

RESEARCH

Open Access

Geometry descriptors of irregular microphone arrays related to beamforming performance

Jingjing Yu* and Kevin D Donohue

Abstract

Performance analysis for microphone arrays with irregular geometries typically requires direct computation of beamforming gains over the spatial and frequency ranges of interest. However, these computations can be very consuming and limit synthesis methods for applications that require rapid answers, as in the case of surveillance and mobile platforms. A better understanding of microphone arrangements and their impact on performance can result in more efficient objective functions for optimizing array performance. This article, therefore, analyzes the relationship between irregular microphone geometries and spatial filtering performance with Monte Carlo simulations. Novel geometry descriptors are developed to capture the properties of irregular microphone distributions showing their impact on array performance. Performance metrics are computed from three-dimensional beam patterns through a delay and sum beamformer with a fixed number of microphones for irregular arrays and comparable regular arrays. Statistical analysis and Multi-way Analysis of Variance establish relationships between key performance metrics and proposed geometry descriptors. It is demonstrated that in conjunction with array centroid offset and dispersion, statistics of the microphone differential path distance can explain variations of performance metrics when steering at targets for immersive or near-field microphone applications.

Keywords: Irregular microphone arrays, array geometry, beamforming

1 Introduction

Microphone arrays use spatial diversity of element positions to capture acoustic signals and reduce degradation brought on by reverberation and noise. It is widely applied in speech enhancement, teleconferencing, talker tracking, hands-free human-machine interfaces, and acoustic surveillance systems [1,2]. Because most of these applications involve separating desired signals from noise and estimating acoustic parameters, array performance is usually assessed by its ability to locate, track, and separate sound sources in the field of view (FOV) [1]. Critical factors affecting performance include acoustic environment, source spectral content, processing algorithm, and microphone geometry. For a fixed number of microphones it has been demonstrated that the array geometry is the dominant factor for performance [3-5]. However,

previous studies have largely focused on regular geometries in far-field. These results are not as useful for immersive geometries that typically occur for surveillance and smart room applications. This article, therefore, focuses on the relationship between microphone distribution properties and spatial filtering performance that is more suited for cases when the focal point is close to the arrays and the arrays have irregular placements. Classes of irregular geometries for immersive environments are statistically analyzed through Monte Carlo simulations to identify key geometric characteristics related to array performance.

Regular arrays (elements arranged under a regular spacing constraint) have been considered in previous research, such as uniformly spaced linear, planar, and circular arrays [2]. Due to the regularity of element arrangements, their geometries are specified by a small parameter set, such as aperture and number of elements or their spacings, which are directly related to aspects of

* Correspondence: J.YU@ieee.org
Department of Electrical and Computer Engineering, University of Kentucky,
Lexington, KY 40506, USA

performance [2,6]. In general, most of these analyses have been done for narrow-band far-field cases where spatial aliasing is directly related to microphone spacing and resolution to aperture. Irregular arrays, which diversify microphone positions, can potentially achieve better performance, as demonstrated in [3,4,7]. Instead of limited optimal range of signal frequency for regular arrays, irregular arrays can result in a more consistent performance over a broader range of frequencies, such as those associated with speech [6].

Although special arrays that deviate from simple Cartesian arrangements have been studied for better performance, such as spherical arrays to capture and render sound fields [8,9], and minimum redundancy arrays to achieve maximum spatial resolution with fixed number of microphones [10-12], they still retain certain regularity of element placements and restricted by previous limitations of regular arrays. The study in this article constrains the microphone geometries to a plane, but allows for any arrangement of elements and compares geometries with similar relationships to the focal point. For example, Figure 1 shows three planar arrays with the same centroid and dispersion focused on a point 0.2 m below the array centroid. (Dispersion is analogous to aperture.) Array gains over the FOV were computed via simulation by moving a colored noise source of unit power with speech-like frequency distribution over the grid points of the FOV and then computing the received power from the beamformed focal point as described in [7,13]. The microphone positions are superimposed over their array gains showing the irregular array in Figure 1(b) having larger gains at the non-focal points than regular array in Figure 1(a), while the irregular array in Figure 1(c) shows lower gains at non-focal points. These performance differences cannot be explained by previous analyses of regular geometries. Geometric descriptions for classes of irregular geometries have not been considered for arrays lacking a regular structure. The studies [14,15] introduced optimization approaches for irregular

geometries by minimizing the residues between desired gain pattern and actual pattern computed from each microphone position. However, it is still not clear what geometric properties are crucial for the superior beamforming performance of irregular arrays. Therefore, this article proposes novel geometry descriptors with a relationship to performance that are useful for explaining the performance differences between irregular arrays (as shown in Figure 1), and provides guidelines and insight for the irregular microphone cluster design.

To relate the geometric descriptions to performance, experiments are performed using Monte Carlo simulations to analyze three-dimensional beam patterns by uniformly distributed microphones over a planar design space. Since the main applications considered for the irregular arrays involve speech (as in the case of surveillance in a cocktail party environment), the excitation of the arrays need to compute the performance metrics is colored noise with the same spectral distribution as the band importance function used in the speech intelligibility index (SII) [5]. This provides a compact summary statistic that is relevant for application where speech intelligibility is important. Results show the primary geometry factor that explains array performance is the differential path distance (DPD) distribution between microphone pairs to the target/noise locations. Statistics are derived to assess the performance and array geometry parameters with fixed number of microphones and constant target source location. Delay and sum beamforming (DSB) using an inverse distance weighting is applied to generate the array gains.

This article is organized as follows: Section 2 presents the formulations for computing array gains based on DSB and indicates its relationship with microphone distribution. Section 3 introduces geometry descriptors to characterize array geometries and derives statistics related to the DPD distribution of microphone array. Section 4 analyzes the relationship between proposed geometry descriptors and key performance metrics

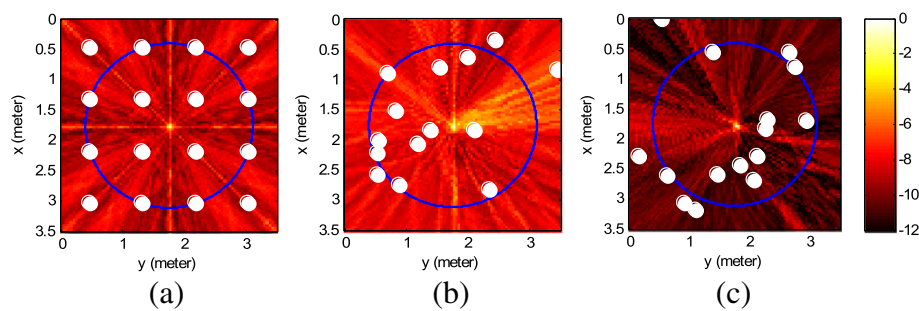


Figure 1 Gain patterns of 16 microphone arrays with the same centroid and dispersion, focused on the center of FOV where white dots represent microphone positions. Array centroid is in the center of microphone plane. Blue circle represents dispersion from array centroid. (a) Regular planar array. (b) Irregular array with inferior performance. (c) Irregular array with superior performance.

based on Monte Carlo experiments, and demonstrates a strong correlation between these descriptors and array performance. Finally, "Conclusions" section summarizes the results and presents conclusions.

2 Problem formulation

In order to reveal the impact of microphone distributions on beamforming performance, this section presents the formulations to compute the three-dimensional array gains for microphone arrays relative to a given focal point. Parametric performance metrics are directly computed from these gain patterns.

Consider microphones and sound sources distributed in a three dimensional space. Let $u(t; \mathbf{r}_s)$ be the sound source located at position \mathbf{r}_s , where \mathbf{r}_s is a vector denoting the x , y , and z coordinates. The waveform received by the p^{th} microphone can be expressed as:

$$v(t; \mathbf{r}_s, \mathbf{r}_p) \int u(t; \mathbf{r}_s, \mathbf{r}_p) = \int_{-\infty}^{\infty} u(\tau; \mathbf{r}_s) h(t - \tau; \mathbf{r}_s, \mathbf{r}_p) d\tau, \quad (1)$$

where $h(\cdot)$ represents the impulse response of propagation path from \mathbf{r}_s to \mathbf{r}_p . For a reverberant room, the impulse response can be given by:

$$h(t; \mathbf{r}_s, \mathbf{r}_p) = a_{sp0}(t - \tau_{sp0}) + \sum_{n=1}^{\infty} a_{spn}(t - \tau_{spn}), \quad (2)$$

where $a_{spn}(t)$ is the response related to the n^{th} propagation path, τ_{spn} is the corresponding time delay, and $n=0$ represents the direct path from source to microphone. In frequency domain the received signal of Equation (1) can now be expressed as:

$$\hat{V}(\omega; \mathbf{r}_s, \mathbf{r}_p) = \hat{U}(\omega; \mathbf{r}_s) \hat{A}_{sp0}(\omega) \exp(-j\omega\tau_{sp0}) + \hat{U}(\omega; \mathbf{r}_s) \sum_{n=1}^{\infty} \hat{A}_{spn}(\omega) \exp(-j\omega\tau_{spn}), \quad (3)$$

where the hat notation expresses the Fourier transform of corresponding time-domain quantity. Denote the desired focal point as \mathbf{r}_i and express the DSB output as:

$$\hat{G}(\mathbf{r}_i, \mathbf{r}_s) = \sum_{p=1}^P B_{ip} \hat{V}(\omega; \mathbf{r}_s, \mathbf{r}_p) \exp(j\omega\tau_{ip}); \quad (4)$$

where P is the total number of microphones, B_{ip} is a scalar representing the filter coefficient related to focal point \mathbf{r}_i and microphone position \mathbf{r}_p , and τ_{ip} is the corresponding time delay. For results in this article the coefficient was set to the inverse distance to the focal point as $B_{ip} = 1/d_{ip}$, where d_{ip} denotes the

distance from \mathbf{r}_i to \mathbf{r}_p . The total output power of this filtered sum is computed by:

$$S(\mathbf{r}_i, \mathbf{r}_s) = \int \sum_{p=1}^P \sum_{q=1}^P B_{ip} B_{iq} \hat{V}(\omega; \mathbf{r}_s, \mathbf{r}_p) \hat{V}^*(\omega; \mathbf{r}_s, \mathbf{r}_q) \exp(j\omega(\tau_{ip} - \tau_{iq})) d\omega. \quad (5)$$

In order to obtain the simplified formulation that is useful for analysis and understanding the geometric relationship, consider only the direct paths in Equation (3). With the assumption that the beamformer coefficients and propagation attenuation product factors are uncorrelated with the path differentials, $S(\mathbf{r}_i, \mathbf{r}_s)$ can be rewritten as:

$$S(\mathbf{r}_i, \mathbf{r}_s) = P^2 \int |\hat{U}(\omega; \mathbf{r}_s)|^2 E \left[B_{ip} B_{iq} \hat{A}_{sp}(\omega) \hat{A}_{sq}^*(\omega) \right] E \left[\exp(j\omega((\tau_{sq} - \tau_{sp}) + (\tau_{ip} - \tau_{iq}))) \right] d\omega, \quad (6)$$

where $E[\cdot]$ denotes the expected value operator over all microphone pairs generated by the double summation of Equation (5), and $S(\mathbf{r}_i, \mathbf{r}_s)$ is the output power of beamformer targeting \mathbf{r}_i with actual sound source at \mathbf{r}_s . For multi-source applications, the total output power of beamformer can be obtained from the superposition of $S(\mathbf{r}_i, \mathbf{r}_s)$ from each source. To investigate the beamforming performance in relation to array geometry, the time delays are expressed in terms of spatial distances and signal wavelengths:

$$S(\mathbf{r}_i, \mathbf{r}_s) = P^2 \int |\hat{U}(\omega; \mathbf{r}_s)|^2 E \left[B_{ip} B_{iq} \hat{A}_{sp}(\omega) \hat{A}_{sq}^*(\omega) \right] E \left[\exp \left(j2\pi \left(\frac{d_{sq} - d_{sp}}{\lambda} + \frac{d_{ip} - d_{iq}}{\lambda} \right) \right) \right] d\omega, \quad (7)$$

Where d_{sp} denotes the distance from sound source \mathbf{r}_s to microphone position \mathbf{r}_p , and d_{ip} denotes the distance from focal point \mathbf{r}_i to microphone position \mathbf{r}_p . Therefore, the formulation of beamforming power for sources in FOV is separated into three parts; the source power, propagation effects and beamforming weights, and the microphone distribution. For arrays with a fixed number of microphones and constant beamformer coefficients, $S(\mathbf{r}_i, \mathbf{r}_s)$ only depends on the exponential terms averaged over all microphone pairs, which is directly related to the microphone positions and source signal frequencies. For the case where a signal source is located at the beamformer focal point, $\mathbf{r}_s = \mathbf{r}_i$, the arguments of the exponents are all 0, and the signal is enhanced by the coherent addition of complex exponential terms. Sources not located at the focal point, $\mathbf{r}_s \neq \mathbf{r}_i$, will have reduced power due to the incoherent phases of

exponential terms. The objective in selecting a microphone distribution is to minimize the average value of the exponential terms in Equation (7) when $r_s \neq r_i$ while maximizing the average when $r_s = r_i$ for all possible target and noise positions in the FOV. Since summations will always be maximized when $r_s = r_i$ (exponential arguments are all zero), the optimization strategy can be reduced to minimize the maximum value of $S(r_i, r_s)$ when $r_s \neq r_i$ for all r_s and r_i in the FOV. A metric based on this notion is the Mainlobe-to-peak-sidelobe Ratio (MPSR), which is used in later simulations to assess performance.

Equation (7) identifies the phase terms responsible for minimizing the power gain when $r_s \neq r_i$, and is related to the source wavelength and the DPD distribution over all (p, q) microphone pairs, given by:

$$\Delta_{pq}(r_i, r_s) = (d_{sq} - d_{sp}) + (d_{ip} - d_{iq}), \quad (8)$$

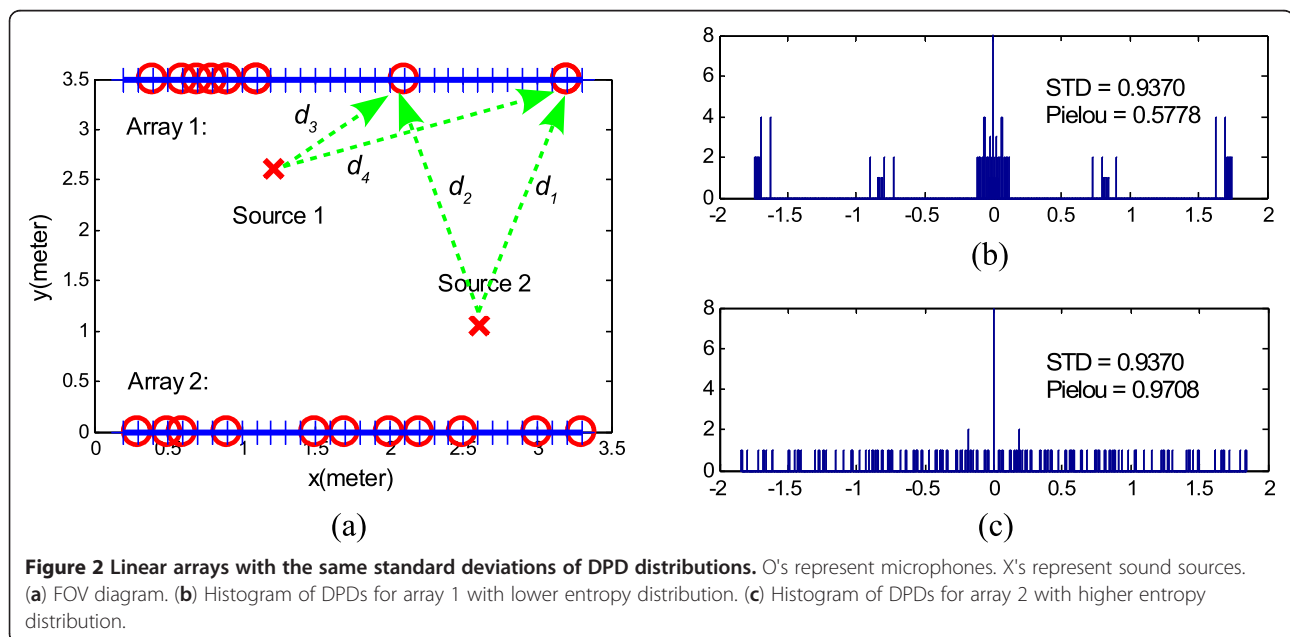
where r_i is the focal point of beamformer (target position), and r_s is the interfering source position. Note that $\Delta_{pq}(r_i, r_s)$ is the exponential argument in Equation (7) without the wavelength scaling. Take the Array 1 in Figure 2 as the example, DPD from the right microphone pair to the sources is defined as $(d_1 - d_2) + (d_3 - d_4)$.

Ideally, if the DPDs of a given microphone geometry and wavelength result in the complex exponential arguments distributed uniformly from $-\pi$ to π over all pairwise microphones, the expected power is zero when targeting r_i [16]. That is to say, in order to minimize gains for the interference/noise sources ($r_s \neq r_i$), the

corresponding DPDs should be distributed as widely as possible relative to the source wavelength (incoherence). For the case of beamforming at the source, all the phase terms in Equation (7) will be close to zero (coherent), and result in a maximum power gain in the target position. Even if the sound source localization errors result in small dislocations between the true target source position and the beamformer focal point, $r_s = r_i + \Delta r_{error} \approx r_i$, as long as the DPD variance is much smaller than the wavelengths of significant speech signal frequencies, the phases of exponential arguments are still limited to a small range and result in significant coherent sums. Therefore, Equation (7) demonstrates the impact of the DPD distribution over all microphone pairs on the array's ability to enhance target and suppress noise signals. The optimal microphone geometry should provide a widely spread and even distributed DPDs relative to the source wavelengths for the noise source positions to decorrelate the noise from target signals. Statistics assessing the uniformity of DPD distributions are proposed in the next section as the novel geometry descriptors to explain the variations of array beamforming performance, especially for irregular arrays.

3 Proposed geometry descriptors

Analysis in previous section suggests a correlation between array beamforming gains and DPD distributions. This section proposes several geometric characterizations applicable to irregular arrays and related to array performance. In addition, descriptors for regular arrays, such as the aperture size and microphone spacings, will be generalized for irregular array geometries.



The array centroid offset is defined as the distance between array focal point $\mathbf{r}_i = (x_i, y_i, z_i)$ and the centroid of array elements given by:

$$L = \sqrt{(x_0 - x_i)^2 + (y_0 - y_i)^2 + (z_0 - z_i)^2}, \quad (9)$$

where $\mathbf{r}_0 = (x_0, y_0, z_0)$ denotes array centroid:

$$\mathbf{r}_0 = (x_0, y_0, z_0) = \left(\frac{1}{P} \sum_{p=1}^P x_p, \frac{1}{P} \sum_{p=1}^P y_p, \frac{1}{P} \sum_{p=1}^P z_p \right), \quad (10)$$

where P is the number of microphones and $\mathbf{r}_p = (x_p, y_p, z_p)$ denotes the position of the p^{th} microphone. Array dispersion, analogous to the aperture size, is a measure of average microphone spread about the centroid, computed by:

$$a = \sqrt{\frac{1}{P} \sum_{p=1}^P \left[(x_p - x_0)^2 + (y_p - y_0)^2 + (z_p - z_0)^2 \right]} \quad (11)$$

Note that L and a can be applied to characterize both regular and irregular geometries, as shown in Figure 1. For regular arrays a directly impacts resolution (main-lobe width MLW), and determines the microphone spacing in conjunction with P , which affects sidelobe behavior. The distance L indicates whether sound sources are effectively located in the near-field (*small* L for immersive application), or far-field (*large* L), where the terms *small* and *large* are used relative to the source wavelengths. However, as illustrated by the examples in the introduction these descriptors are limited in their ability to explain the beamforming behavior when additional degrees of freedom are allowed as in the case of irregular arrays. Therefore, additional descriptors involving DPD distribution for all microphone pairs to points in the FOV are proposed as metrics.

From the analysis of Section 2 a limited DPD distribution increases the likelihood of unexpected coherence at non-target locations, especially when DPDs are less than a quarter wavelengths at significant signal frequencies. DPD distributions can be examined via histograms or characterized with various statistics. One potentially useful statistic is the standard deviation of the DPDs over all microphone pairs. In [16,17] closed form expressions were presented for the expected value of the exponential terms in Equation (7). For a normal DPD distribution over all

microphone pairs, the expected value of the exponential term is given by:

$$\begin{aligned} & \mathbb{E} \left[\exp \left(j2\pi \left(\frac{\Delta_{pq}(\mathbf{r}_i, \mathbf{r}_s)}{\lambda} \right) \right) \right] \\ &= \exp \left(-2 \left(\pi \frac{\sigma_{\Delta}(\mathbf{r}_i, \mathbf{r}_s)}{\lambda} \right)^2 \right), \end{aligned} \quad (12)$$

where σ_{Δ} represents the DPD standard deviation. If the DPDs are uniformly distributed, the expected value becomes

$$\begin{aligned} & \mathbb{E} \left[\exp \left(j2\pi \left(\frac{\Delta_{pq}(\mathbf{r}_i, \mathbf{r}_s)}{\lambda} \right) \right) \right] \\ &= \text{sinc} \left(\pi \frac{\sqrt{12} \sigma_{\Delta}(\mathbf{r}_i, \mathbf{r}_s)}{\lambda} \right). \end{aligned} \quad (13)$$

In both cases the expected value of the exponential terms approaches zero for increasing σ_{Δ} . When $\mathbf{r}_i = \mathbf{r}_s$, the DPDs are zero for all microphone pairs resulting an a DPD variance of 0. Thus, the scaling provided by the DPD exponential factor of Equation (7) is at a maximum of 1, which is desired when the source and focal point are identical. Consistent with previous conclusion, the more widely spread of DPDs (large σ_{Δ}), the better ability of the array to extract target signal at \mathbf{r}_i and decorrelate signals from noise source at \mathbf{r}_s . Therefore, standard deviation is applied as an effective measure that can describe performance of irregular arrays. For particular focal and noise source locations, the DPD standard deviation is computed as:

$$\sigma_{\Delta}(\mathbf{r}_i, \mathbf{r}_s) = \sqrt{\frac{1}{P^2} \sum_{p=1}^P \sum_{q=1}^P (\Delta_{pq}(\mathbf{r}_i, \mathbf{r}_s))^2}, \quad (14)$$

In addition, with the same standard deviation, the expected value of the exponential terms approaches zero for decreasing λ , representing better noise suppressing ability for the signals in high-frequency bands. Wider spread of DPDs is needed to decorrelate the signal source with low frequencies, such as male voice. In this article, in order to focus on the impact of DPDs derived from array geometry, colored noise generated by SII mode is applied as the excitation of the simulations to compute performance metrics.

From Equation (12) and (13), different DPD distributions can also impact the incoherence level of beamforming. Figure 2 provides a real case example of linear arrays. Figure 2(a) shows two linear arrays in a planar FOV with microphone positions denoted by O markers. Two sound sources represented by X markers are located in the FOV, while one source is considered as the target (focal point of beamformer) and the other is the noise source. Colored noise from each source is recorded

separately, and the received signals of microphones are normalized by the average rms value over all channels and superimposed. The Signal-to-noise Ratio (SNR) is computed as the power ratio of beamformed signal from target source over that from noise source. The DPD histograms of both arrays are shown in Figure 2(b), (c), respectively. The beamforming SNR results are provided in Table 1. An analogous simulation of the array recording was also performed and presented in Table 1. For both the real and simulated recordings it can be seen that although these two arrays have the same σ_{Δ} , array 2 shows a 2 - 3 dB SNR improvement over array 1 for both targets due to the reason that array 2 provides a more uniform DPD distribution over the source spectrum, thus demonstrating a need for another statistic related to DPD diversity. In this article Pielou's evenness index [18], which is a normalized Shannon entropy, is introduced to numerically assess the diversity of DPD distribution as:

$$\begin{aligned}
 J(r_i, r_s) &= \frac{H(r_i, r_s)}{H_{\max}(r_i, r_s)} = \frac{-\sum_{k=1}^K (p_k \ln p_k)}{-\sum_{k=1}^K (\frac{1}{K} \ln(\frac{1}{K}))} \\
 &= \frac{-\sum_{k=1}^K (p_k \ln p_k)}{\ln K}, \quad (15)
 \end{aligned}$$

where K is the total number of DPD bins for the histogram estimate, and p_k is the percentage of DPDs within the k^{th} bin, $H(r_i, r_s)$ is the Shannon entropy, and $H_{\max}(r_i, r_s)$ is the maximum possible entropy for the given number of bins, which represents a ideal uniform distribution of DPDs. This normalization avoids the variations from different ranges of DPD distributions and different numbers of microphones. Note that, the DPD range is binned by constant intervals whose size should be associated with the quarter wavelengths of significant signal frequencies to result in reasonably smooth histograms of DPDs related to the incoherent level of phase terms of beamforming gain. For the results in this article, bin size is set to 0.1m, which is less than a quarter wavelength of the important frequency band around 800 Hz for male voice intelligibility [19].

Therefore, four geometry descriptors $\{L, a, \sigma, J\}$ are proposed to characterize both regular and irregular

Table 1 SNR results of linear arrays

		Target at source1	Target at source 2
Simulations	Array 1	7.70 dB	7.28 dB
	Array 2	10.20 dB	9.11 dB
Real Recordings	Array 1	4.10 dB	3.52 dB
	Array 2	6.33 dB	6.48 dB

Table 2 Dependencies of geometry descriptors

	Mic coordinates	Target space	Noise space
L	√	√	
a	√		
σ	√	√	√
J	√	√	√

microphone distributions and study their impact on array beamforming performance. As summarized in Table 2, these descriptors depend on various geometric aspects of the application environment. Descriptors $\{L, a\}$ are related to microphone coordinates or beamforming focal point. They are usually applied together as a basis for comparing similar arrays. The descriptors $\{\sigma, J\}$ can vary with each array geometry instance and also depend on the characteristics of possible target and noise source distributions. This dependency brings the expectation of stronger correlation with array performance based on different acoustic scenes. In the next section, these proposed geometry descriptors are applied

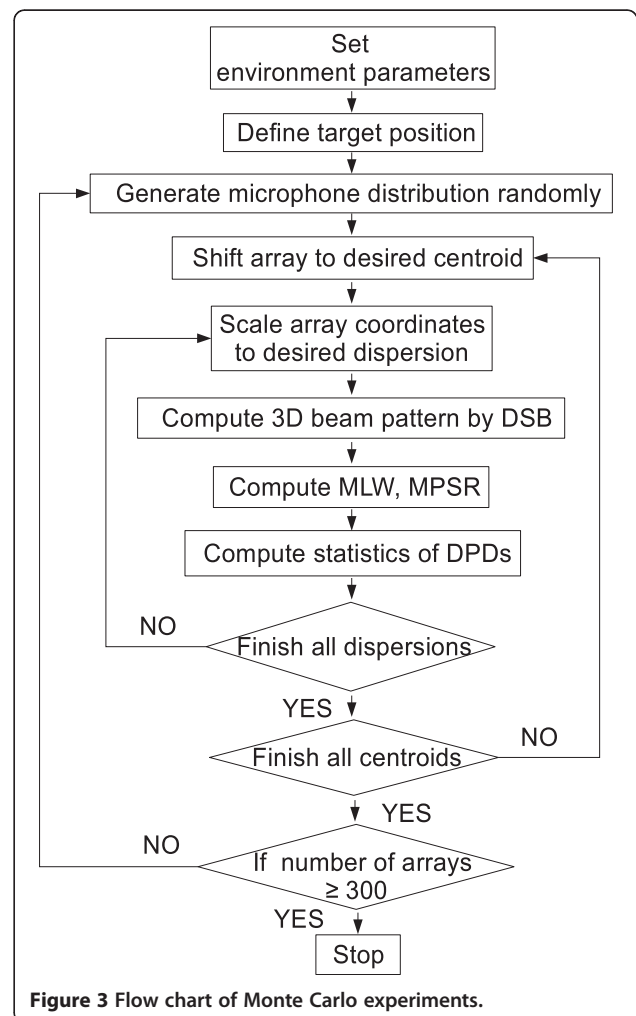


Figure 3 Flow chart of Monte Carlo experiments.

together to characterize different stochastic array geometries, and their relationships with key performance metrics of three-dimensional beam pattern are analyzed with Monte Carlo simulations.

4 Numerical simulations

4.1 Experimental setup

This section applies Monte Carlo experiments to evaluate the performance of irregular arrays based on the proposed geometry descriptors. The simulation flow chart is shown in Figure 3. The FOV is a $10 \times 10 \times 2$ m room and microphone positions are randomly generated with a uniformly distribution on the ceiling plane. Then the microphone coordinates are shifted and scaled to obtain the desired array centroid and dispersion. Array centroid values range from the center of ceiling to the edge at 1m intervals along x-axis, while five levels of dispersion are applied with each centroid. For each combination of centroid offset and dispersion level, 300 independent array distributions are generated by Monte Carlo experiments. The 3D beam pattern of each array is obtained by moving a sound source with constant power over all spatial points in FOV, while the focal point is fixed in the center of room. The DSB output power is computed for the source at each spatial point to form the 3D beam pattern.

As shown in Figure 3, two metrics are applied to assess array performance: (MLW) associated with resolution, and (MPSR) associated with noise suppression ability. In this article, the size of mainlobe is characterized by the dimensions of the surface consisting of spatial points with gains 3 dB below that at the focal point (maximum gain). Let x_δ , y_δ , and z_δ denote the projections of the 3dB mainlobe contour onto the x, y, and z axes, respectively. The MLW can then be expressed as:

$$B_{3\text{dB}} = \sqrt{(x_\delta^2 + y_\delta^2 + z_\delta^2)}. \quad (16)$$

Let $S(\mathbf{r}_i, \mathbf{r}_s)$ denote the power gain of the beamformer focused on \mathbf{r}_i with a unit power source at \mathbf{r}_s . The MPSR for a beamformer focused on \mathbf{r}_i can be expressed as:

$$\Gamma_i = \frac{S(\mathbf{r}_i, \mathbf{r}_i)}{\max_{\mathbf{r}_s \notin \text{ML}} [S_o(\mathbf{r}_i, \mathbf{r}_s)]}, \quad (17)$$

where $S_o(\mathbf{r}_i, \mathbf{r}_s)$ denotes the local maxima of $S(\mathbf{r}_i, \mathbf{r}_s)$ outside the 3dB mainlobe region (ML) in FOV. This metric represents the worst case leakage. In the Monte Carlo experiments, the maximum sidelobe level (denominator) is measured from the maximum local peak of gain pattern outside ML. Because there is normally a tradeoff between $B_{3\text{dB}}$ and Γ_i , the common criterion to decide the optimal array beam pattern is to limit the MLW to a tolerable spatial resolution and maximize the MPSR in

FOV. For given number of microphones, increasing dispersion about array centroid tends to result in higher sidelobe levels while sharpening the mainlobe. However, for a given class of randomly generated arrays with fixed centroid and dispersion, the sidelobe levels will also vary based on the DPD statistics.

In addition to array geometry, the signal frequency and the number of microphones are critical factors impacting performance. To make results more reflective for the performance where the primary sources are speech, the sound sources consist of colored noise with a spectrum equivalent to the band importance function from the SII, which emphasizes the frequency bands most important to human understanding of speech [5]. Because the impact of each geometry descriptor also depends on microphone number, irregular arrays with 16, 25, 36, 49 and 64 microphones are examined with comparable regular arrays and logarithmic arrays. The logarithmic array consists of three superimposed regular subarrays used for octaves from 800Hz to 3200Hz to generate a relative uniform frequency response over the important frequency bands. Statistical analyses of simulation results are presented in the next section to assess the impacts of proposed geometry descriptors and demonstrate their relationship with performance metrics in immersive or near-field applications.

4.2 Results and discussion

Plots from Monte Carlo simulations are presented to reveal relationships between each geometry descriptor and performance metrics. Figures 4, 5, 6 and 7 present the geometry descriptors versus MLW and MPSR, where the error bars span ± 1 standard deviation about the mean. For comparison sake a regular planar array and logarithmically spaced array with the same geometry descriptors are also marked in the figures.

Figure 4 indicates the impact of centroid offset on array performance. From Figures 4(a)(b), it can be seen that for fixed array dispersion, increasing the centroid offset increases the MLW and reduces MPSR, representing degradation of array performance. The standard deviation of MLW increases with the growing of centroid offset, while ± 1 dB variance of MPSR is observed for each centroid offset value with fixed dispersion. Logarithmic arrays show much larger increases in MLW than regular and irregular arrays because the microphone density is high near array centroid causing a longer mainlobe in the direction of the offset. Although better MPSR can be observed for logarithmic arrays with large centroid offset, it does not necessarily represent superior ability to suppress non-target sources. The lower sidelobe levels are primarily the result of FOV being included in a huge mainlobe. Therefore, logarithmic array has a major limitation on target space, and cannot adjust well to focal

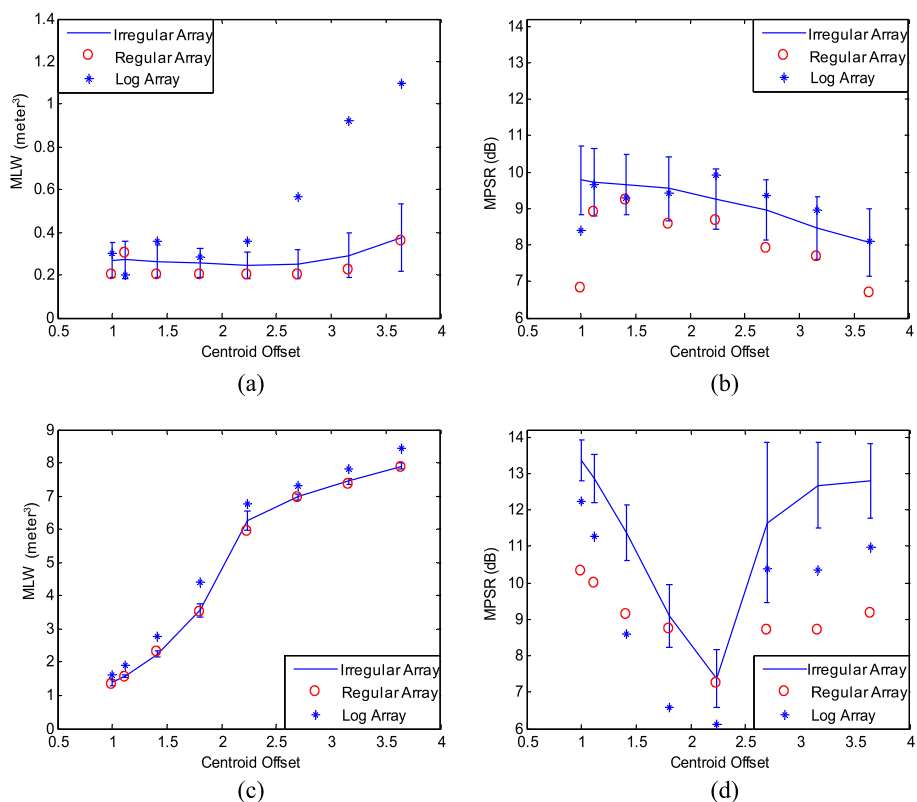


Figure 4 Centroid offset (in meters) and performance metrics with fixed dispersions, showing error bars at ± 1 standard deviation. (a) MLW for a dispersion of 3.5 m. (b) MPSR for a dispersion of 3.5 m. (c) MLW for a dispersion of 0.5 m. (d) MPSR for a dispersion of 0.5 m.

points away from array centroid. Figures 4(c)(d) show variations of performance metrics along centroid offset when dispersion is fixed at a small value. For the centroid offset values below 2.5m, the trends of MLW and MPSR over centroid offset levels are as expected with more sensitivity for arrays with smaller dispersion (microphones closer together on average) when compared to Figures 4(a)(b). For the centroid offset values beyond 2.5m (exceeding five times that of the dispersion), the

MLW becomes very large relative to the size of FOV. The apparent improvement in the MPSR after this is artificial because the mainlobe dominates the FOV pushing the significant sidelobes outside the FOV. The observed high MPSR values, therefore, cannot be associated with superior beamforming performance when the centroid offset is large relative to the dispersion. In every case there is a significant portion of randomly generated arrays that perform better than the logarithmic and

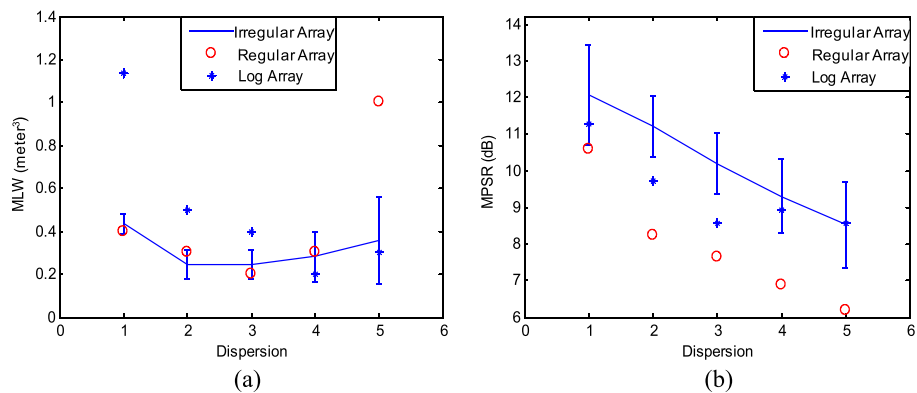
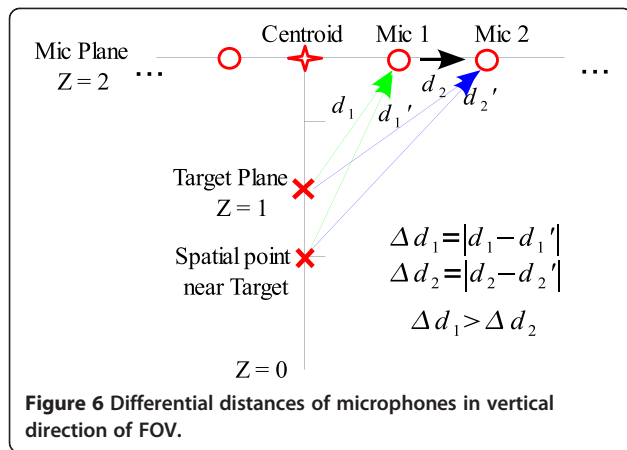


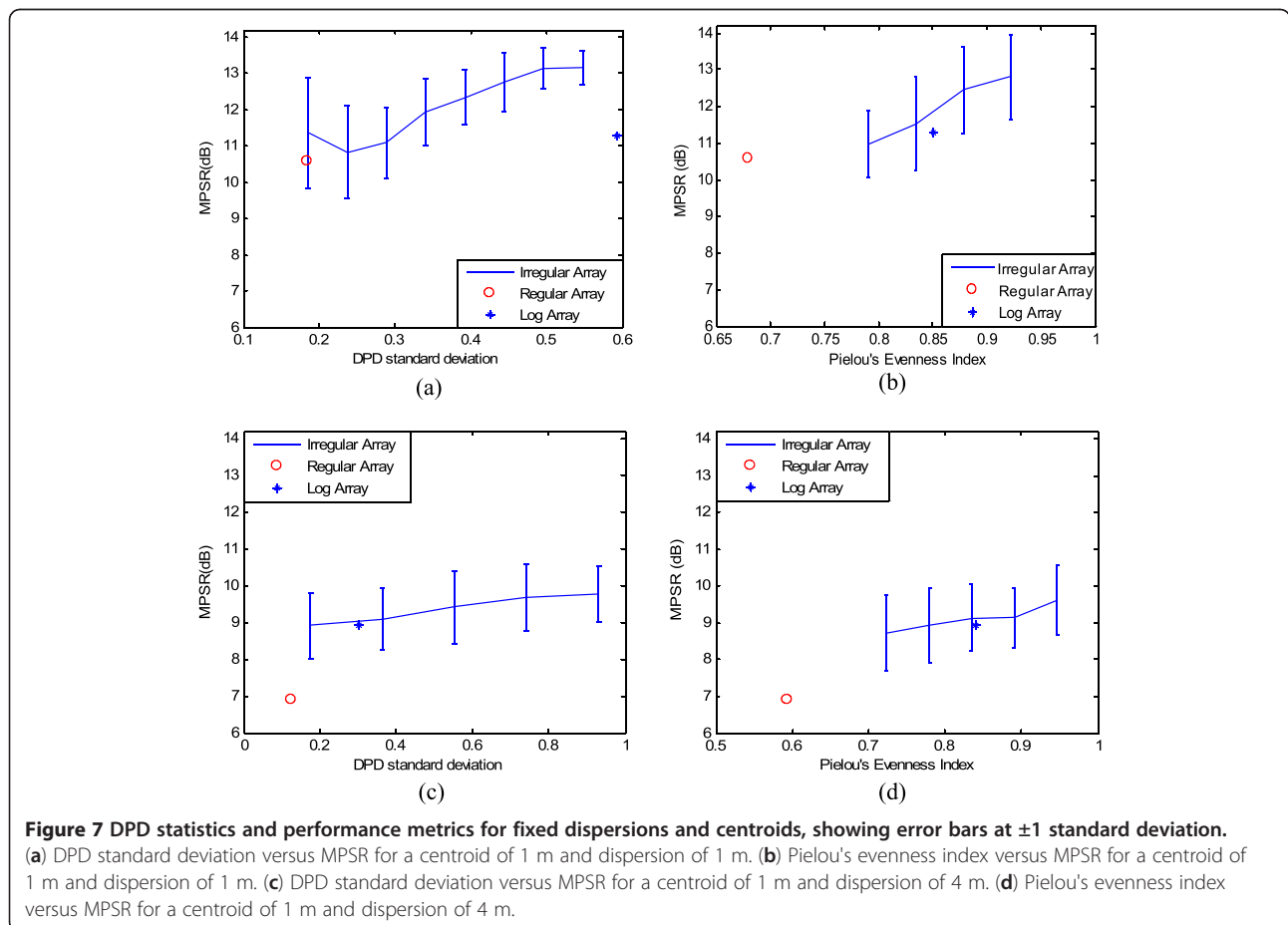
Figure 5 Array dispersion (in meters) and performance metrics with array centroid at the center of ceiling and centroid offset equal to 1 m, showing error bars at ± 1 standard deviation. (a) Dispersion versus MLW. (b) Dispersion versus MPSR.



regular arrays as seen by their marker positions relative to the standard deviation range of the irregular arrays.

Figure 5 presents the impact of array dispersion for a fixed centroid at the center of ceiling. It can be noted that small dispersions result in better MPSR for all geometries (closer average spacings between microphones); however, most of the irregular arrays perform better than either the regular or logarithmic arrays. With the centroid offset fixed, when array dispersion increases in the

horizontal microphone plane, the MLW decreases along the horizontal direction; however, the MLW along vertical direction grows. This phenomenon is illustrated in Figure 6. When moving microphones away from the array centroid/target, the differential distances from each microphone to target point and the nearby locations reduce, resulting in higher coherent power for these points in Z-direction, thus extending the mainlobe. The sensitivity of these variations to dispersion is inversely related to the centroid offset. As the centroid offset becomes large relative to the dispersion, beamforming on a focal point is not practical (no longer an immersive environment). The array takes on more characteristics of a far-field array where the vertical direction MLW is so large that one only considers the angle or look direction instead of a focal point. In summary, for a fixed number of microphones there is tradeoff between MLW and MPSR that is dependent on the dispersion, as would be expected given the similarities between dispersion and aperture. In addition, by inspecting the standard deviation of error bars along each level of dispersion when array centroid is fixed, it can be seen that the variance of MLW increases with growing dispersion. A MPSR variance of $\pm 1 \sim 1.5$ dB is



observed for each dispersion level with fixed centroid. Therefore, additional geometry descriptors based on the DPD distribution are expected to explain part of these variations of array performance.

Results in Figures 4 and 5 demonstrate the impact of geometry descriptors related to aperture and array distance from focal points, which are largely consistent with the expectations. In all cases a portion of the randomly generated irregular arrays was superior to the regular arrays. In order to resolve between classes of irregular arrays, the following paragraphs analyze geometry descriptors based on DPD statistics with fixed centroids and dispersions, and demonstrate their ability to identify classes of irregular geometries with similar performance properties.

For a fixed centroid offset and dispersion, Figure 7 shows a relationship between array geometry DPD statistics and performance. Figure 7(a)(b) presents the results for the arrays with similar centroid offset and dispersion values, while Figure 7(c)(d) presents arrays with small centroid offset and large dispersion. The results for the regular and logarithmic arrays are also plotted for reference. Figures 7(a)(b) demonstrate that larger DPD standard deviations and Pielou's evenness indices result in improved MPSR. These results are consistent with theoretical analysis indicating that wider and more evenly distributed DPDs create more incoherence in the phase terms of Equation (7) and suppress noise better. Pielou's evenness index shows more sensitivity to the MPSR than the standard deviation, primarily because with a fixed dispersion, the standard deviation has limited range. Note that the relative performance of logarithmic array in Figure 7(a) shows it with a very high standard deviation but not consistent with the trends of the irregular array, while for Pielou's index the MPSR of both the regular and logarithmic array are more consistent with irregular array performances.

When the array dispersion becomes much larger than centroid offset in Figure 7(c)(d), improvements of MPSR with increasing standard deviation or Pielou's index are not as dramatic. That is because arrays with large dispersion and small centroid offset typically generate a large DPD distribution spread (demonstrated by the increasing range of DPD standard deviation in Figure 7(c)) extending over many wavelengths in the useful frequency range. In these cases, Pielou's evenness index does not correlate as well with the beamforming gain as in Figure 7(b) because the 2π modularity of the exponential argument. For a frequency of interest, the DPDs scaled by the wavelength are mapped to the $[-\pi, \pi]$ range by the modulo operation. The evenness index can be computed after this operation for frequency specific measures related to beamforming gains. In addition, results of Figure 7(c)(d) show that almost any irregular

distribution will perform better than the regular geometry, and approximately 50% will perform better than the logarithmic array. Also, the relative performance of regular and logarithmic arrays is more consistent with the trends of the irregular array according to Pielou's evenness index than to standard deviation.

When the centroid offset becomes larger than three times of dispersion, the array takes on more characteristics of a far-field application. These cases do not fit with the primary focus of this analysis for immersive environments. The DPD variations are limited and inappreciable over the FOV relative to the signal wavelength and large centroid offset (indicated by the observed dropping range of Pielou's evenness indices). Variations in the microphone distributions will have little impact on performance, unlike for near-field applications. Centroid offset becomes the dominating factor affecting array beamforming performance, and the behavior of microphone array approaches the behavior of a single element in these far-field cases.

The results analyzed above demonstrate the impact of DPD distribution on array beamforming performance. Geometry descriptors based on the statistics of DPD distribution show a correlation with array performance when the focal points and microphone distributions are typical for immersive or near-field applications. These DPD statistics explained the variations in performance when array centroid offset and dispersion were fixed. For a fixed number of microphones, increases in dispersion improved resolution, but degraded noise suppression, while increases in centroid offset degraded both of these performance metrics. However, as shown in Figure 7, with fixed centroid and dispersion, $\pm 0.5 \sim 1$ dB variances of performance metrics are observed for each bin of DPD statistics. Although these variations of performance metrics partly result from the quantization errors of DPD statistics, other geometry parameters may exist that can further reduce these variations.

To further investigate the significance of the proposed geometry descriptors' impact on performance, Analysis of Variance (ANOVA) is applied, which is useful for investigating the effect of independent factors on

Table 3 Three-way ANOVA results of MLW

	Factors	F Value	p value (Pr > F)
Main Effects	Centroid Offset	1955.98	<0.0001
	Dispersion	3327.62	<0.0001
	Pielou's Evenness Index	2.20	0.0031
Interactions	Centroid*Dispersion	2725.46	<0.0001
	Centroid*Pielou	1.63	0.0004
	Dispersion*Pielou	2.20	<0.0001
R-Square		0.9967	

Table 4 Three-way ANOVA results of MPSR

	Factors	F Value	p value (Pr > F)
Main Effects	Centroid Offset	25.60	<0.001
	Dispersion	1.99	<0.0366
	Pielou's Evenness Index	26.21	<0.0001
Interactions	Centroid*Dispersion	52.55	<0.0001
	Centroid*Pielou	3.83	<0.0001
	Dispersion*Pielou	5.50	<0.0001
R-Square		0.8169	

observations [20]. The performance metric variation is partitioned into portions attributed to the effect of independent factor (between-group variation) and portions attributed to random error (within-group variation). An F statistic is computed using the ratio between these variances and tested for significance. Tables 3 and 4 show the three-way ANOVA results for MLW and MPSR values, respectively. Centroid offset, dispersion, DPD statistics, and their interactions are considered as the independent factors impacting the performance metrics. By examining the results, it can be seen that the *p* values for these three geometry descriptors and their interactions are all highly significant (all less than 0.01) for their impact on MLW and MPSR. In addition, high R^2 values indicate that 99.7% of the variation in MLW data can be accounted for by these independent factors, so does 82% data of MPSR. Therefore, it is demonstrated that proposed geometry descriptors, including centroid offset, dispersion and DPD statistics, have strong correlations with array performance.

Finally, through statistical analysis and ANOVA the relationships between proposed geometry descriptors and array performance are established and demonstrated. However, because the number of microphones determines the number of DPDs, the impact of each geometry descriptor varies with the number of microphones. In order to analyze these differences, data collected from Monte Carlo experiments of irregular arrays with 16, 25, 36, 49 and 64 microphones are compared. All the experiments were performed in immersive environments with comparable values of centroid offset and dispersion. Table 5 provides the R^2 results of least squares method by fitting general linear model (GLM) of

selected geometry descriptors on MPSR. It is noted that even with this simplest regression model, over 50% variation of MPSR can be accounted for by $GLM\{a, L, \sigma, \beta\}$. This percentage increases to 70% ~ 90% when applying higher-order fitting functions of geometry descriptors (nonlinear regression models). With increasing microphone number, better R-Square values are obtained.

By comparing the results of $GLM\{a, L\}$ derived from array apertures and positions with $GLM\{a, L, \sigma, \beta\}$ taking account of DPD distributions' impact, at least 10% improvements of R^2 values are observed. Especially for the arrays with microphone density larger than 0.5 mic/m², the impact of $\{a, L\}$ is greatly reduced due to the increasing possibilities of microphone arrangements with fixed centroid and dispersion, while the DPD statistics show stronger correlation with array performance. Furthermore, by comparing the trends of R^2 values of $GLM\{a, L, \sigma\}$ and $GLM\{a, L, \beta\}$ with increasing microphone number, DPD standard deviation assessing the spread of DPD distribution shows a little stronger correlation with MPSR for arrays with microphone density less than 0.2 mic/m², while Pielou's evenness index assessing the diversity of DPD distribution has greater impact on MPSR for array with density larger than 0.2 mic/m². The reason for this phenomenon is that low microphone density cannot provide enough DPD samples to measure the entropy (Pielou's evenness index), and DPD standard deviation representing the average spread of DPDs about zero is more reflective for characteristics of the DPD distribution related to the beamforming gain.

5 Conclusions

This article analyzes and identifies important characteristics for irregular microphone arrays that directly related to beamforming performance. Combined with descriptors analogous to traditional geometry parameters for regular arrays (i.e. array centroid and dispersion), novel geometry descriptors involving DPD statistics describe both regular and irregular arrays. Simulations demonstrated that irregular microphone geometries typically exceed the performance of regular geometries, and arrays with high DPD entropy and wide DPD spread correspond to arrays with better noise suppression ability. These results are primarily applicable for microphone

Table 5 R^2 results for GLMs of geometry descriptors on MPSR

Number of Microphones	Mic Density (mic/m ²)	R-Square of $GLM\{a, L\}$	R-Square of $GLM\{a, L, \sigma\}$	R-Square of $GLM\{a, L, \beta\}$	R-Square of $GLM\{a, L, \sigma, \beta\}$
16	0.16	41.93%	46.72%	46.56%	50.33%
25	0.25	53.84%	58.25%	58.81%	62.72%
36	0.36	60.81%	64.80%	66.49%	69.32%
49	0.49	68.77%	73.20%	74.04%	77.31%
64	0.64	63.61%	80.09%	81.69%	86.97%

arrays in near-field applications, such as in immersive environments.

The relationships between geometry descriptors and beamforming performance developed in this article can be applied directly as the objective functions in optimization procedures to find appropriate microphone distributions for given acoustic environments [7]. The results of this article were based on Monte Carlo experiments with planar microphone distributions, which are more applicable for indoor applications, such as audio surveillance systems. So far, the DPD statistics do not have simple geometric interpretations and must be computed based on all the microphone positions and desired focal points. While these statistics can easily be computed once a geometry is proposed, they cannot directly be used in closed-form analysis and optimizations. Other more direct geometric metrics as they relate to good values of proposed DPD statistics will be needed to guide ad-hoc microphone placements. Future work involving related closed-form relationships between geometry descriptors and key performance metrics could provide a simple and feasible solution for the optimization problems of microphone arrays.

Competing interests

The authors declare that they have no competing interests.

Acknowledgments

This study was supported in part through funding from the FBI Academy.

Received: 9 March 2012 Accepted: 10 September 2012

Published: 27 November 2012

References

1. D Rabinkin, R Renomeron, J French, Optimum sensor placement for array sound capture. *Proceedings of SPIE* **3162**, 227–239 (1997)
2. J Benesty, J Chen, Y Huang, in *Microphone Array Signal Processing*, ed. by J Benesty, W Kellermann, 1st edn. (Springer, Berlin Heidelberg, 2008), p. 39
3. S Shanan, C Pomalaza-Raez, The use of nonuniform element spacing in array processing algorithms. *J. Acoust. Soc. Am.* **86**, 1416–1418 (1989)
4. H Schjaer-Jacobsen, K Madsen, Synthesis of nonuniformly spaced arrays using a general nonlinear minmax optimization method. *IEEE Trans. Antennas Propagat.* **AP-24**, 501–506 (1976)
5. P Townsend, *Enhancements to the generalized sidelobe canceler for audio beamforming in an immersive environment*, MS thesis (Department of Electrical Engineering, University of Kentucky, 2009)
6. M Brandstein, D Ward, in *Microphone Arrays Signal Processing Techniques and Applications*, ed. by A Lacroix, A Venetsanopoulos, M Brandstein, D Ward (Springer, Berlin Heidelberg New York, 2001), p. 3
7. J Yu, K Donohue, *Performance for randomly described arrays* (in IEEE Workshop on Applications of Signal Processing to Audio and Acoustics, 2011)
8. J Meyer, G Elko, *A highly scalable spherical microphone array based on an orthonormal decomposition of the sound field* (Paper presented at the IEEE ICASSP-02, Orlando, Florida, USA, 2002)
9. J Daniel, R Nicol, S Moreau, *Further investigations of high order ambisonics and wavefield synthesis for holophonic sound imaging* (114th Convention of Audio Engineering Society, 2003)
10. A Moffett, Minimum-redundancy linear arrays. *IEEE Trans. Antennas Propagat.* **AP-16**, 172–175 (1968)
11. W Chen, Y Bar-Ness, Minimum redundancy array structure for interference cancellation. *Antennas and Propagation Society International Symposium* **1**, 121–124 (1991)

12. S Pillai, Y Bar-Ness, F Haber, A new approach to array geometry for improved spatial spectrum estimation. *Proc Antennas and Propagation Society International Symposium IEEE* **73**, 1522–1524 (1985)
13. P Townsend, K Donohue, *Beamfield analysis for statistically described planar microphone arrays* (in IEEE Southeastcon, 2009)
14. ZB Li, KFC Yiu, ZG Feng, *A hybrid descent method with genetic algorithm for microphone array placement design* (Applied Soft Computing, Applied Soft Computing, 2012)
15. ZG Feng, KFC Yiu, SE Nordholm, Placement design of microphone arrays in near-field broadband beamformers. *IEEE Trans. on Signal processing* **60**(3), 1195–1204 (2012)
16. K Donohue, K McReynolds, A Ramamurthy, *Sound source detection threshold estimation using negative coherent power* (In IEEE Southeastcon, 2008)
17. K Donohue, S SagharianNejadEsfahani, J Yu, Constant false alarm rate sound source detection with distributed microphones. *EURASIP J. Adv. Signal Process* (2011). doi:10.1155/2011/656494
18. E Pielou, The measurement of diversity in different types of biological collections. *J. Theoret. Biol.* **13**, 131–144 (1966)
19. NTI Audio, Introduce to speech intelligibility. (2008). <http://www.nti-audio.com/Portals/0/data/en/NTI-Audio-AppNote-AL1-Introducing-STIPA.pdf>. Accessed 2011
20. SAS Institute Inc, SAS/STAT(R) 9.22 user's guide. (2010). http://support.sas.com/documentation/cdl/en/statug/63347/HTML/default/viewer.htm#statug_anova_sect022.htm. Accessed 2011

doi:10.1186/1687-6180-2012-249

Cite this article as: Yu and Donohue: Geometry descriptors of irregular microphone arrays related to beamforming performance. *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:249.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com