THE SPHERICAL HARMONICS ROOT-MUSIC

Lalan Kumar, Guoan Bi
School of EEE
Nanyang Technological University, Singapore
Email: {lkumar,egbi}@ntu.edu.sg

Rajesh M. Hegde
Electrical Engineering Dept.
Indian Institute of Technology, Kanpur
Email: r hegde@iitk.ac.in

ABSTRACT

Spherical harmonics root-MUSIC (MUltilple SIgnal Classification) technique for source localization using spherical microphone array is presented in this paper. Earlier work on root-MUSIC is limited to linear and planar arrays. Root-MUSIC for planar array utilizes the concept of manifold separation and beamspace transformation. In this paper, the Vandermonde structure of array manifold for a particular order is proved. Hence, the validity of root-MUSIC in the spherical harmonics domain is confirmed. The proposed method is evaluated by using simulated experiments on source localization. Root mean square error analysis and statistical analysis are presented. The experimental measures at various signal to noise ratios (SNRs) show the robustness of the proposed method. The method is also verified by using experiment on real signal acquired over spherical microphone array.

Index Terms— Root-MUSIC, Spherical microphone array, Spherical harmonics, Manifold separation

1. INTRODUCTION

The use of accurate and search free algorithms for estimating direction of arrival (DOA) has been a very active research area in source localization. Root-MUSIC (MUltilple SIgnal Classification) [1] and Estimation of Signal Parameters using Rotational Invariance Techniques, ESPRIT [2], fall into this category. The root-MUSIC method estimates DOAs as the roots of MUSIC [3] polynomial in terms of spherical harmonics, called SH-MUSIC, in [11] and [12]. The Minimum Variance Distortionless Response (MVDR) spectrum in terms of spherical harmonics, SH-MVDR, was utilized for DOA estimation in [11]. MUSIC-Group delay [13] was formulated for source localization using spherical array in [14] and [15]. Differential geometry was explored for SH domain source localization in [16]. In this work, we have developed the theory of root-MUSIC in SH domain using manifold separation technique. The theory is validated using simulation and real data experiments. The proposed SH-root-MUSIC (SH-RM) technique provides exact solution without the limitation from the discretization issues associated with the SH-MUSIC and SH-MVDR methods for DOA estimation.

2. THE SPHERICAL HARMONICS DATA MODEL

A spherical microphone array of order N, radius r and the number of sensors L is considered. A sound field of L plane-waves is incident on the array with wavenumber k. The lth source location is denoted by \( \Psi_l = (\theta_l, \phi_l) \). The elevation angle \( \theta \) is measured from the positive z axis, while the azimuthal angle \( \phi \) is measured counterclockwise from the positive x axis. Similarly, the i-th sensor location is given by \( \Psi_i = (\theta_i, \phi_i) \).

In spatial domain, the sound pressure at I microphones, \( p(k) = [p_1(k), p_2(k), \ldots, p_I(k)]^T \), is written as

\[
p(k) = V(k)s(k) + n(k)
\]

where \( p(k) \equiv p(k, r, \theta, \phi) \), \( V(k) \) is an \( I \times L \) steering matrix, \( s(k) \) is a \( L \times 1 \) vector of signal amplitudes, \( n(k) \) is an \( I \times 1 \) vector of zero mean, uncorrelated sensor noise and \( (\cdot)^T \) denotes the transpose. The steering matrix \( V(k) \) is expressed as

\[
V(k) = [v_1(k), v_2(k), \ldots, v_L(k)], \text{ where}
\]

\[
v_l(k) = [e^{-jk_1r_1}, e^{-jk_2r_2}, \ldots, e^{-jk_Lr_L}]^T
\]

\[
k_i = -k \sin \theta_i \cos \phi_i, k \sin \theta_i \sin \phi_i, k \cos \theta_i
\]

\[
r_i = (r \sin \theta_i \cos \phi_i, r \sin \theta_i \sin \phi_i, r \cos \theta_i)^T
\]

where \( j = \sqrt{-1} \). The i-th term in (3) refers to the pressure due to l-th unit amplitude planewave with wavevector \( k_i \) at location \( r_i \). This may alternatively be written as [17]

\[
e^{-jk_i r_i} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} b_{m,n}(kr) Y_n^m(\Psi_i)^* Y_n^m(\Phi_i)
\]

where \( b_{m,n}(kr) \) is called mode strength.

The far-field mode strength, \( b_{m,n}(kr) \), is given by

\[
b_{m,n}(kr) = 4\pi j^n j_m(kr), \quad \text{for open sphere}
\]

\[
= 4\pi j^n (j_m(kr) - j_m^*(kr)) h_n^m(kr), \quad \text{for rigid sphere}
\]
where $j_n(kr)$ is the spherical Bessel function, $h_n(kr)$ is $n^{th}$ order spherical Hankel function of second kind and $'$ refers to the first derivative. Figure 1 illustrates mode strength $b_n$ as a function of $kr$ and $n$ for an open sphere. For $kr = 0.1$, the zeroth order mode amplitude is $22$ dB, while the first order has an amplitude of $-8$ dB. It is seen that for an order greater than $kr$, the mode strength $b_n$ decreases significantly. Therefore, the summation in (6) is truncated to a finite value of $N$, which is known as the array order.

With the introduction of spherical harmonics, the spherical harmonics decomposition of the received pressure, $p(k)$, is given as [20]

$$ p_{nm}(k) = \int_0^{2\pi} \int_0^\pi p(k)[Y_{nm}(\Phi)]^* \sin(\theta) d\theta d\phi $$

$$ \cong \sum_{i=1}^I a_i p_i(k)[Y_{nm}(\Phi_i)]^*, $$

(13)

where $p_{nm}(k)$ is spherical Fourier coefficient. The spatial sampling of pressure over a spherical microphone array is captured by using sampling weights, $a_i$ [21]. Re-writing (13) in a matrix form, we have

$$ p_{nm}(k) = Y^H(\Phi) \Gamma p(k), $$

(14)

where $p_{nm}(k) = [p_{00}, p_{1(-1)}, p_{01}, p_{11}, \ldots, p_{NN}]^T$ and $\Gamma = \text{diag}(a_1, a_2, \ldots, a_I)$. Also, under the assumption of (13), we have the orthogonality property of spherical harmonics as follows

$$ Y^H(\Phi) \Gamma Y(\Phi) = I, $$

(15)

For a given array configuration, $B(kr)$ is a constant. Therefore, we get the final spherical harmonics data model by multiplying both sides of (16) with $B^{-1}(kr)$ as

$$ a_{nm}(k) = Y^H(\Phi)s(k) + z_{nm}(k), $$

(17)

where

$$ z_{nm}(k) = B^{-1}(kr)n_{nm}(k). $$

(18)

3. THE SPHERICAL HARMONICS ROOT-MUSIC

Root-MUSIC estimates DOAs as roots of the MUSIC polynomial. Hence, we first write the MUSIC spectrum in spherical harmonics domain. Comparing the spatial data model in (1) with spherical harmonics data model in (17), $[Y^H(\Phi)(1)]_{N+1}^2 \times L$ is the steering matrix in spherical harmonics domain. Hence, the SH-MUSIC spectrum is written as

$$ P_{SH-MUSIC}(\Psi) = \frac{1}{\langle Y(\Phi)S_{nm}^{\text{SH}}(\Phi)S_{nm}^{\text{SH}}(\Phi)\rangle^H Y^H(\Phi) \Psi } $$

(19)

where $Y(\Phi)$ is a steering vector defined in (11). $S_{nm}^{\text{SH}}$ is the noise subspace obtained from eigenvalue decomposition of autocorrelation matrix, $S_{nm} = E[n_{nm}(k)n_{nm}(k)^\dagger]$. Frequency smoothing and whitening of noise should be applied as in [11]. The SH-MUSIC spectrum is shown in Figure 3(a) for two sources at $(20^\circ, 40^\circ)$ and $(20^\circ, 80^\circ)$. The two peaks in the figure correspond to the two sources.

The SH-MUSIC spectrum in (19) results in a peak which corresponds to a source owing to orthogonality between noise eigenvector and steering vector. A comprehensive search algorithm is needed to estimate the DOA of the desired source. The resolution is also limited by the resolution of discretization at which the spectrum is evaluated. The SH-root-MUSIC overcomes these limitations in estimating the DOAs. We first illustrate the Vandermonde structure in the steering vector using manifold separation technique. Utilizing...
(9) and (11), the steering vector for co-elevation \( \theta_0 \) can be written in a more compact form as

\[
y^H(\Psi) = y^H(\theta_0, \phi) = \begin{bmatrix} f_{00}, -f_{11} e^{j\phi}, f_{10}, f_{11} e^{-j\phi}, \cdots, f_{NN} e^{-jN\phi} \end{bmatrix}^T
\]

(20)

where, \( f_{nm} = \sqrt{\frac{(2n + 1)(n - |m|)!}{4\pi(n + |m|)!}} P_n^{|m|}(\cos \theta_0) \).

Then (20) is rewritten in a matrix form as

\[
y^H(\theta_0, \phi) = F(\theta_0) d(\phi)
\]

where, \( F(\theta_0) = \text{diag}(f_{00}, -f_{11}, f_{10}, f_{11}, \cdots, f_{NN}) \)

\[
d(\phi) = [1, e^{j\phi}, 1, e^{-j\phi}, \cdots, e^{-jN\phi}]^T.
\]

The matrix \( d(\phi) \) consists of only the exponent terms containing the azimuth angle and, each submatrix corresponding to a particular order follows the Vandermonde structure with common ratio as \( e^{-j\phi} \).

From (19) and (22), the SH-MUSIC cost function can be written as

\[
P_{SHM}^{-1}(\phi) = d^H(\phi) F(\theta_0) S_{num}^{NS} S_{num}^{NS} d(\phi)
\]

\[
P_{SHM}^{-1}(\phi) = d^H(\phi) F(\theta_0) C F(\theta_0) d(\phi)
\]

(25)

where \( C = S_{num}^{NS} S_{num}^{NS} \).

By defining \( z = e^{j\phi} \), the SH-MUSIC cost function now assumes a polynomial form of degree \( 4N \), given by

\[
P_{SHM}^{-1}(\phi) = \sum_{u=-2N}^{2N} C_u z^u
\]

(26)

where the coefficient \( C_u \) is obtained mathematically. The polynomial has \( 4N \) roots. If \( z \) is a root of the polynomial then \( \frac{1}{z} \) will also be the root. Hence, \( 2N \) roots are within the unit circle while the other \( 2N \) roots are outside the unit circle. Of the \( 2N \) roots within the unit circle, \( L \) roots close to the unit circle correspond to the DOAs. This is illustrated in Figure 3(b) for a fourth order spherical microphone array. The roots are plotted for two sources with co-elevation angle \( 20^\circ \) and azimuth angle \( (40^\circ, 80^\circ) \) at SNR 15dB. All the roots within and near the unit circle are shown in the figure. The DOA can be estimated from the roots by using the relation, \( \phi = \Im(\ln(z)) \), where \( \Im(\cdot) \) is the imaginary part of \( (\cdot) \).

4. PERFORMANCE EVALUATION

Simulation experiments based on source localization were carried out to evaluate the proposed SH-root-MUSIC method. Additionally, experiments were performed on real signal acquired over spherical microphone array to verify the algorithm. The experiments utilized an Eigennike\textsuperscript{®} system [22] which is shown in Figure 4. It consists of 32 microphones, embedded in rigid sphere of radius 4.2cm. The order of the microphone array was taken to be 4. Root mean square error (RMSE) and probability of resolution values were used to evaluate the source localization performance of the proposed method. The performance of the proposed method is compared to SH-MUSIC and SH-MVDR.

Fig. 4. The Eigennike\textsuperscript{®} setup in an anechoic chamber at IIT Kanpur for acquiring a far-field source.

4.1. Simulation Experiments on DOA Estimation

The RMSE analysis and statistical analysis are presented here for 500 independent Monte Carlo trials. The additive noise is assumed to be zero mean Gaussian distributed. Two sources with co-elevation \( 20^\circ \) are considered.

4.1.1. RMSE Analysis

The experiments on source localization are presented as cumulative RMSE (CRMSE) computed by

\[
CRMSE = \frac{1}{2T} \sum_{t=1}^{T} \sum_{l=1}^{2} [(\theta_t - \hat{\theta}_l(t))^2],
\]

(27)

where \( t \) is the trial index while \( l \) denotes the source index. The CRMSE values are plotted in Figure 5(a) with various SNR values
for two sources at \((20^\circ, 40^\circ)\) and \((20^\circ, 80^\circ)\). The CRMSEs values are also plotted in Figure 5(b) for the case where azimuth of one source is fixed as \(40^\circ\), while that of the other source varies at a step size of \(10^\circ\). The SNR in this case is fixed at \(20\text{dB}\). It should be noted that the proposed SH-root-MUSIC method performs reasonably better than SH-MUSIC and SH-MVDR.

4.1.2. Statistical Analysis

Statistical analysis of the proposed method is presented in terms of probability of resolution for various SNRs. Two sources with DOAs \((20^\circ, 40^\circ)\) and \((20^\circ, 80^\circ)\) are considered. The confidence interval of \(\zeta = 5^\circ\) was used for calculating the probability over 500 independent trials. The probability of resolution is given by

\[
P_r = \frac{1}{2T} \sum_{t=1}^{T} \sum_{l=1}^{2} [Pr(|\phi_l - \hat{\phi}_l^{(t)}| \leq \zeta)]
\]

\[
= \frac{1}{2T} \sum_{t=1}^{T} \sum_{l=1}^{2} [sgn(\zeta - |\phi_l - \hat{\phi}_l^{(t)}|)],
\]

where \(Pr(.)\) denotes the probability of an event, and \(sgn(x)\) is defined as

\[
sgn(x) = \begin{cases} 
  1 & \text{if } x \geq 0 \\
  0 & \text{if } x < 0.
\end{cases}
\]

The result is presented in Table 1 in which zero probability indicates inability of the methods to resolve sources in the given confidence interval. It is noted that the proposed method has higher probability of resolution when compared to other methods at all SNR values. It can also be concluded that a higher SNR is required for SH-MVDR to resolve co-elevated sources.

![Fig. 5. Cumulative RMSE for two sources with co-elevation 20°. (a) azimuth (40°, 80°) at various SNRs. (b) azimuth of one source is fixed at 40° and that of other source is varying in steps of 10°. SNR = 20dB.](image)

### Table 1. Probability of resolution performance of various methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>SNR (5dB)</th>
<th>SNR (10dB)</th>
<th>SNR (15dB)</th>
<th>SNR (20dB)</th>
<th>SNR (25dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SH-RM</td>
<td>0.5131</td>
<td>0.7575</td>
<td>0.8386</td>
<td>0.8790</td>
<td>0.9032</td>
</tr>
<tr>
<td>SH-MUSIC</td>
<td>0</td>
<td>0.6198</td>
<td>0.8051</td>
<td>0.8689</td>
<td>0.9013</td>
</tr>
<tr>
<td>SH-MVDR</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0046</td>
<td>0.3168</td>
</tr>
</tbody>
</table>

4.2. Real Data Experiments

The proposed algorithm is also verified by using real signal acquired over spherical microphone array. The experimental set-up for acquiring a source using Eigenmike™ system is shown in Figure 4. A smartphone speaker is utilized as an acoustic source. The source is fixed at location \((90^\circ, 90^\circ)\) in far-field region. A narrowband signal with frequency of 1250Hz is played. The elevation of the source is assumed to be known and the azimuth is estimated using the proposed SH-root-MUSIC method.

All the 2N\((= 8)\) roots within the unit circle are plotted in Figure 6. The root with argument close to 90° corresponds to the source and is represented by red star. It is noted that noisy roots are also competing in magnitude. The DOA estimation mismatch and multiple competing roots are due to the reflection of sound from the tripods, non-point sound source and microphone-source physical placement errors.

![Fig. 6. Azimuth estimation of a source at (90°, 90°) using SH-root-MUSIC. All roots within unit circle are shown for N = 4. The star denotes the actual estimate.](image)

5. CONCLUSION

In this paper, theory of root-MUSIC is established in spherical harmonics domain. The theory is validated using simulation and real data experiments. The SH-root-MUSIC method does not require any search to estimate the DOAs. It provides DOA estimates as direct roots of SH-MUSIC polynomial. The Vandermonde structure of array manifold in spherical harmonics domain is shown using manifold separation technique. The robustness of the method is illustrated by using source localization experiments for various SNRs and angular separations. The RMSE and probability of resolution values indicate the relevance of the proposed method. Additionally, the method is verified with real signal acquired over spherical microphone array. Owing to its robustness and high resolution, a real time implementation for voiced-based camera steering in a meeting room can be explored.
References


