STUDYING ROOM ACOUSTICS USING A MONOPOLE-DIPOLE MICROPHONE ARRAY

Durand R. Begault and Jonathan S. Abel

Poster 1pAAc11

16th International Congress on Acoustics and 135th Meeting of the Acoustical Society of America

(Poster presentation corresponding to pp. 369-370 of the Proceedings)

INTRODUCTION. Many of the current virtual acoustic ("3-D audio") displays for teleconferencing and virtual reality are limited to very simple or non-existent rendering of a diffuse sound field. Perceptual performance and the overall quality of virtual acoustic reality systems have been shown to be improved with the inclusion of simulated reverberation (e.g., Begault 1992). An accurate rendering of the spatial qualities of a measured room impulse response (RIR) is particularly useful for subsequent auralization. Research has been ongoing to devise a technique to measure the spatial aspects of RIR so that individualized head-related transfer functions can then be applied in post-processing, rather than as part of the initial measurement. Intensity measurement techniques have been proposed, and Essert (1996; 1997) recently proposed a method based on sound pressure cross-correlations.

BACKGROUND. Essert (1996) used of a combination of an omnidirectional (W) and three co-located figure-of-eight microphones to examine the directional nature of a room impulse response (oriented left-right X, back-front Y, and down-up Z, respectively; e.g., a B-format output from a *SoundField* MKV microphone). The power in the omnidirectional response $M_{Mw}(t)$ reveals the arrival time of significant early reflections; cross-correlations between the monopole and dipole responses indicate reflection direction of arrival. The ideal dipole microphone sensitivity is $a_i(\theta) = \cos(\theta - \theta_i)$, where θ_i is the axis of the dipole. So-called directional fractions between $M_{Mw}(t)$ and each of the dipole responses $M_{Dx}(t)$, $M_{Dy}(t)$, and $M_{Dz}(t)$

$$C_{DM}(t) = \sum_{t=t-d/2}^{t+d/2} M_D(t) M_M(t) \Bigg/ \sum_{t=t-d/2}^{t+d/2} M_M(t) M_M(t)$$

approximate source direction cosines with respect to the dipole axes for the dominant signal in the analysis window $[t-\delta, t+\delta]$.

EFFECT OF ANALYSIS WINDOW LENGTH. The effect of the cross-correlation/auto-correlation analysis window duration described in Essert (1996) was evaluated for capturing a reflection from a single reflective surface at t = 8 ms. A stereo condensor microphone with a rotatable capsule (Neumann USM 69i) was placed in-between a dodecahedron sound source (Brüel and Kjær 4296) and a 3' x 2' x 5/8'' panel (particle board). The speaker-to-mic distance was 62 in and the mic to the panel was 49 in (t = 4 ms). Reflections from the grass ground surface were minimal. Impulse responses corresponding to the B-format technique were obtained using Golay sequences and high-pass filtering at 500 Hz. Figures 1-2 show results for a single reflective surface, using analysis window durations of 0.5, 2 and 4 ms. Note that the reflection is only clearly detected with the 0.5 ms analysis window. The conclusion to be drawn is that it is important to match the analysis window length to the reflection duration. If the analysis window is too short, estimate variance will be needlessly increased. On the other hand, if the analysis window is too long, noise-only samples or samples including contributions from other reflections will skew the directional fractions and bias the direction of arrival estimate.

FIGURE 1

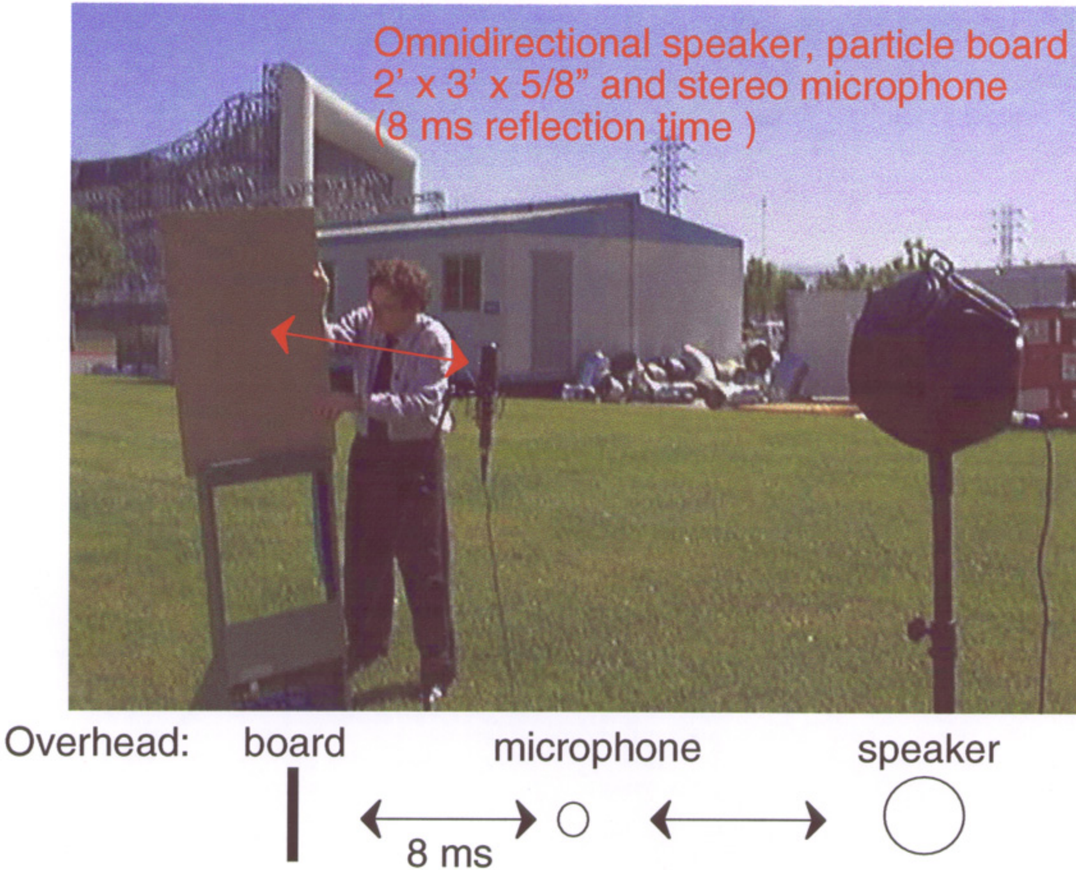
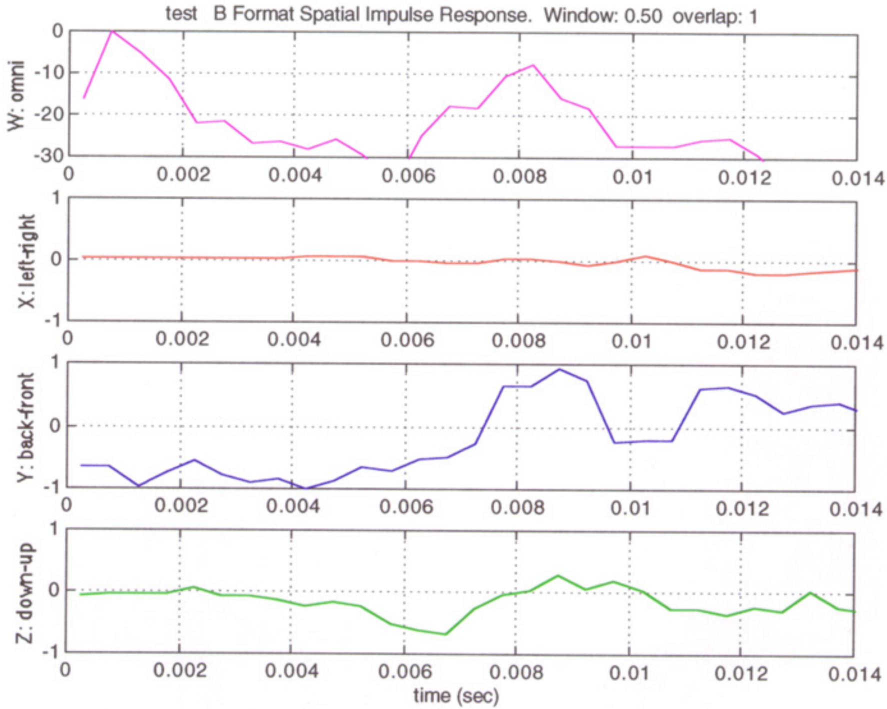
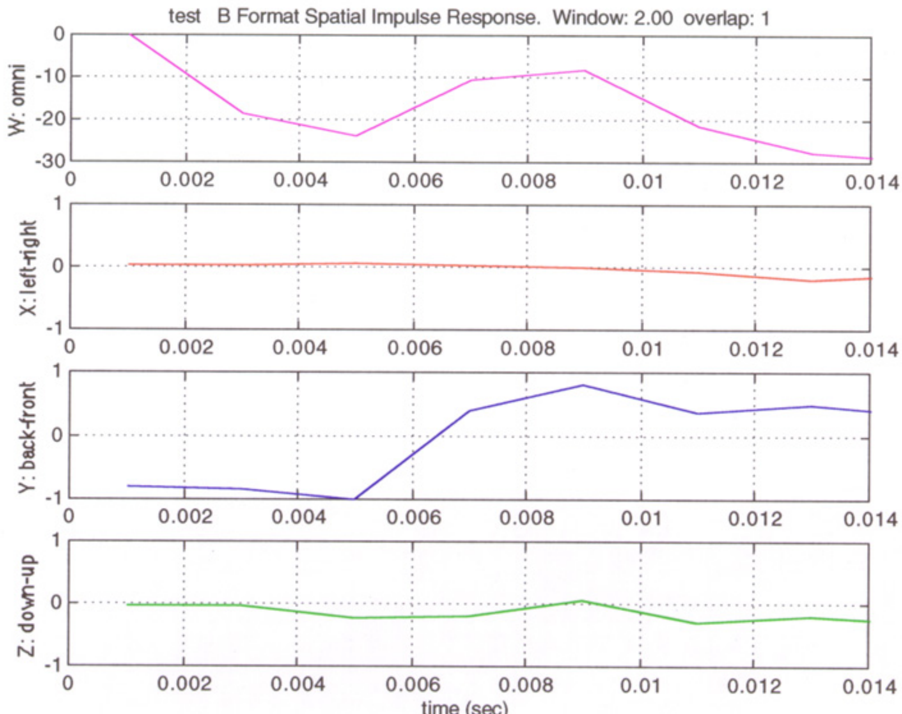

Omnidirectional speaker, particle board
2' x 3' x 5/8" and stereo microphone
(8 ms reflection time)
Overhead:
board
microphone
speaker
8 ms

FIGURE 2

0.5
ms
window

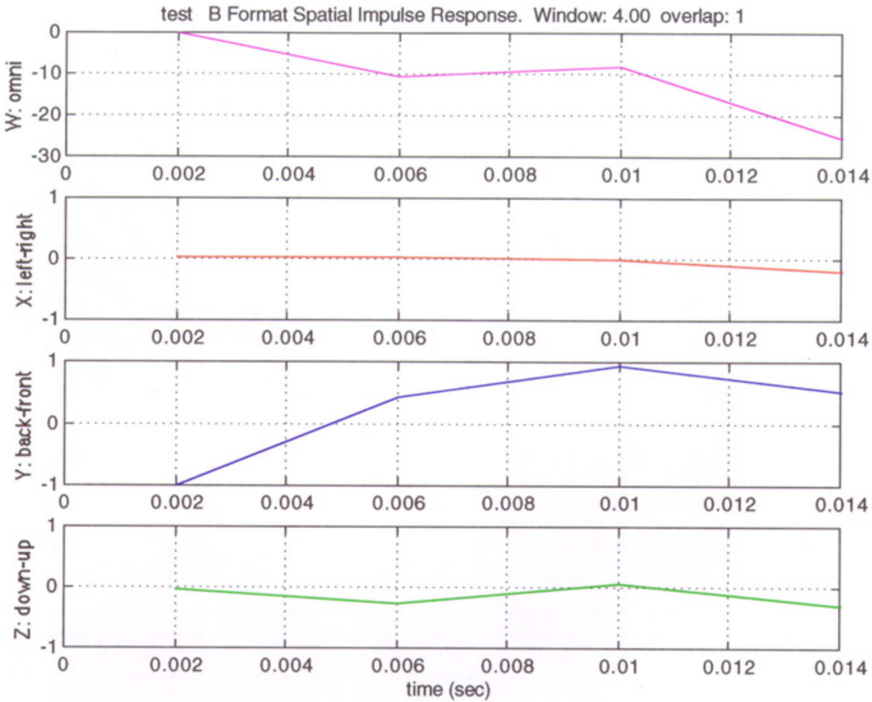


2 ms window



test B Format Spatial Impulse Response. Window: 2.00 overlap: 1
W: omni
X: left-right
Y: back-front
Z: down-up
time (sec)
0
-10
-20
-30
1
0
-1
0.002
0.004
0.006
0.008
0.01
0.012
0.014

4 ms window



test B Format Spatial Impulse Response. Window: 4.00 overlap: 1
W: omni
0
-10
-20
-30
0
0.002
0.004
0.006
0.008
0.01
0.012
0.014
X: left-right
1
0
-1
Y: back-front
Z: down-up
time (sec)

PROPOSED DETECTION ESTIMATION SCHEME. Rather than using a *single number* to describe the pattern of arrivals at any point in time, an approximation to the likelihood function indicates the presence of signal energy *at every direction of arrival* for each analysis window. The log likelihood ratio $\lambda(M; \theta)$ is approximated as the power in the weighted sum of microphone signals, where the weighting is chosen to minimize output power while passing signals arriving from the "look direction" θ unchanged.

The detector is set to a threshold ideally matched to psychoacoustic thresholds for reflections, as a function of angle of incidence and time of arrival (e.g., Olive and Toole, 1989; Begault, 1996).

ADVANTAGES:

- Improved probability of detection for a given false alarm rate (yields fewer "false" reflection detections).
- Can identify multiple arrivals that occur within a given analysis window.
- Additional dipole measurements at intermediate angles can be incorporated to improve accuracy.

MEASUREMENTS COMPARING THE PROPOSED MAXIMUM LIKELIHOOD ESTIMATOR TO THE DIRECTIONAL FRACTION ANGLE OF ARRIVAL ESTIMATOR. To evaluate the performance of the two techniques, a condenser microphone having a rotatable capsule and variable polar response patterns (Neumann USM 69i) was placed in a long narrow hallway along with a small sound source (Bose Acoustimass). The microphone and sound source were placed in the plane perpendicular to the long axis of the hallway, and with the microphone diaphragm facing the short axis. This ensured a cluster of strong early reflections arrivals within a period of 20 msec. Golay sequences were used to measure impulse responses; these were octave-band filtered between 0.5–2 kHz. Figures 3-4 show the experimental configuration. Figures 5-6 shows a comparison of the results, with a circle marking the estimated time of arrival for reflections from a ray tracing algorithm. The proposed technique (Figure 6) can be seen to exhibit less ambiguity in detection than the previous technique (Figure 5). Additional dipole measurements further sharpen the detection.

FIGURE 3


Plan view. Microphone and
speaker are in a long narrow
hallway; specular reflections
from z dimension insignficant

FIGURE 4

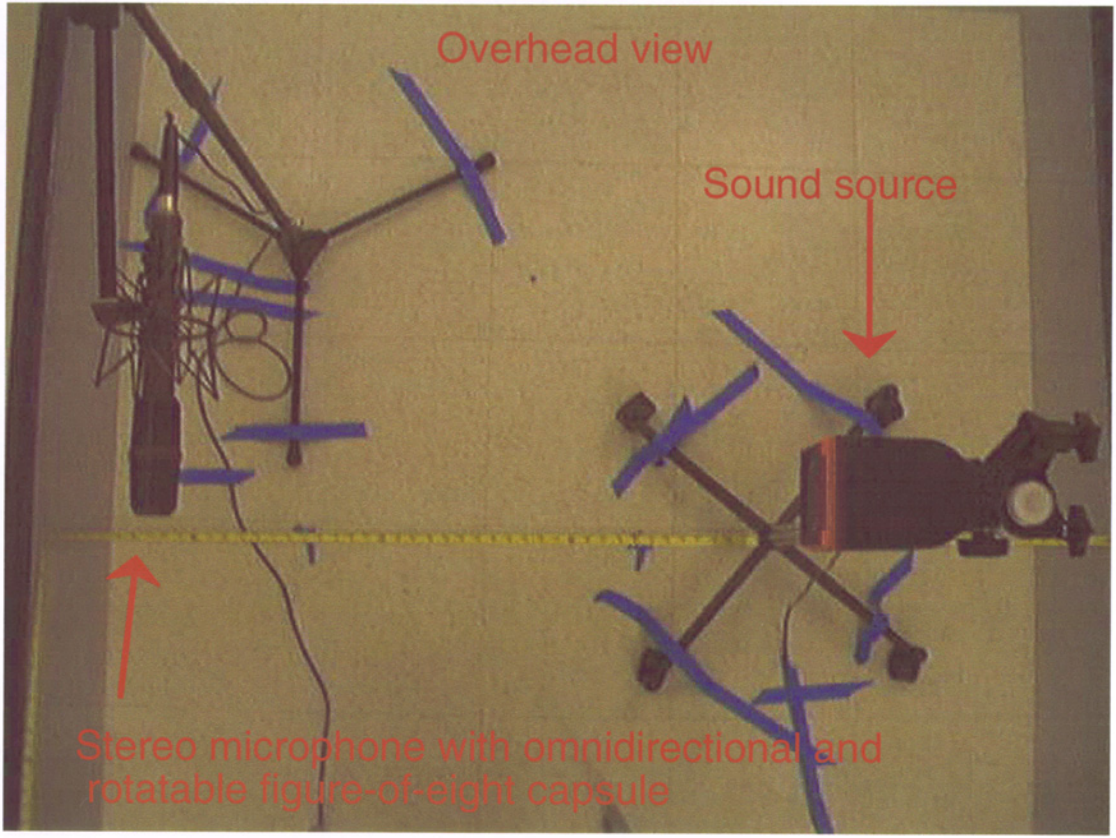
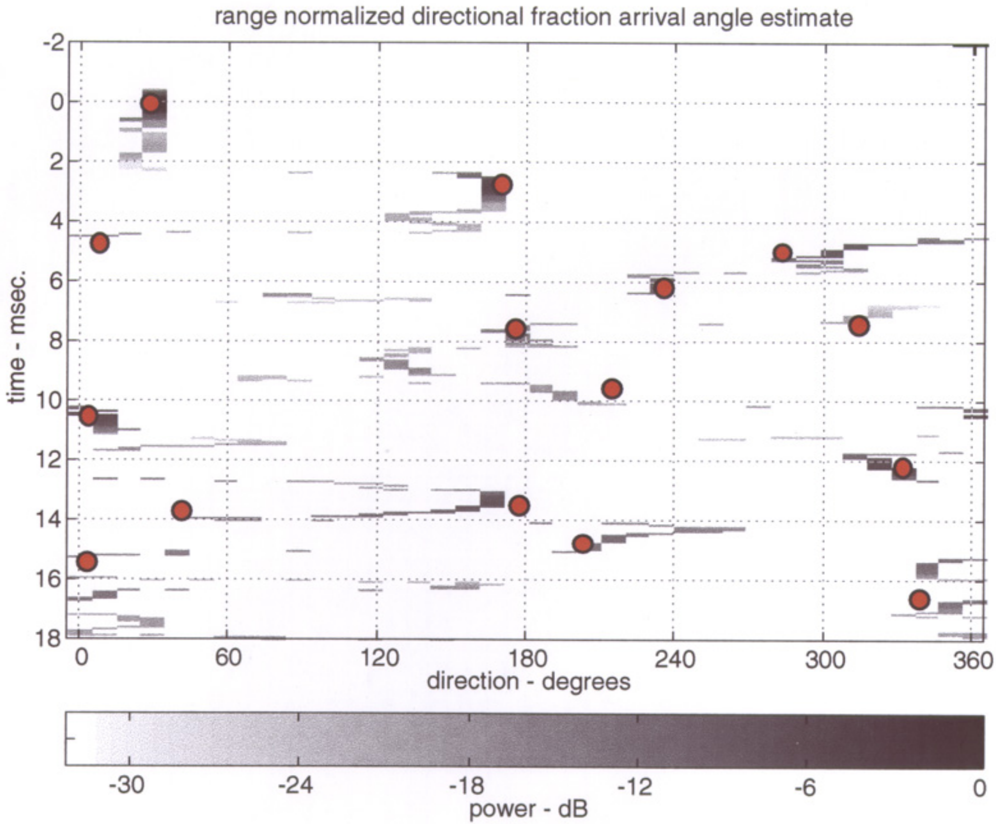

Overhead view
Sound source
Stereo microphone with omnidirectional and
rotatable figure-of-eight capsule

FIGURE 5


range normalized directional fraction arrival angle estimate
time - msec.
-2
0
2
4
6
8
10
12
14
16
18
0
60
120
180
240
300
360
direction - degrees
-30
-24
-18
-12
-6
0
power - dB