MUSIC ALGORITHM TO LOCALIZE SOURCES WITH UNKNOWN DIRECTIVITY IN ACOUSTIC IMAGING

Forooz Shahbazi Avarvand, Andreas Ziehe and Guido Nolte

1Fraunhofer Institute FIRST, Kekuléstrasse 7, 12489 Berlin, Germany
2Humboldt University, Institute of Computer Science, Rudower Chaussee 25, 12489 Berlin, Germany

ABSTRACT
The Multiple Signal Classification (MUSIC) algorithm for acoustic imaging most commonly assumes that all sources have undirectional radiation pattern. Here we propose a modification of this algorithm such that the concept is applicable for arbitrary directional characteristics of the sources. This is accomplished by fitting for each frequency the real valued amplitudes of the acoustic model rather than assuming a fixed functional form. The mathematical problem can be solved analytically resulting in an eigenvalue problem of a real valued Hamiltonian matrix. The performance is illustrated in simulations using pure monopolar, dipolar and quadrupolar sources.

Index Terms— Multiple Signal Classification, Multipoles, Acoustic imaging

1. INTRODUCTION
The task of localizing sound sources can be solved by the Multiple Signal Classification (MUSIC) algorithm [1, 2, 3]. However, it is well known that MUSIC can be quite sensitive with respect to array imperfections and errors in the calibration data [4]. The effect of such modeling errors might be different for phases and amplitudes. Here we propose to adaptively optimize for each frequency the real valued amplitudes of the acoustic model.

In section 2 we first recall in short the theory behind the MUSIC algorithm and then describe the proposed variant. In section 3 we first discuss multipolar sources in the context of the proposed algorithm and then we illustrate the performance of the proposed algorithm in a simulation study.

2. THEORY
2.1. The MUSIC algorithm
In acoustic imaging the MUSIC algorithm is based on an analysis of cross-spectral matrices \( C(f) \) defined as

\[ C(f) = \langle x(f)x^\dagger(f) \rangle \]  

(1)

where \( x(f) \) is a column vector of length \( N \) for \( N \) sensors and denotes the Fourier transforms at frequency \( f \) of the measured signals in a specific segment, and \( \langle \cdot \rangle \) denotes expectation value which is approximated by an average over all segments. In its simplest form, MUSIC assumes that only \( P \) omni-directional sources are active where \( P \ll N \). In that case the rank of \( C(f) \) is \( P \) and \( C(f) \) spans a \( P \)-dimensional subspace of the \( N \)-dimensional sensor space.

If a source is placed at a location with distance \( d_k \) to the \( k \).th sensor and the activity of the source is \( s(f) \) then this source induces a sound pressure

\[ x_k(f) = \exp(-i2\pi f d_k / c) s(f) \equiv v_k(f) s(f) \]  

(2)

in the \( k \).th sensor, where \( i = \sqrt{-1} \) and \( c \) is the speed of sound.

The set of all active source locations defines a set of vectors \( v(f) \) which must span the same space as \( C(f) \) provided that the sources are not perfectly correlated. The MUSIC algorithm now scans a set of locations on a predefined grid, calculates for each grid point the corresponding forward vector \( v(f) \) and measures the deviation of \( v(f) \) to the subspace spanned by \( C(f) \). In practice, \( C(f) \) has always full rank and one considers the space spanned by the eigenvectors of \( C(f) \) of the \( P \) largest eigenvalues as the signal space. Let \( U \) be the \( N \times P \) matrix consisting of these eigenvectors and \( w(l) = v/||v|| \) be the normalized forward vector for the \( l \).th grid point, then

\[ \lambda(l) = \min_{\alpha} ||U\alpha - w||^2 = \sin^2 \Theta \]  

(3)

is a measure of the angle \( \Theta \) between \( w(l) \) and the space spanned by \( U \). Here, \( \alpha \) is a complex valued column vector of length \( P \). If the \( l \).th grid point coincides with one of the source locations, then in the ideal situation \( \lambda(l) = 0 \). We note that alternative formulations are possible. The chosen one is most convenient for the modification presented in the next section.

In practice one searches for maxima of the gain function

\[ h(l) = \frac{1}{\lambda(l)} \]  

(4)

Up to this point, all calculations were made independently for all frequencies \( f \), i.e. specifically \( h(l) = h(l, f) \). To combine different frequencies it was suggested to multiply \( h(l, f) \).
over all frequencies. We here suggest a slightly different approach, namely to combine frequencies as a weighted sum

\[ H(l) = \sum_{f} \frac{h(l,f)}{q(f)} \]

with \( q(f) = \sum_{l} h(l,f) \). The weighting is necessary because at low frequencies phase variations are smaller and topographies are closer to the respective signal subspace such that the low frequencies dominate the estimates. We found that this approach results in somewhat sharper images. However, the results presented below are not substantially different compared to the product approach.

2.2. Phase MUSIC

If the sources are not monopolar but have unclear directions then the amplitude may differ substantially from the monopolar case even though the involved time delays are known.

Let, again, \( U \) be the matrix containing as columns the first \( P \) eigenvectors of the measured cross-spectrum \( C(f) \) at some frequency \( f \). We here consider for each point source only a single forward vector \( v(f) \). We now assume that only the phases of \( v \) are known and not the amplitudes. Specifically, we write the \( i.th \) component of \( v \) in the form

\[ v_i = \Psi_i \sigma_i \]

where \( \sigma_i \) are unknown real numbers and \( \Psi_i \) contain just the (known) phases as \( \Psi_i = \exp(i\Phi_i) \). We denote by \( \Psi \) the diagonal matrix which contains the elements \( \Psi_i \) on the diagonal, i.e. \( \Psi_{ij} = \Psi_i \delta_{ij} \), and similarly for \( \sigma \) and \( v \) as vectors. Then we can express (6) as \( v = \Psi \sigma \) and the angle \( \Theta \) between \( v \) and space \( U \) is given as

\[ \sin^2 \Theta = \min_{\alpha,\sigma} \frac{|U\alpha - \Psi \sigma|^2}{|\alpha|^2} \equiv \min_{\alpha,\sigma}(L(\alpha,\sigma)) \]

First, we minimize over all \( \sigma \) for fixed \( \alpha \). Note, that \( \sigma \) is a vector of real numbers. The minimization is straightforward and the solution reads

\[ \sigma = \Re(\Psi U^* \alpha^*) \]

where \( \Re() \) denotes taking the real part. Inserting this into the cost function \( L \) leads to

\[ L = \frac{|U\alpha - \Psi \Re(\Psi U^* \alpha^*)|^2}{|\alpha|^2} \]

Recall that \( \Psi \) is a diagonal matrix with elements having absolute values of 1, and hence \( \Psi^* \Psi = \Psi \Psi^* = id \) and for arbitrary vectors \( y \) one has \( ||y|| = ||\Psi y|| \). \( L \) can hence be rewritten as

\[ L = \frac{|2U\alpha - \Psi(\Psi U^* \alpha^* + \Psi^* U \alpha)|^2}{4||\alpha||^2} = \frac{|U\alpha - \Psi^2 U^* \alpha^*|^2}{4||\alpha||^2} = \frac{|\Psi^* U \alpha - \Psi U^* \alpha|^2}{4||\alpha||^2} \]

Since \( L \) is real, minimization with respect to the complex \( \alpha \) can be done by writing \( L \) as \( L(\alpha,\alpha^*) \) and then setting the partial derivative with respect to \( \alpha^* \) to zero treating \( \alpha \) as an independent variable. This leads after some calculations to the equation

\[ K\alpha^* = \gamma \alpha \]

where

\[ K = U^\dagger \Psi^2 U^* \]

is a complex symmetric matrix, and the eigenvalue \( \gamma \) relates to the original cost function as

\[ L = \frac{1 - \gamma}{2} \]

As a quick consistency check we consider the case of a one dimensional \( U \) and \( v = U \). That means that \( v \) is exactly in the space spanned by \( U \). In that case \( U^\dagger \Psi^2 U^* = 1 \), the ‘wrong’ phase contained in \( U^* \) on the right, is corrected by \( \Psi^2 \). Then, \( K = 1 = \gamma \) and the cost function vanishes, i.e. \( L = 0 \).

Note, that (11) is not a standard eigenvalue equation because \( \alpha \) on the left side is complex conjugated. It can be transformed into a proper one by decomposing it into real and imaginary parts. Let \( K = K_R + iK_I \) and \( \alpha = \alpha_R + i\alpha_I \), then Eq.11 is equivalent to

\[ H \begin{pmatrix} \alpha_R \\ \alpha_I \end{pmatrix} = \begin{pmatrix} K_R & K_I \\ K_I & -K_R \end{pmatrix} \begin{pmatrix} \alpha_R \\ \alpha_I \end{pmatrix} = \gamma \begin{pmatrix} \alpha_R \\ \alpha_I \end{pmatrix} \]

Note that the matrix \( H \) is a real valued and symmetric matrix of Hamiltonian type. Hence the eigenvalues and eigenvectors are real valued and the eigenvalues always occur in pairs \( \{ \gamma, -\gamma \} \) [5].

We now summarize the theory as an algorithm:

1. For each frequency calculate \( U \) containing as columns the first \( P \) eigenvectors of the measured cross-spectrum.

2. For each source location and frequency calculate the phase delays \( \Phi_i \) to the \( i.th \) channel and construct the vector \( \Psi \) having components \( \Psi_i = \exp(i\Phi_i) \).

3. For each source location and frequency calculate for \( N \) channels an \( N \times P \) matrix \( F \) with elements \( F_{ij} = \Psi_i U_j^* \), and from that the \( P \times P \) matrix \( K = F^T F \).

4. Decompose \( K \) into real and imaginary parts as \( K = K_R + iK_I \) and calculate the largest eigenvalue \( \gamma \) of the real symmetric matrix \( H \) with

\[ H = \begin{pmatrix} K_R & K_I \\ K_I & -K_R \end{pmatrix} \]

5. For each source point \( l \) and frequency \( f \) construct the measure

\[ h(l,f) = \frac{2}{1 - \gamma} \]
6. Eventually combine different frequencies as discussed above for ’standard’ MUSIC using (5).

We refer to this approach as ’Phase-MUSIC’ as we only exploit the phases of the forward model and not the amplitudes.

3. ILLUSTRATION

3.1. Multipolar sources

We tested the method in simulations using approximate multipolar sources. To understand the implication of our approach in this case we will have a closer look at the multipolar expansion which is often only presented in its far field properties. A source monopole at location \( \mathbf{r} \) with amplitude \( s(t) \) induces in the \( k,th \) sensor at location \( \mathbf{R}_k \) a sound pressure \( x_k(t) \)

\[
x_k(t) = \frac{s(t - d_k/c)}{d_k} \equiv x_k^{\text{mon}}(t, r)
\]

where \( d_k = ||\mathbf{r} - \mathbf{R}_k|| \). We here approximate a dipolar source as a difference between two monopolar sources with identical amplitudes but located at slightly different locations

\[
x_k^{\text{disp}}(t, r, \Delta \mathbf{r}_1) = x_k^{\text{mon}}(t, r + \Delta \mathbf{r}_1) - x_k^{\text{mon}}(t, r - \Delta \mathbf{r}_1)
\]

where \( \Delta \mathbf{r}_1 \) is a small displacement vector. The approximation becomes exact in the limit \( \Delta \mathbf{r}_1 \rightarrow 0 \) keeping the total amplitude in the sensors finite.

Analogously, a quadrupole can be constructed from dipoles as

\[
x_k^{\text{quad}}(t, r, \Delta \mathbf{r}_1, \Delta \mathbf{r}_2)
\]

\[
x_k^{\text{disp}}(t, r + \Delta \mathbf{r}_2, \Delta \mathbf{r}_1) - x_k^{\text{disp}}(t, r - \Delta \mathbf{r}_2, \Delta \mathbf{r}_1)
\]

which, similarly, becomes exact in the limit \( \Delta \mathbf{r}_1 \rightarrow 0 \) and \( \Delta \mathbf{r}_2 \rightarrow 0 \).

In the Fourier domain (17) reads

\[
x_k^{\text{mon}}(f) = \exp\left(-i2\pi f d_k/c\right) s(f)
\]

and the dipole can be calculated as a first order Taylor expansion in \( \Delta \mathbf{r}_1 \) giving

\[
x_k^{\text{disp}}(f) = \frac{2(R_k - r) \cdot \Delta \mathbf{r}_1}{d_k^3} g_k(f) \exp\left(-i2\pi f d_k/c\right) s(f)
\]

with

\[
g_k(f) = \frac{1}{d_k} + \frac{i2\pi f}{c}
\]

The crucial point is behavior of the function \( g_k(f) \). For low frequencies \( g_k(f) \approx 1/d_k \) is real valued and can be absorbed into the real valued amplitudes which were fitted and can assume arbitrary values according to our model. For very large frequencies \( g_k(f) \approx i2\pi f/c \) causes a sensor independent phase shift and can be absorbed into \( s(f) \) leading again to an accurate model. For intermediate frequencies the crucial question is whether the phase of \( g_k(f) \) varies with sensor index \( k \). This variation will be small if the distance of the source to the sensors is large compared to the size of the sensor array.

3.2. Simulations

We simulated a single source placed 2m away from a sensor array consisting of 120 microphones places regularly on a spherical surface with radius 60cm. The source could be either a monopole, dipole, or quadrupole, and was located on the z-axis of the coordinate system. The displacement vectors were chosen to be \( \Delta \mathbf{r}_1 = (0 0 0)\ T \) cm and \( \Delta \mathbf{r}_2 = (0.5 0 0)\ T \) cm where the x-, and y-axes correspond to horizontal and vertical direction, respectively. As source activity we chose a speech signal saying ’Good morning’ at a sampling rate of 16 kHz.

As source space we chose a square of size 1m×1m orthogonal to and centered on the z-axis. The distance of the this square to the sensor array was 3m thus deliberately avoiding perfect reconstruction. The square was covered by a grid consisting of 101×101 points corresponding to a distance of 2cm between adjacent grid points.

To estimate cross-spectral matrices the data were divided into segments consisting of 50 samples corresponding to 42msec duration resulting in a frequency resolution of 320Hz. Each segment was windowed with a Hanning function and Fourier transformed to calculate the cross spectral matrices \( C(f) \) with (1). For each frequency we chose \( P=10 \) eigenvectors corresponding to the 10 largest eigenvalues of \( C(f) \). The scans were performed up to 3kHz.

In Fig.1 we show results for MUSIC and Phase-MUSIC scans for multipolar sources. We observe that in contrast to MUSIC, the Phase-MUSIC approach correctly recovers dipolar and multipolar sources. We also observe that Phase-MUSIC scans are typically much sharper than MUSIC scans. That can be explained by the fact that the model mismatch due to wrong distance of the considered source space, i.e. 3m rather than 2m, is larger for the amplitudes than for the phases.

In Fig.2 we show Phase-MUSIC scans for the quadrupolar source for each frequency. The main observation is that the method works excellently for small frequencies but becomes poor for frequencies above around 2.5 KHz. Eventually ghost sources appear at the edges of the considered source space.

4. CONCLUSION

We proposed a modification of the well known MUSIC algorithm to account for cases in which a good model for the phase
delay between sources and sensors are given while the amplitudes are largely or completely unknown. All corresponding equations could be solved analytically leading to an eigenvalue equation of Hamiltonian type. We have shown in simulations that the method works satisfactorily for multipolar sources while the standard MUSIC approach breaks down completely. We also found that the scans for the Phase-MUSIC algorithm are sharper than the ones for MUSIC algorithm.

It is obvious that it is disadvantageous to not use amplitude information of the forward model if that information is accurate. In general, this can be a cause of non-robustness of the method. On the other hand, for sources with amplitude heavily depending on direction a false model may substantially bias the source reconstruction. Whether the proposed method is useful then apparently depends on the studied sources. In the future we will address the question in which situations the trade-off between bias and variance favors the use of the proposed algorithm over the MUSIC approach.

Fig. 1. Results for MUSIC scans (left panels) and Phase-MUSIC scans (right panels) for multipolar sources. The true source location is always in the center of the squares. We display $\log(H(l))$ (see (5)) rather than $H(l)$, because $H(l)$ is in some cases so focal that only a single point would be visible.

Fig. 2. Results for Phase-MUSIC for a quadrupolar source as a function of frequency. Scans at lower frequencies have much higher amplitude (not shown) and dominate the result for frequency averaged scans.

5. REFERENCES


