Experiments with Spatially Distributed Sound Sources in Real and Virtual Environments

Pedro Novo

Institute of Communication Acoustics, Ruhr-University of Bochum, D-44780 Bochum, Germany,
Email: pedro.novo@ruhr-uni-bochum.de

Introduction

The simulation of sounds produced by large crowds has produced an unexpected result. The listeners have reported that the auditory events were mainly concentrated to the left and to right, although the sound sources were equally distributed around the listeners [1]. A similar result was reported with 'dummy head' recordings of applause [2]. To investigate whether this effect was due to the use of non-individual Head-Related Transfer-Functions (HRTFs), a reverberant chamber was used to create a diffuse sound field and a test was performed with 6 listeners. Besides, an analysis of the binaural signals recorded with a 'dummy head' in the reverberant chamber was performed.

3D-Diffuse Sound Fields Tests

In a perfectly diffuse sound field the cross-correlation coefficient (CCC) between the pressure at two points, p_1 and p_2 (eq. 1) follows a *sinc* function (eq. 2),

$$\psi = \frac{\overline{p_1 p_2}}{\sqrt{\left(\overline{p_1^2}\,\overline{p_2^2}\right)}} \tag{1}$$

$$\psi(kd) = \frac{\sin(kd)}{kd} \tag{2}$$

where k is the wave number and d is the distance between the measuring points [3]. To create a diffuse sound field, the set-up described in Figure 1 was installed in the reverberant chamber. The CCC was measured with two microphones located at the listener position (distanced 30 cm apart), for three different directions, and employing decorrelated 1/3 octave band noise signals at each loudspeaker. There was a very good fitting between the measurements and the theoretical curve (eq. 2).

Broadband impulsive signals (rain sounds) were employed to perform the listening test. The 8 loudspeakers were supplied with 8 decorrelated sound signals. The listeners were provided a sheet of paper where a circle of 5 cm radius had been printed. They were asked to indicate the location of the horizontal projection of the auditory events, assuming that the centre of the circle was occupied by the listener. The results confirmed the previous observations: a concentration of the auditory events at the lateral positions (i.e. right/left of the listeners) was observed (Figure 2).

Furthermore, the sound field created in the reverberant chamber was binaurally recorded and subsequently presented to the listeners over headphones. The results of these tests further confirmed the lateral concentration of the auditory events.

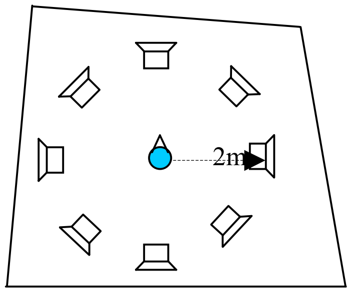

Figure 1: Set-up employed in the reverberant chamber. Eight decorrelated signals are supplied to the eight loudspeakers. The listener seats in the centre

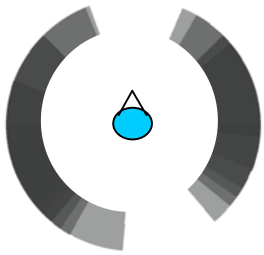

Figure 2: Location of the auditory events as reported by the listeners in the reverberant chamber. The darker regions indicate a higher incidence of answers.

Binaural Signal Analysis

Being assured that the cause of this effect was not due to the use of non-individual HRTFs, we proceeded to the analysis of the recorded binaural signals. They were filtered in critical bands and half-wave rectified, modelling processes occurring in the cochlea and in the transduction to neuronal signals [4]. A Normalized Cross-Correlation Function, NCCF, (eq. 3) was employed for the lateralization prediction within each critical band [4],

$$\Phi_{p_l p_r}(\tau) = \frac{\overline{p_l(t)p_r(t+\tau)}}{\sqrt{\left(\overline{p_l^2}\,\overline{p_r^2}\right)}} \tag{3}$$

where p_l and p_r represent the signals recorded at each ear and τ the interaural delay.

The result is shown in Figure 3 in the form of a 3D correlogram. The τ axis represents the interaural delay (in ms) and the frequency axis represents the frequency range (in Hz) being analysed. The vertical axis represents the binaural signals' NCCF values. These results represent an unweighted average of a binaural signal of 1 second duration.

For frequencies below approximately 500 Hz there is a peak in NCCF at τ = 0.0 ms. From approximately 500 Hz to 1000 Hz there are two peaks of the NCCF at approximately τ = +-0.725 ms and from approximately 1000 Hz to 1500 Hz there are three peaks: one at τ = 0.0 ms and two symmetric at approximately τ = +-0.800 ms. According to this model, a peak at τ = 0.0 ms indicates an auditory event centred at 0 degrees, in front (or in the back) of the listener and peaks at +-0.725 ms and at +-0.800 ms indicate auditory events located near +-90 degrees (i.e. to the left and right sides of the listener).

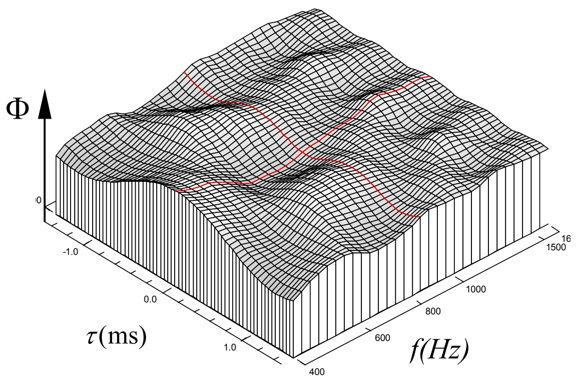

Figure 3: 3D Correlogram. τ axis - interaural delay (associated with lateralization). Frequency axis - frequency range being analysed. Vertical axis - Normalized Cross-Correlation Function, Φ, (eq. 3) of the binaural signals. (calculated and displayed with software described in [5])

Let us now turn the attention to the CCC. When the CCC is calculated using measurements performed at the entrance of the ear canals the first CCC zero occurs at approximately 500 Hz, compared to approximately 1200 Hz when the CCC is measured at the interaural distance but without the head in-between [6]. The second zero crossing occurs at approximately 1000Hz. In between the CCC takes negative values. Therefore, in the range of frequencies where the CCC (measured with head) takes positive values there is a peak of the NCCF at τ =0.0 ms and in the range where the CCC takes negative values there are two NCCF peaks symmetric to τ =0.0 ms (Figure 3). The latter is a very interesting result as it may contribute to the explanation of the lateral concentration of the auditory events previously reported. The absence of a (negative) peak of NCCF at τ = 0.0 ms for the 500 Hz - 1000 Hz range is related to the halfwave rectification neural transduction model applied in the present study.

Further correlograms were calculated employing binaural recordings of crowd sounds in a semi-anechoic environment and employing binaural recordings of noise impulses emitted by eight loudspeakers in anechoic conditions. In both cases the 'dummy head' employed for the recordings was (homogenously) surrounded by sound sources. The correlograms for these two cases follow a similar pattern to that shown in Figure 3.

It is, then, plausible that the two peaks of the NCCF occurring approximately between 500 Hz and 1000 Hz may play a role in the lateral concentration of the auditory events displayed in Figure 2. Besides, this frequency region conforms with the findings of Blauert and Lindemann [7], which report that the low frequency region adds front-back extension to the auditory image. In addition, this range fits within the range reported in Potter et. al [8] as optimal for spaciousness.

Conclusions

An unexpected lateral concentration of the auditory events was observed when simulating crowds in an auditory virtual environment. These results have led to investigations in a reverberant chamber to test if this effect was due to the use of non-individual HRTFs. The reverberant chamber test results were similar to those obtained in the virtual environment. This has led to investigations of the binaural signal employing a cross-correlation based model. The results showed that for frequencies between approximately 500Hz and 1000Hz, a region where the CCC (calculated with signals recorded at the entrance of the ear canals) takes negative values, two peaks in the NCCF occur. These peaks are located at an interaural delay of the order +-0.700/0.800msec, which in this model corresponds to a central location of the auditory events of approximately +-90 degrees, i.e. to the left and right of the listener, which may contribute to explain the lateralization observed.

References

[1] Novo, P., Korany, N. (2003) Simulation of Extended Sound Sources in Virtual Environments. In: Fortschr. Der Akustik- DAGA'03, Dtsch. Ges. Akust., D-Oldenburg

[2] Kleiner, M. Private Communication

[3] Cook, R., Waterhouse, R., Berendt, R., Edelman, S., Thompson, M. (1955) Measurements of Correlation Coefficient in Reverberant Sound Fields. J. Acoust. Soc. Am. 27 (6), pp. 1072-1077

[4] Blauert, J. (1996) Spatial Hearing: The Psychophysics of Human Sound Localization. Rev. Edition. The MIT Press, Cambridge Mass.

[5] Akeroyd, M. (2001) A Binaural Cross-Correlogram Toolbox for MATLAB. www.biols.susx.ac.uk/home/Michael_Akeroyd/

[6] Benade, A., Lindevald, I. (1986) Two-ear correlation in the statistical sound fields of rooms J. Acoust. Soc. Am. 80, pp. 661-664

[7] Blauert, J., Lindemann, W. (1986) Auditory Spaciousness: Some Further psychoachoustic analysis. J. Acoust. Soc. Am., 79, pp.806-813

[8] Potter, J., Bilsen, F., Raatgever, J. (1995) Frequency Dependence of Spaciousness. Acta acustica, vol.3, pp. 417-427