19th INTERNATIONAL CONGRESS ON ACOUSTICS
MADRID, 2-7 SEPTEMBER 2007

VIRTUAL AUDIO REPRODUCED IN A HEADREST

PACS: 43.25.Lj

M.Jones, S.J.Elliott, T.Takeuchi, J.Beer
Institute of Sound and Vibration Research; University of Southampton, University Road, Highfield Southampton S017 1BJ; mj@isvr.soton.ac.uk; sje@isvr.soton.ac.uk

ABSTRACT

When virtual audio reproduction is simultaneously required in many seats, such as in aircraft or cinemas, it may be convenient to use loudspeakers mounted inside each seat's headrest. In this preliminary study, the feasibility of virtual audio reproduction in the headrest of a single seat is explored using an inversion technique to compensate for crosstalk and the synthesis of head related transfer functions. Although large changes in the magnitude of the signals reproduced at the listener's ears are observed as the listener moves their head within the headrest, informal listening tests indicate that the reproduced acoustic images are surprisingly stable in about an eighth of an arc either side of the loudspeaker positions. Not surprisingly, frontal images are more difficult to reproduce with headrest loudspeakers.

INTRODUCTION

A number of methods could be used to provide people with the experience of virtual audio. One class of these methods is that of binaural reproduction, where a small number of sources are used to reproduce the virtual soundfield such that the signals arriving at the ears of a listener match the signals from a binaural recording or generated from HRTF's. In this way, the signals at each of the ears must be controlled separately with minimum interference, and the path from the loudspeakers to the ears must be equalised.

By reproducing the virtual soundfield using different loudspeakers and geometrical arrangements, each arrangement can be shown to have certain advantages and drawbacks. Up until now far field loudspeakers or headphones have been used to reproduce binaural recordings. In this paper the use of loudspeakers mounted in headrests will be investigated for personal binaural reproduction. One advantage of such a system is that it could potentially be scaled to provide binaural reproduction in a large number of seats over a wide area.

For this work, a binaural reproduction system has been developed using the frequency responses measured from loudspeakers installed in the headrests of a pair of airline seats. These measurements are used to infer the strengths and limitations of this arrangement, which are discussed below. Finally, an implementation of a virtual audio system in the seat has been developed, from which some informal subjective results have been obtained.

CROSSTALK CANCELLATION

In this section we discuss the fundamentals of the crosstalk cancellation method of virtual audio. Figure 1 shows a diagram of the problem, where the two outputs of the reproduction system arrive directly at the two ears of the listener, H_{ll} and H_{rr}, along with the crosstalk paths to the opposite ears, H_{lr} and H_{rl}.

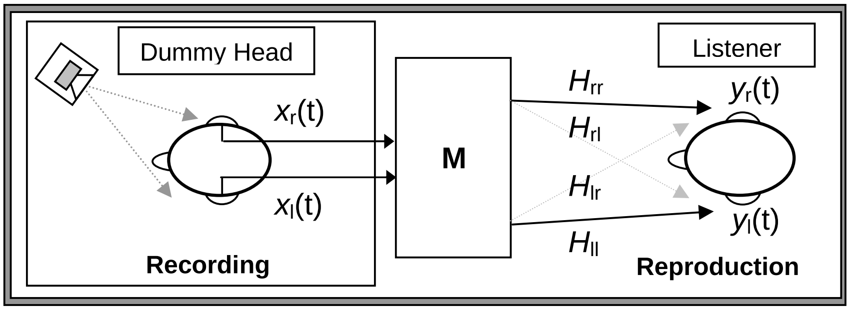

Figure 1 - Diagram of crosstalk cancellation method

In this technique, the objective is to obtain signals y_l(t) and y_r(t) that are as similar as possible to the signals x_l(t) and x_r(t) recorded at the dummy head ears. By representing the mixing that occurs between the reproduction system outputs and the signals arriving at the ears in the frequency domain as the complex matrix **H** which has elements [H_{ll} H_{lr};H_{rl} H_{rr}], and representing the unmixing process as the complex matrix **M**, the required matrix **M** can be calculated as the inverse of the mixing matrix **H**. In this way, the crosstalk paths are cancelled completely, and the frequency response of the direct paths are equalised.

In reality, the design of the filters is not quite as straightforward, since although the filters do not need to be constrained to be causal, the inversion operation results in large peaks where there are nulls in the response and the effects of these peaks need to be mitigated. Also the use of finite length filters means that usually some trade-off is required between the performance in each frequency region. The difficulties in inverting **H** can be quantified by the condition number of the matrix as a function of frequency, and it has been shown that the higher the condition number the smaller the zone will be around the head of the listener where effective crosstalk cancellation occurs.

REPRODUCTION ARRANGEMENTS

Figure 2 shows a diagram of the three different methods of reproducing the virtual audio considered here, which are:

(A) Reproduction using a headset
(B) Reproduction using a pair of loudspeakers located some distance from the listener
(C) Reproduction using the loudspeaker mounted in the headrests of a seat

Due to high isolation between the channels of a headset, crosstalk cancellation is not usually required and so provided the response of the drivers in the headrest are relatively flat, the raw binaural recorded signals may be sent directly to the drivers. When loudspeakers are employed for the reproduction, due to the distance separating the sources and the listener there is significant crosstalk and the crosstalk cancellation method is required to achieve adequate performance. It is not clear however whether crosstalk cancellation is required for headrest reproduction since the left and right ears are somewhat acoustically isolated due to the shielding and absorption of the head and the headrest.

Another aspect to take into consideration when comparing the reproduction systems is the robustness of the virtual audio to movement of the listener's head. The use of headsets gives the best robustness since the drivers move with the listener, although it is known that by removing and putting the headset back on there may be large changes in the frequency response to the ears, despite only a small change in the position of the headset. The acoustic images move with head motion, though, and the images are often perceived as not being well externalised with headset reproduction. When loudspeakers are used, the virtual audio system is only able to cope with small changes in the listener position, and is similarly affected for head rotation due to the small size of the region around the ears within which the pressure is controlled. Recent research [1] has allowed such systems to operate over a larger region of space, through careful positioning of the source array.

A headrest reproduction system poses different problems, since the head is unable to move very far due to the restrictions of the seat and head rotation is likely to be small. However, due to the proximity of the drivers and the head, the small changes in head position can have a large effect on the frequency responses. The effect of typical head movement on the performance of a virtual audio system is considered in this work.

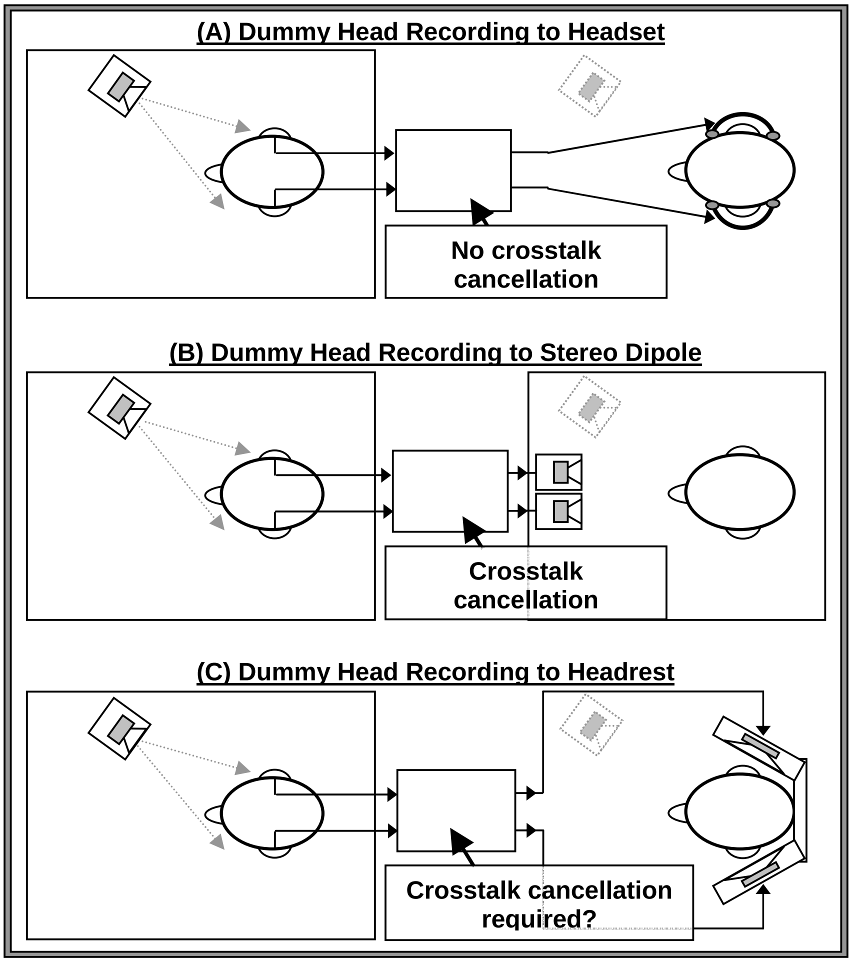

Figure 2 - Diagram of three methods of reproducing the virtual audio

MEASURED RESPONSES

The frequency responses from a pair of loudspeakers mounted in the headrest of a typical airline-style seat to the ears of a dummy head placed in this seat have been measured in anechoic conditions. For this work a small full range loudspeaker [2] was used that was able to fit in the small space surrounding the headrest, and to accurately reproduce frequencies in the range of 150Hz to 16kHz.

Figure 3 shows the passive isolation between the ears nearest and furthest from a source, i.e. the ratio of the measured pressure at the two ears when only one loudspeaker is being driven, averaged from results obtained from both sides of the symmetrical arrangement. The attenuation of sound at the opposite ear is greater than the attenuation due to spherical spreading alone, modelled with a free field monopole, which can be attributed to the presence of reflecting and absorbing objects. However, the isolation is quite poor around 600 to 900Hz where little improvement is seen over the free field result. The frequency averaged value of the isolation is around 20dB. This indicates that crosstalk cancellation may well be necessary, although it should be noted that little attempt has been made to design the headrest to increase isolation, since the headrest has no additional absorption behind the head as would likely be present in a final solution.

The matrix of measured responses from each source to each ear of the dummy head was then used to generate the matrix **H** a range of frequencies. A measure of the difficulty in performing crosstalk cancellation is the condition number of this matrix, which is the ratio of the largest to the smallest eigenvalue. For a virtual audio reproduction system that uses closely spaced loudspeakers located far from the listener, the condition numbers can be large, of the order 10^3, since the paths from a loudspeaker to both ears are very similar. Figure 4 shows the condition number for the matrix **H** when the loudspeakers in the headrest are used for reproduction, which is generally about 2 with a peak at 3.5.

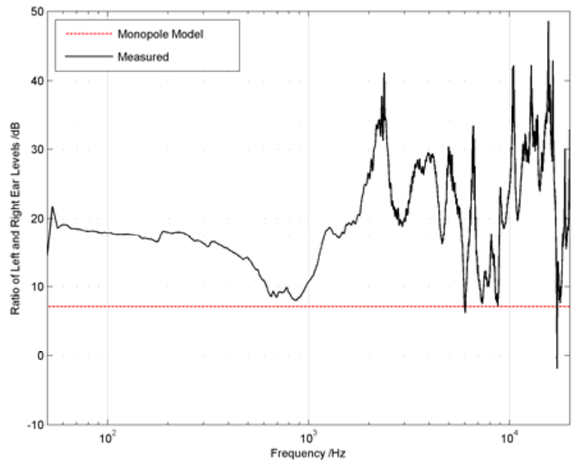

Figure 3 - Passive isolation between the ears nearest and furthest from the source, comparing results from a simple monopole model to the results obtained from measurements made in an anechoic chamber

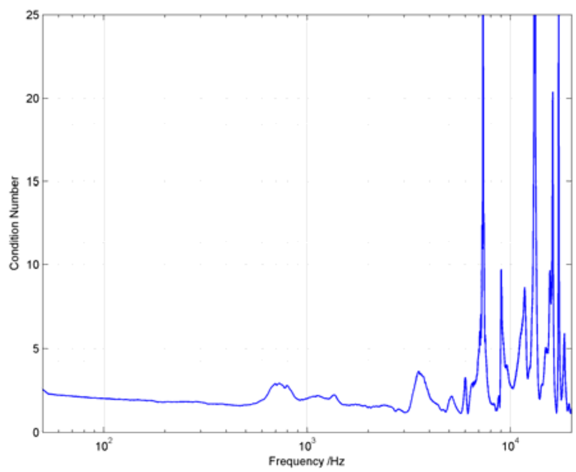

Figure 4 - Condition number of mixing matrix to be inverted for crosstalk cancellation filter, as a function of frequency

These low values for the condition number suggest that the loudspeakers will not have to be driven very hard to achieve crosstalk cancellation.

HEAD MOVEMENT

For accurate virtual audio, there must be as little change as possible in the overall frequency responses after crosstalk cancellation. Any changes in the magnitude or phase response result in non-perfect equalisation and crosstalk cancellation so that the spectra of sound arriving at each ear is not the same as the spectra recorded from the dummy head. Figure 5 shows the magnitude of the matrix of responses from each source to each ear after the crosstalk cancellation stage, with the dummy head in the seat in the central position used to calculate **M**, and positioned about 5cm to the left and to the right. Figure 6 shows the phase of these frequency responses. With the head in the centre the crosstalk cancellation is perfect with the diagonal elements of the matrix having a modulus of 0dB and phase of 0°, and the magnitude of the off-diagonal terms being zero.

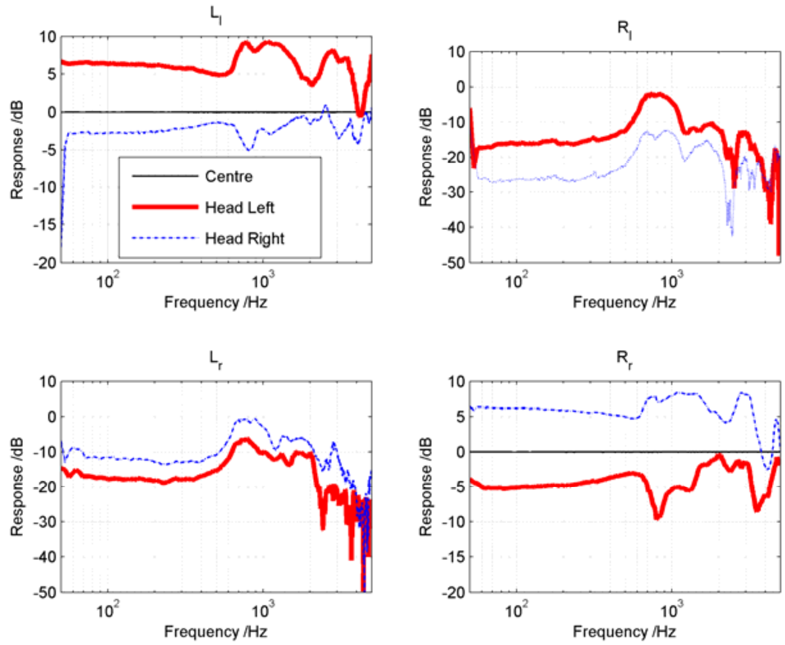

Figure 5 - Change in the magnitude of the frequency responses to the ears after crosstalk cancellation, as a function of frequency when the head is moved to the left and right

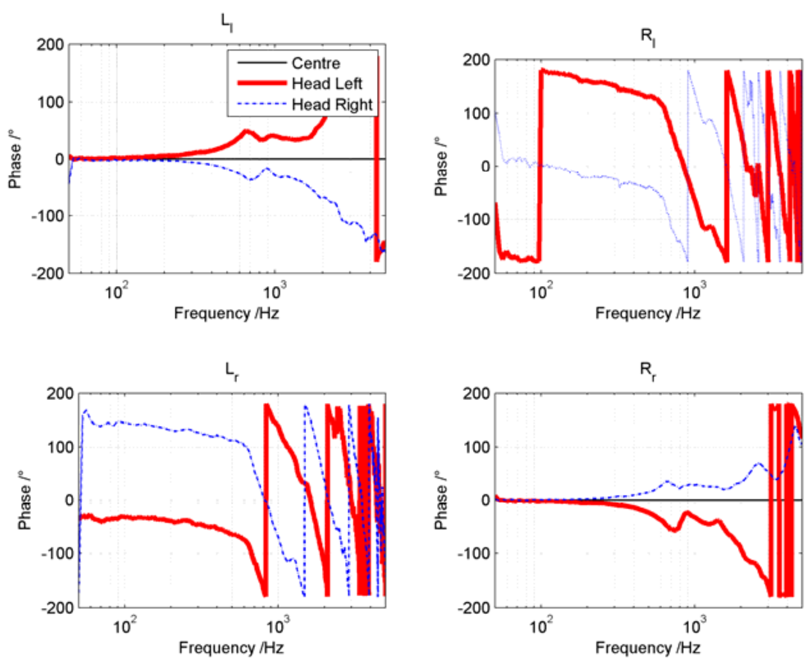

Figure 6 - Change in the phase of the frequency responses to the ears after crosstalk cancellation, as a function of frequency when the head is moved to the left and right

These plots show that the apparent effect of head movement on the crosstalk cancellation is severe, with variations of at least +-5dB in amplitude and 180° in phase in the diagonal terms, and the off-diagonal, cross coupling, terms increasing to about 15dB below the diagonal, direct, terms. The isolation between the ears becomes comparable or even less than the isolation due to passive effects before crosstalk cancellation (Figure 3). By moving the head in one direction, the changes experienced at one ear are the opposite to the changes experienced by the other ear, meaning that the relative difference between the ears is even greater and so the effects on the virtual audio could be expected to be significant. In a fully symmetric arrangement, the results for head movement in one direction would be identical but opposite to the results for head movement in the opposite direction, although this is not seen here since in our case the arrangement is not entirely symmetrical due to the presence of the adjacent seat.

IMPLEMENTATION

To verify whether the headrest reproduction system is capable of providing accurate virtual audio, and to investigate whether head movement deteriorates the performance to the extent predicted from the measurements, an implementation of a crosstalk cancellation virtual audio system has been developed for the headrest.

As before, anechoic measurements of the frequency responses are used to calculate the crosstalk cancellation filters **M**. These filters are converted into time-domain FIR filters with a small degree of truncation, and implemented in software using a Huron [3] system. The Huron system is chosen to operate at a sample frequency of 44.1kHz so that the useable bandwidth of reproduced audio is up to a maximum of 20kHz, and the outputs of the DAC's connected to a high-fidelity audio amplifier before being sent to the drivers.

So far in this work only informal subjective testing has been carried out with a small number of subjects, and therefore only the descriptions of the subjective impressions can be included here. An interesting conclusion from this informal testing was that the effect of head movement was not nearly as drastic as one might expect. Relatively stable images above and to the sides were maintained throughout typical head movement, although the precision of the virtual source locations was affected. The headrest system was particularly convincing for virtual sources located to the sides and behind the listener, presumably due to the location of the loudspeakers relative to the ears, while virtual sources located in front of the listener were harder to reproduce reliably.

CONCLUSIONS

A headrest virtual audio reproduction system has been implemented using crosstalk cancellation, and informal listening tests suggest that it is effective at producing an accurate virtual audio experience. Due to the natural isolation of the ears in this arrangement, the crosstalk cancellation system is able to work effectively without driving the sources too hard. Objective measurements of the effect of head motion on crosstalk cancellation suggests that such head movements severely disrupt its effectiveness, however, and the reasons why the subjective results appear to be so robust to these large changes is the subject of current investigation.

A more extensive study of the virtual audio performance through subjective testing is planned, so that the true performance of the system may be quantified. It is likely that with some optimisation of the headrest design and the inclusion of more absorbing materials the performance could be enhanced further.

Since the design of this virtual audio system has been conceived for use in public areas where more than one seat is present, one interesting area of study would be to investigate the performance of multiple discrete virtual audio systems when placed close together. The interference of the individual systems may be a problem for such scenarios as cinemas, and therefore there may be a requirement for active control (e.g. [4]) to be used in conjunction with the virtual audio to ensure adequate isolation of the seats.

References: [1] T. Takeuchi, P. A. Nelson: Optimal source distribution for binaural synthesis over loudspeakers. Journal of the Acoustical Society of America **112** (2002) pp. 2786.

[2] Skiff Loudspeaker manufactured by Role Audio, *www.roleaudio.com*

[3] Lake Huron, Hardware/Software realtime audio DSP system by Lake Technology, *www.lake.com.au*

[4] S. J. Elliott, M. Jones: An Active Headrest for Personal Audio. Journal of the Acoustical Society of America **119** (2006) pp. 2702-2709.