ESTIMATION OF THE SOUND FIELD AT ARBITRARY POSITIONS IN DISTRIBUTED MICROPHONE NETWORKS BASED ON DISTRIBUTED RAY SPACE TRANSFORM

M. Pezzoli, F. Borra, F. Antonacci, A. Sarti, S. Tubaro

Politecnico di Milano, p.zza Leonardo da Vinci 32, 20133 Milano - Italy

ABSTRACT

In this paper we propose a parametric sound field reconstruction approach. In particular, the technique is based on the estimation of three parameters for each acoustic source (source position, radiation pattern and source signal) given the signals acquired by few arbitrarily placed microphone arrays. This allows us to synthesize the signal of a virtual microphone placed in any point of the acoustic scene.

***Index Terms*—** distributed microphone networks, source localization, virtual microphone

1. INTRODUCTION

An emerging application in space-time processing is that of *Virtual Miking* (VM) . The VM is a procedure which concerns the synthesis of a *virtual microphone* signal starting from the information given by other acoustic sensors. The virtual microphone can be arbitrarily placed in the space, thus changing the listening point. This is possible only if the acoustic field is known or can be reconstructed at the desired location. Therefore, the problem concerns the reconstruction of the sound field at every point in space.

In the literature different approaches to the sound field reconstruction can be found [1, 2, 3, 4]. Here, we propose a parametric approach to the problem, in which the acoustic information is acquired by means of a network of circular compact microphone arrays that can be arbitrarily placed around the sources. This enables the adoption of a geometry-based acoustic representation of the sound field defined as a parametrization of the plenacoustic function [5] in terms of acoustic rays [6]. The mapping of the acoustic information in the ray space domain is here performed by a novel tool called Distributed Ray Space Transform (DRST), which generalizes the Ray Space Transform to arbitrary array geometries [7].

The parametric approach adopted here is similar to the one presented in [4], but in our case the parameters to be estimated are directly related to the source: the source position, the source radiance pattern and the source signal. The DRST enables the estimation of these parameters individually and the inference of the direct sound field in each point in the space. In this work we limit ourselves to a free field scenario, but the DRST can be generalized to reverberant environments.

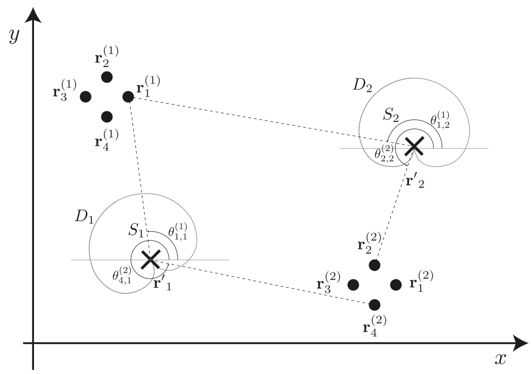


Fig. 1: Graphical representation of the model in (1)

The paper is structured as follows: in Sec. 2 the parametric sound field model is introduced and the problem is formulated. Sec. 3 presents the estimation of the model parameters. In Sec. 4 the synthesis of the VM signal is described. Sec. 5 reports the results obtained during simulations and experiments. Finally, Sec. 6 draws conclusions.

2. SIGNAL MODEL

Consider a distributed microphone network composed by A compact arrays, each composed by M microphones. Let $a = 0, \dots, A-1$ be the array index and $m = 0, 1, \dots, M-1$ the microphone index within a given array. Let us start modeling the signal acquired by the generic mth microphone within the ath array. Where not needed, for reasons of compactness in the notation, we omit the array index a. Assuming sources and microphone to lie on a horizontal plane, this model is defined starting from the far field solution of the Rayleigh's first integral [8], i.e [9]

$$P(\mathbf{r}_m^{(a)}, \omega) = \sum_{n=0}^{N-1} D_n(\theta_{m,n}^{(a)}, \omega) g(\mathbf{r}_m^{(a)}, \mathbf{r'}_n, \omega) S_n(\omega) + e_m^{(a)}(\omega), \tag{1}$$

where ω is the temporal frequency, $\mathbf{r}_m^{(a)}$ is the position vector of the mth microphone in the ath array, $\mathbf{r'}_n$ is the position vector of the nth source, N is the number of sources, $S_n(\omega)$ is the signal emitted by the nth source and $e_m^{(a)}(\omega)$ is a noise term representing the microphone self-noise. The function

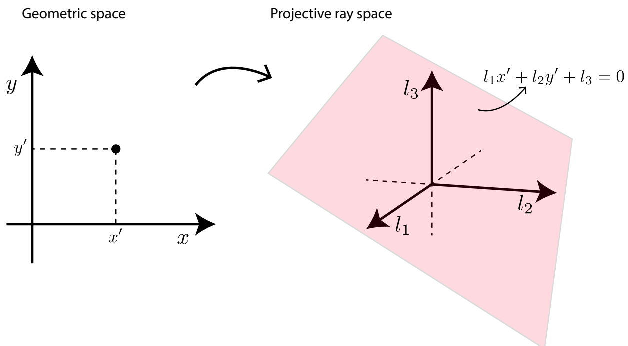


Fig. 2: Projective ray space representation of a point source

$g(\mathbf{r}_m^{(a)}, \mathbf{r'}_n, \omega)$ is the Green's function defined as

$$g(\mathbf{r}_m^{(a)}, \mathbf{r'}_n, \omega) = \frac{e^{-jk\|\mathbf{r}_m^{(a)} - \mathbf{r'}_n\|}}{4\pi\|\mathbf{r}_m^{(a)} - \mathbf{r'}_n\|}, \tag{2}$$

where $k = \omega/c$ with c the speed of sound and $\| \cdot \|$ is the ℓ_2 norm. The term $D_n(\theta_{m,n}^{(a)}, \omega) : \mathbb{R}^2 \to (0, 1)$ in (1) is known as radiation pattern of the nth source and describes the intensity of the sound field emitted toward the direction $\theta_{m,n}^{(a)} = \angle(\mathbf{r}_m^{(a)} - \mathbf{r'}_n)$ (see Fig. 1). It is worth noting that despite we are considering a 3D propagation model, we consider the plane where both sources and microphones lie. We now model the signals of the microphones within the ath array. In vectorial form

$$\mathbf{p}^{(a)} = (\mathbf{D}^{(a)} \otimes \mathbf{G}^{(a)})\mathbf{s} + \mathbf{e}^{(a)}, \quad \forall a = 0, \dots, A-1, \tag{3}$$

where $\otimes$ is the Hadamard product, $\mathbf{p}^{(a)} = [P(\mathbf{r}_0^{(a)}), \dots, P(\mathbf{r}_{M-1}^{(a)})]^T$, $\mathbf{e}^{(a)} = \left[e_0^{(a)}, \dots, e_{M-1}^{(a)}\right]^T$, $\mathbf{s} = [S_0, \dots, S_{N-1}]^T$, $\left[\mathbf{D}^{(a)}\right]_{m,n} = D_n(\theta_{m,n}^{(a)})$ and $\left[\mathbf{G}^{(a)}\right]_{m,n} = g(\mathbf{r}_m^{(a)}, \mathbf{r'}_n)$. The variable ω has been omitted for convenience. In order to reconstruct the signal at a different microphone location, we need the three source parameters that uniquely define (1): location, radiance pattern, and signal.

3. SOUNDFIELD ANALYSIS

3.1. Source localization

Since both the transfer function model (2) and the positions of the microphones are assumed to be known, we need to localize the sources in the acoustic scene. Recently, the Ray Space Transform (RST) [7] has emerged as a convenient tool to describe acoustic information in the ray space domain. Each point in this domain corresponds to an acoustic ray in the geometric space and the main acoustic primitives (sources, arrays and reflectors) are mapped onto lines, whose parameters depend uniquely on their geometric location. Hence, the localization of the sources can be easily performed through pattern analysis techniques. The RST is based on multiple beamforming operations performed on sub-arrays obtained through windowing on an extended linear array.

The Distributed Ray Space Transform introduced in this paper inherits the core idea of the RST but, since we deal with a network of distributed compact arrays, the beamforming operation is performed individually to each array. Moreover, for reasons of generality, here we map the information in the projective ray space [10] in which rays are represented by their projective coordinates $\mathbf{l} = [l_1, l_2, l_3]^T$ that correspond to the parameters of the implicit equation of the line on which the ray lies, i.e. $l_1x + l_2y + l_3 = 0$. In this domain acoustic primitives are mapped onto planes, whose parameters depend uniquely on the geometric location of the acoustic primitives (see Fig. 2). Therefore, for each microphone array a we perform a beamforming operation [11] that results in a pseudospectrum defined as a function of angle $\alpha \in [0, 2\pi]$ and that is mapped in the projective ray space as

$$\begin{aligned} l_1^{(a)} &= \gamma \sin(\alpha^{(a)}); \\ l_2^{(a)} &= \gamma \cos(\alpha^{(a)}); \\ l_3^{(a)} &= \gamma[y^a \cos(\alpha^{(a)}) - x^a \sin(\alpha^{(a)})], \gamma > 0. \end{aligned} \tag{4}$$

where $x^{(a)}, y^{(a)}$ are the array reference point coordinates. In order to localize the sources, similarly to what is done in [10], we search for the $N \times A$ highest peaks and we use RANSAC [12] to perform a robust linear regression to identify the planes relative to the sources. The parameters of the identified planes correspond to an estimate of the source positions $\hat{\mathbf{r}}'_n$ $\forall n = 0, \dots N-1$.

3.2. Radiance pattern estimation

As far as the estimation of the radiance pattern is concerned, we assume the sources to be located at a distance much greater than the size of the arrays. Therefore, all the microphones within each array will "see" the nth source under the same angle $\theta_n^{(a)}$. As a consequence, the elements of the matrix $\mathbf{D}^{(a)}$ are independent on the microphone indices, i.e $\left[\mathbf{D}^{(a)}\right]_{m,n} = D_n(\theta_n^{(a)})$, where $\theta_n^{(a)} = \angle(\mathbf{r}^{(a)} - \mathbf{r'}_n)$ with $\mathbf{r}^{(a)}$ a reference point for the ath array (e.g center of gravity). If we define the matrix $\mathbf{P}^{(a)} = \mathrm{diag}(D_0(\theta_0^{(a)}), \dots, D_{N-1}(\theta_{N-1}^{(a)}))$, the model in (3) reduces to

$$\mathbf{p}^{(a)} = \mathbf{G}^{(a)}(\mathbf{P}^{(a)}\mathbf{s}) + \mathbf{e}^{(a)}, \quad \forall a = 0, \dots, A-1. \tag{5}$$

Since after the localization of the sources we have an estimate $\hat{\mathbf{G}}^{(a)}$ of the matrix $\mathbf{G}^{(a)}$, we can infer each element of the vector $\mathbf{b}^{(a)} = \mathbf{P}^{(a)}\mathbf{s}$ using an LCMV beamformer [13] whose coefficients are the result of the following optimization process

$$\mathbf{h}_n^{(a)} = \underset{\mathbf{h}}{\mathrm{argmin}}\ \mathbf{h}^H\mathbf{h} \qquad \text{s.t. } \mathbf{h}^H\hat{\mathbf{G}}^{(a)} = \mathbf{c}, \tag{6}$$

where $\mathbf{c} \in \mathbb{C}^{1\times N}$ with $[\mathbf{c}]_i = 1$ if $i = n$ and zero otherwise. The solution to (6) is given by [14]

$$\mathbf{h}_n^{(a)} = \hat{\mathbf{G}}^{(a)} \left(\hat{\mathbf{G}}^{(a)^H} \hat{\mathbf{G}}^{(a)}\right)^{-1} \mathbf{c}^H. \tag{7}$$

The estimate $\hat{\mathbf{b}}^{(a)}$ of the vector $\mathbf{b}^{(a)}$ is obtained as

$$\hat{\mathbf{b}}^{(a)} = \mathbf{H}^{(a)^H} \mathbf{p}^{(a)} \quad \forall a = 0, \dots, A-1, \tag{8}$$

where $\mathbf{H}^{(a)} = \left[\mathbf{h}_0^{(a)}, \dots, \mathbf{h}_{N-1}^{(a)}\right]$. If we define the vector $\mathbf{q}_n$ as

$$\begin{aligned} \mathbf{q}_n &= \left[\left| \left[\hat{\mathbf{b}}^{(0)}\right]_n \right|, \dots, \left| \left[\hat{\mathbf{b}}^{(A-1)}\right]_n \right|\right]^T \\ &= |\hat{S}_n| \left[\hat{D}_n(\theta_n^{(0)}), \dots, \hat{D}_n(\theta_n^{(A-1)})\right]^T, \end{aligned} \tag{9}$$

we have an estimate of the directivity of the nth source for the angle $\theta_n^{(a)} \, \forall a = 0, \dots, A-1$ scaled by $|\hat{S}_n|$. In order to reconstruct the sound field we need the directivity for $\theta \in [0, 2\pi]$. Hence, we adopt a model to reconstruct the whole radiance pattern from the estimates at a limited set of angles. In particular, we assume that

$$D_n(\theta_n^{(a)}) = \sum_{l=0}^{L-1} w_{n,l} \cos(l\theta_n^{(a)}) + r_{n,l} \sin(l\theta_n^{(a)}). \tag{10}$$

If we define the matrix $\mathbf{A}_n = [\mathbf{A}_{n,1}, \mathbf{A}_{n,2}]$, with $[\mathbf{A}_{n,1}]_{a,l} = \cos(l\theta_n^{(a)})$, $[\mathbf{A}_{n,2}]_{a,l} = \sin(l\theta_n^{(a)})$ and $\mathbf{y}_n = [w_{n,0}, \dots, w_{n,L-1}, r_{n,0}, \dots, r_{n,L-1}]^T$, we can find an estimate of the expansion coefficients as the result of the optimization problem [15]

$$\hat{\mathbf{y}}_n = \underset{\mathbf{y}_n}{\operatorname{argmin}} \; \|\mathbf{q}_n - \mathbf{A}_n \mathbf{y}_n\|^2 \qquad \text{s.t. } \mathbf{F}\mathbf{y}_n \geq \mathbf{0}, \tag{11}$$

where $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2] \in \mathbb{R}^{I\times 2L}$, $[\,\mathbf{F}_1]_{i,l} = \cos(l\phi_i)$, $[\,\mathbf{F}_2]_{i,l} = \sin(l\phi_i)$ with $\phi_i \in [0, 2\pi]$ $i = 0, \dots, I-1$.

3.3. Signal reconstruction

In order to reconstruct the signals $S_n(\omega)$, we can exploit the estimate of the radiance patterns obtained in Sec. 3.2. In particular, we design an informed spatial filter for each source n as

$$\mathbf{u}_n^\star = \underset{\mathbf{u}}{\operatorname{argmin}} \; \mathbf{u}^H \mathbf{u} \qquad \text{s.t. } \mathbf{u}^H \mathbf{Y} = \mathbf{d}, \tag{12}$$

where $\mathbf{Y} = \left[\left(\hat{\mathbf{G}}^{(0)} \otimes \hat{\mathbf{D}}^{(0)}\right)^T, \dots, \left(\hat{\mathbf{G}}^{(A-1)} \otimes \hat{\mathbf{D}}^{(A-1)}\right)^T\right]^T \in \mathbb{C}^{AM\times N}$ and $\mathbf{d} \in \mathbb{C}^{1\times N}$ with $[\mathbf{d}]_i = 1$ if $i = n$ and zero otherwise. The matrices $\hat{\mathbf{D}}^{(a)}$ contain the directivity values obtained using the expansion coefficients (ref. (11)) in (10). Finally, the estimate $\hat{S}_n(\omega)$ of $S_n(\omega)$ is obtained as

$$\hat{S}_n(\omega) = (\mathbf{u}_n^\star)^H \, \mathbf{p}, \tag{13}$$

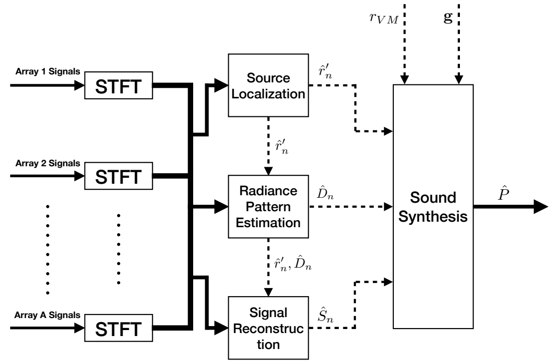


Fig. 3: The sound synthesis procedure block diagram. Each block refers to its corresponding section.

where $\mathbf{p} = \left[\left(\mathbf{p}^{(0)}\right)^T, \dots, \left(\mathbf{p}^{(A-1)}\right)^T\right]^T$. It is worth noting that since the vector $\mathbf{q}_n$ in (9) contains a scaling factor $|\hat{S}_n|$, the matrix $\hat{\mathbf{D}}^{(a)}$ represents a scaled version of the radiance pattern. Hence, also the result of (13) is a scaled version of the signal. However, this is not a problem in our framework, as described in the next section.

4. SOUND SYNTHESIS

Once the parameters are estimated as described in the previous sections, the sound field at the virtual microphone location can be computed through the parametric model in (1) as

$$\hat{P}(\mathbf{r}_{VM}, \omega) = \sum_{n=0}^{N-1} \hat{D}_n(\theta_{VM,n}, \omega) g(\mathbf{r}_{VM}, \hat{\mathbf{r}'}_n, \omega) \hat{S}_n(\omega) \tag{14}$$

where $\mathbf{r}_{VM}$ is the VM position, $\theta_{VM,n} \angle(\mathbf{r}_{VM} - \hat{\mathbf{r}'}_n)$ and $\hat{\mathbf{r}'}_n$ is the estimated source position. Moreover, since $g(\mathbf{r}_{VM}, \hat{\mathbf{r}'}_n, \omega)$ is proportional to $1/\|\mathbf{r}_{VM} - \hat{\mathbf{r}'}_n\|$, in practice it is necessary to limit its value to $g_{\max}$ to avoid amplifying the signal too much when the VM is very close to the nth source. The block diagram of Fig. 3 summarizes the whole VM procedure.

5. SIMULATIONS AND EXPERIMENTS RESULTS

In order to validate the proposed procedure, we test our methodology in two contexts: software simulations and real applicative scenarios. For both simulations and experiments different source signals are adopted. In particular, the white Gaussian noise is useful to observe the behavior of the technique at different frequencies, while speech signals are used to simulate a real use case. In order to be consistent with the adopted model, a Short Time Fourier Transform (STFT) of the microphone signals is performed in both simulations and experiments using a 20ms Hann window with 50% overlap.

Since the technique is divided into different steps, we devise proper metrics in order to evaluate them separately. Specifically, the metrics are

Localization Error:

$$\mathrm{LE}(\mathbf{r}') = \frac{1}{N}\sum_{n=0}^{N-1}\|\hat{\mathbf{r}}'_n - \mathbf{r}'_n\| \tag{15}$$

Directivity Error:

$$\mathrm{DE}_n = \frac{1}{F}\sum_{f=0}^{F-1}\frac{1}{I}\sum_{i=0}^{I-1}(\hat{D}_n(\theta_i,\omega_f) - D_n(\theta_i,\omega_f))^2 \tag{16}$$

where f is the frequency index and i is the angle index.

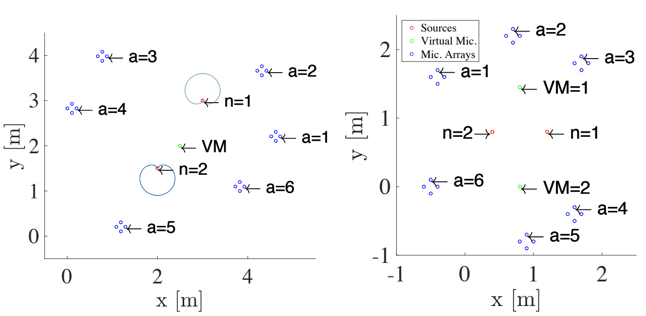
Synthesized Signal Error:

$$\mathrm{SSE}_{VM} = \sqrt{\frac{1}{T}\sum_{t=0}^{T}(x_{VM}(t) - x_{ref}(t))^2} \tag{17}$$

where t is the discrete time index, x_{VM} the VM signal in time domain and x_{ref} stands for the reference signal.

The simulation setup is shown in Fig. 4a. As we can see, six circular microphone arrays (radius= 0.1m) composed by four microphones are employed. The source signals are weighted by a cardioid radiance pattern and propagated to the sensors. I.i.d. white noise for a $\mathrm{SNR} = 40\,\mathrm{dB}$ w.r.t the first microphone in the first array has been added to the microphone signals. The most significant simulation results are reported in Tab. 1a. As far as the localization performances are concerned, we can see that the error is in the order of 10^{-2}m. Moreover, since the synthesized signal directly depends on the estimation of the parameters, we can see that the values of the SSE vary accordingly with the other two metrics. In addition, we can notice that when two sources are active simultaneously the performance slightly decrease.

For what concerns the experiments, we tested the proposed technique in a semi-anechoic chamber characterized by a reverberation time $\mathrm{T}_{60} = 0.05$ s. The deployment of the equipment is shown in Fig. 4b and it is consistent with the one adopted for the simulations. In this case, in order to obtain a reference signal, useful to evaluate our system, two physical microphones are placed as shown in Fig. 4b. To simulate the source signal, two custom loudspeakers are adopted. Since we do not have a reference radiance pattern for these loudspeakers, the DE metric is not evaluated. The most significant



(a) (b)

Fig. 4: Test setup. Sources (red) are surrounded by microphone arrays (blue), while the virtual microphones that coincide with the reference ones are highlighted in green.

Simulations					
Sources	Orientation	Source Signal	LE [m]	DE	SSE
1	$0°$	White Noise	1.12×10^{-2}	6.61×10^{-3}	6.74×10^{-4}
2	$0°$		7.40×10^{-3}	2.66×10^{-2}	5.41×10^{-4}
1, 2	$0°, 0°$		3.50×10^{-2}	$8.56 \times 10^{-2}, 5.24 \times 10^{-2}$	2.42×10^{-2}
1	$-90°$	Speech	1.10×10^{-2}	3.57×10^{-3}	8.18×10^{-4}
2	$90°$		8.00×10^{-3}	1.12×10^{-2}	7.72×10^{-4}
1, 2	$-90°, 0°$		4.39×10^{-2}	$8.97 \times 10^{-2}, 9.94 \times 10^{-2}$	5.75×10^{-3}

(a)

Experiments				
Sources	Orientation	Source Signal	LE [m]	SSE
1	$90°$	White Noise	3.54×10^{-2}	$1.00 \times 10^{-2}, 2.84 \times 10^{-3}$
2	$-90°$		5.33×10^{-2}	$3.44 \times 10^{-3}, 9.30 \times 10^{-3}$
1, 2	$90°, -90°$		6.40×10^{-2}	$1.32 \times 10^{-2}, 1.29 \times 10^{-2}$
1	$90°$	Speech	6.03×10^{-2}	$5.83 \times 10^{-3}, 4.26 \times 10^{-3}$
2	$90°$		3.18×10^{-2}	$9.72 \times 10^{-3}, 6.98 \times 10^{-3}$
1, 2	$90°, 90°$		5.53×10^{-2}	$1.34 \times 10^{-2}, 1.00 \times 10^{-2}$

(b)

Table 1: Tests are referred to the setup in Fig. 4a and Fig. 4b, respectively. We define the orientation of source as the angle of the maximum value of the radiation pattern measured counter-clockwise from the positive x axis.

results from experiments are reported in Tab. 1b. Looking at the LE, we can see that the localization is less accurate, but remains in the order of 10^{-2}m. The considerations made for the simulations hold also for the experiments. Finally, we report in Fig. 5 an example referred to the fourth row in Tab. 1b of a spectrogram of the VM signal w.r.t to the two reference microphones.

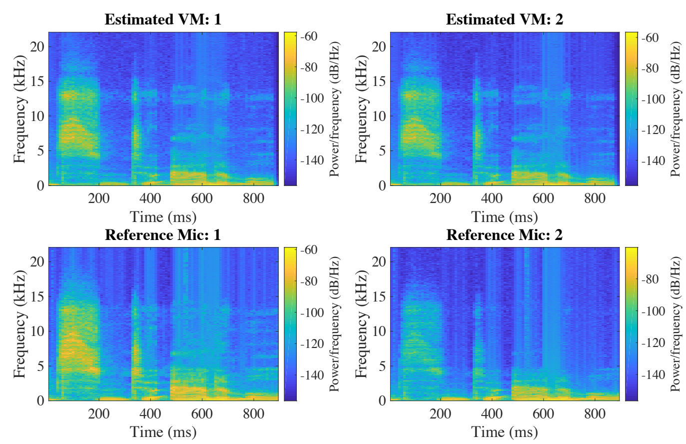


Fig. 5: *VM spectrograms compared to their references from experiment 4.*

6. CONCLUSIONS

In this paper we presented a parametric sound field reconstruction technique used for the estimation of the signal of a virtual microphone arbitrarily placed in the acoustic scene. An analysis framework has been developed in order to estimate the model parameters starting from the signals acquired from a network of compact microphone arrays. The signal model has been developed in a free-field scenario, and can be extended in its current formulation to moderately reflective environments. The procedure has been tested through simulations and experiments to show the effectiveness of the proposed approach. It is in our plans, however, to extend the proposed method also to the case of reverberant environments, possibly with the help of a priori information related to the geometry of the environment [16] [17].

7. REFERENCES

[1] Juha Vilkamo, Tapio Lokki, and Ville Pulkki, "Directional audio coding: Virtual microphone-based synthesis and subjective evaluation," *J. Audio Eng. Soc*, vol. 57, no. 9, pp. 709–724, 2009.

[2] Richard Schultz-Amling, Fabian Kuech, Oliver Thiergart, and Markus Kallinger, "Acoustical zooming based on a parametric sound field representation," in *Audio Engineering Society Convention 128*. Audio Engineering Society, 2010.

[3] Miriam A Doron and Eyal Doron, "Wavefield modeling and array processing. i. spatial sampling," *IEEE Transactions on signal Processing*, vol. 42, no. 10, pp. 2549–2559, 1994.

[4] Oliver Thiergart, Giovanni Del Galdo, Maja Taseska, and Emanuël AP Habets, "Geometry-based spatial sound acquisition using distributed microphone arrays," *IEEE transactions on audio, speech, and language processing*, vol. 21, no. 12, pp. 2583–2594, 2013.

[5] Thibaut Ajdler, Luciano Sbaiz, and Martin Vetterli, "The plenacoustic function and its sampling," *IEEE transactions on Signal Processing*, vol. 54, no. 10, pp. 3790–3804, 2006.

[6] Dejan Markovic, Fabio Antonacci, Augusto Sarti, and Stefano Tubaro, "Soundfield imaging in the ray space," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 12, pp. 2493–2505, 2013.

[7] Lucio Bianchi, Fabio Antonacci, Augusto Sarti, and Stefano Tubaro, "The ray space transform: A new framework for wave field processing," *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5696–5706, 2016.

[8] E. G. Williams, *Fourier Acoustics*, Academic Press, London, UK, 1999.

[9] Antonio Canclini, Luca Mucci, Fabio Antonacci, Augusto Sarti, and Stefano Tubaro, "Estimation of the radiation pattern of a violin during the performance using plenacoustic methods," in *Audio Engineering Society Convention 138*. Audio Engineering Society, 2015.

[10] Dejan Marković, Fabio Antonacci, Augusto Sarti, and Stefano Tubaro, "Multiview soundfield imaging in the projective ray space," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 23, no. 6, pp. 1054–1067, 2015.

[11] Jack Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.

[12] Martin A Fischler and Robert C Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[13] Harry L Van Trees, *Optimum array processing: Part IV of detection, estimation and modulation theory*, vol. 1, Wiley Online Library, 2002.

[14] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Prentice Hall, Upper Saddle River, NJ, USA, 2004.

[15] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, Cambridge, MA, USA, seventh edition, 2009.

[16] D Aprea, Fabio Antonacci, Augusto Sarti, and Stefano Tubaro, "Acoustic reconstruction of the geometry of an environment through acquisition of a controlled emission," pp. 710–714, 2009.

[17] Jason Filos, Antonio Canclini, Fabio Antonacci, Augusto Sarti, and Patrick A Naylor, "Localization of planar acoustic reflectors from the combination of linear estimates," in *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European*. IEEE, 2012, pp. 1019–1023.