COMPARISON OF SUBSPACE-BASED AND STEERED BEAMFORMER-BASED REFLECTION LOCALIZATION METHODS

Edwin Mabande[1], Haohai Sun[2], Konrad Kowalczyk[1], and Walter Kellermann[1]

[1]University of Erlangen-Nuremberg, Multimedia Communications and Signal Processing, Erlangen, Germany
[2]Norwegian Uni. of Sci. and Tech., Acoustics Research Center, Dept. of Elec. and Telecom., Trondheim, Norway
{mabande,kowalczyk,wk}@LNT.de, haohai.sun@ieee.org

ABSTRACT

This paper presents a comparative study of several techniques that are applicable to the localization of the most significant reflectors in an acoustic environment. It focuses on subspace-based and steered beamformer-based localization techniques and includes eigenbeam (EB)-ESPRIT, EB-MUSIC, and EB-minimum variance distortionless response (EB-MVDR) methods. The directions of arrival (DOAs) of the sound sources and reflective surfaces in an acoustic environment are estimated from the acoustic images of rooms, which are obtained using a microphone array. The applicability of the presented methods in practice is confirmed by experimental results from real measurements.

1. INTRODUCTION

When microphone arrays are employed for sampling wavefields, signal processing of the microphone data may be used for extracting parameters characterizing an acoustic environment. Acoustic source localization aims at extracting the localization information of one or several sound sources from signals captured with an array of sensors. Spherical microphone array signal processing has recently become an important technique in the application of 3D DOA estimation [1, 2, 3]. The main reason is that the signal processing can be performed using the elegant spherical harmonics (eigenbeams, EB) for representation of acoustic wavefields.

In this work, we aim to accurately localize one source and several dominant reflections in an enclosure from a single measurement by using a spherical microphone array. In order to obtain accurate DOA estimates, robust and high resolution source localization algorithms are needed which are capable of localizing coherent sources. We present a comparative study of several techniques to localize major reflectors in an enclosure using subspace-based and steered beamformer-based localization techniques based solely on the recorded microphone signals. The knowledge of the DOAs of early reflections is of interest, e.g., for signal enhancement methods such as dereverberation [4], 3D room geometry inference [5], robust adaptive beamforming [6], and matched filter based signal recovery [7].

In general, the steered beamformer-based localization approach relies on scanning the room using, e.g., the EB-MVDR beamformer [5] and localizing acoustic sources as the peaks in an acoustic map that shows the output power for each look-direction. On the other hand, subspace source localization approaches perform an eigenvalue decomposition (EVD) of the covariance matrix constructed from the microphone signals. By observing the eigenstructure of the sensor covariance matrix, a signal or noise subspace can be extracted and the position of the sound sources can be estimated. This class of localization algorithms include the EB-MUSIC [2] and EB-ESPRIT [8, 9] algorithms. Subspace based high resolution DOA estimation methods using spherical arrays have been studied in [1, 10]. A major limitation of these algorithms is that the number of sources must be known a priori in order to obtain accurate results. Focusing matrices and frequency smoothing techniques are employed to improve the robustness of DOA estimation of coherent sources [10].

Experimental results obtained by using an Eigenmike [3] in a real room are used to compare the performance of the various DOA estimation methods.

2. WAVEFIELD DECOMPOSITION AND FREQUENCY SMOOTHING

Three-dimensional (3D) eigenbeam (EB) array processing is based on the transformation of the original microphone signals into the EB domain (i.e., the spherical harmonics domain) using a spherical microphone array. After the transformation, the calculation of array manifold vectors in the EB domain can be much simpler than in the traditional element-space (corresponding to the microphone signals), and the inherently frequency-independent manifold vectors can be obtained by decoupling and removing the frequency-dependent components from the EB domain manifold vectors.

2.1 Spherical harmonic decomposition

In this section, the transformation of the original microphone signals to the spherical harmonics (eigenbeam) domain is described. In the following, a conventional spherical coordinate system (r,θ,ϕ) is used, where r denotes the radius, and θ and ϕ denote the elevation and azimuth angular displacements in radians measured from the positive z-axis and x-axis in Cartesian coordinates, respectively. Consider a spherical microphone array consisting of M microphones located at a radius a, and a plane wave with a unit amplitude and frequency ω arriving at the array from direction $\Omega_d = (\theta_d, \phi_d)$. The sound pressure at any microphone position $\Omega_s = (\theta_s, \phi_s)$, $s = 1 \ldots M$, in the frequency domain can be written as

$$P(ka, \Omega_d, \Omega_s) = \sum_{n=0}^{\infty} b_n(ka) \sum_{m=-n}^{n} [Y_n^m(\Omega_d)]^* Y_n^m(\Omega_s), \quad (1)$$

where $k = \omega/c$ and c is the speed of sound. The corresponding expression in the spherical harmonics domain takes the form

$$P_{nm}(ka, \Omega_d) = b_n(ka)[Y_n^m(\Omega_d)]^*, \quad (2)$$

where Y_n^m is the spherical harmonic of order n and degree m, superscript $*$ denotes complex conjugation, and $b_n(ka)$ depends on the sphere configuration, e.g., rigid or open sphere, as described in [8]. Considering D plane waves impinging on the microphone array from directions $\Omega_1,\ldots,\Omega_D$ and adding uncorrelated noise, the signal captured by the microphone at a position Ω_s can be written as

$$X(ka,\Omega_s)=\sum_{d=1}^{D}P(ka,\Omega_d,\Omega_s)S_d(k)+V(k,\Omega_s) \quad (3)$$

where $S_d(k)$ denotes the d-th source signal spectrum and $V(k,\Omega_s)$ is the additive noise spectrum.

Applying the n-th order spherical Fourier transformation (see e.g., [8] for details), the spherical harmonics domain microphone signal $X_{nm}(ka)$ is given by

$$X_{nm}(ka)=\sum_{d=1}^{D}P_{nm}(ka,\Omega_d)S_d(k)+V_{nm}(k) \quad (4)$$

where $V_{nm}(k)$ denotes the spherical Fourier transform of the noise. Finally, we can write the $(N+1)^2\times 1$ spherical harmonics domain signal vector $\mathbf{X}_{nm}$ (where N is the largest order considered), obtained by applying the spherical Fourier transform to the microphone signal vector, as

$$\mathbf{X}_{nm}(ka)=\bar{\mathbf{P}}_{nm}(ka,\Omega)\mathbf{S}(k)+\mathbf{V}_{nm}(k) \quad (5)$$

where the $(N+1)^2\times D$ associated manifold vector $\bar{\mathbf{P}}_{nm}(ka,\Omega)=[\mathbf{P}_{nm}(ka,\Omega_1),\ldots,\mathbf{P}_{nm}(ka,\Omega_D)]$ with each element defined as $\mathbf{P}_{nm}=[P_{n(-n)},P_{n(-n+1)},\ldots,P_{n(n-1)},P_{nn}]$, $\mathbf{S}(k)=[S_1(k),\ldots,S_D(k)]$ is the $D\times 1$ source signal spectra vector, and $\mathbf{V}_{nm}(k)$ is the $(N+1)^2\times 1$ vector of the additive noise spectrum. The corresponding $(N+1)^2\times(N+1)^2$ spherical harmonics domain covariance matrix $\mathbf{R}_{nm}$ of $\mathbf{X}_{nm}$ is given by

$$\mathbf{R}_{nm}(ka)=E\{\mathbf{X}_{nm}(ka)\mathbf{X}_{nm}^H(ka)\}, \quad (6)$$

where $E\{.\}$ denotes the statistical expectation and $(.)^H$ is the Hermitian transpose.

2.2 Frequency smoothing

Many localization techniques suffer from severe performance degradation in environments where interference sources are highly correlated with the desired source, such as is the case for our scenario where the direct path signal is coherent with its early reflections. For such signals the reduction of the correlation between the desired signals and the interference signals can be achieved by using a frequency smoothing technique [5, 10]. As shown in (2), the sound pressure representation can be separated into the frequency-dependent mode amplitude $b_n(ka)$ and the angular-dependent spherical harmonics, so that we can rewrite the array manifold matrix in (5) as

$$\bar{\mathbf{P}}_{nm}(ka,\Omega)=\mathbf{B}_n(ka)\mathbf{Y}(\Omega) \quad (7)$$

where $\mathbf{B}_n(ka)=\mathrm{diag}[b_0,b_1,b_1,b_1,\ldots,b_N]$ is an $(N+1)^2\times(N+1)^2$ diagonal matrix and

$$\mathbf{Y}(\Omega)=[\mathbf{y}(\Omega_1),\ldots,\mathbf{y}(\Omega_D)] \quad (8)$$

where

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

Since the configuration of the spherical array is known, the following focusing matrix $\mathbf{T}(k_q)$ is obtained [5]

$$\mathbf{T}(k_q)=\mathbf{B}_n(k_qa)^{-1}\mathbf{B}_n(k_0a) \quad (9)$$

where $k_0=\omega_0/c$ and ω_0 is the focusing frequency. The focused and frequency-smoothed covariance matrix $\hat{\mathbf{R}}_{nm}(k_0a)$ is then given by

$$\hat{\mathbf{R}}_{nm}(k_0a)=\frac{1}{Q}\sum_{q=1}^{Q}\mathbf{T}(k_q)\mathbf{R}_{nm}(k_qa)\mathbf{T}^H(k_q) \quad (10)$$

where Q is the number of frequency bins over which such a smoothing is performed.

3. LOCALIZATION ALGORITHMS

The EB-MVDR beamformer, EB-MUSIC, and EB-ESPRIT which were used for estimating the DOAs of the direct signal and its reflections are briefly described in this section.

3.1 EB-MVDR

For the estimation of DOAs corresponding to the original source and reflected signals, the room is scanned using the EB-MVDR beamformer and the output power for each look-direction is plotted to form an acoustic map of the environment. The locations of the peaks of the acoustic map determine the estimated DOAs.

The focused and frequency smoothed EB-MVDR beamformer aims to minimize the output power subject to a distortionless constraint on the beamformer response in the look direction. The cost function is written as [5]

$$\min_{\mathbf{W}_{nm}}\mathbf{W}_{nm}^H(k_0)\hat{\mathbf{R}}_{nm}(k_0a)\mathbf{W}_{nm}(k_0)$$

subject, to

$$\mathbf{W}_{nm}^H(k_0)\mathbf{P}_{nm}(k_0a,\Omega_l)=\frac{4\pi}{M}, \quad (11)$$

where Ω_l is the beamformer look direction. Note that a uniform sampling over the sphere is assumed here, and thus the output amplitude for the spherical harmonics domain array processing is higher than for the conventional element-space domain by a factor of $4\pi/M$ [2]. Using the method of Lagrange multipliers to solve (11), the solutions for the array weight vector and the output power spectrum are

$$\mathbf{W}_{nm}(k_0)=\frac{\hat{\mathbf{R}}_{nm}^{-1}(k_0a)\mathbf{P}_{nm}(k_0a,\Omega_l)}{\mathbf{P}_{nm}^H(k_0a,\Omega_l)\hat{\mathbf{R}}_{nm}^{-1}(k_0a)\mathbf{P}_{nm}(k_0a,\Omega_l)} \quad (12)$$

and

$$Z_{\text{MVDR}}(k_0,\Omega_l)=\frac{1}{\mathbf{P}_{nm}^H(k_0a,\Omega_l)\hat{\mathbf{R}}_{nm}^{-1}(k_0a)\mathbf{P}_{nm}(k_0a,\Omega_l)}, \quad (13)$$

respectively.

3.2 EB-MUSIC

EB-MUSIC is a simple DOA estimation method with high resolution. The EB-MUSIC spectrum can be computed as [2]

$$Z_{\text{MUSIC}}(k_0,\Omega) = \frac{1}{\mathbf{P}_{nm}^H(k_0 a,\Omega)\mathbf{E}(k_0)_V\mathbf{E}_V^H(k_0)\mathbf{P}_{nm}(k_0 a,\Omega)}, \tag{14}$$

where the columns of the matrix $\mathbf{E}_V(k_0)$ are the eigenvectors of matrix $\hat{\mathbf{R}}_{nm}(k_0 a)$ associated with the $(N+1)^2 - D$ smallest eigenvalues, which may be conveniently calculated using an eigenvalue decomposition of $\hat{\mathbf{R}}_{nm}(k_0 a)$, forming the noise subspace. Peaks in the MUSIC spectrum, which is as a function of Ω, indicate the DOAs of the plane waves.

3.3 EB-ESPRIT

The computational cost of EB-MUSIC is still quite high due to the fact that a full 2D/3D spatial spectrum computation and peak searching are necessary. On the other hand, EB-ESPRIT can reduce the computational cost since it automatically provides the closed-form source localization solutions [8]. An EB-ESPRIT matrix is constructed based on the signal subspace of the signal covariance matrix in the EB domain, and an EB-ESPRIT equation can be formulated based on a displacement invariance structure of circular harmonics or a recurrence relation for spherical harmonics, which can be easily solved in the sense of total least squares [1]. Then, by computing the eigenvalues of the solution of the EB-ESPRIT equation, the DOA information can be directly obtained from the phases and amplitudes of the eigenvalues [8] as follows

$$\theta_d = \tan^{-1}|\mu_d|, \tag{15}$$

$$\phi_d = \arg(\mu_d). \tag{16}$$

where μ_d, $d = 1,\ldots,D$, are the eigenvalues.

Improvements to several issues that occur in practical implementations of the spherical array EB-ESPRIT algorithm were proposed in [9]. Such improvements include the use the condition number of the EB-ESPRIT matrix as a robustness measure, the use of the Wigner-D eigenbeam space rotation [2] technique to avoid the ill-conditioning of the matrix and improve the robustness, and the application of the manifold vector extension technique [11] to address the problem of localization of large source numbers.

4. EXPERIMENTAL EVALUATION OF DOA ESTIMATION

In order to compare the performance of the three methods, experimental measurements were carried out in a rectangular room with a T_{60} of about 900ms. Fig. 1 shows the dimensions of the room and the measurement setup. The height of the room is 3.02m and a spherical array was placed centrally in the room at a height of 1.41m. A white noise signal was played back via a loudspeaker that was placed at approximately $(90^\circ, 0^\circ)$ relative to the array. The microphone signals were recorded and then processed offline. The signal to noise ratio (SNR) was approximately 21dB and the sampling rate was 44.1kHz.

The Eigenmike® [3] microphone array was used for the experiments. It has a radius of 0.042m and consists of 32 well-calibrated high-quality microphones placed on the surface of a rigid sphere.

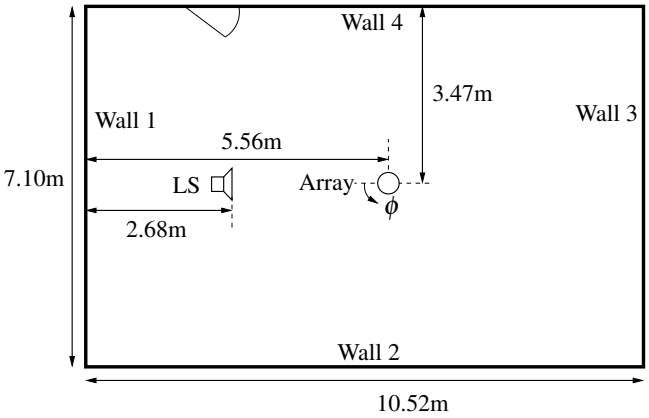

Figure 1: Experimental setup.

To verify the performance of the aforementioned algorithms for reflection localization, the results were compared with the 'ground truth' DOA values. The 'ground truth' is based on the manually measured dimensions of the room, the positions of the loudspeaker and spherical array. These were fed into the CATT acoustics room simulation software [12] which provided the DOAs of the reflections relative to the array axis. Only reflections of up to second order were considered.

Table 1: Ground truth DOAs $(\theta_{\text{ref}}, \phi_{\text{ref}})$ [degrees]

Ref.	Roof	Floor	Wall 2	Wall 3	Wall 4
(90°,0°)	(41°,0°)	(135°,0°)	(90°,69°)	(90°,180°)	(90°,293°)

The EB-MVDR and EB-MUSIC with frequency smoothing and a focusing frequency 4.5kHz ($k_0 a = 3.5$) were used with an angular resolution of 1 degree. The frequency smoothing range used is 1.3kHz to 4.5kHz, i.e., $ka \in [1, 3.5]$. Diagonal loading [13, 14] and noise pre-whitening at low frequencies were employed to improve the algorithm robustness. Additionally, the power spectra (acoustic maps) were also compared with those of the narrowband EB-MVDR obtained by using (6) in (11) for $ka = 3.5$ and the narrowband EB-MUSIC, i.e., the columns of the matrix $\mathbf{E}_V$ in (14) are eigenvectors of (6) for $ka = 3.5$.

Fig. 2a depicts the localization results for the narrowband EB-MVDR and Fig. 2b depicts the results for EB-MVDR with frequency smoothing. The peaks on the 2D acoustic maps correspond to the DOAs of the sound source and its reflections. The EB-MVDR with frequency smoothing has significantly higher resolution than the narrowband MVDR which suffers from the self-cancellation of the sources [6]. Up to six sources, i.e., the direct sound and five first order reflections, in the room can be well localized by the EB-MVDR with frequency smoothing.

Fig. 3a depicts the localization results for the narrowband EB-MUSIC and Fig. 3b depicts the results for EB-MUSIC with frequency smoothing. The peaks on the presented acoustic maps also clearly correspond to the DOAs of the sound source and its reflections. The EB-MUSIC with frequency smoothing has significantly higher resolution than the narrowband EB-MUSIC. Comparing Fig. 2b and 3b the

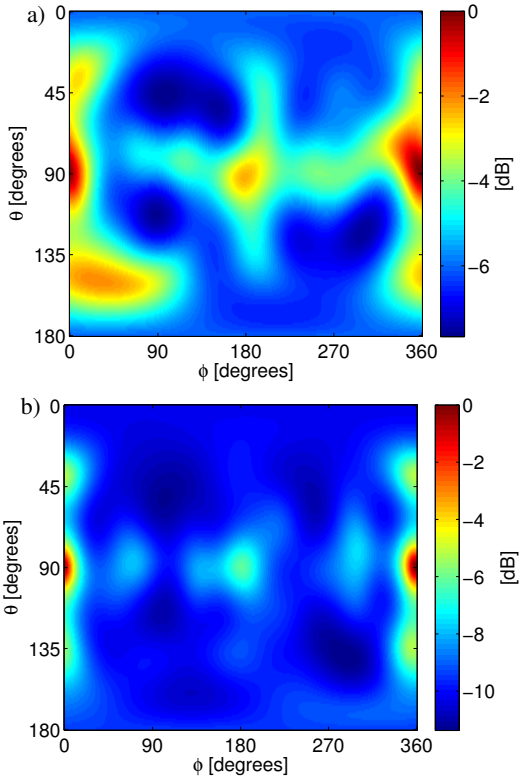

Figure 2: Localization results for a) narrowband EB-MVDR and (b) EB-MVDR with frequency smoothing

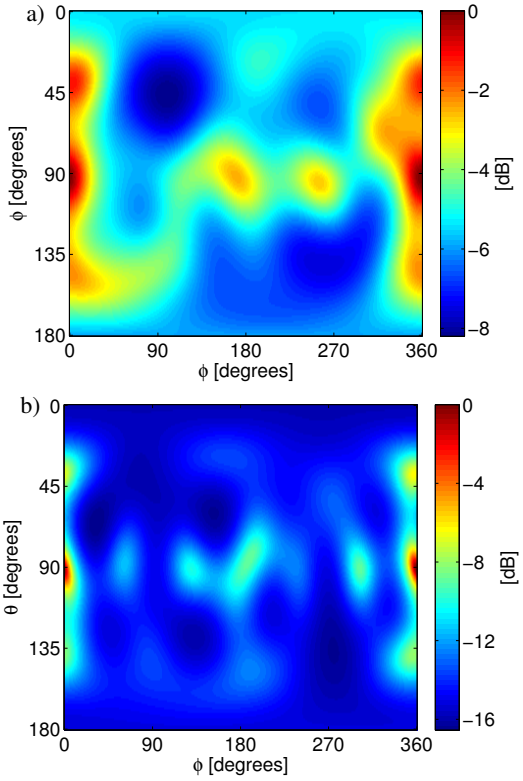

Figure 3: Localization results for a) narrowband EB-MUSIC and (b) EB-MUSIC with frequency smoothing.

results of the algorithms are similar, although the resolution of the EB-MUSIC acoustic map is slightly higher if the number of sources is known. As the exact number of significant reflections is typically unknown in practice, this may lead to a strong performance degradation. It is important to note that EB-MUSIC requires the number of sources as input. Therefore, since the results of the two are comparable, the EB-MVDR with frequency smoothing is the preferable choice since it is robust and does not require any knowledge about the number of sources as input.

Table 2: DOA estimates, $(\hat{\theta},\hat{\phi})$ [degrees], for EB-MVDR and EB-ESPRIT (NB: Narrowband, FS: Frequency Smoothing)

	EB-MVDR		EB-MUSIC	
Ref.	FS	NB	FS	NB
(90°,0°)	(89°,358°)	(88°,358°)	(91°,359°)	(92°,359°)
(41°,0°)	(39°,358°)	(37°,9°)	(38°,356°)	(40°,2°)
(135°,0°)	(134°,355°)	(149°,29°)	(138°,356°)	(144°,359°)
(90°,69°)	(87°,69°)	-	(88°,61°)	-
(90°,180°)	(90°,181°)	(92°,179°)	(85°,189°)	(92°,172°)
(90°,293°)	(81°,298°)	-	(91°,301°)	(95°,255°)

Table 2 shows the localization results for the EB-MVDR and EB-MUSIC (both the narrowband (NB) and frequency smoothed (FS) DOA estimation results). For convenience of comparison the first column again shows the DOAs of the 'ground truth' reference. It is clear that only the algorithms which use frequency smoothing can localize up to six reflections or more and have higher accuracy.

In order to have a clearer view of the accuracy of the algorithms, the angular deviation from the 'ground truth' was computed by $\theta_{\text{dev}} = |\theta_{\text{ref}} - \hat{\theta}|$ and $\phi_{\text{dev}} = |\phi_{\text{ref}} - \hat{\phi}|$. Table 3 shows that the DOA deviations for the localization procedures. Taking the measurement errors and the random errors prevalent in a real microphone array into consideration, the deviation in the DOA estimation is acceptable, especially when one considers the first three to four reflections. It should also be noted that, in general, the angular deviation tends to increase with decreasing power of the reflected signal, as shown in Table 4.

Table 3: DOA deviation, $(\theta_{\text{dev}},\phi_{\text{dev}})$ [degrees], from 'ground truth'

EB-MVDR		EB-MUSIC	
FS	NB	FS	NB
(1°,2°)	(2°,2°)	(1°,1°)	(2°,1°)
(2°,2°)	(4°,9°)	(3°,4°)	(1°,2°)
(1°,5°)	(14°,29°)	(3°,4°)	(9°,1°)
(0°,1°)	(2°,1°)	(5°,9°)	(2°,8°)
(9°,5°)	-	(1°,8°)	(5°,38°)
(3°,0°)	-	(2°,8°)	-

EB-ESPRIT was also used to localize sources, for which the frequency smoothing range was 2kHz to 3kHz. This resulted in two dominant coherent signals, i.e., the direct sound

Table 4: DOA deviation, $(\theta_{\text{dev}}, \phi_{\text{dev}})$ [degrees], and output power Z_{MVDR} of the EB-MVDR with frequency smoothing, normalized to the power of the direct path

$(\theta_{\text{dev}}, \phi_{\text{dev}})$	(1°,2°)	(2°,2°)	(1°,5°)	(0°,1°)	(9°,5°)	(3°,0°)
Z_{MVDR} [dB]	0	-5.14	-5.2	-6.16	-7.55	-7.85

and the reflection from wall 2. Since most of the sources are distributed at the $\pi/2$-plane, it is difficult to localize them due to the inherent ill-conditioning issue explained in [9]. By tilting the coordinate system by 45° using the Wigner-D weighting $(\theta + 45, \phi)$, it is possible to find DOAs which correspond to the direct sound and one reflection (from the front wall) as shown in Table 5. By tilting the spherical coordinate, the EB-ESPRIT algorithm is able to localize the direct sound and one reflection fairly accurately, but still finds it difficult to localize more reflections due to the high reverberation in the room. Therefore, this localization algorithm is not well suited for localization of multiple reflections in a reverberant room.

Table 5: DOA [degrees] results for EB-ESPRIT

Ref.	EB-ESPRIT	Deviation
(45°, 0°)	(44°, 356°)	(1°,4°)
(135°, 180°)	(136°, 175°)	(1°,5°)

5. CONCLUSIONS

In this paper, a comparative study of beamforming-based and subspace-based localization techniques for DOA estimation of room reflections has been presented. EB-MUSIC and EB-MVDR techniques were shown to be applicable to localizing multiple room reflections from just one measurement using a spherical microphone array. Furthermore, frequency smoothing increases both accuracy and robustness of the presented algorithms. The EB-MVDR was shown to give comparable resolution and accuracy as the well tuned EB-MUSIC. Moreover, it does not require any knowledge about the number of sources as input. Although EB-ESPRIT has been shown to successfully localize the direct signal and one reflection fairly accurately, it is not suitable for the localization of many reflections.

6. ACKNOWLEDGEMENT

This work was partially supported by the Future and Emerging Technologies (FET) programme within the Seventh Framework Programme for Research of the European Commission, under FET-Open grant number: 226007 SCENIC and by the QUEVIRCO project, financed by the Research Council of Norway, TANDBERG (now part of CISCO), and Statoil AS.

REFERENCES

[1] H. Teutsch, *Modal array signal processing: Principles and applications of acoustic wavefield decomposition*, Springer-Verlag, Berlin/Heidelberg, 2007.

[2] B. Rafaely, Y. Peled, M. Agmon, D. Khaykin, and E. Fisher, "Spherical microphone array beamforming," in *Speech Processing in Modern Communication: Challenges and Perspectives*. I. Cohen and J. Benesty and S. Gannot, Eds. Springer, Berlin, 2010.

[3] J. Meyer and G. Elko, "A highly scalable spherical microphone array based on an orthonormal decomposition of the soundfield," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1781–1784, May 2002.

[4] Y. Peled and B. Rafaely, "Method for dereverberation and noise reduction using spherical microphone arrays," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 113–116, March 2010.

[5] H. Sun, E. Mabande, K. Kowalczyk, and W. Kellermann, "Joint DOA and TDOA estimation for 3D localization of reflective surfaces using eigenbeam MVDR and spherical microphone arrays," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011.

[6] H. Sun, S. Yan, and U. P. Svensson, "Space domain optimal beamforming for spherical microphone arrays," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 117–120, March 2010.

[7] A. O'Donovan, R. Duraiswami, and D. Zotkin, "Automatic matched filter recovery via the audio camera," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 2826–2829, March 2010.

[8] H. Teutsch and W. Kellermann, "Detection and localization of multiple wideband acoustic sources based on wavefield decomposition using spherical apertures," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 5276–5279, 2008.

[9] H. Sun, H. Teutsch, E. Mabande, and W. Kellermann, "Robust localization of multiple sources in reverberent environments using EB-ESPRIT with spherical microphone arrays," *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011.

[10] D. Khaykin and B. Rafaely, "Coherent signals direction-of-arrival estimation using a spherical microphone array: Frequency smoothing approach," *Proc. IEEE WASPAA*, pp. 221–224, Oct. 2009.

[11] R. Goossens and H. Rogier, "Unitary spherical ESPRIT: 2-D angle estimation with spherical arrays for scalar fields," *IET Signal Processing*, vol. 3, no. 3, pp. 221–231, 2009.

[12] B. Dalenback, *CATT-Acoustic*, CATT, Svanebacksgatan 9B, S-41452, Gothenburg, Sweden.

[13] J. Bitzer and K. U. Simmer, "Superdirective microphone arrays," in *Microphone Arrays: Signal Processing Techniques and Applications*. M.S. Brandstein and D.B. Ward, Eds. Springer-Verlag, Berlin, Germany, 2001.

[14] H. L. Van Trees, *Optimum Array Processing: Detection, Estimation, and Modulation Theory*, Wiley, New York, 2002.