ROBUST INFERENCE OF ROOM GEOMETRY FROM ACOUSTIC MEASUREMENTS USING THE HOUGH TRANSFORM

Jason Filos[1], *Antonio Canclini*[2], *Mark R. P. Thomas*[1], *Fabio Antonacci*[2], *Augusto Sarti*[2] *and Patrick A. Naylor*[1]

[1] Electrical and Electronic Engineering,
Imperial College London, UK
email: {jf203, mrt102, p.naylor}@imperial.ac.uk

[2] Dipartimento di Elettronica ed Informazione,
Politecnico di Milano, Milano, Italy
email: {canclini, antonacc, sarti}@elet.polimi.it

ABSTRACT

The problem of localizing reflective boundaries and obstacles in an acoustic environment from acoustic measurements is considered. Specifically, localization of multiple two-dimensional (2-D) line reflectors is achieved by estimation of the time of arrival (TOA) of reflected signals by analysis of acoustic impulse responses (AIRs). The estimated TOAs are used in conjunction with the source and receiver locations to find the loci of solutions whose common tangents correspond to the location of a reflector. The solution to the common tangent estimation is a nonlinear and non-convex problem that can yield local sub-optimal solutions using existing approaches. We therefore propose an analytic method, based on a closed-form estimator, that is guaranteed to converge to the global minimum in an error-free scenario. We further improve the robustness of the approach when errors are introduced in the estimated TOAs by using the Hough transform to find the optimal solution. The proposed approach is evaluated through Monte Carlo runs, using simulated rooms, that demonstrate the feasibility of the proposed approach.

1. INTRODUCTION

Acoustic scene reconstruction is a process that aims to establish an understanding of the acoustic environment in which space-time processing algorithms operate. Specifically the location of reflecting boundaries and obstacles can be estimated from acoustic measurements, allowing the geometry of the room to be reconstructed. In recent years, interesting solutions have appeared that make use of microphone arrays as acoustic cameras [1] aimed at reconstructing location, principal dimensions and shape of obstacles in the environment [2] and for scene reconstruction [3].

Recent interest has been shown in the localization of line reflectors in the acoustic space by analysis of the times of arrival (TOAs) of first-order reflections. Under the hypothesis of optical acoustics, it is known that the total time of flight for a reflected path consists of two components: the sum of the path from source to reflector and reflector to microphone. It was shown in [4] that for a single source-microphone pair, the estimate of the TOA of the reflected signals results in a locus of solutions that lies on an ellipse, and that by considering the multiple-input-single-output (MISO) case there exists a common tangent that corresponds to the reflective boundary. The problem was extended in [5] in the single-input-multiple-output (SIMO) case.

However, the existing approaches concerned with estimating the line reflector rely on solving a nonlinear and non-convex optimization problem. It is known that such approaches can converge to non-optimal solutions in certain practical environments. This paper aims to address some of these problems by first deriving a closed-form solution for the single-reflector case, which yields a more robust solution when compared to existing approaches. The algorithm provides all local minimum solutions ranked in order of cost, and in an error-free scenario is guaranteed to find the global optimum solution. In other words, in the case where TOAs are known exactly, the global minimum of the closed-form estimator also corresponds to the true reflector. However, in the cases where the TOAs contain errors, the true reflector may not correspond to the global minimum but to one of the set of local minima. For this reason we propose a second-stage correction that estimates a collection of meaningful coordinate points, that are geometrically related to the estimated line reflector from the closed-form solution and the set of ellipses related to the TOAs. The points that exhibit greatest collinearity, and therefore most likely to correspond to the true reflector, are found by application of the Hough Transform. In addition to providing greater robustness to error in the SIMO case, this approach can be used in future multiple-input-multiple-output (MIMO) cases, with minor additional computational overhead. This permits measurements from multiple source positions to be considered sequentially in order to achieve better localization results in adverse conditions where errors are introduced in the TOAs, particularly when considering more complex room geometries. The proposed algorithm is evaluated by Monte Carlo simulations of acoustic environments, aimed at reconstructing the geometry of the acoustic enclosure.

The remainder of this paper is organized as follows. In Section 2, we formulate the problem by deriving a cost function for the common tangent. In Section 3, we propose a closed-form solution and robustness enhancement by means of the Hough transform. Experimental validation with simulated Monte Carlo runs is presented in Section 4 and finally Section 5 summarizes the paper and suggests directions for future work.

2. PROBLEM FORMULATION

Assume that there are M microphones distributed arbitrarily in a 2-D plane located at positions

$$\mathbf{r}_i \triangleq [x_i\, y_i]^T, \quad i = 0, \ldots, M-1, \tag{1}$$

with the reference microphone ($i = 0$) placed at the origin of the coordinate system, $\mathbf{r}_0 = [0\,0]^T$, and a source located at $\mathbf{r}_\mathrm{s} \triangleq [x_\mathrm{s}\, y_\mathrm{s}]^T$. The difference in the distances of microphones i and j from the source is the *range difference*, $d_{i,j}$, and is proportional to the TDOA of the direct-path between the ith and jth microphone, $\Delta_{i,j}$. If the speed of sound is η, then

$$d_{i,j} = \eta \cdot \Delta_{i,j}. \tag{2}$$

Each microphone receives the signal $x_i(t),\ i = 0, \ldots, M-1$ which is the sum of the direct-path signal and scaled replicas of the source signal. The delay of each replica is determined by the respective positions of reflectors, source and receivers. The observed signals $x_i(t)$ are therefore given by the convolution of the source $s(t)$ with the corresponding acoustic room impulse responses $h_i(t),\ i = 0, \ldots, M-1$,

$$x_i(t) = \int_0^\infty h_i(t')s(t-t')\,dt' + n_i(t),\ i = 0, \ldots, M-1, \tag{3}$$

where $n_i(t)$ is additive environmental noise. Accordingly, the AIRs are given by

$$h_i(t) = \sum_{q=0}^{Q} \alpha_{i,q}\delta(t - \tau_{i,q}), \tag{4}$$

where Q is the total number of reflections of all orders, $\alpha_{i,q}$ is an attenuation term and $\tau_{i,q}$ is defined as the TOA associated with the ith

microphone and the qth reflection. Note that the TOA of the direct-path is defined with respect to the *null* reflector, $q = 0$. However, for the remainder of this paper, we will not consider all reflections Q but only TOAs that are related to direct-paths and N first-order reflections, equal to the number of actual reflectors present in the acoustic environment. We can obtain estimates of the TOAs related to the direct path and first-order reflections by analyzing $h_i(t)$. For this we note that the first peak in $h_i(t)$ is related to the time of flight of the direct-path from $\mathbf{r}_\mathrm{s}$ to $\mathbf{r}_i$ given by $\tau_{i,0}$. Any subsequent peak in $h_i(t)$ is related to the composite time of flight of the sound due to reflection. By defining $\mathbf{r}_{\mathrm{p},i}$ as the reflection point on any reflector, we obtain $\tau_{i,k}$, $k = 1,\ldots,N$ as the sum of the two times of flight from $\mathbf{r}_\mathrm{s}$ to $\mathbf{r}_{\mathrm{p},i}$, and then from $\mathbf{r}_{\mathrm{p},i}$ to $\mathbf{r}_i$ for any N reflectors present in our environment. Under the hypothesis of optical acoustics, the angle of reflection and incidence are assumed equal such that the locus of possible solutions for the reflector forms an ellipse. We therefore establish the distance D_i from the ith microphone to the source such that

$$\hat{D}_i = \sqrt{x_\mathrm{s}^2 + y_\mathrm{s}^2} + \hat{d}_{i,0}. \tag{5}$$

The aim is to then find the parameters of the ellipse given the foci $\mathbf{r}_i = [x_i\ y_i]^T$ and $\mathbf{r}_\mathrm{s} = [x_s\ y_s]^T$ and the major axis, which will be defined in the following. In homogenous coordinates a conic in two dimensions using the parameters $\{a,b,c,d,e,f\}$ can be expressed as [6]

$$\mathscr{C} = \left\{(x,y) \in \mathbb{R}^2 | ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0\right\}. \tag{6}$$

By setting $\mathbf{x} = [x\ y\ 1]^T$ and $\mathbf{C} = \begin{bmatrix} a & b & d \\ b & c & e \\ d & e & f \end{bmatrix}$ this can be written as

$$\mathbf{x}^T \mathbf{C} \mathbf{x} = 0, \tag{7}$$

which parameterises an ellipse after constraining

$$\det(\mathbf{C}) \neq 0, \quad \begin{vmatrix} a & b \\ b & c \end{vmatrix} > 0, \quad \det(\mathbf{C})/(a+c) < 0. \tag{8}$$

From [5] we can define the ellipse associated with the ith microphone ($i \in \{0,\cdots,M-1\}$), and the kth reflector ($k \in \{1,\cdots,N\}$)

$$\mathbf{C}_{i,k} = \mathbf{T}_i^{-T} \mathbf{R}_i^{-T} \mathbf{S}_{i,k}^{-T} \mathbf{C}_\mathrm{I} \mathbf{S}_{i,k}^{-1} \mathbf{R}_i^{-1} \mathbf{T}_i^{-1}, \tag{9}$$

where we can define translation, rotation, scaling and unit circle matrices such that

$$\mathbf{T}_i = \begin{bmatrix} 1 & 0 & \Delta x_i \\ 0 & 1 & \Delta y_i \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{R}_i = \begin{bmatrix} \cos\gamma_i & -\sin\gamma_i & 0 \\ \sin\gamma_i & \cos\gamma_i & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{S}_{i,k} = \begin{bmatrix} Q_{i,k}^{\mathrm{maj}} & 0 & 0 \\ 0 & Q_{i,k}^{\mathrm{min}} & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{C}_\mathrm{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

The quantities Δx_i, Δy_i, γ_i, $Q_{i,k}^{\mathrm{maj}}$ and $Q_{i,k}^{\mathrm{min}}$ are defined as follows. The point at $(\Delta x_i, \Delta y_i)$ can be seen as the geographic midpoint between $\mathbf{r}_\mathrm{s}$ and $\mathbf{r}_i$ and is defined by

$$\Delta x_i \triangleq x_\mathrm{s} + \frac{D_i \cos(\gamma_i)}{2}; \ \Delta y_i \triangleq y_\mathrm{s} + \frac{D_i \sin(\gamma_i)}{2},$$

with $\gamma_i \triangleq \tan^{-1}\left(\frac{y_\mathrm{s} - y_i}{x_\mathrm{s} - x_i}\right)$. If we assume N reflectors in our environment, then every channel estimate contains information about the N TOAs due to the reflective sound path. The scaling of the semi-major and semi-minor axes of each ellipse is then given by

$$Q_{i,k}^{\mathrm{maj}} \triangleq \frac{\sqrt{(\eta\,\tau_{i,k})^2 - D_i^2}}{2}; \ Q_{i,k}^{\mathrm{min}} \triangleq \frac{\eta\,\tau_{i,k}}{2},$$

respectively. Furthermore we define a line in homogeneous coordinates as

$$\mathscr{L} = \left\{(x,y) \in \mathbb{R}^2 | l_1 x + l_2 y + l_3 = 0\right\}, \tag{10}$$

which after setting $\mathbf{l} = [l_1\ l_2\ l_3]^T$ can be written as

$$\mathbf{l}^T \mathbf{x} = 0. \tag{11}$$

If we consider N reflectors in our acoustic environment, then every channel of our identified AIRs should contain information about the N reflective path TOAs. By estimating these TOAs we can construct a set of $M \cdot N$ ellipse representations from which the line parameters of the reflectors can be obtained. If we group together M ellipses, extracted from every channel estimate and associated with a particular reflector, then the line parameters of that particular reflector can be estimated. Consequently for every $k \in \{1,\cdots,N\}$ reflectors we can define the following cost function,

$$J_\mathrm{e}\left(\mathbf{l}, \left\{\mathbf{C}_{i,k}^*\right\}_{i=0}^{M-1}\right) = \sum_{i=0}^{M-1} \left\| \mathbf{l}^T\ \mathbf{C}_{i,k}^*\ \mathbf{l} \right\|^2, \tag{12}$$

where $M \geq 3$ and $\mathbf{C}_{i,k}^* = \det(\mathbf{C}_{i,k})\mathbf{C}_{i,k}^{-1}$ is the adjoint of the conic-matrix $\mathbf{C}_{i,k}$.

3. PROPOSED METHOD

In this section we show how a line reflector can be estimated using an analytical minimization technique. In Section 3.1 we first present the closed-form solution considering only one reflector. In Section 3.2 we again consider all reflectors but follow on by grouping them together on a per-reflector basis in Section 3.2.2. Finally in Section 3.3 we introduce a second stage correction technique that improves the robustness and the numerical precision using the Hough transform.

3.1 Closed-form solution for a single reflector

The cost function in (12) is a multivariate fourth-order polynomial in l_1, l_2, l_3. We notice that the cost function admits the trivial solution $\mathbf{l} = \mathbf{0}$. In order to find the global minimum we resort to an analytical minimization technique. More specifically, we slice the homogeneous coordinates space (l_1, l_2, l_3) with the two planes $l_1 = 1$ and $l_2 = 1$. On these two planes the cost function $J(\mathbf{l})$ is not homogeneous and the set of local minima can be found in an analytical way. By merging the minima found on the two planes, we can obtain the global solution.

If we consider the case in which we iteratively estimate the line reflectors one at a time (i.e. $\mathbf{C}_{i,k}$ in (12) reduces to $\mathbf{C}_i$) then we can denote the coefficients of the adjoint conic associated to the ith ellipse with the matrix

$$\mathbf{C}_i^* = \begin{bmatrix} \alpha_i & \beta_i/2 & \delta_i/2 \\ \beta_i/2 & \gamma_i & \varepsilon_i/2 \\ \delta_i/2 & \varepsilon_i/2 & \zeta_i \end{bmatrix}.$$

Using this notation the cost function can be expanded as

$$\begin{aligned} J(\mathbf{l}) \quad = \quad & \textstyle\sum_{i=1}^{M} [\alpha_i^2 l_1^4 + \gamma_i^2 l_2^4 + \zeta_i^4 l_3^4 + 2\alpha_i\beta_i l_1^3 l_2 + 2\alpha_i\delta_i l_1^3 l_3 + 2\beta_i\gamma_i l_1 l_2^3 + \\ & + 2\gamma_i\varepsilon_i l_2^3 l_3 + 2\delta_i\zeta_i l_1 l_3^3 + 2\varepsilon_i\zeta_i l_2 l_3^3 + (2\alpha_i\gamma_i + \beta_i^2) l_1^2 l_2^2 + \\ & + (2\alpha_i\zeta_i + \delta_i^2) l_1^2 l_3^2 + (2\gamma_i\zeta_i + \varepsilon_i^2) l_2^2 l_3^2 + 2(\alpha_i\varepsilon_i + \beta_i\delta_i) l_1^2 l_2 l_3 + \\ & + 2(\beta_n\varepsilon_n + \gamma_n\delta_n) l_1 l_2^2 l_3 + 2(\beta_n\zeta_n + \delta_n\varepsilon_n) l_1 l_2 l_3^2]\,. \end{aligned} \tag{13}$$

Slicing $J(\mathbf{l})$ with the planes $l_1 = 1$ and $l_2 = 1$ means computing $J(\mathbf{l})|_{l_1=1}$ and $J(\mathbf{l})|_{l_2=1}$, respectively. Notice that $J(\mathbf{l})|_{l_1=1}$ and $J(\mathbf{l})|_{l_2=1}$ are no longer homogeneous. Fig. 1 shows an example of slices $J(\mathbf{l})|_{l_1=1}$ and $J(\mathbf{l})|_{l_2=1}$ (right-hand side), for the configuration of microphones and sources on the left-hand side.

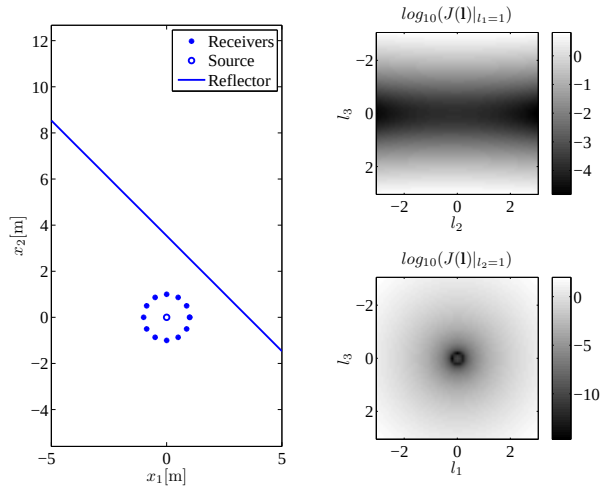

Figure 1: Example of cost functions $J(\mathbf{l})|_{l_1=1}$ and $J(\mathbf{l})|_{l_2=1}$ for a specific configuration of microphones and sources.

We proceed by finding the zeros of the gradient of $J(\mathbf{l})|_{l_1=1}$ and $J(\mathbf{l})|_{l_2=1}$, which yields the sets

$$L_1 = \left\{ \mathbf{l} : \left.\frac{\partial J(\mathbf{l})}{\partial l_2}\right|_{l_1=1} = 0 \ \wedge \ \left.\frac{\partial J(\mathbf{l})}{\partial l_3}\right|_{l_1=1} = 0 \right\}, \qquad (14)$$

and

$$L_2 = \left\{ \mathbf{l} : \left.\frac{\partial J(\mathbf{l})}{\partial l_1}\right|_{l_2=1} = 0 \ \wedge \ \left.\frac{\partial J(\mathbf{l})}{\partial l_3}\right|_{l_2=1} = 0 \right\}. \qquad (15)$$

Notice that the partial derivatives of the slices $J(\mathbf{l})_{l_1=1}$ and $J(\mathbf{l})_{l_1=1}$ are polynomials of order 3, and therefore L_1 and L_2 contain 9 solutions each. Some of them are in the complex domain and do not admit solution. We denote with $\bar{L}_1$ and $\bar{L}_2$ the subsets of purely real solutions of L_1 and L_2, respectively. We then define

$$\bar{L} = \bar{L}_1 \cup \bar{L}_2 = \{\mathbf{l}_1 \dots \mathbf{l}_K\}, \qquad (16)$$

which contains $K \leq 18$ candidate solutions. The global non-trivial minimum of $J(\mathbf{l})$ is selected as

$$\hat{\mathbf{l}} = \arg\min_{\mathbf{l}_m} J(\mathbf{l}_m), \ \mathbf{l}_m \in \bar{L}. \qquad (17)$$

3.2 Geometrical relation between line estimates and ellipses

Having obtained an estimate of the reflector line using (17) our aim is to estimate a number of points on the ellipses that are geometrically related to the line, in order to numerically improve the solution by use of the Hough transform. Generally speaking, given M ellipses and N reflector lines, our aim is to establish a collection of points

$$\mathbf{p}_j \triangleq [x_j \ y_j]^T, \quad j = 0, \dots, P, \qquad (18)$$

where $MN - 1 \leq P \leq 2MN - 1$, that are related to both the ellipse $\mathbf{C}$ and the line $\mathbf{l}$ in the following way:

1. If $\mathbf{l}$ goes through $\mathbf{C}$ then we obtain two points of intersection.
2. If $\mathbf{l}$ touches $\mathbf{C}$ at one point, or in other words if $\mathbf{l}$ is tangent to $\mathbf{C}$, then we obtain one point of tangency.
3. If $\mathbf{l}$ does not go through $\mathbf{C}$ then we need to calculate the closest point on the line with respect to the conic.

In the following, we will outline a methodology to check which of the above three relationships is valid for any candidate reflector line.

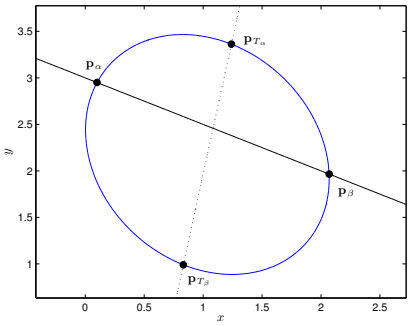

Figure 2: The line $\mathbf{l}$ (solid) intersects the ellipse at points $\mathbf{p}_\alpha$ and $\mathbf{p}_\beta$. The two parallel lines of $\mathbf{l}$ that are tangential to the ellipse at points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$ are represented by $\mathbf{l}_\mathrm{T}$ (dotted) that intersects the ellipse at the two tangential points.

3.2.1 Analytical framework

Consider two lines that are parallel to $\mathbf{l}$ and tangential to the ellipse at points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$. With reference to Fig. 2, we can construct a line $\mathbf{l}_\mathrm{T}$ that goes through both points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$, and consequently compute the points of intersection with the ellipse. For this we note that the slope of $\mathbf{l}$ is given by $m = -\frac{l_1}{l_2}$. Therefore the problem is constrained to finding the points on the ellipse for which the tangents have slope m. This can be achieved by implicit differentiation of the ellipse defined in (6)

$$\frac{d}{dx}(\mathscr{C}) = 2ax + 2bx\frac{dy}{dx} + c2y\frac{dy}{dx} + 2d\frac{dy}{dx} + 2by = 0. \qquad (19)$$

After setting $\frac{dy}{dx} = m$ the line that goes through both tangential points can be expressed as

$$\mathbf{l}_\mathrm{T} = [(a + bm)\,(b + cm)\,(d + em)]^T. \qquad (20)$$

For any line $\mathbf{l}$ it is possible to find the two points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$ at which two lines are both parallel to $\mathbf{l}$, i.e. with slope m, and also tangential to the ellipse. Since we can construct the line $\mathbf{l}_\mathrm{T}$ that goes through both points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$ from (20), all we need to do is to compute the points of intersection of $\mathbf{l}_\mathrm{T}$ and the ellipse. First, we will elaborate the methodology used to find the general intersection points of a line and an ellipse and then show how this can be used to compute points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$.

Given a line $\mathbf{l}$ that goes through the ellipse $\mathbf{C}$, the points of intersection $\mathbf{p}_\alpha \triangleq [x_\alpha \ y_\alpha]^T$ and $\mathbf{p}_\beta \triangleq [x_\beta \ y_\beta]^T$ are given by

$$x_\alpha = \frac{l_2\sqrt{(A+B+C)}+D}{E}, \quad y_\alpha = -\frac{-l_3 + l_1 x_\alpha}{l_2}; \qquad (21)$$

$$x_\beta = -\frac{l_2\sqrt{(A+B+C)}-D}{E}, \quad y_\beta = -\frac{-l_3 + l_1 x_\beta}{l_2}; \qquad (22)$$

with

$$\begin{aligned} A &= b\left(bl_3^2 - 2dl_2l_3 - 2el_1l_3 + 2fl_1l_2\right), \\ B &= d\left(dl_2^2 - 2el_1l_2 + 2cl_1l_3\right), \\ C &= e^2l_1^2 + 2ael_2l_3 - cfl_1^2 - afl_2^2 - acl_3^2, \\ D &= bl_2l_3 - dl_2^2 - cl_1l_3 + el_1l_2, \\ E &= cl_1^2 - 2bl_1l_2 + al_2^2. \end{aligned}$$

Instead of using $\mathbf{l}$ in (21) and (22) to find the general solutions $\mathbf{p}_\alpha$ and $\mathbf{p}_\beta$, we can replace $[l_1 \ l_2 \ l_3]^T$ with $[l_{T_1} \ l_{T_1} \ l_{T_1}]^T$, as given by (20), such that

$$\mathbf{l} \triangleq [(a + bm)\,(b + cm)\,(d + em)]^T, \qquad (23)$$

in order to find $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$. Since any line $\mathbf{l}$ will have two parallel lines that are tangential to the ellipse, (21), (22) and (23) can be used to check whether $\mathbf{l}$ goes through the ellipse, is tangential to the ellipse or does not go through the ellipse. If $\mathbf{l}$ cuts through the ellipse, then the parallel line touching the ellipse at point $\mathbf{p}_\alpha$ will be either to the left or right, above or below $\mathbf{l}$. In other words $\mathscr{L}(x_\alpha, y_\alpha)$ will either be positive or negative. Therefore if $\mathscr{L}(x_\alpha, y_\alpha)$ is positive and $\mathbf{l}$ does indeed go through the ellipse, then by definition $\mathscr{L}(x_\beta, y_\beta)$ must be negative. Consequently if $\mathscr{L}(x_\alpha, y_\alpha) < 0$ then $\mathscr{L}(x_\beta, y_\beta) > 0$. If $\mathbf{l}$ is tangential to the ellipse, then $\mathscr{L}(x_\alpha, y_\alpha)$ is either equal to zero, or not equal to zero. Therefore if $\mathscr{L}(x_\alpha, y_\alpha) = 0$, then $\mathscr{L}(x_\beta, y_\beta) \neq 0$. A similar argument holds for the case when $\mathscr{L}(x_\beta, y_\beta) = 0$. If $\mathbf{l}$ neither intersects or is tangential to the ellipse, then the two parallel lines touching the ellipse at points $\mathbf{p}_\alpha$ and $\mathbf{p}_\beta$ are either both below, above, left or right of $\mathbf{l}$. In other words if $\mathscr{L}(x_\alpha, y_\alpha) > 0$ then $\mathscr{L}(x_\beta, y_\beta) > 0$. If $\mathscr{L}(x_\alpha, y_\alpha) < 0$ then $\mathscr{L}(x_\beta, y_\beta) < 0$. Consequently in order to determine the relationship between $\mathbf{l}$ and the ellipse, it is sufficient to compute

$$\left|\operatorname{sgn}\left\{\mathscr{L}(x_\alpha, y_\alpha)\right\} + \operatorname{sgn}\left\{\mathscr{L}(x_\beta, y_\beta)\right\}\right|, \tag{24}$$

where $\operatorname{sgn}\{\cdot\}$ is defined as

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

If the result of (24) is 0, then $\mathbf{l}$ goes through the ellipse. If the result is 1, then $\mathbf{l}$ is tangential to the ellipse. Finally, if the result is 2 then the line does not intersect the ellipse.

3.2.2 Choosing candidate points

Given the analytical framework outlined in Section 3.2.1 it is possible to estimate a number of points $\mathbf{p}_i$ that are related to each ellipse in (12) and the candidate lines given by (17). Instead of estimating $\mathbf{p}_i$ exhaustively for each ellipse and candidate lines, it is more efficient to group candidate points on a per-reflector basis, denoted $\mathbf{p}_{j,k}$. We define

$$\mathbf{p}_{j,k}, \quad j = 0, \ldots, P^\dagger;\ k \in \{1, \cdots, N\}, \tag{25}$$

where $M - 1 \leq P^\dagger \leq 2M - 1$, as the jth candidate point for every kth reflector. Consequently for every kth reflector there will be M related ellipses. For each of the M ellipses $\left(\mathbf{C}_{i,k}\right)$ and each candidate solution $\left(\hat{\mathbf{l}}_k\right)$, related to the kth reflector, we first use (24) to check whether the estimated line either intersects the ellipse, is tangential to the ellipse or does not intersect the ellipse. In the first case, when $\hat{\mathbf{l}}_k$ goes through the ellipse, it is sufficient to use (21) and (22) to obtain the two points of intersection. In the second case, when $\hat{\mathbf{l}}_k$ is tangential to the ellipse, we simply use either (21) or (22), since the solutions will be equivalent, to obtain one point of tangency. In the final case, where $\hat{\mathbf{l}}_k$ does not intersect the ellipse, we use (21), (22) and (23) to obtain the two points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$: the tangential points of the two parallels of $\hat{\mathbf{l}}_k$ on the ellipse. The reason why we compute these two points is because one of them will be the closest point on the ellipse and the other the furthest point on the ellipse, with respect to the line. In order to then choose the closest point it is sufficient to compute the distance of points $\mathbf{p}_{T_\alpha}$ and $\mathbf{p}_{T_\beta}$ and the line, by projecting them both onto the line and selecting the shortest distance such that

$$\min\left\{\frac{\left|l_1 x_{T_\alpha} + l_2 y_{T_\alpha} + l_3\right|}{\sqrt{l_1^2 + l_2^2}}, \frac{\left|l_1 x_{T_\beta} + l_2 y_{T_\beta} + l_3\right|}{\sqrt{l_1^2 + l_2^2}}\right\}. \tag{26}$$

3.3 Improved reflector localization using the Hough transform

The Hough transform is a method for estimating the parameters of a shape from its boundary points [7]. The idea can be generalised to estimate the line parameters of $\mathscr{L}$ based on a selection of candidate points. In a noise-free scenario, and neglecting the effects of machine precision, the global minimum of (17) will also be the true solution, so that all ellipses will be perfectly aligned and yield a single solution that is the common tangent to all ellipses considered. However, in practice neither will the ellipses align perfectly to yield the optimal line estimate nor will the minimum of (17) always represent the true global solution. For this reason we estimate a number of candidate points $\mathbf{p}_{j,k}$, using the method outlined in Section 3.2.2, which represent a best match between the estimated line $\left(\hat{\mathbf{l}}_k\right)$ for each of the M ellipses $\left(\mathbf{C}_{i,k}\right)$ and every kth reflector. The Hough transform method considers the following normal parametrization [7]

$$\rho = x\cos\theta + y\sin\theta, \tag{27}$$

which specifies a straight line by the angle θ of its normal and its algebraic distance ρ from the origin. A point in the cartesian space corresponds in the Hough parameter space to all the lines passing through it, i.e. a sinusoid. Conversely, points in the parameter space are transformed into lines in the cartesian coordinate space. Given two points lying on a line with parameters ρ, θ, in the Hough parameter space the sinusoids corresponding to these two points intersect at ρ, θ. Therefore, given a collection of points $\mathbf{p}_j$ in the coordinate space, it is possible to estimate the line parameters of a line which is seen as a best-fit to that collection of points. If the points $\mathbf{p}_j$ lie on a straight or quasi-straight line, then by computing the intersection of the sinusoids in the Hough space it is possible to obtain values for θ and ρ from (27) that can be used to estimate the line parameters of the best-fit. The Hough Transform for a single-reflector case is shown in Fig. 3. Let $\rho \in \{-\mathbb{R}, \mathbb{R}\}$ and $\theta \in \{0, \pi\}$. For each edge point $[x_j\ y_j]^T$ we calculate

$$\hat{\rho} = x_j\cos\hat{\theta} + y_j\sin\hat{\theta}, \quad \forall\hat{\theta} \in \{0, \pi\}. \tag{28}$$

The results are stored in an accumulator $\mathscr{A}$, initially set to zero, which is incremented at every step such that:

$$\mathscr{A}\left(\hat{\rho}, \hat{\theta}\right) = \mathscr{A}\left(\hat{\rho}, \hat{\theta}\right) + 1. \tag{29}$$

The largest maximum of the accumulator given by

$$\left[\hat{\theta}_{\max}, \hat{\rho}_{\max}\right] = \max\left\{\mathscr{A}\left(\hat{\rho}, \hat{\theta}\right)\right\}, \tag{30}$$

is then picked, which finally leads to the line parameters of the best-fit:

$$\hat{\mathbf{l}} = \left[\cos(\hat{\theta}_{\max})\ \sin(\hat{\theta}_{\max})\ (-\hat{\rho}_{\max})\right]^T. \tag{31}$$

4. EXPERIMENTAL VERIFICATION

To evaluate the performance of the proposed algorithm, Monte Carlo simulations were performed for geometry reconstruction of an arbitrary rectangular room through iterative reflector localization. An experimental environment was created to evaluate the performance of the proposed algorithm in two variations. Experiment A considers the case where the line parameters are estimated using the method outlined in [5] and experiment B considers the case where the closed-form estimator with the second stage Hough transform correction, outlined in this paper, is used. In both cases the same data-set was used.

Simulated AIRs were obtained with the source-image method [8] for random source and receiver placement in a rectangular room of random dimensions 3–5 × 4–6 m with perfectly absorbing floors and ceilings. The sound source ($\mathbf{r}_s$) and the microphones ($\mathbf{r}_i$) were uniformly distributed inside the room, constraining the positions to be at a distance of at least 0.5 m from each wall and with each microphone being kept at a minimum distance of 0.5 m

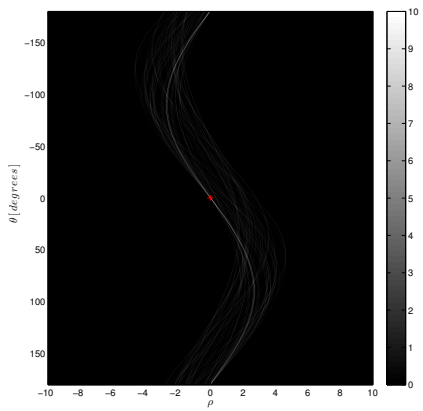

Figure 3: The Hough transform space graph showing the largest maximum (red +).

from the source. Note that although the source and receiver placements were randomized, the cases in which neighbouring peaks in the RIR are inseparable are excluded from the data set. Additionally the simulations were limited to only include first-order reflections. Note that unlike the results presented in [5] we consider the source location ($\mathbf{r}_s$) to be known.

4.1 Evaluation Criteria

The performance of the proposed algorithm was assessed by averaging the results of 100 Monte Carlo runs. The evaluation of the reflector localization algorithm is as follows. Let $\mathbf{l}$ and $\hat{\mathbf{l}}$ be the reflector line and the estimated reflector line, respectively. From the reference reflector line $\mathbf{l}$ and its estimate $\hat{\mathbf{l}}$ we can evaluate the distance κ from $\mathbf{r}_0$ to a point on each line and the orientation α. The distance and orientation can be evaluated by projecting $\mathbf{r}_0$ onto the line such that

$$\kappa = \frac{|l_1 x_0 + l_2 y_0 + l_3|}{\sqrt{l_1^2 + l_2^2}}, \quad \omega = 2\arctan\frac{\sqrt{l_1^2 + l_2^2} - l_1}{l_2}. \qquad (32)$$

The accuracy of the reflector localization is measured in terms of:

- distance error $\varepsilon_d = |\kappa - \hat{\kappa}|$;
- angular error $\varepsilon_a = |\omega - \hat{\omega}|$;
- alignment error $\varepsilon_l = \frac{\hat{\mathbf{l}}^T \mathbf{l}}{\|\hat{\mathbf{l}}\|\,\|\mathbf{l}\|}$, where values closer to 1 indicate the angle between the lines is small.

The mean and variance of ε_d, ε_a and ε_l were calculated for each experiment considering all located walls and individual walls ranked in order of error. In some cases not all walls are identified with the same degree of accuracy; ranking the error in this way provides insight into the distribution of errors as a function of the number of identified walls.

4.2 Results & Discussion

The results of the alignment, distance and angular error are given in Tables 1 & 2. In terms of alignment the four walls are very well localized using the original algorithm outlined in [5]. However the closed-form estimator yields a perfect alignment in all cases, $\mu(\varepsilon_l) = 1$. In terms of distance and angular error the closed-form estimator outperforms the original line estimator with lower mean and considerably smaller variance. Averaged across all walls the analytic line estimator presented in this paper achieves a $\mu(\varepsilon_d)$ and $\mu(\varepsilon_a)$ of less than a cm and degree respectively. Note that the improvement of the second stage correction is not explicitly stated in the table, but the gain in accuracy is calculated at about 21% for $\mu(\varepsilon_d)$ and 18% for $\mu(\varepsilon_a)$, when compared to the values estimated by the closed-form estimator alone. It can play an even more significant role and provide considerable advantages when considering a MIMO case instead of the SIMO example outlined here.

Table 1: Alignment error results.

Exp.	Walls	$\mu(\varepsilon_l)$	$\sigma(\varepsilon_l)$
A	**All**	**0.996**	**0.055**
A	Best	1.000	0.000
A	2nd best	1.000	0.005
A	2nd worst	0.996	0.033
A	Worst	0.987	0.105
B	**All**	**1.000**	**0.000**
B	Best	1.000	0.000
B	2nd best	1.000	0.000
B	2nd worst	1.000	0.000
B	Worst	1.000	0.000

Table 2: Distance and angular error results.

Exp.	Walls	$\mu(\varepsilon_d)$ [cm]	$\sigma(\varepsilon_d)$ [cm]	$\mu(\varepsilon_a)$ [°]	$\sigma(\varepsilon_a)$ [°]
A	**All**	**3.720**	**16.580**	**0.799**	**3.258**
A	Best	0.290	0.230	0.046	0.074
A	2nd best	1.390	6.130	0.264	1.606
A	2nd worst	3.820	12.300	0.956	3.544
A	Worst	9.400	29.490	1.931	5.048
B	**All**	**0.926**	**1.169**	**0.215**	**0.426**
B	Best	0.206	0.210	0.034	0.030
B	2nd best	0.505	0.295	0.091	0.057
B	2nd worst	0.884	0.421	0.179	0.138
B	Worst	2.109	1.756	0.555	0.737

5. CONCLUSIONS

An analytic solution using a closed-form estimator for localizing 2-D line reflectors using TOAs from SIMO acoustic impulse responses has been presented, circumventing the problem of converging to local sub-optimal solutions found with existing techniques. Robustness is further improved for those cases in which the TOAs contain errors with an approach based on the Hough transform. The improvements gained using the closed-form estimator are confirmed by Monte Carlo simulations. The approach may also be useful in developing future algorithms in which sequential SIMO measurements are made with multiple source locations, by providing a computationally-efficient means of combining multiple estimates of reflector location.

REFERENCES

[1] A. O'Donovan and R. Duraiswami, "Microphone arrays as generalized cameras for integrated audio visual processing," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, 2007.

[2] U. Castellani, A. Fusiello, V. Murino, L. Papaleo, E. Puppo, and M. Pittore, "A complete system for on-line 3d modelling from acoustic images," *Signal Processing: Image Communication*, vol. 20, no. 9–10, pp. 832–852, 2005, special issue on European projects on visual representation systems and services.

[3] P. Aarabi and B. Mungamuru, "Scene reconstruction using distributed microphone arrays," in *Proc. Intl. Conf. on Multimedia and Expo (ICME)*, vol. 3, 2003, pp. 53–56.

[4] F. Antonacci, A. Sarti, and S. Tubaro, "Geometric reconstruction of the environment from its response to multiple acoustic emissions," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2010, pp. 2822 –2825.

[5] J. Filos, E. A. P. Habets, and P. A. Naylor, "A two-step approach to blindly infer room geometries," in *Proc. Intl. Workshop Acoust. Echo Noise Control (IWAENC)*, Tel Aviv, Israel, Sep. 2010.

[6] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2001.

[7] R. O. Duda and P. E. Hart, "Use of the hough transformation to detect lines and curves in pictures," *Commun. ACM*, vol. 15, pp. 11–15, January 1972.

[8] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *Journal Acoust. Soc. of America*, vol. 65, no. 4, pp. 943–950, Apr. 1979.