DESIGN OF ROBUST STEERABLE BROADBAND BEAMFORMERS INCORPORATING MICROPHONE GAIN AND PHASE ERROR CHARACTERISTICS

Chiong-Ching Lai, Sven Nordholm, Yee-Hong Leung

Department of Electrical and Computer Engineering
Curtin University
Kent Street, Bentley, WA 6102, Australia

ABSTRACT

Beamformers are known to be sensitive to errors and mismatches in their array elements. This paper proposes a robust steerable broadband beamformer design using the Farrow structure and for arbitrary array geometry. The design formulation includes stochastic models describing the microphone characteristics as random variables, thus allowing flexibility for microphone errors. This method establishes a direct relationship in controlling the robustness specification in the design from given microphone characteristics. The robust design procedure optimises the mean performance of the beamformer. Design examples show significant reduction in error sensitivity in the robust design formulation.

Index Terms—steerable, broadband beamformer, robust, stochastic, microphone characteristics

1. INTRODUCTION

Compared to fixed beam broadband beamformers, steerable broadband beamformers (SBBFs) offer an extra ability to steer its main beam on-the-fly. This steering capability offers dynamic beamforming which is extremely useful in applications that involve dynamic environments. Some examples of commercial applications where SBBFs are becoming more and more prevalent include audio-video conferencing, hands-free communication systems and audio surveillance systems, where the speaker is likely to move around [1].

Initial works on SBBFs include decoupling the beampattern spectro-spatial dependencies in order to steer the main beam using a Wigner rotation matrix [2], and using Farrow filters [3] for designing SBBFs [1, 4]. The SBBF design using Farrow filters is interesting as the main beam can be steered online with only a single parameter. However, these SBBFs, which resemble superdirective beamformers for low frequencies, are extremely sensitive to spatial white noise and errors in the array elements [5]. Although a robust design using norm constraint on the resulting filter weights is proposed in [4], there is no clear indication on the choice of the norm constraint parameter and the choice is made rather intuitively.

In this paper, we propose a new robust SBBF design formulation. In the new formulation, the Farrow structure is used again as in [4], but it now admits any arbitrary array geometry. As for robustness, we follow one of the methods described in [6], which incorporates the probability density function of the microphone error characteristics into the design formulation and sets as the design objective, the mean performance. The proposed design formulation thus provides a direct relationship between the error characteristics and robustness control, resulting in better robustness specification decision. The robust SBBF design is formulated in both time domain (TD) and frequency domain (FD).

This paper is organised as follows. Section 2 describes the spiral array geometry and SBBF structure with Farrow filters. Both the TD and FD design of robust SBBFs are formulated in Section 3, followed by design examples in Section 4. Finally, the conclusion is provided in Section 5.

2. FAR-FIELD SBBFS CONFIGURATION

2.1. Array geometry

Consider the spiral array [4] shown in Fig. 1. It consists of P concentric rings with K microphones uniformly spaced in each ring. Successive rings are offset by 2π/(KP) relative to each other. Under the far-field signal model, the array response vector d(ω, φ) for a signal with frequency ω and impinging the array at angle φ in the plane of the array is given by

\[d(\omega, \phi) = d_{pK+k}(\omega, \phi) = \exp \left( j \frac{\omega r_p c}{\omega} \cos \left( \phi - \frac{2\pi k}{K} - \frac{2\pi p}{PK} \right) \right)\]  

(1)

where \(p \in \{0, ..., P-1\}\) is the index of the \(p^{th}\) ring, \(k \in \{0, ..., K-1\}\) is the index of the \(k^{th}\) microphone in each ring, \(c\) is the speed of sound in the propagating medium (\(c = 345 m/s\) in air) and \(\psi = pk + k \in \{0, ..., PK-1\}\) is the index of the \(k^{th}\) microphone in the \(p^{th}\) ring. \(r_p\) is the radius of the \(p^{th}\) ring and is given by

\[r_p = \frac{f_p}{f_p \sin(\pi/K)}\]  

(2)

where \(0 < \alpha \leq 0.5\) to prevent spatial aliasing, and \(f_p\) is the highest frequency in Hz received by the \(p^{th}\) ring. Note that the spacing between successive ring radii \(r_p\) can be uniform or non-uniform, depending on the choice of \(f_p\). In (2), it is assumed that there are sufficient rings in the spiral array. For a spiral array with only a few rings (\(P = 2, 3\)), \(r_0\) is obtained by setting \(f_0\) to be the maximum frequency of the received signal. Subsequent ring radii are then selected by trading off between having narrow beamwidth at low frequency and frequency invariant response. One example is \(r_p = 2r_{p-1}\) for \(p \in \{1, ..., P-1\}\).

2.2. Beamformer structure

For the TD Farrow filters shown in Fig. 2, its beampattern, steered to azimuth angle \(\psi\), is given by

\[G(\psi, \omega, \phi) = \sum_{p=0}^{P-1} \sum_{k=0}^{K-1} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} a_{pK+k}(\omega, \phi) \times h_{p, k, m}[n] e^{-jnT_\psi} \psi^m d_{pK+k}(\omega, \phi)\]  

(3)

\(1\) All vectors used in this paper are column vectors.
Fig. 1: Spiral array geometry

Fig. 2: TD Farrow filter structure for SBBF

Fig. 3: FD Farrow filter structure for SBBF

where \( M - 1 \) is the order of Farrow filters indexed by \( m \in \{0, \ldots, M - 1\} \), \( n \in \{0, \ldots, N - 1\} \) is the time index of the \( N \)-tap FIR filters, and \( T_s \) is the sampling period. \( a_n(\omega, \phi) \) specifies the microphone characteristics given by

\[
a_n(\omega, \phi) = \gamma_n(\omega, \phi) e^{j \gamma_n(\omega, \phi)} \quad (4)
\]

where both the gain \( \gamma_n(\omega, \phi) \) and phase \( \gamma_n(\omega, \phi) \) can be frequency and angle dependent. We keep the array response \( a_n(\omega, \phi) \) arbitrary (in the azimuth plane) so that the robust design formulation in the next section is valid for any microphone array geometry. We can write (3) in matrix form given by

\[
G(\psi, \omega, \phi) = h^T g(\psi, \omega, \phi) \quad (5)
\]

where

\[
g(\psi, \omega, \phi) = (a(\omega, \phi) \odot d(\omega, \phi)) \odot f(\psi) \odot e(\omega),
\]

\[
[a(\omega, \phi)_v]_n = a_n(\omega, \phi), \quad f(\psi) = [\psi^0, \ldots, \psi^{M-1}]^T, \quad h^T = h_{p,k,m}[n].
\]

\[
[h] = h_{p,k,m}[n], \quad e(\omega) = \left[e^{-j \omega T_s (0)}, \ldots, e^{-j \omega T_s (N-1)}\right]^T. \quad (8)
\]

The symbol \( \odot \) denotes element-wise product and \( \otimes \) denotes Kronecker product. Index \( q = pK + kM + mN + n \) describes how the real elements in \( h \) are stacked. Note that without modelling the microphone characteristics, we have

\[
g_f(\psi, \omega, \phi) = g_f(\psi, \omega, \phi) = d(\omega, \phi) \otimes f(\psi) \otimes e(\omega). \quad (9)
\]

For the FD Farrow structure shown in Fig. 3 (\( \omega \)) is the \( k^{th} \) frequency bin), we have

\[
G_f(\psi, \omega, \phi) = h_f^T(\omega) \cdot g_f(\psi, \omega, \phi) \quad (10)
\]

where

\[
g_f(\psi, \omega, \phi) = (a(\omega, \phi) \otimes d(\omega, \phi)) \otimes f(\psi),
\]

\[
[h_f(\omega)]_{q'} = H_{p,k,m}(\omega) = \sum_{n=0}^{N-1} h_{p,k,m}[n] e^{-j \omega T_s} \quad (12)
\]

Index \( q' = pK + kM + m \) describe how the transfer functions \( H_{p,k,m}(\omega) \) are stacked to form \( h_f(\omega) \). Without microphone characteristics modelling, we have

\[
g_f(\psi, \omega, \phi) = g_f(\psi, \omega, \phi) = d(\omega, \phi) \otimes f(\psi). \quad (13)
\]

3. ROBUST SBBF DESIGN

3.1. Time domain formulation

Following [6], the least-squares error function for specific microphone characteristics can be written as

\[
J(h, \zeta, a_0, \ldots, a_{PK-1}) = [G(\zeta) - H_d(\zeta)]^2
\]

\[
= h^T Q(\zeta) h - 2h^T g(\zeta) + |H_d(\zeta)|^2 \quad (14)
\]

where

\[
Q(\zeta) = g(\zeta) g^H(\zeta), \quad (15)
\]

\[
g_a(\zeta) = \text{Re} \{ g(\zeta) H_d^*(\zeta) \}, \quad (16)
\]

\[
H_d(\zeta) = \begin{cases} e^{-j \omega N_d}, & \omega \in \Omega_{pb}, \phi \in \Phi_{pb}(\psi) \quad (17) \\ 0, & \text{otherwise} \end{cases}
\]

\( N_d \) gives the desired delay, \( \Omega_{pb} \) specifies the spectral passband and \( \Phi_{pb}(\psi) \), which depends on steering angle \( \psi \), defines the spatial passband. Note that \( \omega \) and \( \phi \) are dropped from \( a, \kappa \) and \( \gamma \) (their dependencies are understood from the context), and \( (\psi, \omega, \phi) \) is replaced with \( \zeta \) for notational convenience. The total cost function, which is defined as the sum of the cost functions for all feasible microphone characteristics, weighted by the probability density functions (PDFs) of the microphone characteristics is given by

\[
J_{tot}(h, \zeta) = \int \cdots \int J(h, \zeta, a_0, \ldots, a_{PK-1}) f_{A_0}(a_0) \times \cdots f_{A_{PK-1}}(a_{PK-1}) da_0 \cdots da_{PK-1} \quad (18)
\]

where \( f_A(a) = f_{a,\rho}(\kappa, \gamma) \) is the PDF for the random variable \( A \) (microphone characteristics), which is a joint PDF for the random variables \( a \) (gain) and \( \rho \) (phase). To simplify the design model, we assume all microphones have similar characteristics, such that \( a_v \) for \( v \in \{0, \ldots, PK-1\} \) can be described by the same PDF \( f_A(a) \). Furthermore, we assume \( a \) and \( \rho \) are independent such that \( f_A(a) = f_a(\kappa) f_\rho(\gamma) \), where \( f_a(\kappa) \) and \( f_\rho(\gamma) \) are the PDFs of the gain \( \kappa \) and phase \( \gamma \) respectively. Hence, (18) reduces to

\[
J_{tot}(h, \zeta) = \int \cdots \int J(h, \zeta, a_0, \ldots, a_{PK-1}) f_A(a_0) \times \cdots f_A(a_{PK-1}) da_0 \cdots da_{PK-1}
\]

\[
= h^T \tilde{Q}_H(\zeta) h - 2h^T \tilde{g}_H(\zeta) + |H_d(\zeta)|^2 \quad (19)
\]
\[ \mathcal{G}_R(\zeta) = \mu_c \left[ \mu_v \mathcal{G}_v \mathcal{G}_r(\zeta) - \mu_v \mathcal{G}_m(\zeta) \right] \]  
with \( \mathcal{G}_v(\zeta) \) and \( \mathcal{G}_m(\zeta) \) the real and imaginary parts of \( \mathcal{G}_v(\zeta) = \mathcal{G}_v(\zeta) H_2(\zeta) \), and 
\[ \mu_v = \int \kappa_s f_s(\kappa_s) d\kappa_s , \quad \mu_c = \int \cos(\gamma_{\theta}) f_p(\gamma_{\theta}) d\gamma_{\theta} , \] 
\[ \mu_i = \int \sin(\gamma_{\theta}) f_p(\gamma_{\theta}) d\gamma_{\theta} . \]

As for \( \mathcal{Q}_R(\zeta) \), we have
\[ \mathcal{Q}_R(\zeta) = \int \int \cos(\gamma_{\theta} - \gamma_{\phi}) \int \left[ \mathcal{G}_v(\zeta) \mathcal{G}_v(\zeta) + \mathcal{G}_m(\zeta) \mathcal{G}_m(\zeta) \right] Q_{\mathcal{Q}_R}(\zeta) \]
\[ + \mathcal{Q}_{\mathcal{Q}_R}(\zeta) \int \sin(\gamma_{\theta} - \gamma_{\phi}) \int \left[ \mathcal{G}_v(\zeta) \mathcal{G}_v(\zeta) + \mathcal{G}_m(\zeta) \mathcal{G}_m(\zeta) \right] \]
\[ = \mathcal{Q}_{\mathcal{Q}_R}(\zeta) \int \kappa^2 f_s(\kappa) d\kappa = \sigma^2_{\mathcal{Q}_R}(\zeta) \mathcal{Q}_{\mathcal{Q}_R}(\zeta) \]
where \( \sigma^2_{\mathcal{Q}_R}(\zeta) \) is the second moment of gain PDF, i.e.
\[ \sigma^2_{\mathcal{Q}_R}(\zeta) = \int \kappa^2 f_s(\kappa) d\kappa . \]

On the other hand, for \( \nu \neq w \), we have
\[ \mathcal{Q}_R(\zeta) = \mu_v^2 \left[ \sigma^2_\mathcal{Q}_R(\zeta) + \sigma^2_\mathcal{Q}_I(\zeta) \right] \]
where \( \sigma^2_\mathcal{Q}_R(\zeta) = (\mu_v^2)^2 + (\mu_i^2)^2 \), \( \sigma^2_\mathcal{Q}_I(\zeta) = (\mu_v^2)(\mu_i^2) - \mu_v^2 \mu_i^2 = 0 \).

Combining the results from (23)-(28), we have
\[ \mathcal{Q}_R(\zeta) = A \odot \mathcal{Q}_{\mathcal{Q}_R}(\zeta) \]
where \( A = (\mu_v^2 \sigma_\mathcal{Q}_R(1_{L_2} - 1_{L_2}) + \sigma_\mathcal{Q}_I 1_{L_2}) \odot 1_{L_2} \)  
where \( I \) is the identity matrix and \( 1 \) is a square matrix with all 1 elements to 1. Their dimensions are given by the subscripts \( L_1 = P \cdot K \) and \( L_2 = M \cdot N \) (\( L_2 = M \) for FD design). We are now ready to define the cost function for designing robust SBBFs in the weighted least-squares (WLS) sense. Weighting the cost function in (19) with \( F(\zeta) \) and integrating it across steering angle \( \Psi \), frequency \( \Omega \) and azimuth angle \( \Phi \) results in 
\[ J_{WLS}(\mathbf{h}) = \int_{\Omega} \int_{\Phi} \int_{\Psi} F(\zeta) J_{\text{tot}}(\mathbf{h}, \zeta) d\Phi d\Omega d\Psi \]
\[ = h^T WLS \mathbf{Q}_{WLS} \mathbf{h} - 2 h^T WLS \mathbf{g}_{WLS} + d_{WLS} \]
where \( \mathbf{Q}_{WLS} = \int_{\Phi} \int_{\Omega} \int_{\Psi} F(\zeta) \mathcal{Q}_R(\zeta) d\Phi d\Omega d\Psi \), \( \mathbf{g}_{WLS} = \int_{\Phi} \int_{\Omega} \int_{\Psi} F(\zeta) \mathcal{Q}_I(\zeta) d\Phi d\Omega d\Psi \), \( d_{WLS} = \int_{\Omega} \int_{\Phi} \int_{\Psi} F(\zeta) H_2(\zeta)^2 d\Phi d\Omega d\Psi \).

Minimising (31), we have the robust SBBF weights given by 
\[ \mathbf{h}_{WLS} = \mathbf{Q}_{WLS}^{-1} \mathbf{g}_{WLS} \]  

3.2. Frequency domain formulation

Consider a least-squares error function similar to (14), except we now formulate the design in FD. We have 
\[ J(\mathbf{h}, \zeta, \omega) = \| \mathbf{G}_f(\zeta) - H(\zeta) \|^2 \]
\[ = h^T(\omega) Q_f(\omega) h(\omega) - 2 \text{Re}(h^T(\omega) Q_f(\omega)) + |H_2(\zeta)|^2 \]  
where \( Q_f(\zeta) = \mathcal{G}_v(\zeta) \mathcal{G}_v(\zeta) \) and \( g_f(\omega) = g_v(\omega) H_2(\zeta) \). Following [7], we express the matrices and vectors in (36) as the direct sum of their corresponding real and imaginary parts such that (36) can be expressed in real matrices and vectors as follows.

\[ J(\mathbf{h}, \zeta, \omega) = \| \mathbf{G}_f(\zeta) - H(\zeta) \|^2 \]
\[ = 2 \text{Re}(h^T(\omega) Q_f(\omega)) + |H_2(\zeta)|^2 \]  
where \( Q_f(\zeta) = \begin{bmatrix} \text{Re}(Q_f(\omega)) & -\text{Im}(Q_f(\omega)) \\ \text{Im}(Q_f(\omega)) & \text{Re}(Q_f(\omega)) \end{bmatrix} \), \( h_f(\omega) = \text{Re}(h(\omega)) \), \( |H_2(\zeta)|^2 \), \( \Psi = [-0.2\pi, 0.2\pi], \Phi = [-\pi, \pi] \), \( \alpha = 0.45, 0.15 \) beam width, 64 frequency, 64 azimuth angle and 16 steering angle points. The
ring radii are \( r_p = [0.033, 0.089, 0.242, 0.657] \) m which correspond to \( f_p = [4000, 1474, 543, 200] \) Hz. For illustration purposes, we assume (without loss of generality) that the microphone characteristics are independent of azimuth angle. The microphone gain and phase characteristics are assumed to be Gaussian, \( N(\mu_\kappa(\omega), \sigma_\kappa^2(\omega)) \) and \( N(\mu_\gamma(\omega), \sigma_\gamma^2(\omega)) \), with \( \mu_\kappa(\omega) = (\pi - \omega + 9)/10 \), \( \mu_\gamma(\omega) = (\pi - \omega - 1)/10 \), and \( \sigma_\kappa(\omega) = \sigma_\gamma(\omega) = (\pi - \omega + 10)/20 \).

Figs. 4, 5 and 6 show the beampatterns, steered to 20°, for the designed non-robust TD, robust TD and robust FD SBBF respectively. In order to demonstrate the designs’ robustness, we have injected both frequency dependent gain and phase errors into the designed SBBFs. Both gain and phase errors follow zero mean Gaussian distributions with standard deviation \( \sigma(\omega) = (2 - \omega/\pi)/10 \).

From Figs. 4 and 5, it is clear that the beampatterns of the TD robust design, especially at low frequency, are maintained even in the presence of errors. Similar improvement is obtained from the FD design. Significant reduction towards error sensitivity is clear from Table 1, which shows the values of the cost function in (31) and (46), evaluated for the non-robust and robust designs, with and without microphones errors.

<table>
<thead>
<tr>
<th>Table 1: Cost function values for different SBBFs designs</th>
</tr>
</thead>
<tbody>
<tr>
<td>TD design</td>
</tr>
<tr>
<td>Non-robust</td>
</tr>
<tr>
<td>Without error</td>
</tr>
<tr>
<td>With error</td>
</tr>
</tbody>
</table>

5. CONCLUSION

A robust SBBF design using the Farrow structure and for arbitrary microphone array is proposed. Robustness is achieved by modelling the microphone characteristics as random variables, which provides flexibility for errors and mismatches in the microphones. The design procedure optimises the mean performance of the SBBF in order to achieve a desired level of robustness. This error modelling method allows the SBBF robustness level to be controlled during design to match the given microphone characteristics. The design has been formulated in both TD and FD. Lastly, design examples show that the proposed robust design has significantly reduced the sensitivity of SBBFs towards microphone errors.

6. REFERENCES