GEOMETRIC RECONSTRUCTION OF THE ENVIRONMENT FROM ITS RESPONSE TO MULTIPLE ACOUSTIC EMISSIONS

F. Antonacci, A. Sarti, S. Tubaro

Dipartimento di Elettronica ed Informazione, Politecnico di Milano,
p.zza Leonardo da Vinci, 32, 20133 Milano, Italy

ABSTRACT

In this paper we propose a method for reconstructing the 2D geometry of the surrounding environment based on the signals acquired by a fixed microphone, when a series of acoustic stimula are produced in different positions in space. After estimating the Times Of Arrival (TOAs) of the reflective paths, we turn each TOA into a projective geometric constraint that can be used for determining the locations of the reflectors. The result consists of a collection of planar surfaces that correspond to the reflectors' locations. In this paper we present the whole processing chain and prove its effectiveness through experimental results.

Index Terms— Geometrical acoustics, acoustic environment, projective geometry, active sensing.

1. INTRODUCTION

A wide variety of space-time processing algorithms can greatly benefit from the knowledge of the environment's geometry. For example, in [1] we showed that the interferer cancelation can be improved when we have some hints on the Directions Of Arrival of the interferer signal as well as its dominant reflections.

The literature is rich with solutions that enable the reconstruction of the geometry of the environment based on various methodologies ranging from computer vision ([2]) to radar tomography ([3]). Acoustic acquisitions conducted in a reverberant environment, in fact, carry information not only on visible but also on non visible surfaces. If advantageously used, this information may enrich our knowledge on the environment surrounding the acquisition device. Spherical arrays ([4, 5]) have stirred a certain interest in the research community thanks to their remarkable resolution and compactness. In [6], [7] and [8] spherical arrays are used for inferring the temporal sequence of reflections in the environment as well as the three-dimensional directions of arrival. In order to achieve the desired resolution, the spherical array needs to have numerous microphones in a compact case.

In this paper we assume that reflectors are piece-wise linear and two-dimensional. Our acquisition device is composed by a single microphone standing in a known position and a loudspeaker rotating around the microphone. The emitter generates a noise-like sequence which propagates in the environment. The microphone receives, along with the direct signal, delayed and dimmed replicas of the signal associated to wall reflections. Wavefront propagation is assumed to be subject to the laws of optical geometry, which set specific constraints on the reflective path. In this paper we devise a methodology that exploits such constraints in order to efficiently and robustly infer the environment's geometry.

We begin with measuring the impulse response of the environment from the loudspeaker's positions to the microphone, using the cross-correlation of the emitted and the acquired signals. The Times of Arrival (TOAs) of the reflected signals are then extracted through a simple peak-picking algorithm. The time of flight of the path from the loudspeaker's positions to the reflector is the sum of two components: one from the loudspeaker to the reflector; and one from the reflector to the microphone. The locus of candidate reflection points is, therefore, an ellipse. If we consider a different location of the loudspeaker, the reflection point changes and, in turn, the Time Of Arrival changes. However, the reflection points are bound to be aligned with the reflector. The reflector, in fact, is the line of tangency to all the ellipses that are traced, each one for a different position of the loudspeaker. We will show how a projective geometric representation can be seamlessly used for estimating such lines of tangency.

The manuscript is organized as follows: Section 2 describes the data model for the localization of a single reflector; Section 3 describes how to localize a single reflector; Section 4 extends the result to the case of multiple reflectors; Section 5 presents some experimental results; finally Section 6 draws some conclusions on the method.

2. DATA MODEL AND PROBLEM FORMULATION

In this Section we introduce the data model used throughout the paper and we formulate the problem of reflector localization. Let us consider a loudspeaker that moves on a pre-determined discrete trajectory denoted by the points $\mathbf{x}_k = [x_{1k}, x_{2k}]^T,\ k = 1, \ldots, K$. On each point of the trajectory a known signal $s_l(t)$ of duration $T[s]$ is produced. We use the same probing signal $s_l(t)$ for all locations.

A synchronized microphone receives the signal $s_k(t),\ k = 1, \ldots, K$. With no loss of generality we assume that the origin of the reference frame is centered in the microphone position M. Together with the direct signal, the microphone receives replicas reflected by walls and obstacles. Our goal is to infer the geometry of the environment from the analysis of $s_k(t)$. We will first focus on the problem of localizing a single reflector, which means that only a reflection is present in $s_k(t)$. We will then consider the case of multiple reflectors in Sec. 4.

Given $s_l(t)$ and $s_k(t)$, we can estimate the channel response from $\mathbf{x}_k$ to M as the cross-correlation of $s_l(t)$ and $s_k(t)$. Our estimate is:

$$h_k(t) = \alpha_k^{(d)}\delta(t - \tau_k^{(d)}) + \\ +\alpha_k^{(r)}\delta(t - \tau_k^{(r)}) + \nu(t)\ , k = 1, \ldots, K\ , \tag{1}$$

where $\tau_k^{(d)}$ and $\tau_k^{(r)}$ are the times of flight of the direct and reflected paths, respectively, when the loudspeaker is in $\mathbf{x}_k$; $\alpha_k^{(d)}$ and $\alpha_k^{(r)}$ are the corresponding attenuations. The term $\nu(t)$, finally, models

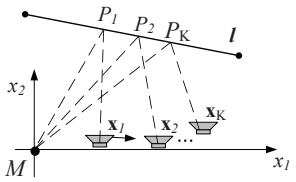

Fig. 1. Geometry of the problem: a reflector lies on the line $\boldsymbol{l}$, the microphone is fixed in M, the loudspeaker is moved on different locations $\mathbf{x}_k$. Correspondingly, the reflection point on $\boldsymbol{l}$ changes, therefore the Time of Arrival $\tau_k^{(r)}$ changes as well.

an additive noise. We assume that the laws of optical geometry are valid. Figure 1 illustrates the geometry of the problem: the obstacle lies on the line $\mathbf{l} = [l_1, l_2, l_3]^T$, the microphone is still in M; as we move the loudspeaker, the reflection points P_k , $k = 1, \ldots, K$ that satisfy the Snell's law moves as well and, consequently, the Time of Arrival $\tau_r^{(k)}$ changes. The Time of Arrival (TOA) $\tau_r^{(k)}$ is the sum of the time of flight from $\mathbf{x}_k$ to P_k, and from P_k to M. The locus of possible points P_k that share the same TOA $\tau_r^{(k)}$ is an ellipse of foci M and $\mathbf{x}_k$. Let us consider the tangent to that ellipse in P. Basic geometry tells us that its perpendicular is also the bisector of the angle $\widehat{M\ P\ \mathbf{x}_k}$. This means that every tangent to the ellipse is a putative reflector line $\mathbf{l}$, as it honors Snell's law. Each one of the positions $\mathbf{x}_k$ $k = 1, \ldots, K$ of the source corresponds to an ellipse of possible points of reflection, and just as many reflectors (tangents to such point). We notice, however, that the line $\mathbf{l}$ is tangent to all the ellipses. This line is therefore the intersection of all the bundles of the tangent lines (one bundle per ellipse) and represents the actual reflector.

In order to determine the intersection of all the bundles of tangent lines, we adopt a notation borrowed from projective geometry [9]. A point $\mathbf{x} = [x_1, x_2, 1]^T$ (in homogeneous - scalable - coordinates) belongs to a conic of parameters $[a_k, b_k, c_k, d_k, e_k, f_k]^T$ iff (if and only if):

$$a_k^2 x_1^2 + b_k^2 x_2 x_1 + c_k x_2^2 + d_k x_1 + e_k x_2 + f_k = 0 \ . \qquad (2)$$

Eq. (2) can be rewritten in matrix form as

$$\mathbf{x}^T \mathbf{C}_k \mathbf{x} = 0 \ , \qquad (3)$$

where

$$\mathbf{C}_k = \begin{bmatrix} a_k & b_k/2 & d_k/2 \\ b_k/2 & c_k & e_k/2 \\ d_k/2 & e_k/2 & f_k \end{bmatrix} \qquad (4)$$

is the matrix of the "point conic". The parameter vector $[a_k, b_k, c_k, d_k, e_k, f_k]^T$ is scalable, therefore the knowledge of five points on the ellipse is sufficient to determine it (see [9] for reference on this point). As far as the determination of points on the ellipse, the knowledge of the Time of Arrival $\tau_k^{(r)}$ and of the foci M and $\mathbf{x}_k$ allow us to determine an arbitrarily large number of points on the ellipse. More useful for our purposes is the definition of the "line conic" associated to the above-defined point conic. The homogeneous representation of a line is given by the scalable vector $\mathbf{l} = [l_1\ l_2\ l_3]^T$. The line $\mathbf{l}$ passes through the point $\mathbf{x}$ iff (if and only if) $\mathbf{l}^T \mathbf{x} = 0$. The line $\mathbf{l}$ is therefore tangent to the point ellipse $\mathbf{C}_k$ if

$$\mathbf{l}^T \mathbf{C^*}_k \mathbf{l} = 0 \ , \qquad (5)$$

where $\mathbf{C^*}_k = \det(\mathbf{C}_k)\mathbf{C}_k^{-1}$ is the adjoint of the conic-matrix $\mathbf{C}_k$ [9]. As we are dealing with homogeneous coordinates, we can drop the scale factor $\det(\mathbf{C}_k)$ and use the inverse of $\mathbf{C}_k$ as the adjoint. The matrix $\mathbf{C^*}_k$ is known as the "line conic" of $\mathbf{C}_k$, i.e. the bundle of all lines that are tangent to $\mathbf{C}_k$. This bundle corresponds to a manifold in the parameter space $[l_1\ l_2\ l_3]^T$, as shown in Fig. 2. In the next Section we devise an algorithm that nails the position

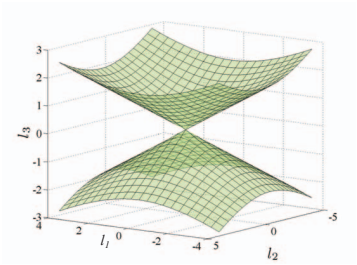

Fig. 2. The surface represents the set of lines tangent to an ellipse in the line parameter space.

of the reflector combining multiple constraints derived for different positions of the loudspeaker.

3. LOCALIZATION OF A SINGLE REFLECTOR

As shown in the last paragraph, if we acquire the channel response for different positions of the loudspeaker, the actual reflector line $\mathbf{l}$ must satisfy eq.(5) for $k = 1, \ldots, K$. Figure 3 depicts an example:

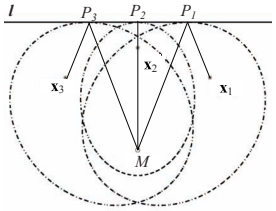

Fig. 3. The reflector line is the common tangent to the ellipses traced for $\mathbf{x}_1$, $\mathbf{x}_2$ and $\mathbf{x}_3$.

the actual reflector line is the common tangent to the ellipses traced for $\mathbf{x}_k, k = 1, 2, 3$.

If the points $\mathbf{x}_k$, $k = 1, \ldots, K$ are known in advance and the TOA $\tau_k^{(r)}$ is estimated by inspection of $h_k(t)$, we are able to trace the ellipses $\mathbf{C}_k, k = 1, \ldots, K$. The actual reflector line is the solution of the equations:

$$\begin{cases} \mathbf{l}^T \mathbf{C^*}_1 \mathbf{l} & = 0 \\ \mathbf{l}^T \mathbf{C^*}_2 \mathbf{l} & = 0 \\ \quad \ldots & \\ \mathbf{l}^T \mathbf{C^*}_K \mathbf{l} & = 0 \end{cases} \qquad (6)$$

Since we have three unknowns (the parameters l_1, l_2, l_3) we need at least three equations in the system (6).

The solution of a nonlinear system as in (6) is not trivial as our measures of $\tau_k^{(r)}$ are affected by estimation error. We call the algorithm as COmmon Tangent Algorithm (COTA). The cost function to be minimized is:

$$\hat{\mathbf{l}} = \arg\min_{\mathbf{l}} \sum_{k=1}^{K} ||\mathbf{l}^T \hat{\mathbf{C}}_k^* \mathbf{l}||^2 \ . \qquad (7)$$

The matrices $\hat{\mathbf{C}}_k^*$ are affected by estimation error with respect to $\mathbf{C}_k^*$. Moreover, we observe that $\mathbf{C}_k^*$ are neither negative nor positive definite. Therefore the objective function in (7) is not granted to have a global minimum: in fact we notice that if the line $\mathbf{l}$ is a solution of (6), then also $k\mathbf{l}$ is. For this reason, in the next paragraphs we will consider a subspace of $\mathbb{R}^3$ in which the minimization problem in (7) is granted to have a unique solution. More specifically, we impose that l_1 and l_2 lie on a circle of unitary radius:

$$\begin{aligned} l_1 &= \cos(\alpha), \\ l_2 &= \sin(\alpha). \end{aligned} \tag{8}$$

If we perform our minimization on $\mathbf{l}_\alpha = [\cos(\alpha), \sin(\alpha), l_3]^T$, ($\rho = 1$), the COTA now becomes:

$$\hat{\mathbf{l}}_\alpha = \arg\min_{\mathbf{l}_\alpha} \sum_{k=1}^{K} ||\mathbf{l}_\alpha^T \hat{\mathbf{C}}_k^* \mathbf{l}_\alpha||^2 \ . \tag{9}$$

Fig.4 depicts the geometry of the problem in the line parameter space: the solution of eq.(7) in $\mathbb{R}^3$ corresponds to finding the common solution of the manifolds of tangent lines for $\mathbf{x}_k \ k = 1, 2, 3$. When we impose eq.(9) we are finding the intersection of the manifolds along the cylinder of unitary radius. $[\alpha, l_3]$. The problem of

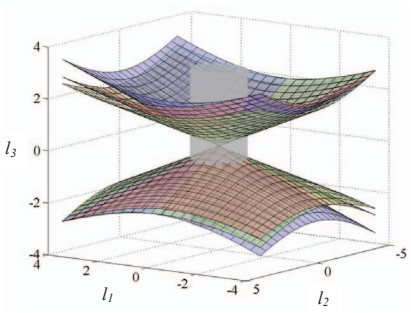

Fig. 4. Eq.(9) corresponds to finding the solution of the equation system in (6) along a cylinder of unitary radius whose axis coincides with the l_3 axis.

estimating the parameters $[l_1, l_2, l_3]^T$ has now turned into the problem of estimating the parameters $[\alpha, l_3]^T$ that enable the intersection of the manifolds of tangent lines on the cylinder.

Since we are considering a system of second order equations, we expect two solutions for the system in (9). Figure 5 presents the cylindrical section of the manifolds of tangent lines corresponding to the situation in Figure 4. We observe, as expected, two solutions for the problem: $[\alpha = 45°, l_3 = 0.5]$ and $[\alpha = 225°, l_3 = -0.5]$.

4. LOCALIZATION OF MULTIPLE REFLECTORS

In this Section we will extend the approach described above to the case of multiple reflectors. In this situation the k-th impulse response becomes:

$$\begin{aligned} h_k(t) \quad &= \alpha_k^{(d)} \delta(t - \tau_k^{(d)}) + \textstyle\sum_{g=1}^{G} \alpha_k^{(g)} \delta(t - \tau_k^{(g)}) + \\ &+ \textstyle\sum_{h=1}^{H} \alpha_k^{(h)} \delta(t - \tau_k^{(h)}) + \nu(t) \ . \end{aligned} \tag{10}$$

The meaning of the terms in eq.(10) is analogous to eq.(1), but now G reflective paths appear. The first summation in eq.(10) concerns

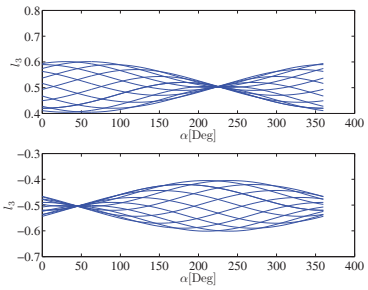

Fig. 5. Cylindrical section of the manifolds of tangent lines traced for $\sqrt{l_1^2 + l_2^2} = 1$.

the first-order reflections that link $\mathbf{x}_k$ and M, while the last summation in eq.(10) models reflections involving multiple bouncings on obstacles.

For each point of the trajectory, the peak picking algorithm returns multiple local maxima in the impulse response. The variable τ_{kj} represents a maximum found in $h_k(t)$. These maxima, however, are not organized in a labeled dataset, as each element τ_{kj} is not assigned to a reflector. In order to use the COTA for multiple reflectors, we have to find TOAs in τ_{kj} that correspond to the same reflector. At each step of the trajectory, the loudspeaker is located at a distance r and at an angle Θ_k in a reference frame centered at the microphone. After some passages (omitted here for reasons of space) the TOAs relative to a reflector positioned at distance ρ and seen from the microphone under the angle θ are aligned on the function $\tau_k(\theta, \rho) = \rho/c - r/c \sin(\Theta_k - \theta)$. The geometric notation is described for convenience in Fig.6. In order to find the correspon-

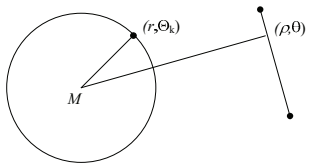

Fig. 6. Geometry used for the TDOA assignment

dences, we look in the space (θ, ρ) for the TOAs in the set τ_{kj} that belong to the same curve $\tau_k(\theta, \rho), k = 1, \ldots, K$. In order to do so we use the generalized Hough transform [10]. We do not describe the assignment algorithm for reasons of space. Figure 7 shows an example of labeling. Fig.7(a) shows the unlabeled dataset, while (b) shows the labeled one. We can observe that some items in the set τ_{kj} are not assigned as they are not aligned on any curve $\tau_k(\theta, \rho)$. The proposed labeling method, therefore, is robust to the presence of outliers in τ_{kj} due to environmental noise.

5. EXPERIMENTAL RESULTS

In this Section we will present experimental results for the reconstruction of the geometry of two environments. Experiments have been conducted in an acoustically dry room in which wood panels have been placed to alter the impulse response. The first experiment concerns the problem of localizing two reflectors that are not mutually visible. In a later stage we will consider the case of mutually visible reflectors. The points $\mathbf{x}_k, k = 1, \ldots, 24$ are equally spaced on a circle whose radius is $R = 0.16m$. The sampling frequency

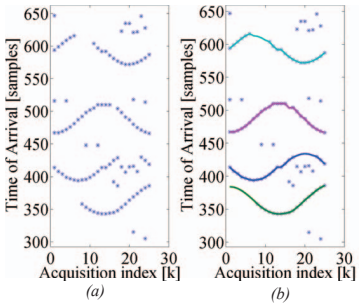

Fig. 7. An example of labeling of the dataset τ_{kj}. (a) is the unlabeled dataset, while (b) is the result of the labeling process.

is $Fs = 44,1kHz$ and the excitation signal is a white noise in the bandwidth $[0, 5kHz]$. The accuracy of localization of reflectors (real and virtual ones) is measured with two parameters, partially interdependent:

- if $\hat{\mathbf{l}}$ and $\mathbf{l}$ are, respectively, the actual and estimated reflector lines, we measure $|\mathbf{l} \cdot \hat{\mathbf{l}}|/(||\hat{\mathbf{l}}||||\mathbf{l}||)$: a value close to 1 reveals that the angle between lines is small.
- each line is observed from the center of the reference frame (the microphone) at a distance $\hat{\rho}$ and under an angle of view $\hat{\theta}$. We compare these estimates with the measures (ρ,$\hat{\theta}$) done with rulers.

The geometry of the system for the first environment is depicted in Figure 8. Numbers next to the walls refer to indexes in Table 1, where localization results for both reflectors are shown according to the criteria defined above. In the second experiment we have tested

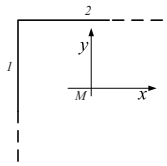

Fig. 8. Geometry of the first test environment

Table 1. Experimental results for the environment in Fig.8): $(\theta[^\circ], \rho[m])$ and $(\hat{\theta}[^\circ], \hat{\rho}[m])$ are the hand-measured and estimated positions, while $|\mathbf{l} \cdot \hat{\mathbf{l}}|/(||\hat{\mathbf{l}}||||\mathbf{l}||)$ measures the alignment of the estimated and theoretical lines.

Index	$(\theta[^\circ], \rho[m])$	$(\hat{\theta}[^\circ], \hat{\rho}[m])$	$\lvert\mathbf{l} \cdot \hat{\mathbf{l}}\rvert/(\lvert\lvert\hat{\mathbf{l}}\rvert\rvert\lvert\lvert\mathbf{l}\rvert\rvert)$
1	(180,1.805)	(178,1.81)	0.99
2	(90, 2.24)	(89, 2.24)	1

the system in the environment depicted in Figure 9. Dashed lines refer to "virtual reflectors", originated from second-order reflections. Table 2 shows the corresponding localization results using the metrics defined at the beginning of this Section. We notice that even "virtual" reflectors are localized with a good accuracy.

6. CONCLUSIONS

In this paper we have shown an efficient and effective methodology for the reconstruction of the geometry of an environment

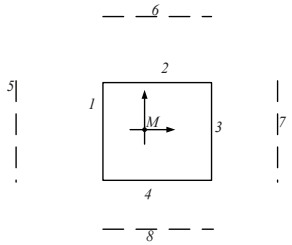

Fig. 9. Geometry of the second test environment

Table 2. Experimental results for the environment in Fig.9): $(\theta[^\circ], \rho[m])$ and $(\hat{\theta}[^\circ], \hat{\rho}[m])$ are the hand-measured and estimated positions, while $|\mathbf{l} \cdot \hat{\mathbf{l}}|/(||\hat{\mathbf{l}}||||\mathbf{l}||)$ measures the alignment of the estimated and theoretical lines.

Index	$(\theta[^\circ], \rho[m])$	$(\hat{\theta}[^\circ], \hat{\rho}[m])$	$\lvert\mathbf{l} \cdot \hat{\mathbf{l}}\rvert/(\lvert\lvert\hat{\mathbf{l}}\rvert\rvert\lvert\lvert\mathbf{l}\rvert\rvert)$
1	(270,1.18)	(270,1.17)	1
2	(180,1.37)	(179,1.38)	0.99
3	(90,1.68)	(89,1.68)	0.99
4	(360,2.08)	(358,2.08)	0.99
5	(270,2.82)	(271,2.84)	0.99
6	(90,2.82)	(85,2.85)	0.97
7	(180,3.44)	(181,3.45)	0.99
8	(360,3.44)	(357,3.44)	0.98

through acoustic acquisitions by using projective geometry constraints. Some experimental results show the feasibility of the approach. In the next months, we will work on the problem of discerning real and virtual reflectors.

7. REFERENCES

[1] F.Antonaccci, D.Riva, G.Prandi, M.Tagliasacchi, A.Sarti, and S.Tubaro, "Efficient interferer cancellation based on geometrical information of the reverberant environment," in *Proc. of European Signal Processing Conference 2008, EUSIPCO 2008*, Sept. 2008.

[2] M. Han and T. Hanade, "Multiple motion scene reconstruction with uncalibrated cameras," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, July 2003.

[3] P. La Riviere, J. Bian, and P. Vargas, "Penalized-likelihood sinogram restoration for computed tomography," *Medical Imaging, IEEE Transactions on*, vol. 25, no. 8, pp. 1022–1036, Aug. 2006.

[4] J.Meyer and G.Elko, "Spherical harmonic modal beamforming for an augmented circular microphone array," in *Proceedings of IEEE ICASSP 2008*, 2008.

[5] Z. Li and R. Duraiswami, "Flexible and optimal design of spherical microphone arrays for beamforming," *IEEE Transactions on Audio, Speech, and Language Processing*, Feb. 2007.

[6] A.O'Donovan, R. Duraiswami, and D. Zotkin, "Imaging concert hall acoustics using visual and audio cameras," in *Proceedings of IEEE ICASSP 2008*, 2008.

[7] A.O'Donovan, R.Duraiswami, and J. Neumann, "Microphone arrays as generalized cameras for integrated audio visual processing," in *Proc. IEEE CVPR*, vol. 1, 2007, pp. 1–8.

[8] S. Negahdaripour, H. Sekkati, and H. Pirsiavash, "Opti-acoustic stereo imaging, system calibration and 3-d reconstruction," in *Proceedings of IEEE Conference Computer Vision and Pattern Recognition CVPR '07.*, June 2007, pp. 1–8.

[9] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. New York, NY, USA: Cambridge University Press, 2001.

[10] L.Shapiro and G.Stockman, *Computer Vision*. Prentice Hall, 2001.