardware is required. Fig. 1 shows an example of a complete ARA system. The lower part includes blocks for creating and rendering virtual sounds to a user. One of the key concepts is position data for keeping virtual audio objects in place while the user moves. The upper block is for transmitting the binaural signals, e.g., for communication purposes. The headset and the ARA mixer in the middle are the user-worn parts. The mixer takes care of routing and mixing all the signals involved in the system. The headset part is similar to common headphones that are widely used with portable players, with added microphones.

Ideally, a user should not have to wear any extra hardware for ARA applications, just a pair of wireless headsets including all the necessary hardware. Unfortunately, with current technology this is not yet possible. However, hardware can already be miniaturized enough to make practically portable ARA devices.

Although many people are fairly used to wear earplugs or headphones throughout the day in everyday-life situations, there are still very few studies on how people would perceive and accept an ARA headset when worn for longer periods of time in everyday-life situations. People wearing hearing aids naturally have experience in listening to surrounding sounds through a "headset". However, there the primary design target is to provide enough sound pressure to make sounds audible and to maximize speech clarity in front of the user, whereas for an ARA headset the primary design target is to keep the sound environment as unaltered as possible. Of course, an ARA headset can be set to fit many purposes, including usage as a hearing aid.

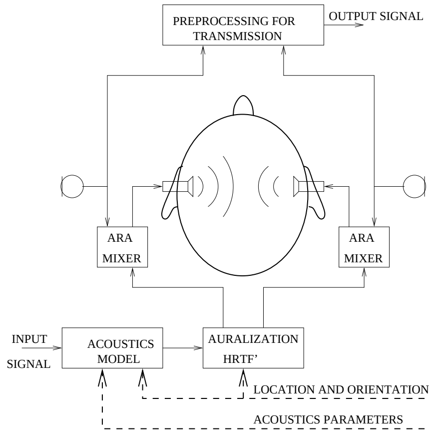

Figure 1: ARA system diagram (after [1]).

One of the key factors in acceptability of an ARA headset is sound quality. If people are expected to wear a headset for longer periods of time, the pseudo-acoustics should not differ from natural acoustics too much. Therefore, the ARA mixer combined with the headset should be equalized to provide a transparent illustration of the surrounding sounds [2].

Another usability issue comes from the hardware portability. Nowadays people are fairly used to carry small mobile items like mobile phones, mp3-players, and PDAs. Also, many people listen to music daily for long periods of time, thus carrying and wearing a headset should not be a problem. If the usability of the headset is acceptable, the headset would offer a practical platform for all kinds of applications [3] from work to entertainment. In general the headset could be used as an all-around audio interface for personal appliances, as well as for other services.

Among the most potential ARA applications are full audio quality binaural telephony and audio meetings with distant subjects panned around the speaker, information services and object browsing based on position and orientation information (such as in stores product information could be given directly to the headset), virtual tourist guide giving location-dependent object information by voice, audio memos and contact management, as well as audio-based games and making of music. The pseudo-acoustic environment can also be processed in many ways such as adding audio effects, speech enhancement, noise cancellation, hearing protection, and hearing aid functions.

Figure 2: Left: Headset used in the evaluation. Microphone is on the left and headphone outlet on the right. Middle: Headset fitted in ear. Right: Prototype ARA mixer.

Although multimodal communication is in most cases preferred over a single modality, one of the strengths of ARA techniques is that they can be used practically anywhere, anytime, hands-free, and eyes-free.

This paper introduces results of a pilot study on the usability of an ARA headset when used for longer periods of time in everyday-life situations. In addition to general usability, the goal was also to find out different usability aspects of such headsets. A group of four subjects wore an ARA headset in everyday-life conditions and reported the observations in a diary.

2. ARA MIXER AND HEADSET

An ARA mixer and headset was constructed and evaluated from the point of view of pseudo-acoustic sound quality and usability in practice. The headset was constructed from a noise-cancelling headphone (Philips SHN2500) that contains insert type earphones and electret microphones integrated together as shown in Fig. 2. The original functionality was changed by replacing the external electronics box for noise cancellation by an ARA mixer designed in the project, shown also in Fig. 2.

The earphones fit quite tightly to the ear canal entrance, while the microphones remain about 1 cm out from that point, which could mean degradation of spatial perception and inconvenience of using a telephone. The microphones have also some directivity at highest frequencies, which means that sound coloration is dependent on sound source direction. The main technical problem was, however, the acoustic effects inside the ear canal and how to compensate them as well as acoustic leakage of external sound to the ear canal.

2.1. Coloration of Pseudo-acoustics due to the ARA Headset

The external ear modifies the sound field in many ways while transmitting sound waves through the ear canal opening to the ear drum. Normally, the ear canal is open and acts as a quarter wavelength resonator with one end being closed by the eardrum and the other end open to the air. For an open ear canal, the first ear canal resonance occurs at around 2-4 kHz depending on the length of the canal [4]. When a headphone blocks the ear canal, this resonance disappears, and the sound field is perceived unnatural. In this case, the ear canal is closed from both ends and it starts to act more like a half-wavelength resonator [5]. The lowest resonance now occurs at around 5-10 kHz depending on the length of the ear canal and fitting of the earplug.

In order to make the headset acoustically more transparent, equalization is needed to recreate the quarter-wave resonance and to damp the half-wave resonance. Also the frequency response of the headset causes coloration, which has to be equalized.

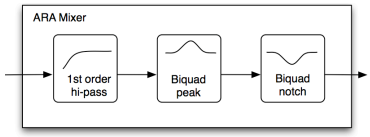

Figure 3: Filter sections in the equalizer.

The headset (especially an insert-type) attenuates the mid-range and high frequencies coming from outside quite efficiently. However, there is always some leakage through the headset and also between the skin and the cushion of the headphone [6]. Low frequencies can leak through the headphone quite effectively. The leaking from the real environment sums up in the ear canal with the pseudo-acoustic representation produced by the transducer. This summing causes coloration especially at low frequencies and deteriorates the pseudo-acoustic experience [1]. The amplification of low frequencies has to be equalized.

2.2. Equalization Properties of the ARA Mixer

The ARA mixer of this study is presented in more detail in [2]. The mixer includes a mixing section for transmitting the microphone signals to the earphones, and also for mixing external sound sources to the pseudo-acoustic environment. The mixer includes also an adjustable equalization section to make the pseudo-acoustics sound as natural as possible. For lowest latency possible the mixer was constructed with analog electronics. This is important since the addition of the acoustic leakage and the delayed pseudo-acoustic sound creates a comb filtering effect resulting in coloration. Even a fraction of millisecond of latency in processing can be disturbing. Typical digital audio hardware and software can make several milliseconds of latency, and are therefore not suitable for the task.

The equalizer section has two parametric resonators and a high-pass filter. The resonators can be used to recreate the missing quarter-wave resonance, and to damp the added half-wave resonance. The high-pass filter is used to compensate for boosted low frequency reproduction due to sound leakage between the headphone and skin. Figure 3 shows a block diagram for the equalization section of the mixer.

During the evaluation described in Section 3, two of the testees used individually tuned equalization, and the other two used generic equalization. For individual equalization the headset responses were measured in an anechoic chamber. The transfer function from a loudspeaker in front of the testee into the ear canal of the testee was first measured without a headset, and then the measurement was repeated with a headset on. The lower plot in Fig. 4 shows the measurement results without a headset (black line) and with a headset (grey line). There is no equalization used in the headset in this case. The measurement data clearly demonstrate the quarter wave resonance around 2.5 kHz when there is no headset, and how it disappears when the headset is put on. Also, pronounced low frequencies and the added half-wave resonance at around 8 kHz are clearly visible. The upper plot in Fig. 4 shows measurement data for the same subject when the equalization is accurately tuned and switched on. For the generic (non-individual) equalization, shown in Fig. 5, the equalization curve was computed as the average of four different individual measurement data.

It should be noted that due to the simple implementation of the equalization circuit, even the individual equalization setting differ noticeably from the measured curve. Therefore, the individual equalizer setting is still an approximation of the real measured curve. A small-scale quality evaluation with four listeners was

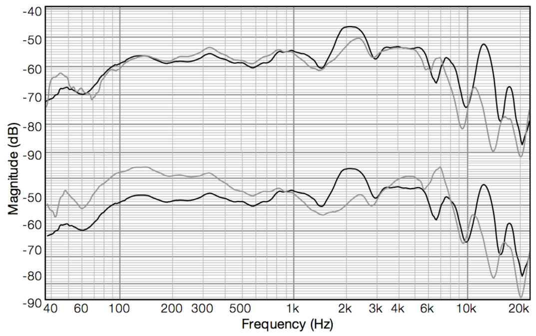

Figure 4: A transfer function from a loudspeaker into the ear without a headset (black line in both plot pairs), with an equalized headset (grey line in upper plot pair), and with an unequalized headset (gray line in lower plot pair).

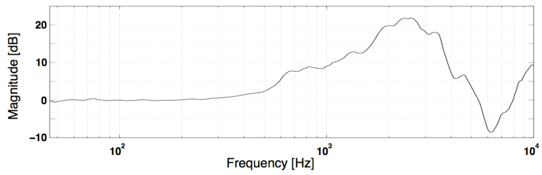

Figure 5: Generic target equalization curve for ARA headset [2].

conducted, in which the subjects compared real acoustic sound, the headset without equalization, the headset with generic equalization, and with individual equalization [2]. There was a clear quality difference between the unequalized and any equalized case. In some cases the lack of equalization accuracy caused that the generic equalization curve sounded better than the individual one, or at least, the difference was very small [2].

3. USABILITY EVALUATION

The usability of the ARA headset was evaluated in real-life situations with a set of subjects. A group of four testees where given an ARA headset and mixer and they were instructed to wear the headset for long periods in everyday-life situations. The main question for the testees was: *"Would you be willing to wear this kind of a headset in everyday-life situations?"*.

The testees were instructed to keep a diary and mark down all observations they find concerning the usage of the headset. Additionally, the test subjects were given a list of guidance questions to ease the diary keeping. Between testees, the evaluation period lasted from four days to a week, and the total time of usage of the headset varied between 20 - 35 hours. The continuous usage periods varied from few minutes to eight hours. The test subjects wore the headset in many different situations. Situations included for example dining, conversations, walking, light exercises, driving, public transportation, listening to music, watching movies, using mobile phones, listening to lectures, and so on.

4. EVALUATION RESULTS

The overall comments from the users were very positive. By sound quality the headset was found very good and applicable for most cases of practical everyday-life situations. The most criticism arouse from the wires that transmitted mechanical noise to the headsets, and also the wires tend to get stuck in clothes. The following chapters summarize and comment the diary reports from the testees. A more compact summary of the diaries is shown in Table 1.

4.1. Audio quality

Audio quality of the headset was found very good. Sense of space and directional hearing was reported to be very close to natural hearing. The sound color of pseudo-acoustics was found sufficiently good for most practical everyday-life situations. The main complaint was that the higher frequency range was slightly attenuated. For critical listening situation, e.g., listening to music, this might be slightly annoying.

It was surprising how well sound sources were externalized also in front of the subject, because this is very difficult in HRTF-based auralization. The reason to very natural externalization in front is probably the co-operation of hearing and vision, as it happens in normal life. When binaural signals were recorded and listened to later without visual information, frontal externalization was not as good anymore.

There was some audible noise present from the electret microphones. At first it was noticeable but the testees reported to get used to it very fast, and after a while the noise was not perceived. When listening to quiet sounds in a silent environment the noise could be heard. This phenomenon can be considered as a slightly raised hearing threshold due to the noise. In any case, whenever the noise was audible, it was not reported to be annoying.

One test subject reported that after a long wearing period the sound color of the natural environment was forgotten, and the pseudo-acoustics felt natural. Though, removing the headset revealed all the possible artifacts in pseudo-acoustics, mainly inherent (and mostly inaudible) noise, and spectral colorations.

4.2. Ergonomics

Some critique emerged about the handling noise and limitations caused by the wires of the headset. When they made contact with objects (for example the collar of the shirt or the zip of a jacket) the noise was mechanically transmitted to the headset. The wires tend to get stuck and limit the head motion. One way to overcome this problem would be to use softer and more flexible wires. Bluetooth or other wireless techniques could be used to route the signals between the headset and the mixer, thus eliminating the wire. Another issue for all test subjects was the unnecessarily large size of the ARA mixer, which was found cumbersome to carry. However, the electronic circuits in the mixer could be easily fitted in a smaller circuit layout, and thus the problem would be easily fixed.

Placing and removing the headset was found easy. Wearing the headset did not cause ear ache even after hours of usage. Only one testee, with smaller ear canals, had troubles fitting the headset, and also experienced ear ache while wearing the headset.

The headset extended slightly outside the ear canal. This had effect on using a mobile phone. All the testees reported that using a mobile phone was a little troublesome. The mobile phone touching the headset created unpleasant clicks in the headset. Furthermore, due to the headset the mobile phone could not be properly coupled to the ear and this resulted in a thin sound in the mobile phone. A smaller and deeply inserted headset might solve this problem. Another, and more preferred solution would be to use the headset as a hands-free device for the mobile phone.

4.3. Communications

All of the testees reported that they had no problems in normal conversation situations. For speech the sound quality was very

Table 1: *Summary of four diarys. ++++ = all testees found an attribute working, +++ - = all but one testee found the attribute working,... - - - - = all testees found the attribute not working. Please, refer to text for more detailed explanation on results.*

++++	+++-	++ - -	+ - - -	- - - -
Directional hearing, Spatial impression, Speech intelligibility, Walking	Sound quality, Wind, Inherent noise, Ease of fitting, Ear ache, Watching TV, Concert, Traffic noise, Listening to music	Distortion (impulses)	Distortion (continuous)	Talking to a mobile Running/jumping Strong wind Size of the mixer

good. One of the big complaints was that the users own voice sounded very boomy and the sound was localized inside the head. However, most of the test subjects mentioned to got accustomed to this fairly fast and, in the end, this was not found annoying. One test subject reported that sometimes he had trouble determining the level of his own voice.

Wearing an ARA headset in some situation where it is not appropriate or polite to wear headphones, e.g., lectures and conversations, created some social uncomfort. Of course, if other people would know that the headset transmits the sound unaltered, then this would be no problem.

4.4. Other observations

Eating and drinking was reported to be one of the most irritating situations with the headset. Also, other similar situations like washing teeth was found unpleasant.

Loud bangs, e.g. from doors, overloaded either the microphones or the amplifier resulting in distorted sound in the headset. Some of the test subjects found this very irritating while others less annoying. Listening to continuous loud sounds, e.g., music, resulted also in distorted sound, and was reported to be annoying. Distortion performance could be improved by better electric design and by choosing mechanically less sensitive microphones. The connectors of the wires could cause undesired scratchy noises when there were contact problems due to movement of the wires.

At first, walking sounded uncomfortable due to boomy sound from the foot step but test subjects reported to get accustomed to this fairly fast. However, running and jumping was still reported to be unpleasant due to boomy foot step sounds. Also, heavy breathing, e.g., after some exercise, sounded boosted and unpleasant. In a very windy situation the microphones got mechanically overloaded and this resulted in unpleasant sound in the headset. Mild wind was not found a problem.

Some test subjects tried listening to music by using the ARA mixer to mix music to pseudo-acoustics. Just like listening to conventional headphone, the music was located inside the head. However, this was found positive as it separated the music from the pseudo-acoucstics. Also, the ability to hear surrounding sound was found useful.

4.5. Feature suggestions

The most desired lacking feature was the ability to adjust the level of the pseudo-acoustics, either boost or attenuate. However, the system should have some sort of automatic gain control to prevent hearing damage due to surrounding loud sounds, and also due to user's own sound sources.

Another suggested feature was the ability to use the headset as a hands-free device for a mobile phone. In addition the headset could function as an audio interface for any audio device. The connection could be offered wirelessly, e.g., with bluetooth.

As for the main question whether the test subjects would be willing to use the headset in everyday-life situation the answer was mainly positive. The testees were willing to use the headset, even for longer periods, in the everyday-life situations, if the headset would offer some added value or other advance.

5. CONCLUSIONS AND DISCUSSION

As a conclusion it seems that a generically equalized ARA headset is applicable in everyday-life situations. The sound quality is sufficient and the usability was good enough in most cases of practical situations. However, the pseudo-acoustics as such did not bring any added value to listening to surrounding sounds. If there was any extra advantage on using the headset, the test group would be willing to use the headset, even for longer periods of time.

An ARA headset is naturally designed for ARA applications. Therefore, the real usability, or rather, usefulness of the headset comes with the applications. There are many application areas where it is essential that a user's perception of surrounding sounds should not be interfered. For example, while walking in the streets with an augmented audio guide, for sake of safety, a user should be well aware of surrounding sounds. Analysis of surrounding sound environment could be one attractive added value application for the ARA headset. The system would continuously analyze the microphone signals and, e.g., warn of approaching vehicles, or based on user preferences give information on surrounding environment.

The number of testees in the evaluation was fairly small and an evaluation with a bigger test group would be required for more thorough analysis on the subject. Based on the findings and suggestions from this evaluation we are planning a larger scale usability evaluation of ARA-headsets.

6. ACKNOWLEDGMENTS

This work has been supported by the Academy of Finland (project no. 106956) and the Nokia Research Center as part of KAMARA+ project (*Killer Applications of Mobile Augmented Reality Audio*).

7. REFERENCES

[1] A. Härmä et. al., "Augmented reality audio for mobile and wearable appliances," *J. Audio Eng. Soc.*, vol. 52, no. 6, pp. 618–639, June 2004.

[2] V. Riikonen, M. Tikander, and M. Karjalainen, "An augmented reality audio mixer and equalizer," in *124th AES Convention*, Amsterdam, May 17-20 2008.

[3] T. Lokki et. al., "Application scenarios of wearable and mobile augmented reality audio," in *116th AES Convention*, Berlin, Germany, May 8-11 2004.

[4] C. Poldy, "Headphone fundamentals - Tutorial," in *120th AES Convention*, Paris, France, May 20-23 2006.

[5] H. Møller et.al., "Design criteria for headphones," *J. Audio Eng. Soc.*, vol. 43, no. 4, pp. 218–232, 1995.

[6] M. Tikander, "Modeling the attenuation of a loosely-fit insert headphone for augmented reality audio," in *in Proc. AES 30th Int. Conf.*, Saariselkä, Finland, Macrh 15-17 2007.