Usability Issues in Listening to Natural Sounds with an Augmented Reality Audio Headset*

MIIKKA TIKANDER, *AES Member*
(miikka.tikander@hut.fi)

Helsinki University of Technology, Department of Signal Processing and Acoustics, Espoo, Finland

An augmented reality audio (ARA) headset is described with regard to the equalization of its response and, in particular, 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 as well as any problems with extended use. Furthermore it was tested how, during the test period, users would adapt to certain practical activities and to the electrically transmitted and reproduced sound environment. It was found that acoustically the headset worked well for most occasions. The main problems with usage were related to boosted user sounds due to occlusion effects, handling inconveniences, and mechanical noise from the wires. To some degree the test subjects adapted to the slightly altered sound environment, whereas adaptation to physical activities was very small or insignificant.

0 INTRODUCTION

In augmented reality audio (ARA) the natural surrounding sound environment is enriched with interactive virtual sounds. Virtual sounds can be played to the user with a special ARA headset, which consists of a pair of earphones with integrated microphones. In normal usage the microphone signals are transmitted to the earphones, exposing the user to natural surrounding sounds. To distinguish this situation from the normal situation of listening without a headset, this is called pseudoacoustics [1]. Ideally the ARA headset should be acoustically transparent, with no difference between pseudoacoustics and listening to the environment without a headset. Such a headset is seen as a basic component of future personal audio and voice communications.

For creating realistic augmented reality audio, different kinds of external hardware are required. Fig. 1 shows an example of a complete ARA system. The lower part includes blocks for creating virtual sounds and rendering them to a user. One of the key concepts is location and orientation data for keeping virtual audio objects in place while the user moves. The upper block is for transmitting the binaural microphone signals, such as 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 earphones, which 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 wireless headset that includes all the necessary hardware. Unfortunately with current technology this is not yet possible in practice. However, hardware can already be miniaturized enough to make practically portable ARA devices.

Although many people are fairly used to wearing earplugs or headphones throughout the day, 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. Based on a prestudy [2], the overall usability of the headset used in this study was found acceptable on most occasions. People wearing hearing aids naturally have experience in listening to surrounding sounds through a "headset." However, with hearing aids 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 natural sound environment as unaltered as possible. Of course an ARA headset can be set to fit many purposes, including use as hearing aid or hearing protector. Furthermore, as found in the prestudy, good sound quality as such, without any applications, does not motivate people to wear the headset. With hearing aids the useful application is built in, and thus people are motivated to wear the hearing aid, even if the sound quality or user experience is far from optimal.

Among the most promising 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, where

*Manuscript received 2008 December 18; revised 2009 May 17.

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 music [3]. The pseudoacoustic environment can also be processed in many ways, such as adding audio effects, speech enhancement, noise cancellation, hearing protection, and hearing-aid functions.

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

One of the key factors in the acceptability of an ARA headset is sound quality. If people are expected to wear a headset for longer periods of time, especially in everyday-life situations, the pseudoacoustics should not differ too much from natural acoustics. Therefore the headset should be equalized to provide a transparent reproduction of the surrounding sounds [4], [5].

In addition to tonal transparency, spatial hearing should be preserved as close to normal as possible. This means that the headset should provide the necessary spatial information to the user. A common way to construct an ARA headset is to place microphones outside the earphones. If the microphones can be placed close to the ear canal entrance, all the necessary spatial information can be provided to the user [6], [7].

Another usability issue comes from the requirement for hardware portability. Wearing an ARA system should not interfere with normal life practices. However, nowadays people are fairly used to carry small mobile items such as 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 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.

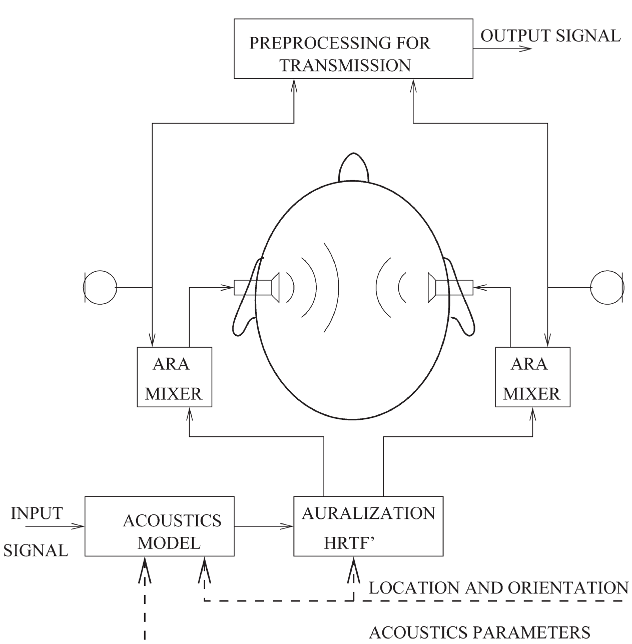


Fig. 1. ARA system diagram [1].

Building an ideal ARA system is a great challenge, and therefore a practical ARA system will be more a less compromised in some areas. The headset used in the study closes the ear canal fairly tightly, and this results in an occlusion effect where all internal sounds (such as own voice, biting, foot steps) are boosted in the ear canal. This is the first, and in most cases the most annoying, effect that is noticed when the headset is worn for the first time. One interesting question is how well, and in what conditions, users adapt to a slightly degraded version of natural acoustics, and to other issues arising from wearing the headset.

This paper presents the results of a study on the usability of an ARA headset when worn for longer periods of time in everyday-life situations. In addition to general usability, the adaptation to certain practical activities and to sound quality of pseudoacoustics was tested. A group of ten subjects wore an ARA headset in everyday-life conditions and reported the observations in a diary. Annoyance of eating, drinking, speaking, and chewing a bubble gum with the headset on was evaluated, as well as the overall sound quality of the headset. The evaluation was performed before and after the field test period to investigate whether users adapt to any of the actions tested, or to the sound quality of the headset. In addition to the particular ARA headset, the results are interesting from the point of view of wearing any insert-type earphones.

1 ARA MIXER AND HEADSET

The ARA mixer and the headset used in the evaluation were constructed and evaluated from the point of view of pseudoacoustic sound quality and usability in practice [4]. The headset was constructed from a noise-canceling 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, also shown in Fig. 2.

The earphones fit quite tightly to the ear canal entrance, while the microphones protrude about 10 mm from that point, with potential degradation of spatial perception and inconvenience when using a telephone. The microphones have also some directivity at the highest frequencies, which means that sound coloration is dependent on the sound source direction. The main technical problem was, however, the acoustic effects inside the ear canal and how to compensate for this, as well as acoustic leakage of external sound to the ear canal.

1.1 Coloration of Pseudoacoustics Due to 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. When an earphone blocks the ear canal, this resonance disappears and, as such, the sound field is perceived as unnatural. In this case the ear canal is closed from both ends, and it starts to act more like a half-wavelength resonator. 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 dampen the half-wave resonance. Also the frequency response of the headset causes coloration, which has to be equalized.

The insert earphone used in the study attenuates the surrounding sounds quite efficiently in the mid and high frequencies. However, there is always some leakage through the headset and also between the skin and the cushion of the earphone [8]. Low frequencies can leak through the earphone quite effectively. The leaking from the real environment sums up in the ear canal with the pseudoacoustic representation produced by the transducer. This summing causes coloration, especially at low frequencies, and deteriorates the pseudoacoustic experience [1]. The amplification of low frequencies has to be equalized, as well as other spectral colorations caused by the headset design itself.

1.2 Equalization Properties of ARA Mixer

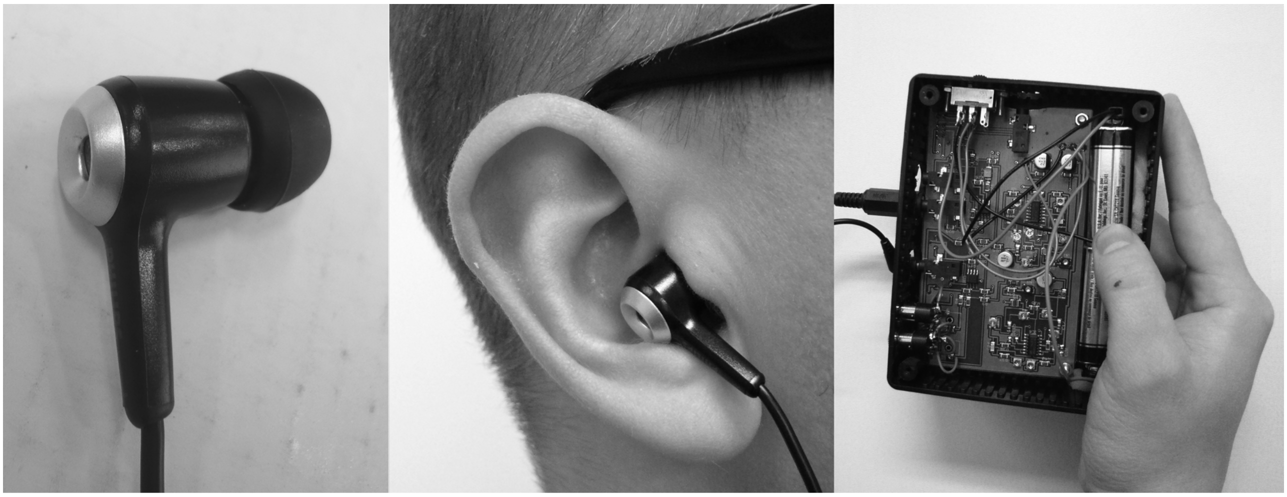
The ARA mixer of this study is presented in more detail in [4], [9]. The mixer includes a mixing section for transmitting the microphone signals to the earphones, and also for mixing external sound sources to the pseudoacoustic environment. It includes also an adjustable equalization section to make the pseudoacoustics sound as natural as possible. To avoid latency the mixer was constructed with analog electronics. This is important since the leakage happens without delay, and any delay in the electrically transmitted sound would create a comb-filter effect. Typical digital audio hardware and software can make several milliseconds of latency, and therefore is not suitable for the task.

The equalization section of the mixer is shown in Fig. 3. The equalizer has two parametric resonators and a high-pass filter. The resonators can be used to recreate the missing quarter-wave resonance, and to dampen the added half-wave resonance. The high-pass filter is used to compensate for boosted low-frequency reproduction. Low frequencies are transmitted to the ear canal electronically via the microphone and also as leakage between the headset enclosure and skin. Furthermore the prototype headset itself has a slightly boosted low-frequency response, which must be compensated for as well.

The equalization target curve for the headset was measured as follows. First a transfer function from a loudspeaker in front of a subject into the ear canal (15 mm from the entrance) of the subject was measured without a headset using a miniature microphone (upper and lower dark lines in Fig. 4). Next the measurement was repeated with the headset on and without any equalization (lower gray line in Fig. 4). The difference of the measured responses was used as target equalization curve. The average of measurements on four subjects was used to create a generic equalization curve target, which is shown in Fig. 5. The equalization curve was found to work well in the prestudy [2], and therefore it was chosen for this study as well. Another option would have been to use individually tuned equalization, but because of the simplicity of the equalization circuit there was no guarantee that it would work any better when compared to generic equalization. Also, from a practical point of view it was more interesting to see how generic equalization would perform for a larger group of people.

2 USABILITY EVALUATION

The usability of the prototype headset was tested in real-life situations. The evaluation consisted of two parts,



(a) (b) (c)

Fig. 2. (a) Headset used in evaluation. Microphone is at left and earphone outlet at right. (b) Headset fitted in ear. (c) Prototype ARA mixer.

the actual field test, where the test subjects wore the headset in everyday-life situations, and a sound quality and adaptation test, where the overall sound quality and the usability of the headset in certain situations were tested before and after the field test evaluation period.

2.1 Participants

A group of ten subjects (nine male, one female) took part in the experiment. None of the participants reported having any hearing-related problems. The whole group had a technical background, and they had experience in evaluating sound quality. Every test subject had a positive attitude on new technical devices, including the prototype headset.

One of the main goals of the study was to evaluate the practical usability of this kind of headset, and it was important that the test subjects were willing to wear (tolerate) and evaluate the headset in situations where it did not perform optimally. It should be noted, however, that the device is designed to be used by ordinary users, and in all kinds of practical real-life situations. In this respect a more naive test group would have been justified. However, the current prototype is not portable nor practical enough for a naive test group, and the impracticalities of the device could have drawn too much attention from the actual usage of the device.

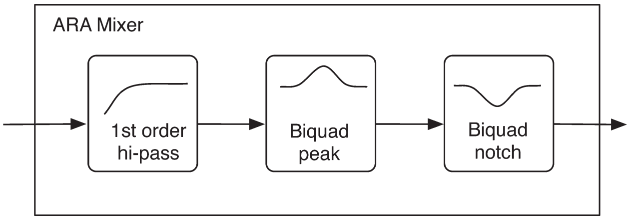


Fig. 3. Filter sections in equalizer.

2.2 Field Test

The field test part of the usability evaluation was performed in real-life situations. The subjects were given an ARA headset and a mixer, and they were instructed to wear the headset for long periods in their daily routines. The main task for the subjects was to wear the headsets in as many practical situations as possible, and to write down all their observations in a diary. A guiding question for the subjects to keep in mind during the test was, "Would you be willing to wear this kind of headset in real-life situations?" Furthermore the test subjects were encouraged to find and try all kinds of new approaches to using the device.

For the subjects the evaluation period lasted from four days to a week, and the total usage of the headset varied between 20 and 40 hours. The continuous usage periods, within a test period, varied from a few minutes to eight hours. The test subjects wore the headsets in many different situations, such as dining, conversations, walking, light exercises, driving, public transportation, listening to music, watching movies, using mobile phones, and listening to lectures.

2.3 Annoyance Factors

Before and after the field test period the test subjects evaluated the overall sound quality of the headsets, and also the usability of the headsets for certain practical activities. Prior to the first evaluation the subjects were given some time to get used to wearing the headsets and to pseudoacoustics as well. During this trial period subjects were given information and instructions on the forthcoming evaluation.

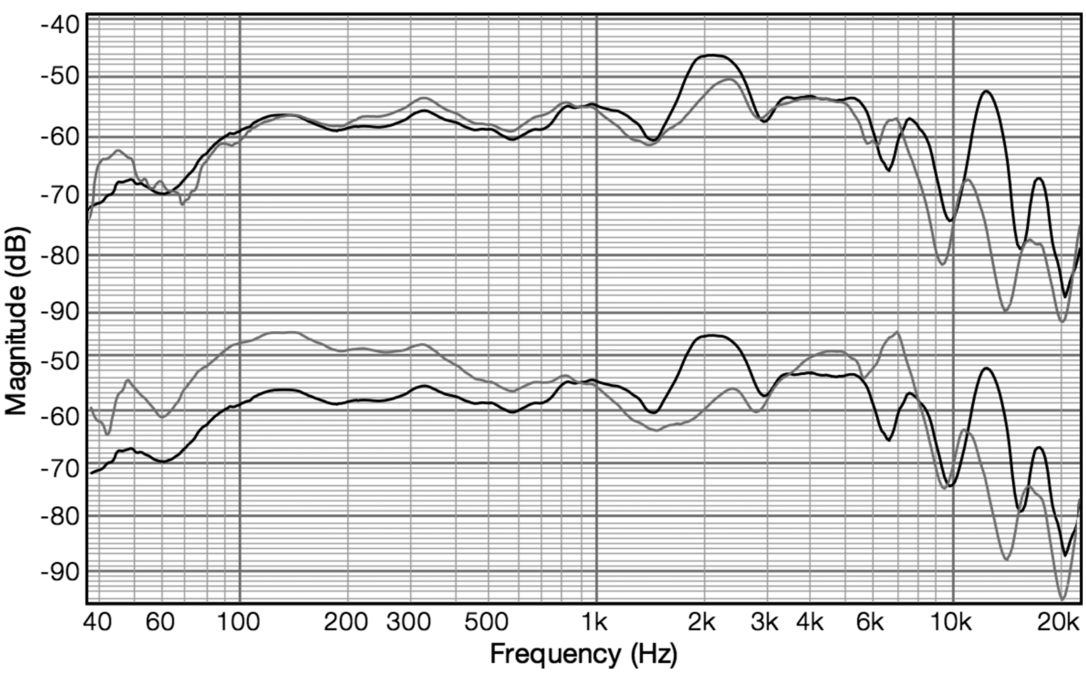


Fig. 4. Transfer functions from loudspeaker into ear without a headset (black lines), with an equalized headset (grey line in upper plot), and with an unequalized headset (gray line in lower plot).

The usability test consisted of four tasks:

T1) Reading text out loud
T2) Eating crisp bread
T3) Drinking water
T4) Chewing bubble gum.

For the text reading the test subject was given a sheet of text, which he or she had to read out loud paragraph by paragraph. Each paragraph (about 35 words) was read twice, first with the headset on and then without the headset. The subjects were free to read as long as was necessary to evaluate the annoyance of the headset in this task. The annoyance of reading the text with the headset on compared to reading without the headset was evaluated with the following mean opinion score (MOS) scale:

5 Imperceptible
4 Perceptible, but not annoying
3 Slightly annoying
2 Annoying
1 Very annoying.

The grading was done with pen and paper by marking the grade on a continuous scale from 1 to 5.

In crisp bread eating the subject was evaluating the annoyance of eating with a headset on compared to eating without the headset. The subjects were allowed to eat as much as was necessary to form an opinion. Water drinking and bubble gum chewing cases were evaluated with a similar procedure. The subjects were instructed to grade the annoyance as if they had to perform the activities in every-day life situations.

2.4 Sound Quality

The sound quality of the pseudoacoustics was evaluated with real-world sounds (speech and finger snaps) and with two different stereo recordings in a standard listening room. Subjects were asked to evaluate the following attributes (using the same MOS scale as before):

A1) Spatial impression
A2) Timbre
A3) Location accuracy.

Spatial impression was described to include the overall sense of space, especially room size and reverberation. Timbre was instructed to mean tonal coloration, and location accuracy was described to measure spatial hearing, including sound source direction and distance.

The sound quality evaluation started with a human voice. The experimenter spoke out loud in different locations in the room, and the subject was instructed to listen without the headset and with the headset on. The subject was allowed to ask the experimenter to move to different positions, if needed. Next the experimenter made finger snaps in different locations in the room. The locations are marked with small circles in Fig. 6. In addition the subject could ask the experimenter to snap fingers in other positions as well. After this the subject listened to two different music sound samples, played through a stereo loudspeaker setup, as shown in Fig. 6. The subject was free to listen to sound samples as long as needed to form an opinion. Finally, based on the whole test sound material, the subject evaluated the sound quality, points A1), A2), and A3) in the list, of the pseudoacoustics, compared to natural acoustics. The grading was done by using the same MOS scale as was used with the usability evaluations.

2.5 Adaptation Evaluation after Field Test Period

The sound quality and adaptation evaluation was repeated after the field test period. The subjects were instructed to wear the headset continuously for at least 1.5 hours before performing the second evaluation test. This time the test subject had already become adapted to the pseudoacoustics and forgotten the natural acoustic reference, at least to some degree.

The evaluation was first done with no comparison to an unoccluded (open ear) case, thus the subjects did not remove the headset during the test. The grading was done

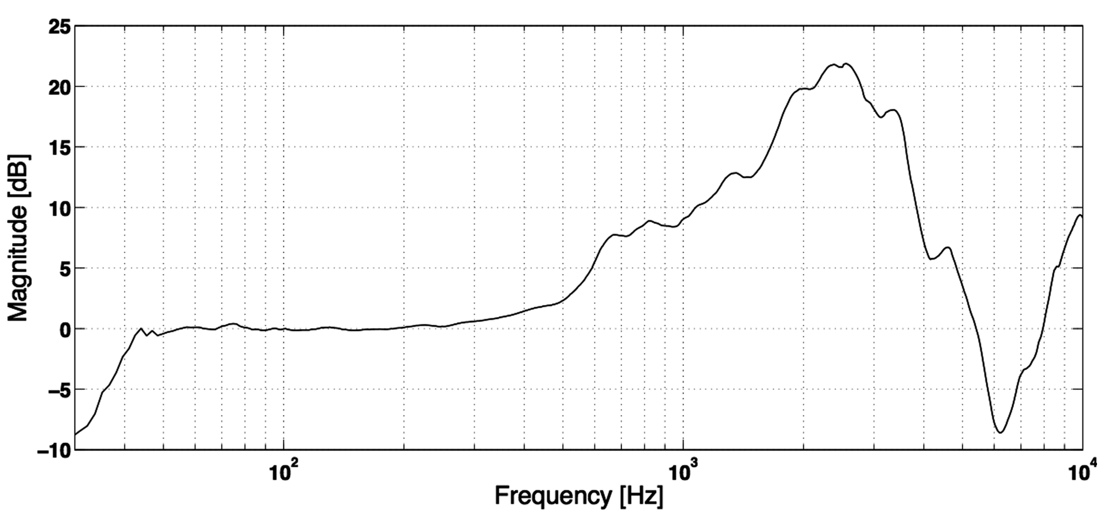


Fig. 5. Generic target equalization curve for ARA headset [4].

with no direct reference to natural perception without the headset.

Then the evaluation was repeated similarly as before the field usage period, that is, the test cases were evaluated alternating wearing and not wearing the headset. The grading and the test procedure were identical to the evaluation before the field usage period.

2.6 Postevaluation Interview

After the field usage period and final evaluations each subject was interviewed. During the interview the diary was examined and discussed. Furthermore the subjects were asked to evaluate numerically the annoyance of the following usability-related attributes:

1) Cords (usability and mechanical noise)
2) Artifacts in general (such as distortion, occasional clip pings, pops, radio interference)
3) Using mobile phone (handset on the ear)
4) Walking.

The numerical scale for the oral evaluation was the same as that used with the sound quality and usability evaluations.

3 RESULTS

3.1 Usability Evaluation and Adaptation

Figs. 7 and 8 show the results of the usability and sound quality evaluations. Each attribute has three bars corresponding to evaluation before the field test period (leftmost bar), evaluation after the field test without reference to the open ear (middle bar), and evaluation after the field test with reference to the open ear (rightmost bar). Median values are indicated by the horizontal lines inside the boxes. The upper and lower ends of the boxes correspond to upper and lower quartile values, respectively, and the whiskers extending from the ends of the boxes show the extent of the rest of the data. Plus signs inside the boxes denote the mean for each case, and plus signs outside the boxes are data points considered to be outliers.

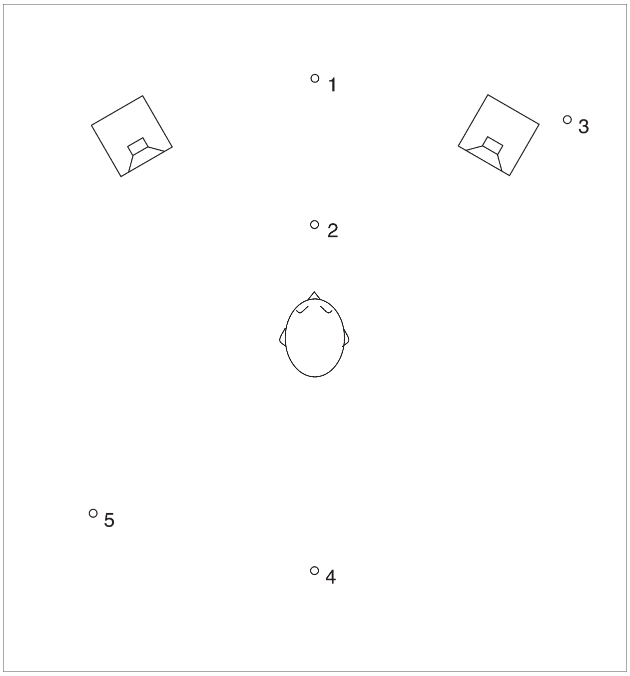


Fig. 6. Listening test setup for sound quality evaluation. Numbered circles show locations where experimenter made finger snaps.

Figs. 9 and 10 show the intrasubject differences between different evaluation times for usability and sound quality. The leftmost bars for each attribute, in both figures, correspond to differences in intrasubject results between evaluations before and after the field test period without reference to the open ear. The rightmost bars correspond to differences between evaluations before and after the field test period when both evaluations were done with reference to the open ear. Positive values correspond to adaptation, that is, the test subject evaluated an attribute to be less annoying after the field test period compared to the evaluation before the field test period.

Results for the usability evaluation tests (Fig. 7) show a wide individual spread of the scores, ranging from very annoying to almost imperceptible, but the main trends are visible. The most annoying activities were "eating crisp bread" and "chewing bubble gum," whereas "drinking water" was judged less annoying. Based on the results, no strong conclusions can be drawn about whether or not adaptation occurred. Fig. 9 shows the intrasubject score differences between the usability evaluations before and after the field test period. Although the test subjects evaluated the annoyance with a broad scale, the intrasubject evaluations seem to be more consistent. The results show some tendency to adaptation to the usage of the headset with the attributes tested. However, there was no significant difference in the results between the evaluations with or without reference to the open ear. One possible reason for this might be that the user sounds of eating, drinking, and speaking are so familiar that 1.5 hours is not enough to fully forget the feeling or sense of these physical activities.

Based on the test subjects' comments during the evaluations, the crisp bread eating was found noisy, as it meant biting hard bread, and every bite was amplified by the occlusion effect. The bubble gum eating was not considered as noisy as the crisp bread eating, but the awareness of the longer duration of the activity led test subjects to grade the activity more on the annoying side. Similarly, drinking was found mostly not annoying because the duration of the event was short.

Sound quality evaluation results are given in Fig. 8. The results show that the overall sound quality of the pseudoacoustics was found good. Especially the spatial impression and location accuracy were found very convincing. One test subject reported that the sounds were nicely lateralized but very poorly externalized during the the first evaluation round before the field test period. However, in the evaluations after the field test period the sound scene was well externalized by the same test subject.

Fig. 10 shows the intrasubject differences between the sound quality evaluations before and after the field test period. Compared to results with physical activities (eating, drinking, speaking) there was more adaptation to the sound quality of the pseudoacoustics. Especially when there was no reference to open ear listening, test subjects graded the quality clearly better after the field test period (the leftmost bars for each attribute). Even when the evaluation was done against a reference (open ear), the test subjects perceived the pseudoacoustics less annoying after the field test period than before.

After the evaluations the test subjects were interviewed for overall comments on the usage of the headset. In addition to verbal comments the test subjects were asked to evaluate some usability aspects with the same MOS scale as used in the evaluations described. The results are shown in Fig. 11. Overall the cords were found very annoying and impractical. The cords tend to get mixed and stuck in clothes and in other items. Furthermore the cords conducted all the mechanical noise to the headset. This was one of the main reasons why the headset was noticed all the time because any movement was heard in the headset. Overall, occasional artifacts such as clipping due to very loud bangs, occasional radio interference, or pops when turning switches were not perceived as annoying.

Using a mobile phone (holding the phone on the ear) was found very annoying. The headset extended slightly outside the ear canal, and this made it hard to couple a mobile phone to the pinna. This resulted in a very thin and low sound to be heard from the phone. Also, the contact noise from the phone touching the headset was found unpleasant.

Walking, in general, was found fairly normal. Thumping sounds from foot steps (during normal and faster walking) were not perceived as annoying. However, this only applies to just walking, without considering the cords, and without

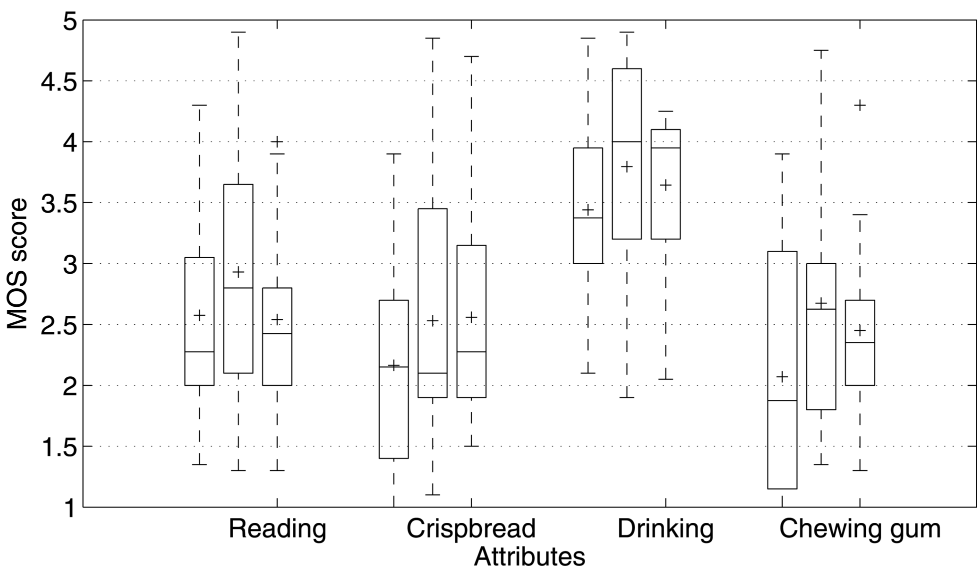


Fig. 7. Usability evaluation results from each evaluation run. Leftmost bars in each triplet correspond to evaluation before field test period, middle bars to evaluation after field test period without reference to open ear, and rightmost bars with reference to open ear.

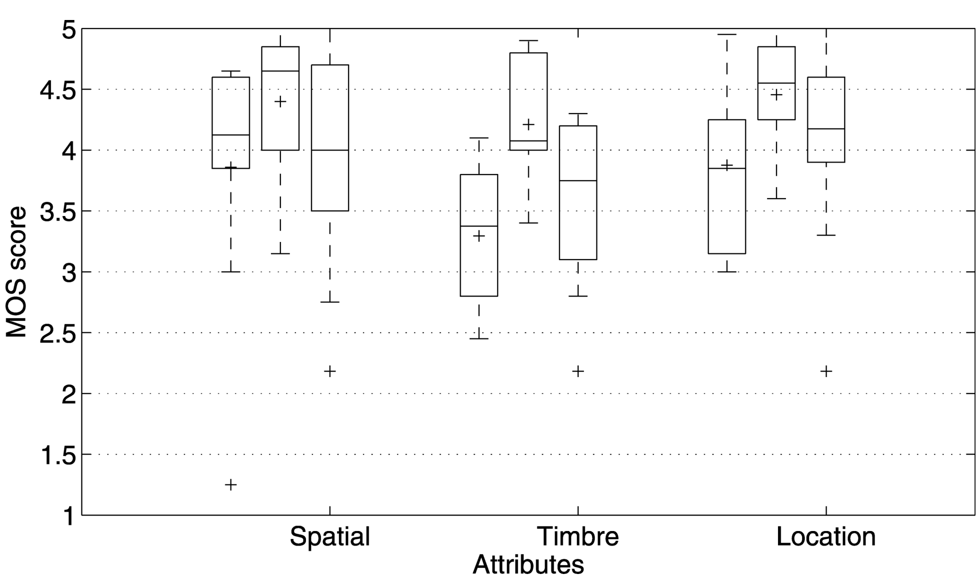


Fig. 8. Sound quality evaluation results from each evaluation run. Leftmost bars in each triplet correspond to evaluation before field test period, middle bars to evaluation after field test period without reference to open ear, and rightmost bars with reference to open ear.

communication with other people (such as conversation while walking). More comments on walking-related problems are described in the following section.

3.2 Interviews and Diaries

The overall comments from the users were very positive. The sound quality of the headset was found to be very good and applicable for most of the practical daily situations. The following sections review and summarize the diary reports from the subjects.

3.2.1 Audio Quality

The audio quality of the headset was found very good. Sense of space and directional hearing were reported to be very close to natural hearing. The sound color of pseudoacoustics was found sufficiently good for most practical everyday-life situations. Many test subjects complained that the higher frequency range was slightly attenuated. Some commented that due to attenuated high frequencies it was harder to understand in "coctail party"-like events. Also, many test subjects reported that when listening to familiar wide-band sounds, such as air-conditioners, frying sounds, or running water, the spectral coloration at higher frequencies was noticed. However, this was not found disturbing. For a critical listening situation, such as, listening to music pseudoacoustically, this might be slightly annoying.

At the end of the test every test subject commented that the sound sources were very well externalized, even in front. The reason for natural externalization in front is probably the real-time cooperation of hearing and vision, as it happens in normal life. When binaural signals were

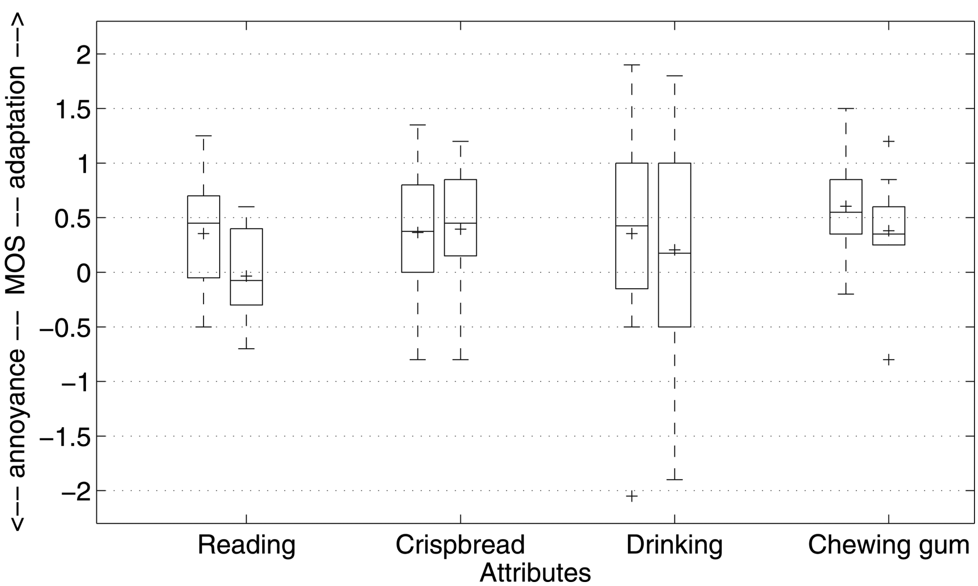


Fig. 9. Intrasubject differences between usability evaluations before and after field test period. Left bars evaluation without reference to open ear; right bars—evaluation with reference to open ear case.

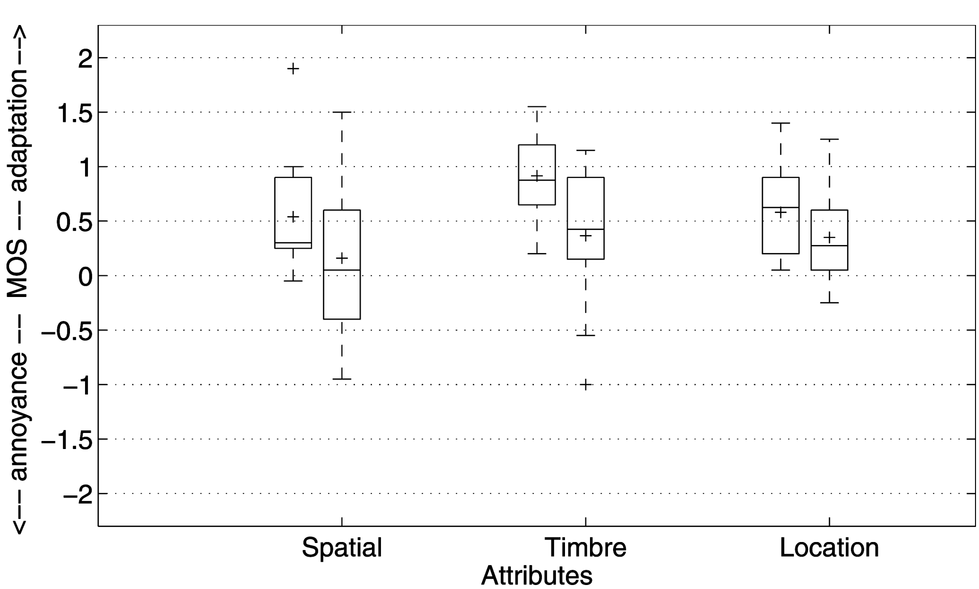


Fig. 10. Intrasubject differences between sound quality evaluations before and after field test period. Left bars evaluation without reference to open ear; right bars—evaluation with reference to open ear case.

recorded and listened to later without visual information, frontal externalization was no longer as good. This could be expected because the head movement cues are lost in the recordings. One subject noted that sometimes nonvisible and unfamiliar sounds where heard inside the head.

There was some audible noise present from the electret microphones. At first it was noticeable, but the subjects reported getting used to it very fast, and after a while no attention was paid to the noise. The measured equivalent inherent noise level of the microphones together with the mixer was 27 dB (A-weighted). 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 frequent comment in the diaries and the interviews was that wearing the headset made users pay more attention to surrounding sounds. Some familiar sounds, such as a car passing by, might have sounded different with the headset on, but upon removing the headset the user noticed that the sound sounded the same with and without the headset. Thus wearing the headset had, if not a tonal, but some effect on the listening experience. On the other hand some subjects reported that during the first days of the field test period they had missed their mobile phone ringing because the ringing tone sounded different with the headset on.

One test subject reported that after a long wearing period the sound color of the natural environment was forgotten, and the pseudoacoustics felt fully natural. For a short while, after removing the headset, the normal acoustics sounded strange.

3.2.2 Ergonomics

Every test subject commented something about the handling noise and limitations caused by the wires of the headset. Some test subjects noted that the headset itself could be forgotten and go unnoticed in the ears but the mechanically transmitted noise from the wires kept reminding about the headset, which was found annoying. Bluetooth or other wireless techniques could be used to route the signals between headset and mixer, thus eliminating the wires. Another criticized issue for all test subjects was the unnecessarily large size of the ARA mixer box, which was found cumbersome to carry. Technically the electronic circuits in the mixer could be easily fitted in a smaller circuit layout, and thus the problem would be fixed.

Placing and removing the headset was found easy. The majority of the test subjects commented that they had some ear ache in the beginning of the field test period. This commonly occurred after about 1.5 hours of usage on the first day of the field test period. However, thereafter most subjects got used to the headsets and no ear ache was reported. Nevertheless some test subjects reported that because of ear ache they had trouble wearing the headset any longer than 20–50 minutes at a time throughout the test period.

Some test subjects commented that after a certain period of usage the earphones started to itch in the ear canals, and some reported that they had the itchy feeling after wearing the headset for a long period.

As already noted, all the test subjects commented that using a mobile phone was found troublesome because the headset slightly extended outside the ear canal. A smaller and more deeply inserted headset might solve this problem. Another, and more obvious solution would be to use the headset as a hands-free device for the mobile phone. In fact, in preliminary tests, using the headset as a hand-free headset for a mobile phone has been found very practical and useful.

3.2.3 Communications

All of the subjects reported that they had no problems in normal conversation situations. For speech the sound quality was very good. One of the big complaints was that the user's own voice sounded very boomy due to the occlusion effect, and the user's own speech was localized inside the head. However, most of the test subjects men-

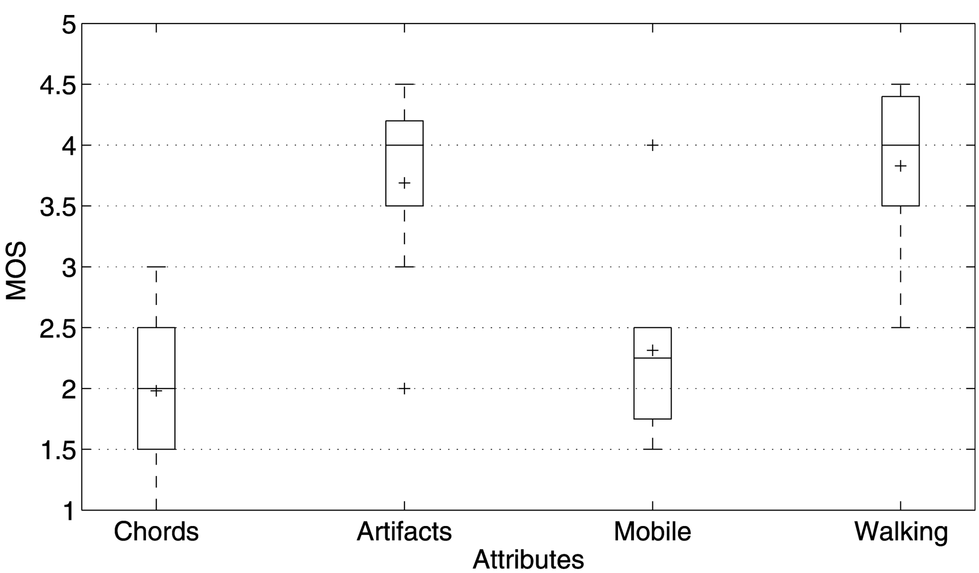


Fig. 11. MOS values given in oral interviews after field test period.

tioned that they got accustomed to this fairly fast and, in the end, this was not found annoying. Some test subjects reported that sometimes they had trouble determining the level of their own voice.

One common complaint was that conversations were troublesome during dinners and while walking. Due to the occlusion effect, eating sounds were boosted inside the subject's ear canals, which masked the surrounding conversations. The same phenomena occurred while walking as the wires created mechanical noise when moving against the clothes, and the foot steps created a thumbing sound in the ear canal. The added noise in the ear canal masked the sounds coming from surroundings, thus interfering with listening to conversations. This phenomenon was found somehow annoying, and also was mentioned to be somewhat exhausting as it required more mental focusing for listening.

Most of the test subjects reported to have felt some social discomfort when wearing the ARA headset in situations where it is not appropriate or polite to wear headphones, such as during lectures and conversations. Some subjects commented that they had a feeling that they had to prove somehow that they were paying attention, such as by replying faster during conversations. One test subject reported that other people had felt he was slightly isolated from the others when wearing the headset. However, one test subject who had been used to wearing similar types of earphones to listen to music extensively, had not perceived any social pressure. Of course, if other people would know that the headset transmits the sound unaltered and enables normal communication, then this would not be a problem.

3.2.4 Other Observations

Eating and drinking was reported to be practically tolerable, especially if there was no need to pay attention, for example, to a conversation. Most of the subjects commented that they got used to eating during the field test period. The results in Fig. 9 show that with crisp bread eating there was very little, if any, adaption to the annoyance factor with eating. However, crisp bread is a fairly hard dish compared to most dishes, and a common lunch would compare more to bubble gum chewing, which showed some adaptation among the test group.

Some test subjects commented that brushing teeth was found unpleasant due to boosted mechanical noise caused by the occlusion effect. On the other hand, one test subject reported that shaving felt more efficient, in a positive way, because the sound was boosted due to the occlusion effect.

Loud bangs, as from closing 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 for others it was less annoying. Also listening to continuous loud sounds (listening to a live rock concert) resulted in distorted sound and was reported to be annoying. Distortion performance could be improved by better electric and acoustic design. With some test subjects the connectors of the wires caused undesired scratchy electric noise when there were contact problems due to movement of the wires. This was found very unpleasant.

At first, walking sounded uncomfortable due to boomy sound from the foot steps, 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, such as after some exercise, sounded boosted and was reported to be unpleasant. Furthermore, strong wind resulted in unpleasant sound in the headset, whereas mild wind was not found to be a problem and sounded quite natural.

Some test subjects experimented with listening to music by using the ARA mixer to mix music to pseudoacoustics. It was commented that it was nice to be able to hear the full frequency range from the music, as well as the surrounding sounds. However, the downside was that the level of the music had to be set fairly high in order to be heard in noisy environments. Just like listening to normal headphones, the music was located inside the head. This was reported to be positive as it helped separating the music from real surrounding sounds. However, one test subject complained that he had trouble determining which sounds were part of the recording and which were part of the natural surrounding sound environment.

3.3 Feature Suggestions

The most desired missing feature, according to the entire test group, was wireless operation. Without the cords the overall usability of the headset would increase considerably. Another feature suggestion from many subjects was the ability to adjust the level of the pseudoacoustics, either to boost or to attenuate. This way a user could attenuate pseudoacoutics when walking in a noisy environment, and still be able to hear, and to react to surrounding sound events.

Many of the subjects suggested using the device as a mobile phone hands-free headset. In addition the headset could function as a general audio interface for any audio device. The connection could be offered wirelessly, for example, with Bluetooth.

To the main question of whether the test subjects would be willing to use the headset in everyday-life situations the answer was mainly positive. The subjects were willing to use the headset, even for longer periods, if the headset would offer some added value or other advance; for now the system was used only as a hear-through device.

4 CONCLUSIONS AND DISCUSSION

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

As the results clearly indicate, when designing an ARA headset a great deal of effort should be put into designing a mechanically silent system. This, as well as most of the results shown, applies to a normal headphone design process, as well.

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 not be interfered with. For example, while walking in the streets with an augmented audio guide, for the sake of safety, a user should be able to hear the surrounding sounds. Analysis of the surrounding sound environment could be one attractive added value application for the ARA headset. The system would analyze the microphone signals continuously and, for example, warn of approaching vehicles or, based on user preferences, give information on the surrounding environment.

When these kinds of headsets are more common, there will still be all types of practical issues to be solved. One socially important issue is how to notify other people that the user is able to hear other people normally. This was already noted among the test group. Furthermore, binaural recording can be made easily and unnoticed with the headset. For courtesy, or even in some cases for legal reasons, there should be a practical way to let other people know when you are recording an event. A visionary idea would be that people's personal appliances would communicate wirelessly, notify surrounding people (devices), and even ask automatically for permission for recording.

One common trend, and challenge, in augmented reality audio has been how to embed virtual sounds convincingly in the real sound environment. However, another challenge is how to embed virtual sound sources into the environment so that the sounds are convincingly part of the environment but still are distinguishable as virtual. For example, in the streets it is essential that the user can separate virtual and real sound sources.

This work studied the case where the subjects were listening only to natural sounds in their environment. For future work the next obvious and interesting step will be to augment the surroundings with virtual sounds and study the effect it has on the usability and what kind of issues will come up in this scenario.

5 ACKNOWLEDGMENT

This work was supported by the Academy of Finland (project 106956), Helsinki University of Technology (UI-ART project), and the Nokia Research Center as part of the KAMARA+ project (Killer Applications of Mobile Augmented Reality Audio).

6 REFERENCES

[1] A. Härmä, J. Jakka, M. Tikander, M. Karjalainen, T. Lokki, J. Hiipakka, and G. Lorho, "Augmented Reality Audio for Mobile and Wearable Appliances," *J. Audio Eng. Soc.*, vol. 52, pp. 618–639 (2004 June).

[2] M. Tikander, M. Karjalainen, and V. Riikonen, "An Augmented Reality Audio Headset," in *Proc. 11th Int. Conf. on Digital Audio Effects (DAFx-08)*, (Espoo, Finland, 2008 Sept. 1–4).

[3] T. Lokki, H. Nironen, S. Vesa, L. Savioja, A. Härmä, and M. Karjalainen, "Application Scenarios of Wearable and Mobile Augmented Reality Audio," presented at the 116th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 52, p. 786 (2004 July/Aug.), convention paper 6026.

[4] V. Riikonen, M. Tikander, and M. Karjalainen, "An Augmented Reality Audio Mixer and Equalizer, presented at the 124th Convention of the Audio Engineering Society, *(Abstracts)* www.aes.org/events/124/124thWrapUp.pdf, (2008 May), convention paper 6026.

[5] M. Tikander, "Sound Quality of an Augmented Reality Audio Headset," in *Proc. 8th Int. Conf. on Digital Audio Effects (DAFx-05)*, (Madrid, Spain, 2005 Sept. 20–22).

[6] D. S. Brungart, A. J. Kordik, C. S. Eades, and B. D. Simpson, "The Effect of Microphone Placement on Localization Accuracy with Electronics Pass Through Earplugs," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, (New Paltz, NY, 2003 Oct. 19–22).

[7] D. Hammershøi and H. Møller, "Sound Transmission to and within the Human Ear Canal," *J. Acoust. Soc. Am.*, vol. 100, pp. 408–427 (1996).

[8] P. Hiselius, "Acoustic Modeling of Earplugs," *Acta Acustica*, vol. 92, pp. 135–138 (2006).

[9] V. Riikonen, "User-Related Acoustics in a Two-Way Augmented Reality Audio System," Master's thesis, Helsinki University of Technology, Espoo, Finland (2008).

THE AUTHOR

Miikka Tikander was born in Riihimaki, Finland, in 1973. He received an M.Sc. degree in electrical and communication engineering from the Helsinki University of Technology, Espoo, Finland, in 2003. His major subject was acoustics and audio signal processing.

Since 2001 he has worked as a research assistant and, after graduation, a researcher in the Laboratory of Acoustics and Audio Signal Processing at the Helsinki University of Technology. Currently he is working on a Ph.D. degree in the field of augmented reality audio. His research topics include audio signal processing and transducer technologies.

Since 2001 Mr. Tikander has been secretary of the Acoustical Society of Finland.