Capturing blocked-entrance binaural signals from open-entrance recordings

Dorte Hammershøi, Pablo Hoffmann, Søren Olesen and Per Rubak

Acoustics, Aalborg University, Fredrik Bajers Vej 7 B5, 9220 Aalborg Ø, Denmark
dh@es.aau.dk

Binaural recordings enable us to capture all sound attributes including spatial information, room effect, and source characteristics in a given environment. It has been shown that blocked-entrance binaural recordings provide advantages over open-entrance recordings, primarily because the blocked-entrance recordings are not influenced by the ear canal acoustics of the individual. However, blocking the ear canal imposes an obvious disruption to normal hearing conditions, which may be unacceptable for applications in which binaural audio capturing is desired but without interfering the individual's hearing and doing. In this work we propose a strategy for the recording of binaural audio with minimal hearing interference, and for transforming these recordings to blocked-entrance versions that are more suitable for analysis and reproduction of binaural audio in a more general context.

1 Introduction

The motivation of this work originates from a project that deals with the acquisition of skills using multimodal virtual reality technology [1]. The underlying idea is that models of skills can be derived from the analysis of multimodal information captured on experts performing the task for which the skill is defined, e.g. rowing in sports. Capturing of multimodal information includes motion capturing, i.e. experts' movements, as well as capturing of visual and auditory characteristics of the environment that may be relevant for skilled performance. In relation to the auditory modality there is particular interest to obtaining a representation of what the expert may be listening to, and thus, sound capturing at the ears is essential.

Binaural recordings can capture all sound attributes such as source characteristics, environment effect and spatial information of a given acoustic environment [2]. As mentioned before, this is important in the context of skill acquisition because it allows to obtain a direct representation of the acoustic events that may be used by the expert as cues for skilled performance. It has been shown that binaural recordings made at the blocked ear canal, or blocked-entrance recordings, are well suited because they are not influenced by the ear-canal acoustics of the individual from whom the recordings are made. This allows for the analysis and reproduction of binaural audio in a more general context. However, blocking the ear canal unavoidably impairs normal hearing conditions, and this has been shown to cause a loss in the sense of presence [3], which may be detrimental for performance. Therefore, there is a need to find solutions that enable binaural audio capturing without interfering the individual's hearing. In addition, because task performance involves movements, it is also important not to interfere with the individual's doing, and thus, the target sound-capturing system should be wearable system.

The alternative to blocked-entrance recordings are recordings made with the open ear canal, or open-entrance recordings. Binaural recordings of this type include the effect of the ear-canal acoustics, and because this effect is highly individual [4] it precludes us from having a more general representation of sound. It seems that a strategy that allows to eliminate the effect of the ear-canal acoustics without affecting the characteristics of the recorded sound field may be beneficial. Here, we propose a strategy to transform open-entrance recordings to blocked-entrance versions. The strategy is based on deriving an equalization filter from transfer function measurements made for the open- and blocked-entrance ear canals. In this study we report on the measuring system and aspects of the measurements. The potential advantages and disadvantages of the system are also discussed.

2 Measurements

Impulse responses for the open- and blocked-entrance ear canals were measured on three subjects. Measurements were carried out using the maximum-length-sequence (MLS) technique [5]. Miniature microphone were placed at the ears, and the MLS signals were played back over headphones. The ratio between the Fourier transforms of the open- and blocked-entrance impulse responses was defined as the target frequency response for equalization.

2.1 Subjects

Three paid subjects participated. Subjects' age ranged from 22 to 23, and none of the subjects reported any ear problems that might have affected the measurements.

2.2 Microphone

An FG23629 miniature microphone (Knowles Acoustics) was employed for the measurements. This is a cylindrical microphone with 2.59 mm in diameter and 2.59 mm long. The frequency response of this microphone is shown in Figure 1. This frequency response was obtained by dividing the Fourier transform of the measured FG23629 impulse response by the Fourier transform of a reference 1/4 inch pressure-field B&K 4136 microphone measured simultaneously. The frequency response of the reference microphone is flat in the audio frequency range.

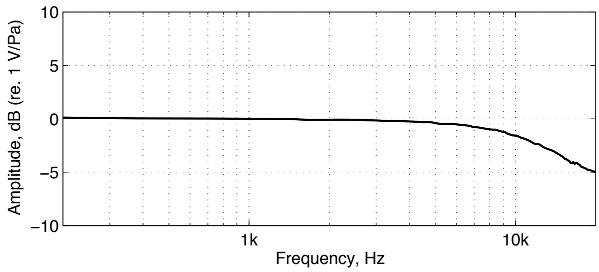

Figure 1: Frequency response of Knowles FG23629 microphone capsule.

2.3 Measuring procedure

The measuring system consisted of a laptop computer equipped with a USB audio interface (Edirol UA-25). Measurements were done using the windows-based software WinMLS 2004 (Morset Sound Development). An MLS signal of order 12 (4095 points) at a 48-kHz sampling frequency was used. This length is sufficient to capture the complete impulse responses. The digitally generated MLS signal was sent to the D/A converter of the audio interface. The analog output signal was delivered to the listener over a pair of Beyerdynamic DT-990 circumaural headphones. The voltage input to the headphones was set to ± 140 mV which produced a sound pressure level (SPL) of 78-80 dB at the ears. According to the standard ISO 11904-1 [6], the measured at-the-ear SPL corresponds to an equivalent free-field SPL of about 75-78 dB(A), depending on subject. At this level the stapedius is assumed to be relaxed. Microphones were connected to a custom-made portable power supply and 20-dB amplifier. The output signal from the amplifier was sent to the A/D converter of the audio interface.

During a measuring session the subject was seated in a comfortable chair that was placed in a standard listening room. Four measurements were obtained for each ear, and subjects were asked to reposition the headphones between measurements. Two microphones were used for the measurements. Microphones were embedded to the end of a metal strap that was flexible enough to be used as a holder that helped to fixate the position of the microphone in the ear. This was particularly important for measurements carried out with the open ear canal. For the blocked-entrance measurement the microphone was mounted on an earplug that was compressed for insertion into the ear canal. After the earplug expanded, care was taken to ensure that the end of the earplug and the microphone were flush with the ear canal's entrance. To assess the reproducibility of the measurements three sessions conducted on different days were completed for each subject.

2.4 Microphone position calibration

A crucial aspect of these measurements was that for both blocked- and open-entrance measurements the position of the microphone had to be the same. To calibrate the microphone position a procedure was developed that is explained as follows. Blocked-entrance measurements were always performed first, and immediately after the measurements ended a photograph of the blocked entrance was taken. During this moment the subject was seated with the head fixed to a head rest, and was instructed to be as quiet as possible. Then, the earplug was carefully removed, the microphone was put back, and a new photograph was taken. Again, the subject was instructed to remain quiet while the photograph was taken. The camera used to take the photographs was at the exact same positions for both photographs. These two photographs were compared on a computer screen via a custom-made graphical user interface, and in case the microphone's position for the open entrance did not match that of the blocked entrance, the microphone was repositioned and a new photograph was taken for comparison. Figure 2 shows an example of the microphone aligned for both blocked- and open-entrance measurements. To corroborate that the subjects' head did not move between the two photographs, subjects were wearing a headband with a head-tracker, thus, it was possible to check that for both pictures the position and orientation of the head were the same within a range of ±0.5 inches for position and ±0.5 degrees for orientation. Note that this procedure could control the position in two dimensions, and thus the depth of the microphone's position was visually inspected.

Figure 2: Example of the microphone positioning for the blocked-entrance measurements (left picture) and the open-entrance measurements (right picture).

2.5 Signal-to-noise ratio

It is well known that at low frequencies it is difficult to have a good signal-to-noise ratio (SNR) due to the inability of the system to measure accurately at low frequencies. This is particularly important for the design of inverse filters, and thus the noise of the measurement system was measured. Noise measurements were conducted with all settings equal to those for the transfer function measurements with the exception that the input to the headphones was set to zero via a software-based digital level controller available in WinMLS. For comparison Figure 3 shows transfer function measurements from one subject (s1) and the measured noise. The SNR is approximately 50 dB at low frequencies, and for those frequencies where notches in the transfer functions are observed the SNR is about 45 dB.

3 Results and Discussion

Results of the measurements were given as impulse responses. The initial 27 samples of the impulse responses were removed because they correspond to the delay of the measuring system (loop-back measurement). The resulting responses were truncated using a 512-points square window, and Fourier transforms of the truncated responses were computed. This provided a frequency resolution of 93.75 Hz. Open-to-blocked ratios for all subjects and days are shown in Figure 4 and Figure 5 for the left and right ears respectively. It can be observed that the reproducibility of the measurements is

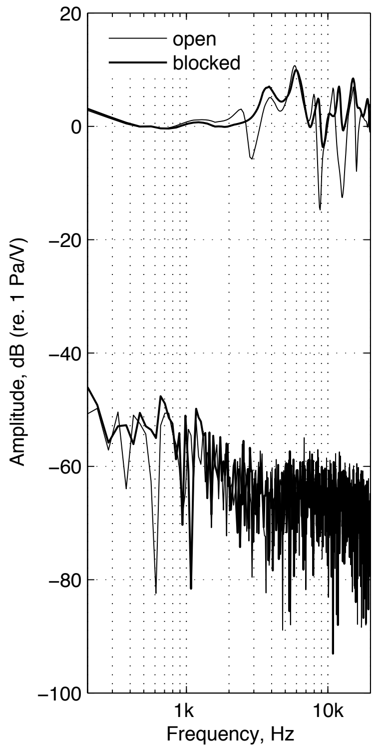

Figure 3: Signal-to-noise ratio measurements are compared to open- and blocked-entrance measurements (subject s1).

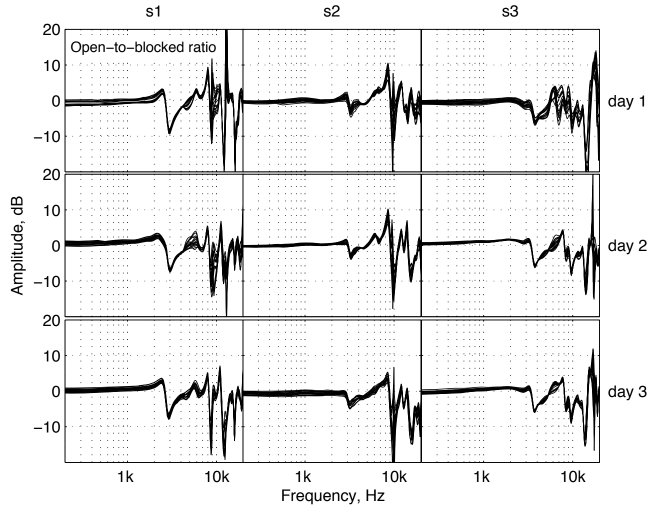

Figure 4: Open-to-blocked entrance ratio for the left ear of three subjects. Repeated measures were performed on three consecutive days.

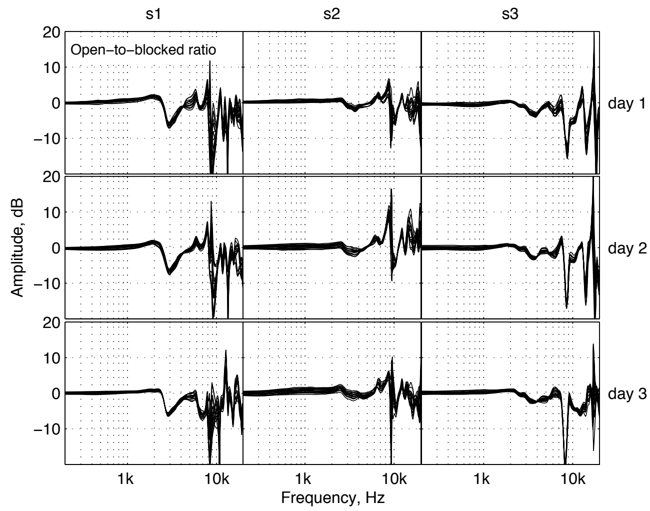

Figure 5: Same as Fig. 4 for the right ear.

good for frequencies up to 6–7 kHz. Within this range the general shapes of the open-to-blocked ratios are in agreement with measurements reported by Hammershøi and Møller [4, p. 418, Fig. 13]. In that study a probe microphone was used. The mean open-to-blocked ratio from 12 subjects shows a shallow peak of about 5 dB at 2 kHz and a notch of about -12 dB at 4.5 kHz. Our results exhibit a similar pattern but with the center frequency of the notch somewhat shifted to 3–4 kHz. Also, the corresponding levels for the peak and notch are less pronounced than those reported in [4], implying that the open-entrance and blocked-entrance transfer functions obtained here are more similar. A possible explanation for these observations is that even though the microphone used in this work was small, it still blocked the entrance to the ear canal to some extent.

Results showing that open-entrance measurements are of a more semi-blocked nature suggests that for open-entrance recordings the microphone may interfere with normal hearing conditions. Therefore, further examination is required in the design of the binaural recording system.

3.1 Inverse filter considerations

As mentioned, one goal of this study is to design a binaural recorder, which besides its recording capability, compensates for the physical open-entrance recording situation and delivers an output that corresponds to a blocked-entrance recordings. This conversion requires an equalization of the transfer functions given in Figure 4 and Figure 5.

The open-to-blocked entrance ratios vary from person to person, and this means that we shall focus on binaural recordings for the individual. Hence it is important to look at the changes over repetitive measurements on the same person. A measurement which is the base for designing an equalization filter must be representative at any time, i.e. it must be adequately repeatable otherwise the designed equalization filter will – no matter how exact it may turn out for a single measurement – be useless.

For example, a plain inversion of the open-to-blocked ratio for subject s2 and day 1 can be computed as

$$H_{2,1}[k] = \frac{Hb_{2,1}[k]}{Ho_{2,1}[k]} \tag{1}$$

where Hb and Ho indicate the blocked- and open-entrance transfer functions respectively. Assuming this is a valid inverse filter it must also be so when applied to the ratio measurements carried out on day 2 and day 3. See Figure 6 where the thick line is the result of the equalization of day 2 and the thin line is the result of the equalization of day 3. In agreement with the frequency range of measurement reproducibility, this equalization

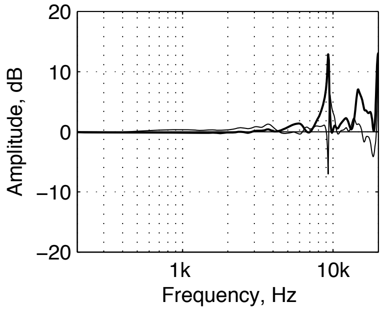

Figure 6: Frequency responses after applying an inverse filter based on measurements on day 1, to the measurements on day 2 (thick line) and day 3 (thin line).

strategy works fine for frequencies up to 6–7 kHz. For higher frequencies, we need to find another strategy than just plain inverse filtering. That is, the exact strategy on how to design the complete inverse filter is still to be investigated.

4 Acknowledgments

This work was conducted with financial support from the European Integrated Project SKILLS "Multimodal Interfaces for Capturing and Transfer of Skill".

References

[1] EU Integrated Project SKILLS, http://www.skills-ip.eu, (2006)

[2] D. Hammershøi, H. Møller, "Binaural Technique, Basic Methods for Recording, Synthesis, and Reproduction", In J. Blauert, editor, *Communication Acoustics*, 223–254, Springer Verlag, Berlin, Germany (2005)

[3] C. D. Murray, P. Arnold, B. Thornton, "Presence Accompanying Induced Hearing Loss: Implications for Immersive Virtual Environments", *Presence: Teleoperators and Virtual Environments*, 9(2), 137–148 (2000)

[4] D. Hammershøi, H. Møller, "Sound transmission to and within the human ear canal", *J. Acoust. Soc. Am.*, 100, 408–427 (1996)

[5] D. D. Rife, J. Vanderkooy, "Transfer function measurement with maximum-length sequence", *J. Aud. Eng. Soc.*, 37, 419–444 (1989)

[6] ISO 11904-1: "Acoustics – determination of sound immissions from sound sources placed close to the ears – Part 1: Technique using microphones in real ears (MIRE-technique)." International Standardization Organization, Geneva, Switzerland, (2002)