Evaluation of Spatial Information from Artificial Head and Four-Microphone Array Measurements.

Jörg Becker and Markus Sapp

Institute of Communication Engineering (IEN), Aachen University of Technology (RWTH), 52056 Aachen, Germany

Abstract: Spatial information can be evaluated with multi-microphone arrays or artificial heads. On the one hand it is easier to get the spatial information by using microphone arrays, on the other hand aurally adequate recordings can only be made with artificial heads. This paper focuses on the further development of algorithms for detecting virtual sound sources and compares evaluations of artificial head recordings with those made by a four-microphone array.

INTRODUCTION

Examinations on the acoustic qualities of enclosures are mostly restricted to the evaluation of single-number quantities like 'reverberation time' and 'center of gravity time' or criteria which rate the perceptual qualities, for example 'definition' or 'clarity index'. Most of these criteria with exception of the lateral energy fraction and the interaural cross correlation (I) are measured monaurally, although the spatial distribution of the reflections is essential for the acoustic quality of rooms. Maximum-length sequences and digital measurement allow broadband measurement with great accuracy. According to this, correlation-postprocessing to evaluate spatial information becomes more important (2), (3), (4).

EVALUATION OF SPATIAL INFORMATION FROM ARTIFICIAL HEAD MEASUREMENTS

The human sense of hearing is not able to locate a single reflection within a diffuse sound field. On the other hand a single echo can be located if it is strong enough. However in the sense of roomacoustic planning the knowledge of the origin of some acoustically important reflections is useful. Because of the broadband character of the measurement signal there is no division into critical bands (5), and simultaneous masking is not taken into account. The algorithm is divided into three steps (6). At first a preprocessing is performed by using function (1), where \( s_1(t) \) is the measured time signal of the left or right ear, \( p = 2 \) ms, \( q = 200 \) ps and \( \tau_c = 500 \) ps.

\[
 d(t) = c \cdot \frac{\int_{\tau=-p}^{\tau=q} |s_1(t + \tau)| \, d\tau}{\int_{\tau=-p}^{\tau-q} |s_1(t + \tau)| \cdot e^{-|\tau|/\tau_c} \, d\tau} \tag{1}
\]

This function takes into consideration, that the sense of hearing uses the increasing edges of the signals for locating sound sources. For \( \tau_c = 2.5 \) ms it also shows a good correlation with the performance of the auditory nerves. At each maximum of function \( d(t) \) a single reflection is extracted from the time signal \( s_1(t) \) with a window of the length 128 Samples (\( \approx 3 \) ms). Figure 1 shows that this is the smallest window length without losing characteristic structures of the interaural magnitude. The azimuth of each reflection is given at the maximum of cross correlation of the two windowed ear signals. The elevation is determined by comparing the calculated interaural magnitude of the ear signals with a set of reference interaural magnitudes, measured in an anechoic chamber at the positions (+) shown in Figure 2. The solid lines depict the orbits of constant group delay between left and right ear signal. According to the localization precision of human hearing the reference measurements are made in steps of 10° on these orbits with a distance of 10° in azimuth between the orbits, although the localization blur of the system in the azimuth is better than 10° due to the correlation-algorithm. In changing this reference data the algorithm can be adjusted to other artificial heads. The comparison between the reference interaural magnitude and the calculated magnitude is made with the method of least squares.
EVALUATION OF SPATIAL INFORMATION FROM FOUR-MICROPHONE ARRAY MEASUREMENTS

For the multi-microphone measurements a regular tetrahedron is chosen. From spherical equations the coordinates of the virtual sound sources are calculated according to (3). The microphone distance can be varied from 3 cm to 17 cm. Good results were achieved for microphone distances of 5 cm. The room impulse responses are measured with maximum-length sequences.

A preprocessing is made by lowpass filtering with a cut off frequency corresponding to the wavelength of the microphone distance. Good results can also achieved using equation (1), where \( s_i(t) \) are the time signals of each microphone. This has the advantage of selecting the aurally important reflections. To get better localization accuracy the time signals are oversampled 16 times.

After this preprocessing 6 short time correlation functions with windows from 0.5-1ms with an overlap corresponding to the microphone distance are calculated to get the delays between the signals. Backwards the maxima of the reflections referring to the calculated delays are searched in the time signals \( s_i(t) \). The redundancy of the 6 cross correlation function is used to eliminate correlation between two different reflections.

CONCLUSIONS

An algorithm for detecting virtual sound sources from artificial head measurements with a location precision true to human sense of hearing has been developed. The results in the azimuth are similar to results made with a four-microphone array. For a better determination of elevation alternative methods to least squares have to be examined. In the case of multi-microphone measurements the detection of virtual sound sources is improved by an effective algorithm for eliminating the contributions of two different reflections and adequate preprocessing.

REFERENCES

5. Zwicker, Psychoakustik, Springer-Verlag, 1982, pp. 119