The Spherical Harmonics root-MUSIC

Lalan Kumar[1], Guoan Bi[2], and Rajesh M. Hegde[3]
[1]IIT Bhubaneswar, [2]NTU Singapore, and [3]IIT Kanpur

ICASSP 2016

Introduction

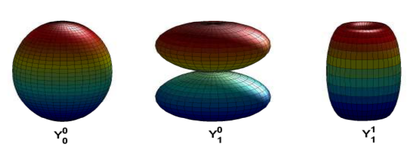
- The root-MUSIC (RM) method estimates DOAs as the roots of the MUSIC polynomial owing to Vandermonde structure of array manifold.
- Beamspace transformation based on phase mode excitation is applied for UCA to get the Vandermonde structure in array manifold with respect to azimuth angle.
- Sparse UCA root-MUSIC and manifold separation techniques were further utilized for extending ULA root-MUSIC for UCA.
- Recently, various existing DOA estimation techniques were reformulated in the spherical harmonics (SH) domain utilizing spherical microphone array.
- In this work, we have developed the theory of root-MUSIC in SH domain using manifold separation technique.

The Spherical Harmonics

- $Y_n^m(\theta,\phi)$ is called spherical harmonic of order n and degree m. It is expressed as

$$\begin{aligned} Y_n^m(\Psi) &= \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}}P_n^m(cos\theta)e^{jm\phi}, \\ &\forall 0\le n\le N, 0\le m\le n \\ &= (-1)^{|m|}Y_n^{|m|*}(\Psi), \forall -n\le m<0, \end{aligned} \quad (12)$$

where P_n^m is the associated Legendre function.



SH-RM using Manifold Separation

- Manifold separation means writing steering vector (Manifold vector) as a product of a characteristic matrix of the array and a vector with Vandermonde structure depending on the azimuth angle.
- Utilizing (14) and (12), the steering vector for co-elevation θ_0, can be written in more compact form as

$$\begin{aligned} \mathbf{y}^H(\Psi) &= \mathbf{y}^H(\theta_0,\phi) \\ &= [f_{00}, -f_{1(-1)}e^{j\phi}, f_{10}, f_{11}e^{-j\phi}, \cdots, f_{NN}e^{-jN\phi}]^T \end{aligned} \quad (19)$$

$$\text{where, } f_{nm} = \sqrt{\frac{(2n+1)(n-|m|)!}{4\pi(n+|m|)!}}P_n^{|m|}(cos\theta_0). \quad (20)$$

- Re-writing (19) in matrix form,

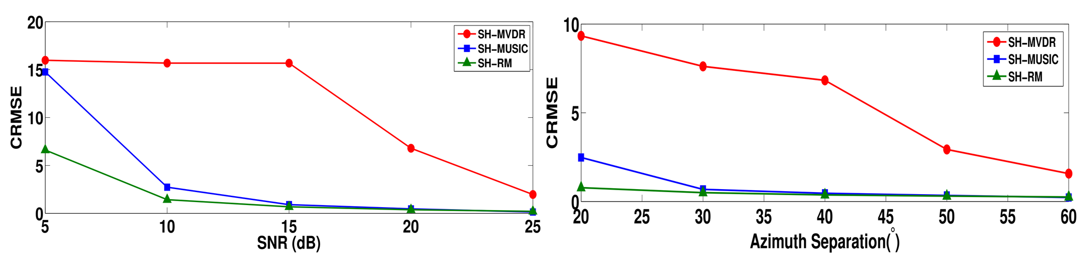
$$\mathbf{y}^H(\theta_0,\phi) = F(\theta_0)d(\phi) \quad (21)$$

$$\text{where, } F(\theta_0) = \text{diag}(f_{00}, -f_{1(-1)}, f_{10}, f_{11}, \cdots, f_{NN}) \quad (22)$$

$$d(\phi) = [1, e^{j\phi}, 1, e^{-j\phi}, \cdots, e^{-jN\phi}]^T \quad (23)$$

RMSE Analysis

- CRMSE vs (a) SNR for two sources at $(20^\circ, 40^\circ)$ and $(20^\circ, 80^\circ)$, (b) azimuth separation, azimuth of one source is fixed at 40° and that of other source is varying in steps of 10°. SNR$=20dB$.



(a) (b)

The Data Model in Spatial Domain

- A spherical microphone array of order N, radius r and the number of sensors I is considered. A sound field of L plane-waves is incident on the array with wavenumber k.
- The l^{th} source location is denoted by $\Psi_l=(\theta_l,\phi_l)$ and i^{th} sensor location is given by $\Phi_i=(\theta_i,\phi_i)$.
- In spatial domain, the sound pressure at I microphones, $\mathbf{p}(k)=[p_1(k), p_2(k), \ldots, p_I(k)]^T$, is written as

$$\mathbf{p}(k) = \mathbf{V}(k)\mathbf{s}(k)+\mathbf{n}(k), \text{ where} \quad (1)$$

$$\mathbf{V}(k) = [\mathbf{v_1}(k), \mathbf{v_2}(k), \ldots, \mathbf{v_L}(k)] \quad (2)$$

$$\mathbf{v}_l(k) = [e^{-j\mathbf{k}_l^T\mathbf{r}_1}, e^{-j\mathbf{k}_l^T\mathbf{r}_2}, \ldots, e^{-j\mathbf{k}_l^T\mathbf{r}_I}]^T \quad (3)$$

$$\mathbf{k}_l = -(k\sin\theta_l\cos\phi_l, k\sin\theta_l\sin\phi_l, k\cos\theta_l)^T \quad (4)$$

$$\mathbf{r}_i = (r\sin\theta_i\cos\phi_i, r\sin\theta_i\sin\phi_i, r\cos\theta_i)^T \quad (5)$$

- $e^{-j\mathbf{k}_l^T\mathbf{r}_i}$ is plane wave solution to the wave equation in Cartesian co-ordinates.

Steering Vector Matrix in SH Domain

- Substituting (6) and (3) in (2), the expression of steering matrix becomes

$$\mathbf{V}(k) = \mathbf{Y}(\Phi)\mathbf{B}(kr)\mathbf{Y}^H(\Psi) \quad (13)$$

- $\mathbf{Y}(\Phi)$ is $I\times(N+1)^2$ matrix whose i^{th} row is given as

$$\mathbf{y}(\Phi_i) = [Y_0^0(\Phi_i), Y_1^{-1}(\Phi_i), Y_1^0(\Phi_i), Y_1^1(\Phi_i), \ldots, Y_N^N(\Phi_i)]. \quad (14)$$

- The $(N+1)^2\times(N+1)^2$ matrix $\mathbf{B}(kr)$ is given by

$$\mathbf{B}(kr) = diag(b_0(kr), b_1(kr), b_1(kr), b_1(kr), \ldots, b_N(kr)). \quad (15)$$

SH-RM using Manifold Separation

- $d(\phi)$ consists of only exponent terms containing azimuth angle. Each submatrix corresponding to a particular order, follows Vandermonde structure.
- Utilizing (21), the SH-MUSIC cost function can be written as

$$\begin{aligned} P_{SHM}^{-1}(\phi) &= d^H(\phi)F^H(\theta_0)\mathbf{S}_{\mathbf{a_{nm}}}^{\mathbf{NS}}[\mathbf{S}_{\mathbf{a_{nm}}}^{\mathbf{NS}}]^H F(\theta_0)d(\phi) \\ &= d^H(\phi)F^H(\theta_0)\mathbf{C}F(\theta_0)d(\phi) \end{aligned} \quad (24)$$

$$\text{where, } \mathbf{C} = \mathbf{S}_{\mathbf{a_{nm}}}^{\mathbf{NS}}[\mathbf{S}_{\mathbf{a_{nm}}}^{\mathbf{NS}}]^H$$

- Utilizing (23) and $z=e^{j\phi}$ in (24), the SH-MUSIC cost function now, assumes a form of polynomial of degree $4N$, given by

$$P_{SHM}^{-1}(\phi) = \sum_{u=-2N}^{2N} C_u z^u \quad (25)$$

where the co-efficients C_u are obtained mathematically.

- If z is root of the polynomial then $\frac{1}{z^*}$ will also be the root.

Statistical Analysis

- Confidence interval of $\zeta=5^\circ$ was used for probability of resolution given by

$$P_r = \frac{1}{2T}\sum_{t=1}^{T}\sum_{l=1}^{2}[Pr(|\phi_l-\hat{\phi}_l^{(t)}|\le\zeta)] \quad (27)$$

Method	SNR (5dB)	SNR (10dB)	SNR (15dB)	SNR (20dB)	SNR (25dB)
SH-RM	0.5131	0.7575	0.8386	0.8790	0.9032
SH-MUSIC	0	0.6198	0.8051	0.8689	0.9013
SH-MVDR	0	0	0	0.0046	0.3168

Finite Order Mode Strength

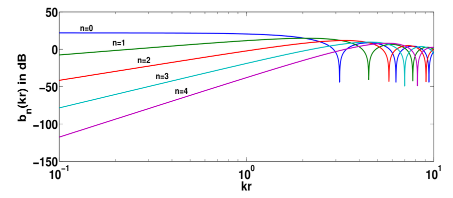
- Writing $e^{-j\mathbf{k}_l^T\mathbf{r}_i}$ in spherical co-ordinates, we have

$$e^{-j\mathbf{k}_l^T\mathbf{r}_i} = \sum_{n=0}^{\infty}\sum_{m=-n}^{n} b_n(k,r)[Y_n^m(\Psi_l)]^* Y_n^m(\mathbf{\Phi}_i) \quad (6)$$

- The far-field mode strength $b_n(k,r)$ is given by

$$b_n(kr) = 4\pi j^n j_n(kr), \quad \text{for open sphere} \quad (7)$$

$$= 4\pi j^n\big(j_n(kr) - \frac{j_n'(kr)}{h_n'(kr)}\big), \quad \text{for rigid sphere} \quad (8)$$



- b_n decreases significantly for $n>kr$. The summation in (6) can be truncated to some finite $N\ge kr$, called array order.

Final Data Model in SH Domain

- Substituting (13) in (1), then multiplying both sides with $\mathbf{Y}^H(\mathbf{\Phi})\mathbf{\Gamma}$ and utilizing the relations in (10) and (11), the data model in spherical harmonics domain can be written as

$$\mathbf{p_{nm}}(k) = \mathbf{B}(kr)\mathbf{Y}^H(\Psi)\mathbf{s}(k)+\mathbf{n_{nm}}(k). \quad (16)$$

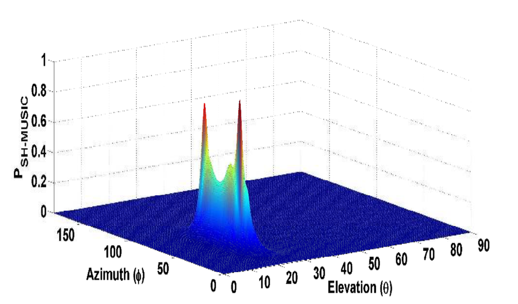
- $\mathbf{B}(kr)$ is a constant based on the array geometry. Multiplying both side by $\mathbf{B}^{-1}(kr)$, we have

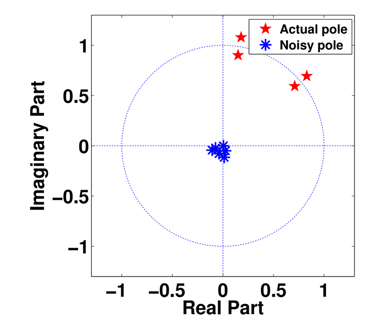
$$\mathbf{a_{nm}}(k) = \mathbf{Y}^H(\Psi)\mathbf{s}(k)+\mathbf{z_{nm}}(k), \text{ where,} \quad (17)$$

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

SH-RM using Manifold Separation

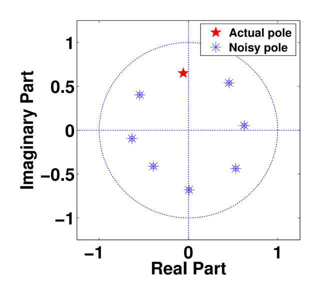
- Out of $4N$ roots, $2N$ roots will be within the unit circle and $2N$ outside the unit circle. Of the $2N$ roots within the unit circle, L roots close to unit circle correspond to the DOAs.
- As $z=e^{j\phi}$, the DOA can be estimated from the roots by using the relation, $\phi=\Im(ln(z))$, where $\Im()$ is imaginary part of $()$.
- SH-MUSIC and SH-root-MUSIC plots are illustrated in the following Figure for two sources at (20°,40°) and (20°,80°), N = 4, SNR = 15dB.





Validation with Real Data

- Eigenmike was utilized in anechoic chamber to acquire data.
- A sound with frequency 1250Hz was played using smartphone speaker fixed at $(90^\circ, 90^\circ)$ in far-field region.
- All the $2N(=8)$ roots within the unit circle are plotted in the Figure.



Spherical Fourier Transform

- The Spherical Fourier Transform (SFT) of the received pressure, $p(k)$, is given as

$$\begin{aligned} p_{nm}(k) &= \int_0^{2\pi}\int_0^{\pi} p(k)[Y_n^m(\Phi)]^*\sin(\theta)d\theta d\phi \\ &\cong \sum_{i=1}^{I} a_i p_i(k)[Y_{nm}(\Phi_i)]^*, \end{aligned} \quad (9)$$

- In matrix form for all $n\in[0,N]$, $m\in[-n,n]$ and I, the SFT becomes

$$\mathbf{p_{nm}}(k) = \mathbf{Y}^H(\mathbf{\Phi})\mathbf{\Gamma}\mathbf{p}(k), \quad (10)$$

where $\mathbf{\Gamma}=\text{diag}(a_1, a_2, \cdots, a_I)$ is matrix of sampling weights.

- Under the assumption of (9), we have the orthogonality property of spherical harmonics as

$$\mathbf{Y}^H(\mathbf{\Phi})\mathbf{\Gamma}\mathbf{Y}(\mathbf{\Phi}) = \mathbf{I}, \quad (11)$$

The Spherical Harmonics MUSIC

- Comparing the spatial data model in (1) with spherical harmonics data model in (17), $[\mathbf{Y}^H(\Psi)]_{(N+1)^2\times L}$ is the steering matrix in spherical harmonics domain.
- The SH-MUSIC spectrum can thus be written as

$$P_{SHM}(\Psi) = \frac{1}{\mathbf{y}(\Psi)\mathbf{S}_{\mathbf{a_{nm}}}^{\mathbf{NS}}[\mathbf{S}_{\mathbf{a_{nm}}}^{\mathbf{NS}}]^H\mathbf{y}^H(\Psi)}, \quad (18)$$

where $\mathbf{y}^H(\Psi)$ is a steering vector and can be written as (14).

- $\mathbf{S}_{\mathbf{a_{nm}}}^{\mathbf{NS}}$ is the noise subspace obtained from eigenvalue decomposition of autocorrelation matrix, $\mathbf{S_{a_{nm}}}=E[\mathbf{a_{nm}}(k)\mathbf{a_{nm}}(k)^H]$.

Performance Evaluation

- The experiments utilized an Eigenmike® system, consisting of 32 microphones, embedded in a rigid sphere of radius 4.2cm.
- The RMSE analysis and statistical analysis are presented for two sources at $(20^\circ, 40^\circ)$ and $(20^\circ, 80^\circ)$ using 500 independent Monte Carlo trials.
- Cumulative root mean square error (CRMSE) and probability of resolution were used to evaluate the performance of the proposed method.
- The CRMSE is computed using

$$CRMSE = \frac{1}{2T}\sum_{t=1}^{T}\sum_{l=1}^{2}[(\phi_l-\hat{\phi}_l^{(t)})^2], \quad (26)$$

Conclusions

- Theory of root-MUSIC is established in spherical harmonics domain. The theory is validated using simulation and real data experiments.
- 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.