Optimization of Microphone Array Geometry with Evolutionary Algorithm

Jingjing Yu
University of Kentucky, Lexington, USA
Email: J.YU@ieee.org

Fashan Yu and Yinglei Li
Henan Polytechnic University, Jiaozuo, CHINA; University of Kentucky, Lexington, USA
Email: yufs@hpu.edu.cn; yli228@uky.edu

Abstract—The complex relationship between array spatial gain pattern and microphone distribution limits the effective use of traditional methods to optimize the element placements of microphone arrays. This paper presents a genetic algorithm (GA) for microphone array optimization problems with the purpose of maximizing beamforming SNR for given possible distribution of sound sources. Functions based on the statistical geometry descriptors are applied as the objective functions of GA. Three optimization experiments involving linear and planar arrays in continuous and discrete design spaces are performed under various acoustic scenes. Results demonstrate that proposed algorithm sorts out these superior geometries with significant SNR improvement over randomly generated arrays and regular arrays. High successful rate, rapid convergence speed, and fast processing time observed in all the experiments demonstrate the feasibility of this algorithm as a practical tool for microphone array optimization.

Index Terms—microphone array optimization, genetic algorithm, beamforming

I. INTRODUCTION

Distributed microphone systems are widely applied in signal enhancing applications such as teleconference, speech recognition, sound source localization and acoustic surveillance systems [1][2]. Array performance usually assessing by its ability to detect, locate, track and capture desired target signals is greatly affect by microphone distribution, acoustic conditions, and processing algorithms, while the distribution of microphones is demonstrated to be the critical one to limit the potential improvement of performance derived from other factors [3]. Therefore, this paper focuses on the optimization problem of microphone arrays with fixed number of elements and specified acoustic scenes (possible distribution of target and noise sources) to search for the optimal geometry providing superior noise suppressing ability.

Because the spatial gain pattern of microphone array is a complex nonlinear function of microphone positions, traditional analytical optimization methods usually apply linear approximation and spatial perturbations to simply this problem, whose performance may be limited by the prerequisite of perturbation levels and unpredictable approximation errors [4][5]. Random or exhaustive searching methods are also used in search of optimal geometry by evaluating each candidate via Monte Carlo simulation [2]. However, it is time-consuming and not feasible for large design space of element placements and complex acoustic scene. Other numerical searching methods based on natural algorithms, such as genetic algorithm, can be considered as an alternation to overcome these limitations [6][7].

Derived from natural selection theory, genetic algorithm (GA) is a heuristic searching method exploiting the historical information of evolution procedure to predict new generation with expected better performance, which has been demonstrated as an effective tool for nonlinear optimization problems [8][9][10][11]. This paper applies GA to the microphone array optimization problem with the purpose to obtain superior ability of noise reduction for speech applications. Instead of computing the gain pattern of every candidate via Monte Carlo simulation, a function based on the relationship between statistical descriptors of microphone geometry and array performance metrics is applied as the objective function of GA. Experiments of linear and planar arrays are performed to validate the effectiveness and feasibility of this algorithm in search of optimal microphone geometries with pre-known knowledge of acoustic scene (possible target and noise space). The results in terms of SNR are compared to comparable uniform-spaced arrays and randomly generated arrays via Monte Carlo simulation.

II. PROBLEM FORMULATION

Consider the acoustic scene with microphone array and sound sources distributed in a three dimensional space. Signal received by the p-th microphone can be expressed as:

$$v_p(t; r_s, r_p) = \int_{-\infty}^{\infty} u(t - \tau; r_s) h(t - \tau; r_s, r_p) d\tau$$  \hspace{1cm} (1)
where \( u(t; r_s) \) is the pressure wave transmitted from the source located at \( r_s \), \( h(.) \) represents the impulse response of propagation from \( r_s \) to \( r_r \) as

\[
h(t; r_s, r_r) = a_{qp}(t - \tau_{qp}) + \sum_{n=1}^{P} a_{qn}(t - \tau_{qn}).
\]

(2)

where \( a_{qp} \) is the response related to the \( n^{th} \) path propagation, and \( n=0 \) represents the direct path from source to microphone. \( \tau_{qp} \) is the corresponding time delay. Then, the power of delay and sum beamformer focused on \( r_r \) with a sound source at \( r_s \) can be computed as [12]

\[
S(r_s, r_r) = \int \int \left[ \hat{U}(\omega; r_s) \right]^* \left( B_{pq} B_{rq} \hat{V}_{pq}(\omega; r_s, r_r) \right) \hat{V}_{rq}(\omega; r_s, r_r) \exp\left( j \omega (\tau_{pq} - \tau_{rq}) \right) d\omega.
\]

(3)

where \( B_{pq} \) is the beamformer coefficient related to desired target position \( r_s \) and microphone position \( r_r \). Normally in order to emphasize signals received by microphones closer to target position, let \( B_{pq} = 1/d_{ip} \), where \( d_{ip} \) denotes the distance form \( r_i \) to \( r_p \). Assuming coefficients of beamformer and propagation attenuation are independent with microphone positions and only considering direct path propagation, by taking mean value over all microphone pairs, (3) can be rewritten as below [12]:

\[
S(r_s, r_r) = P \int \int \left[ \hat{U}(\omega; r_s) \right]^* \left( \hat{V}_{pq}(\omega; r_s, r_r) \right) \exp\left( j \omega (\tau_{pq} - \tau_{rq}) \right) d\omega.
\]

(4)

where \( P \) is the number of microphones, and angular brackets denote mean value over all microphone pairs. For the sound source different with desire target, \( r_s \neq r_i \), which is considered as an interference or noise, \( S(r_s, r_r) \) should be as small as possible to indicate superior ability of noise suppression. From (4), it can be seen that large span of exponential phase terms from \(-\pi \) to \( \pi \) can results in significant incoherence and near zero power gains for non-target positions, while limited delay values over the phase terms make partial coherence more likely for signals received from non-target sources. With the assumption of constant signal power and beamformer coefficients, array ability to decorrelate sound sources at non-target positions is directly related to the differential distances derived from microphone positions. Therefore, (4) demonstrates the significant impact of microphone placements on array beamforming performance.

In order to further formulate this relationship, paper [13] proposed several statistical descriptors to characterize microphone geometry, which are proved to have strong correlation with key performance metrics, such as Mainlobe Width (MLW) associated with spatial resolution or array ability to distinguish signals from close sources, and Mainlobe-to-peak-sidelobe Ratio (MPSR) associated with array ability to suppress noise, as shown in Table I.

**TABLE I. GEOMETRY DESCRIPTORS**

<table>
<thead>
<tr>
<th>Descriptors</th>
<th>Illumination</th>
<th>Relationship with performance metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Centroid offset</td>
<td>The distance from array centroid to focal point of beamformer. It has a direct impact on the mainlobe resolution and shape.</td>
<td>With fixed array dispersion, increased centroid offset brings large MLW and small MPSR, representing degradation of array performance.</td>
</tr>
<tr>
<td>Array dispersion</td>
<td>The standard deviation of microphone coordinates about array centroid.</td>
<td>Small dispersion results in better MPSR. For MLW, when array dispersion increases, the mainlobe area on horizontal plane of target decreases, while the mainlobe region along vertical direction grows.</td>
</tr>
<tr>
<td>Entropy based statistics of inter-path distances</td>
<td>Inter-path distance is the differential distance from microphone pair to target and noise source pair. For specified target and noise source, the distribution of inter-path distances over all microphone pairs is directly related to the spatial gain for the signals received from the noise source when steering at target.</td>
<td>Inter-path distances with rich diversity and large spread for non-target positions represent superior array ability to suppress noise.</td>
</tr>
</tbody>
</table>

The close form relationship functions between geometry descriptors and performance metrics provided by [13] can be applied as the objective function of GA. For a specified acoustic scene, possible source distribution can be depicted by the probability density functions of source locations, related to the behavior patterns of speakers and interested target space. Then, the objective function of GA can be defined as

\[
F'(G) = \int_{r_s \in TargetSpace} \int_{r_n \in NoiseSpace} \left\{ F(G, r_s, r_n) p(r_n) dr_n \right\} p(r_s) dr_s.
\]

(5)

where \( F(G, r_s, r_n) \) is the relationship function provided by [13] for the focal point \( r_s \) and noise source at \( r_n \). Smaller values of \( F(G, r_s, r_n) \) represents better ability to
decorrelate the signals received from this noise source when steering at target. \( \mathbf{G} \) is the set of geometry descriptors to identify a class of microphone arrays with similar performance. \( p(r_t) \) and \( p(r_n) \) are the probability density functions representing expected locations of focal points for the target and noise source occurrences, respectively. Therefore, the criterion to search for optimal array geometry can be written as

\[
\mathbf{G}_{\text{opt}} = \text{argmin}\left\{ F'(\mathbf{G}) \right\}.
\]

III. GENETIC ALGORITHM

As a heuristic searching method, which uses empirical information to search for better solutions and reduce the computational complexity, GA has been demonstrated to be effective for solving nonlinear optimization problems [6][7][8][9][10][11]. The main ideal of this algorithm is to use historical information of evolution procedure to guide searching direction by predicting new generation with higher fitness values. Following the rules of “survival of the fittest”, genes of the individuals with higher fitness values will have more chance to be inherited by offspring, while perturbations are introduced randomly to the population to enhance diversity of evolution. In our case, all coordinates of microphone array are considered as an individual. The fitness value is assessed by the objective function defined in (5). Then parents are selected based on fitness values undergoing crossover and mutation to give birth to new generation. The evolution procedure continues until reaching acceptable fitness value or the limitation of generation number. General flow chart is shown in Fig. 1 and details of GA are applied as below:

Coding: individual of GA is defined as \( \{x_1, x_2, ..., x_P, y_1, y_2, ..., y_P, z_1, z_2, ..., z_P \} \), where \( \{x_p, y_p, z_p\} \) is the coordinates of \( p \)th microphone, \( P \) is the number of microphones.

Selection: evaluate and rank the fitness of individuals in the population by the objective function. According to the ideal of “survival of the fittest”, usually top 20% individuals are selected as elites to give birth to the offspring by crossover and mutation.

Crossover: as shown in Fig. 2, crossover is implemented by randomly choosing one coordinate of offspring from the corresponding parents’ coordinates in each dimension. In our case, 60% of elites are devoted into crossover process to generate new individuals.

Mutation: 40% of elites are used for mutation by adding random spatial perturbations to each dimension of parents, as shown in Fig. 2. These perturbations are generated from a zero mean normal distribution. The standard deviation of this normal distribution can control the levels of perturbations, and further affect the convergence of GA. Generally, large perturbation can increase the diversity of evolution, which is good for avoiding GA trapped in local optimum. However, it will reduce the convergence speed by distracting the evolution from current optimal direction. In this paper, the standard deviation of perturbations is derived from [10]

\[
\sigma(n) = \sigma_0 \rho^{(n-1)}.
\]

where \( n \) is the iteration number, \( \sigma_0 \) is the initial standard deviation related to the wavelength of important signal bands. \( \rho \) is a constant to control the shrinkage rate of perturbations along generations. The idea here is to add large perturbation at the beginning of iteration to start global search of optimum. Along with the evolution procedure, when the best searching direction becomes more specified, the perturbation level is reduced to speed up the convergence to the optimal solution.

Replacement: in order to insure the survival of high-fitness individuals, offspring are ranked together with the old population based on fitness. Illegal individuals, which are outside design space of microphone positions or repeat with the other individuals, are replaced by random generated ones. Then the new generation is sorted out with the same size of initial population.

After the termination of iteration, to verify the actual performance of the optimal individuals, Monte Carlo simulations are performed over the last generation of GA to pick three arrays with top SNR results as the output of optimization procedure. Note that maintaining the balance between inheritance and exploration is critical for the success of GA optimization. It means that the tradeoff between searching diversity to ensure global optimum (computing complexity) and convergence rate needs to be...
taken into account when selecting GA parameters, such as the size of initial population, ratio of parent selection, ratio of crossover and mutation, and the level of perturbations.

IV. EXPERIMENTS

In this section, optimization experiments for 1D linear array and 2D planar arrays with various acoustic scenes are performed to evaluate the performance of GA. The SNRs of the optimal geometries derived from the last generation of GA are computed and compared to comparable regular arrays via Monte Carlo simulations.

A. Setting

All the experiments are performed in a 10×10×2 m room. Colored noise generated by the band importance function from the SII model [3] is considered as the source signal for both target and noise, which emphasizes the most important frequency bands for speech intelligibility. As shown in Table II, these three experiments involve different microphone design spaces and source distributions. To assess the performance in an independent manner for all the experiments, Monte Carlo simulations using delay and sum beamformer are performed to evaluate the SNRs of the optimal geometries from the last generation of GA, compared with randomly generated geometries from the first generation of GA and comparable regular arrays.

Table III gives the GA parameters used for each optimization problem. A robust optimization algorithm needs to maintain the balance between inheritance and exploration, which could be greatly affected by the relative GA parameters. Therefore, in our experiments, the initial population size, ratio of parent selection and the maximum iteration number are adjusted according to the knowledge of acoustic environment and design space.

B. Result Analysis

Table IV shows SNR results of 3 best arrays in the last and first generation of GA, representing the optimal geometries and random generated geometries, respectively. The mean SNRs of each generation are also provided with ± one standard deviation. By comparing the mean SNRs of the first and last generation, significant improvement is observed to demonstrate the effectiveness of GA iterations and the objective function rules. It is noted that by providing a moderate mean value of SNR the first generation of GA includes superior arrays outperforming regular array and inferior arrays showing lower SNRs than regular array. Through the GA optimization procedure including the objective functions of geometry descriptors, those superior arrays are sorted out to compose the last generation with much higher SNR results. For problem 1 and problem 2, where the target space is completely overlapped with the noise space, the improvement of mean SNR for the last generation of GA is over three standard deviations greater than the other arrays. For problem 3, where the continuous spaces for possible target and noise distributions do not overlap, the improvement of SNR is even more significant, which almost doubles the SNRs of randomly generated irregular arrays and regular array.

### TABLE II.

<table>
<thead>
<tr>
<th>Acoustic scenes</th>
<th>Design spaces of microphone</th>
<th>Mic numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>1D linear array</td>
<td>5 discrete sources are randomly distributed in the room. Each source has equal opportunity to be selected as the target, while the others act as interferences with ¼ chances to make noise.</td>
<td></td>
</tr>
<tr>
<td>2D planar array</td>
<td>Ceiling of the room. (continuous design space)</td>
<td>64</td>
</tr>
<tr>
<td>2D planar array</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### TABLE III

<table>
<thead>
<tr>
<th>GA PARAMETERS</th>
<th>Problem 1</th>
<th>Problem 2</th>
<th>Problem 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Initial population</td>
<td>40</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Ratio of elites</td>
<td>40%</td>
<td>20%</td>
<td>20%</td>
</tr>
<tr>
<td>Crossover probability</td>
<td>60%</td>
<td>60%</td>
<td>60%</td>
</tr>
<tr>
<td>Mutation probability</td>
<td>40%</td>
<td>40%</td>
<td>40%</td>
</tr>
<tr>
<td>Iteration threshold</td>
<td>Iter = 100</td>
<td>Iter = 200</td>
<td>Iter = 200</td>
</tr>
</tbody>
</table>

The gain patterns of the optimal arrays and comparable regular arrays for all the problems are provided in Fig. 3 and Fig. 4. It is appeared that through the optimization of GA, these high sidelobes which will severely degrade array performance are successfully steered away from the noise source regions, representing superior ability for noise suppression. Although for problem 1 the main lobe area of GA optimal array is larger than the regular array, it circumvents the noise sources positions. With lower power gains for all the noise sources when steering at target, the optimal array still shows better SNR performance than the regular array. Furthermore, by visually inspecting the optimal microphone distribution in
each problem, it concludes that for discrete sources the microphones are clustered near the source positions, while for the continuous spaces of target and noise distributions the microphone densities are very high near the target space but sparsely distributed over the noise space.

Table IV. SNR results of array geometries

<table>
<thead>
<tr>
<th>Problem</th>
<th>Top 3 SNR</th>
<th>Mean SNR</th>
<th>Top 3 SNR</th>
<th>Mean SNR</th>
<th>Regular array SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.33</td>
<td>12.93</td>
<td>11.48</td>
<td>11.40</td>
<td>6.59 (±0.47) 8.60</td>
</tr>
<tr>
<td></td>
<td>12.06</td>
<td>12.93</td>
<td>11.48</td>
<td>11.40</td>
<td>6.59 (±0.69) 8.60</td>
</tr>
<tr>
<td>2</td>
<td>56.38</td>
<td>55.79</td>
<td>46.07</td>
<td>43.96</td>
<td>40.67 (±3.85) 35.51</td>
</tr>
<tr>
<td></td>
<td>55.31</td>
<td>55.31</td>
<td>46.07</td>
<td>43.96</td>
<td>40.67 (±3.85) 35.51</td>
</tr>
<tr>
<td>3</td>
<td>31.27</td>
<td>31.22</td>
<td>22.17</td>
<td>20.54</td>
<td>15.47 (±2.88) 14.02</td>
</tr>
<tr>
<td></td>
<td>31.13</td>
<td>31.13</td>
<td>22.17</td>
<td>20.54</td>
<td>15.47 (±2.88) 14.02</td>
</tr>
</tbody>
</table>

As mentioned before, the size of initial population and maximum iteration number of the optimization strategy are the important factors to influence the searching diversity (reliability) and convergence of iteration (efficiency). Poor balance between the reliability and efficiency of GA will make the evolution procedure easily trapped in local optimum or distracted away from the optimal searching direction. Fig. 5 shows the evolution of minimum and average values of the objective functions along generations. Three different initial populations are applied to compare the convergence of iteration. It can be seen that for problem 2 and problem 3, the objective functions with different initial populations converge to different values, which cannot be eliminated by the iteration. And for problem 1 with less possible solutions, large initial population doesn’t result in better convergence values. Therefore, it concludes that insufficient size of initial population cannot be compensated by the increase of maximum iteration number. It will restrict the evolution in local optimum from early iterations. On the other hand, excessive initial population only results in limited improvement of optima, but greatly increases the computational complexity during the searching procedure.

C. Reliability and Convergence Speed

In this section, successful rate and convergence speed are provided to validate the reliability and efficiency of GA. 30 runs for each problem are applied to compute these parameters. Results are shown in Table V. Successful rate is a good measure to evaluate the reliability or robustness of optimization strategy. High successful rate is necessary for the real-case applications.
In this paper, 30 experiments for each problem are performed, while the source locations are randomly shifted inside the room. The criterion for the success of optimization is that SNR results of GA optimal arrays outperform regular array. It can be seen that GA shows satisfying successful rates (≥80%) in all the optimization problems.

Convergence speed to global optimum reflects the efficiency of evaluation. It is usually measured by the average iteration number needed to achieve convergence.

Figure 4. Top view gain patterns of regular geometries when targeting at the center source of the field of view. Circles represent microphone positions. Cross or square area represents source space. (a) Problem 1. (b) Problem 2. (c) Problem 3 with the target space in the center of the field of view.

Figure 5. Evolution of minimum (in red) and average (in blue) values of objective functions along generations. (a) Problem 1. (b) Problem 2. (c) Problem 3.
As shown in Table V, rapid convergence speeds are observed in all the experiments, demonstrating the feasibility of this algorithm as a practical tool for microphone array optimization. In addition, the time consuming features of GA is compared to previous random search method via Monte Carlo simulation. As shown in Table V, it can be seen that by applying the objective function of array geometry descriptors and evolutionary algorithm during optimization procedure, instead of directly searching via Monte Carlo simulation for every candidate, 99.5% running time is saved, while GA showing much higher probability to catch these superior arrays. This advantage of fast running time would be particularly applicable to dynamic situations where the acoustic scene or environment cannot be known well in advance, such as audio surveillance system.

### Table V

<table>
<thead>
<tr>
<th>Problem</th>
<th>Successful rate</th>
<th>Convergence speed</th>
<th>Time consuming (compared with random search method)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Problem 1</td>
<td>83%</td>
<td>74</td>
<td>0.51%</td>
</tr>
<tr>
<td>Problem 2</td>
<td>100%</td>
<td>108</td>
<td>0.50%</td>
</tr>
<tr>
<td>Problem 3</td>
<td>80%</td>
<td>49</td>
<td>0.50%</td>
</tr>
</tbody>
</table>

As shown in Table V, rapid convergence speeds are observed in all the experiments, demonstrating the feasibility of this algorithm as a practical tool for microphone array optimization. In addition, the time consuming features of GA is compared to previous random search method via Monte Carlo simulation. As shown in Table V, it can be seen that by applying the objective function of array geometry descriptors and evolutionary algorithm during optimization procedure, instead of directly searching via Monte Carlo simulation for every candidate, 99.5% running time is saved, while GA showing much higher probability to catch these superior arrays. This advantage of fast running time would be particularly applicable to dynamic situations where the acoustic scene or environment cannot be known well in advance, such as audio surveillance system.

### V. CONCLUSION

This paper presented a genetic algorithm to optimize element placements of microphone array with the purpose of maximizing beamforming SNR for given knowledge of acoustic scene or environment. Three optimization experiments related to corresponding actual real-world problems was performed. Simulation results of Monte Carlo experiments demonstrate that this algorithm effectively sorts out superior geometries with significant SNR improvements over randomly generated irregular arrays and regular arrays. Furthermore, high successful rates and rapid convergence speeds are observed in all the experiments, as well as fast processing time which is 99.5% less than the traditional random searching method, demonstrating the feasibility of this algorithm as an effective tool for microphone array design.

This optimization algorithm can be especially useful for immersive environment applications, such as audio surveillance systems where the field of interest is fixed and the situations of sound sources cannot be known well in advance.

Future work will extend this algorithm to the three dimensional design space and complex acoustic environment with reverberation. The impact of GA parameters will also be studied to improve the effectiveness and robustness of this algorithm for different optimization problems.

### REFERENCES

Jingjing Yu was born in Henan, China. She received the B.S. and M.S. degrees in communication engineering from Beijing Jiaotong University, Beijing, China, in 2005 and 2007, respectively. She is currently a Ph.D. candidate in Electrical Engineering Department at the University of Kentucky, Lexington, KY. And she is also a research assistant in the Center for Visualization & Virtual Environments at the University of Kentucky. Her current research interests include communication and information system, audio signal processing, and engineering optimization. Ms. Yu is a student member of the IEEE.

Fashan Yu was born in Hubei, China. He received the B.S. degree in automation from Henan Polytechnic University, Jiaozuo, China, in July 1977. He is currently a Professor and Dean of the School of Electrical Engineering & Automation at Henan Polytechnic University, Jiaozuo, China. He is also appointed as the Director of National Electrical & Electronic Experiment Center at Henan Polytechnic University. He is the author and coauthor of numerous papers and of several books in the field of Control Theory and Control Engineering. His current research interests include industrial process control, PLC control, computer simulation, AC–DC speed control system, etc.

Prof. Yu is a member of the Mine Automation Committee of China, and a member of the teaching guide committee on Automation organized by the Ministry of Education of China.

Yinglei Li was born in Zhejiang, China. She received the B.S. degree in textile trading from Donghua University, Shanghai, China, in 2003, the M.S. degree in economics and management from Tongji University, Shanghai, China, in 2008, and the M.S. degree in statistics from the University of Kentucky, Lexington, USA, in 2010. She is currently a Ph.D. candidate in Statistics Department at the University of Kentucky, Lexington, USA. Her current research interests include biostatistics, multivariate statistics and statistical modeling.

Ms. Li is a member of American Statistical Association.