VISUALIZING THE PERFORMANCE OF LARGE-APERTURE MICROPHONE ARRAYS

Harvey F. Silverman and William R. Patterson III

Laboratory for Engineering Man/Machine Systems (LEMS)
Division of Engineering, Brown University
Providence, RI 02912

ABSTRACT

The use of arrays of microphones having a large number of elements (hundreds) is now in place in research laboratories, and will soon be practical for real applications. In rooms of auditorium or conference size, a large number of microphones will virtually always imply that the aperture will be large compared to the focal distance. This requires understanding the volume selectivity of irregular, widely-distributed sets of microphones. In this paper, several pictures of the beamforming performance of large-aperture distributions of microphones are presented. The purpose is to illustrate some of the pitfalls in large-aperture microphone-array design.

1. INTRODUCTION

Microphone-array systems have potential for high-quality audio reception for many applications. Until recently, however, it has only been feasible to use a small number of microphones (2 - 21) in real-data experiments in the research laboratory[1]. Currently, real-time support for arrays having a large number of microphones is available in the research laboratory [2, 3] and an important issue is how best to use such an array in a particular environment for a specific application.

There is little experience with these large arrays in a large room or conference room environment, although the properties of a uniformly-spaced line or two-dimensional grid of microphones are well understood[4, 5]. Even more is known when the longest linear dimension of that array is much less than the focal distance, which implies that the angle subtended by the array at the source is much less than one radian. This angle is a traditional measure of aperture size and, in the limit of a small angle, the array spatial resolution is usually the wavelength divided by this angle. However, in an interior environment, there is likely some advantage in placing microphones throughout the room, spread in three rather than two dimensions. This implies that for any two widely separated elements of the array, the angle subtended by the array may approach π radians because the array surrounds the source. This order-of-magnitude increase in aperture angle is a principal characteristic of what we call **large-aperture** systems. Another important attribute of such systems is that they are usually sparse, that is, there are too few sensors to reconstruct the acoustic wavefield within the array without spatial aliasing. This is because even with several hundred elements, the room area is still too large to be covered with element spacing comparable to $\frac{\lambda_{min}}{2}$. In all cases, spherical wavefronts must be considered, and there are tradeoffs between a smooth beampattern function and volume selectivity.

The purpose of this paper is to illustrate some of the properties of these large-aperture arrays by comparing volume selectivity performance through the beampattern function.

2. THE BEAMPATTERN FUNCTION

Consider an array of M microphones, each of unity gain and located at $\vec{r}_m = \{x_m, y_i, z_m\}$.[2] A point source at $\vec{r}_s = \{x_s, y_s, z_s\}$ radiates as

$$s(r,t) = \frac{A}{r} \cdot \exp\left[j\omega(t - r/c)\right]. \quad (1)$$

where c is the speed of sound, ω a wave's temporal frequency, and r the radial distance from the source. The signal received by the m^{th} microphone from the source alone is

$$s(\vec{r}_m,t) = \frac{A}{d_m} \cdot \exp\left[j\omega(t - d_m/c)\right]. \quad (2)$$

where $d_m = |\vec{r}_m - \vec{r}_s|$ is the Euclidean distance from the source to the m^{th} microphone. The spatial transfer function from the source to the m^{th} microphone is,

$$S(\vec{r}_m) = \frac{1}{d_m} \cdot \exp\left[-j\omega d_m/c\right]. \quad (3)$$

Delay-and-sum beamforming is the basic tool for the combination of array signals. In linear-weighted delay-and-sum beamforming, the signals received at each microphone are multiplied by an arbitrary weight, W_m, and shifted in time appropriately for coincident arrivals from a particular source. The beampattern function is defined for the near-field array as the magnitude (dB) of the weighted delay-and-sum transfer function, with each microphone specifically delayed by d_m, but evaluated at various points in the room $\vec{r}_e$. This results in the beampattern equation,

$$B(\vec{r}_e, \vec{r}_s, \omega) \equiv 10 \log \left| \sum_{m=1}^{M} \frac{W_m}{e_m} \cdot \exp\left[j\frac{\omega}{c}(e_m - d_m)\right] \right|^2, \quad (4)$$

where $e_m = |\vec{r}_m - \vec{r}_e|$ is the Euclidean distance from the evaluation point to the m^{th} microphone.

Two cases for the weights W_m are commonly used. The first weights all microphones equally with a normalizing constant to

This work partially funded by NSF grants MIP-9509505 and MIP-9314625.

[1] A 408 microphone system was installed at Murray Hill, Bell Laboratories in the mid-80's[1]. This system used hardware for its delays and, therefore, was fixed in its application.

[2] This treatment follows that in [4].

make the output at the aiming point $0dB$,

$$W_m = [\sum_{j=1}^{M} \frac{1}{d_j}]^{-1}. \tag{5}$$

The second weighting condition has weights inversely proportional to their distance to the aiming point, Equation 6. As is shown in [4], these weights maximize the array signal-to-noise ratio for equal random noise in each channel.

$$W_m = [d_m \cdot \sum_{j=1}^{M} \frac{1}{d_j^2}]^{-1}. \tag{6}$$

All computations in this paper use this second weighting method, although the small variation in $\{d_m\}$ for the focal point makes the choice relatively unimportant.

3. ROOM AND ARRAY DEFINITIONS

A top-view of a 10Mx10M room definiting three different arrays is shown in Figure 1. A top view is shown because the vertical direction, Z, is to be treated as somewhat less important; most interfering sources are likely to be side-to-side, rather than one over another. Each array system has the same number of microphones, 512, the number supported by the HMA system [2, 3, 6]. In fact, a set of 16 flat panels with 32 microphones each is being built for the HMA; learning how to distribute the microphones and panels is the motivation for this paper. All arrays are assumed to be mounted on walls on flat panels using omnidirectional microphones[3].

Array I consists of a single 4Mx0.5M flat panel in a grid with 8 rows of 64 microphones each on 6cm centers. The rows are oriented along the X axis. With this spacing, no spatial aliasing is expected for frequencies up to 2850Hz.

Array II has four panels of size 1.25Mx0.6M, a convenient size near that of the panels being built for the HMA. Each panel has 128 microphones distributed randomly on a grid to ensure that no separation is less than 3cm. The idea is to have orthogonal, separated panels similar to the arrangement in [7], with each panel giving highest resolution in the direction parallel to its major axis. Placement on the centers of the four walls of a rectangular enclosure also distributes the microphones making coverage of the room more uniform.

Array III encircles the room with a band of 1.25Mx0.6M panels with no gaps between them. Thirty-two panels of 16 microphones each are used, with a random distribution on each panel constrained to a 3cm minimum separation.

4. BEAMPATTERN RESPONSES

Figure 2 shows the beampattern along a line parallel to the X axis through the aiming point at (0.0M,5.0M,0.0M) for array I. As expected, the response is relatively smooth, the main lobe beamwidth changing inversely proportional to frequency. The height of the main lobe is about 12 dB above that of any side lobes. No spatial aliasing is evident. However, one disadvantage of a single panel along the X axis, is its response in the Y direction. This is shown in Figure 3 in which the locus of the response 6dB down from

[3]Work is ongoing to study the use of multiple channels to form directional microphones, but this is not a topic for this paper.

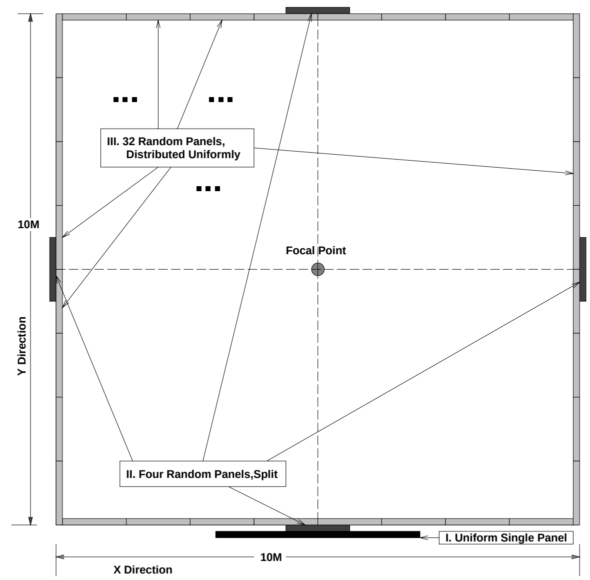

Figure 1: Top View of Room with Three Array Configurations

the peak is shown for the array at a frequency of 2kHz. The three-dimensional plot is for a room that is also ten meters in height. This height is exaggerated to preserve the aspect ratio of the response. While the beampattern of the uniform array is very narrow along the X direction, it is very wide in the Y direction. In Figure 4, an X-Y overview of Figure 3, one can see that the -6dB down region extends for about 5M in the Y direction. This is a significant problem when trying to obtain true volume selectivity.

In [7] widely separated panels were placed on orthogonal walls. Array II is a symmetric extension of this idea. The beampattern response along a line parallel to the X axis through the focal point is shown in Figures 5 and 6. Spatial selectivity is poor. These curves are best understood from the properties of the four individual panels. The magnitude of the response of any panel by itself is similar to that of Array I but is broader because the subarrays are smaller. The broad superposition of the subarray magnitudes is evident in the sidelobe loci of Figure 7. However, within the broad envelope, there is rapid spatial modulation of the response caused by interference among panels. Within the central region where all four panels contribute to the sum, Figure 5 shows variations with periodicity λ caused by the changing phase differences between orthogonal pairs of panels. Away from the center, the periodicity is $\lambda/2$ because the phase difference between panels parallel to the Y axis is proportional to twice the X distance away from the focal point. Extended sidelobes and interpanel interference will affect most arrangements that have a few discrete separated panels.

Array III addresses this problem with a uniform distribution of microphones in a band around the perimeter of the room. The beampattern along a line parallel to the X axis through the focal point is shown in Figures 8 and 9. While responses are not as regular as for Array I, the pattern near the focal point is well behaved. Moreover, the response is qualitatively circularly symmetric as can be seen in the 3D view of all points over -12dB at 1kHz shown in Figure 10. Sidelobes of this level only exist very near the walls, a region normally excluded from the working volume of the array.

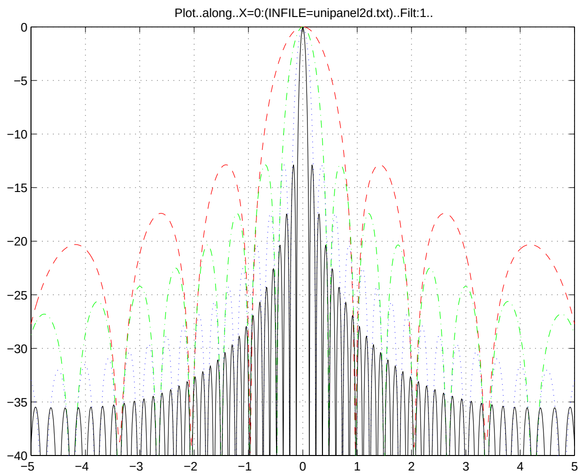

Figure 2: Response for Array I on 10M Line Parallel to X Axis through Aiming point for 500Hz, 1kHz and 4kHz

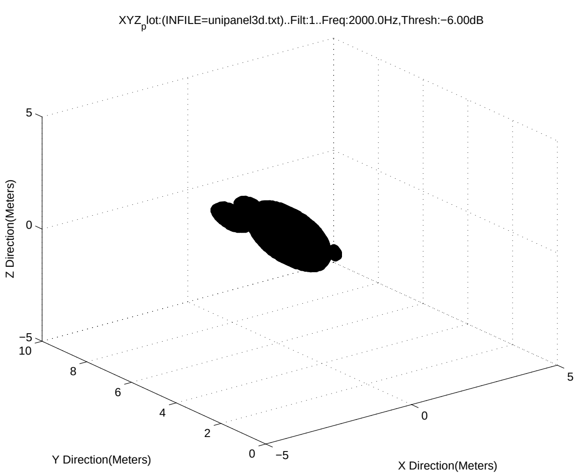

Figure 3: Locus of Points with Response Greater than -6dB for F=2kHz for Array I

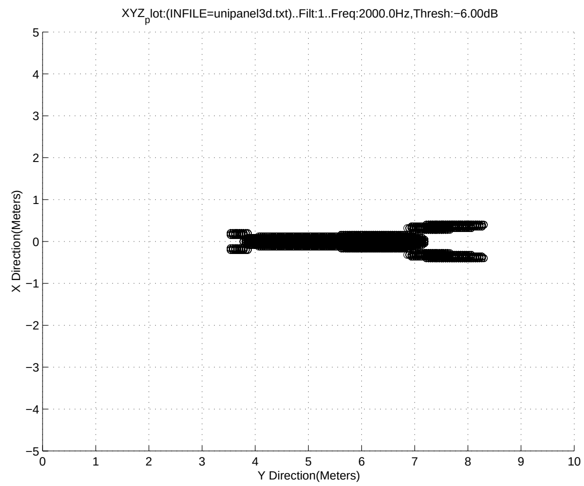

Figure 4: X-Y View of Points with Response Greater than -6dB at F=2kHz for Array I

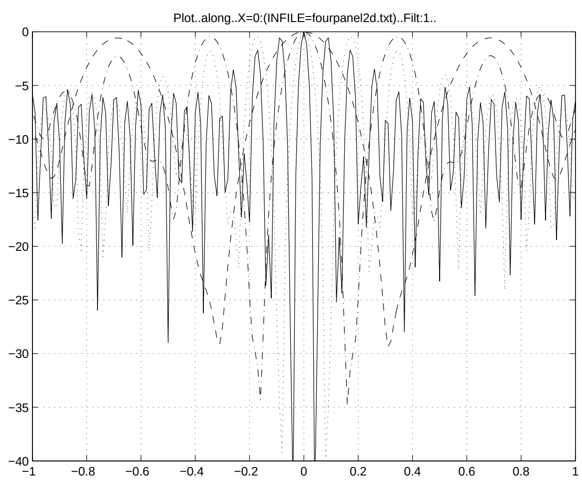

Figure 5: Response for Array II on 2M Line Parallel to X Axis through Aiming point for 500Hz, 1kHz and 4kHz

5. CONCLUSION

Given a large number of microphones, neither a single flat panel array nor a set of separated flat panels has a beampattern that is appropriate for true volume selectivity. A large-aperture system having panels or microphones uniformly distributed along the walls of a room appears to be a promising paradigm for microphone arrangements. However, an important issue remains in that the beampattern at the aiming point narrows inversely proportional to frequency and thus the volume selectivity is frequency dependent. At the higher frequencies of interest for speech ($> 1.5kHz$), the focal volume may be impractically small. Broadening it, without degrading the other properties of the array is a matter of continuing study.

6. REFERENCES

[1] J. L. Flanagan, J. D. Johnson, R. Zahn, and G. W. Elko. Computer-steered microphone arrays for sound transduction in large rooms. *J. Acoust. Soc. Amer.*, 78(5):1508–1518, November 1985.

[2] H. F. Silverman, W. R. Patterson III, J. L. Flanagan, and D. Rabinkin. A digital processing system for source location and sound capture by large microphone arrays. In *Proceedings of ICASSP-1997*, pages I–251, I–254, Munich,Germany, April 1997.

[3] H. F. Silverman, W. R. Patterson III, and J. L. Flanagan. The huge microphone array (HMA). *IEEE Transactions on Concurrency(Accepted for Publication)*, 1998.

[4] Don H. Johnson and Dan E. Dudgeon. *Array Signal Processing: Concepts and Techniques*. Prentice-Hall, Englewood Cliffs, NJ, 1993.

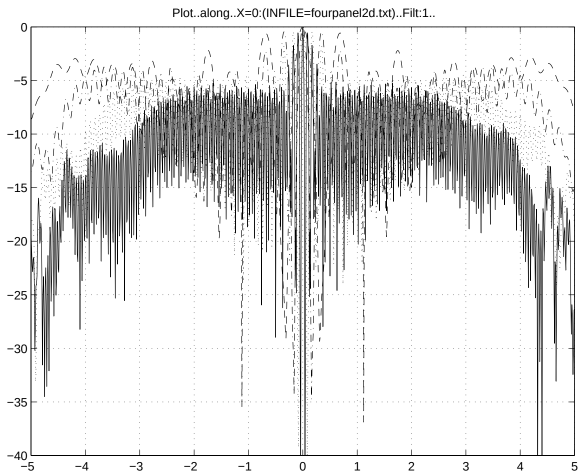

Figure 6: Response for Array II on 10M Line Parallel to X Axis through Aiming point for 500Hz, 1kHz,2kHz and 4kHz

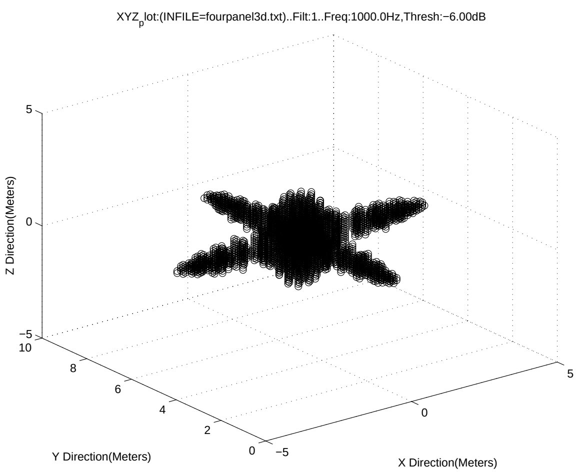

Figure 7: Locus of Points with Response Greater than -6dB at F=1kHz for Array II

[5] J. L. Flanagan. Use of acoustic filtering to control the beamwidth of steered microphone arrays. *J. Acoust. Soc. Amer.*, 78(2):423–428, August 1985.

[6] H. F. Silverman, W. R. Patterson III, and J. L. Flanagan. The huge microphone array (HMA)(abstract). *Journal of the Acoustical Society of America*, 101(5):3119, May 1997.

[7] J. Adcock, Y. Gotoh, D. J. Mashao, and H. F. Silverman. Microphone-array speech recognition via incremental MAP training. In *Proceedings of ICASSP-1996*, pages 897–900, Atlanta, GA, May 1996.

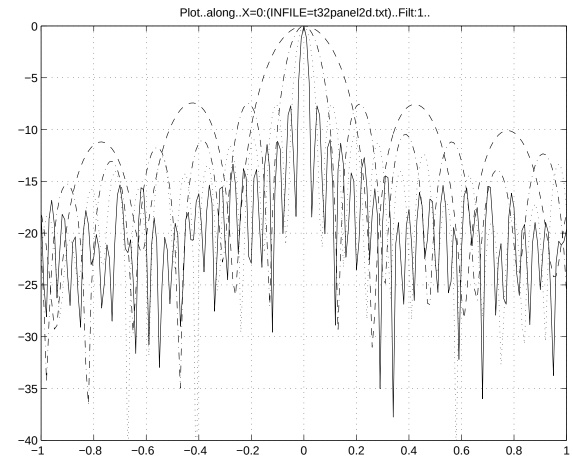

Figure 8: Response for Array III on 2M Line Parallel to X Axis through Aiming point for 500Hz, 1kHz and 4kHz

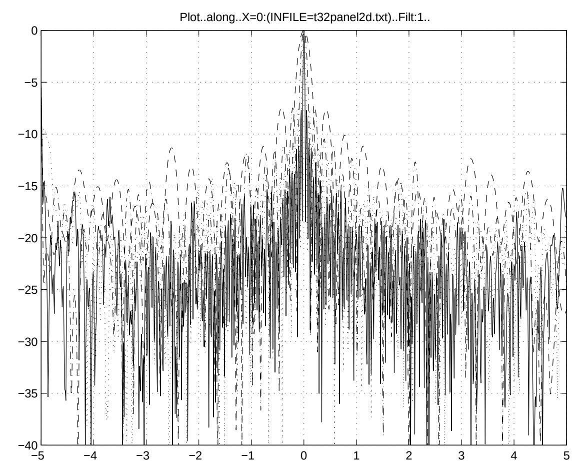

Figure 9: Response for Array III on 10M Line Parallel to X Axis through Aiming point for 500Hz, 1kHz,2kHz and 4kHz

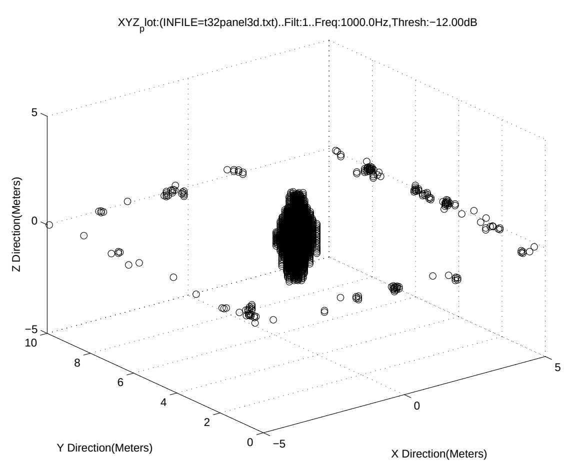

Figure 10: Locus of Points with Response Greater than -12dB at F=1kHz for Array III