

Reflector Localization based on Multiple Reflection Points

Youssef El Baba

International Audio Laboratories
Erlangen[†], Germany
youssef.elbaba@audiolabs-erlangen.de

Andreas Walther

Fraunhofer Institute for Integrated Circuits
Erlangen, Germany
andreas.walther@iis.fraunhofer.de

Emanuël A. P. Habets

International Audio Laboratories
Erlangen[†], Germany
emanuel.habets@audiolabs-erlangen.de

Abstract—Reflector localization has been the subject of growing research interest in recent years. This paper outlines an approach that performs reflector localization based on loudspeaker and microphone positions and their images. The positions of the latter are computed using pre-grouped sets of times of arrival (TOAs) estimated from room impulse responses. First, the TOA sets are used to estimate the microphone positions. Second, these are used with knowledge of the array geometry to determine the locations of reflection points on the available reflectors. Finally, the reflection points are used to obtain the reflector locations. It is shown that the proposed approach facilitates solving the reflector localization problem in ill-conditioned setups.

Index Terms—Image model, reflection point localization, reflector localization, room geometry inference.

I. INTRODUCTION

Room geometry inference (RGI) is concerned with the localization of reflective boundaries in an acoustic environment, and is of interest in several applications, such as 3D sound analysis and reproduction [1], robust acoustic source localization [2] and de-reverberation [3]. RGI is equivalent to reflector localization (RL) applied to all reflectors in a given room. RL methods use times of arrival (TOAs) of the direct-path and reflections — estimated from the room impulse responses (RIRs) between different microphone and loudspeaker positions — to infer the reflector locations. First-order reflections are especially of interest because they can be used to locate physical walls bounding the room. Two prevalent families of RL methods are those relying on ellipse-based geometry [4]–[8], or on hyperbola-based geometry [9], [10]. Others rely on steered response power [11], [12] or directivity-based schemes [13].

Most of the current RL methods assume 2D setups and use cylindrical/planar microphone arrays or randomly distributed microphones [14] accompanied by a single omnidirectional loudspeaker. First, TOAs need to be separated into sets belonging to a single reflector [9]. Each set is then used with the loudspeaker position to define multiple constraints which together determine the location of their reflector. The loudspeaker position is either given or estimated using sound source localization (SSL). An ideal loudspeaker-microphone setup exhibiting a sufficiently high spatial diversity is commonly assumed. For example, in the 2D case for a shoebox

room, a 2D microphone array and a loudspeaker at a fixed position can be used to infer up to all four reflectors.

In this paper, we consider in contrast a setup with a 1D linear loudspeaker array and a single omnidirectional microphone. This setup involves a change of perspective, namely using a loudspeaker array for RL instead of a microphone array. This change is viable considering the reciprocity of microphones and speakers [15]. More importantly, the 1D loudspeaker array, which is lacking one dimension in the 2D case, makes the RL problem suffer from ill-conditioning and prevents the direct use of existing RL methods. Ill-conditioning is especially noticeable when i) SSL estimates become severely inaccurate, and ii) ellipse-based algorithms [4]–[8] provide erroneous, yet plausible, solutions. Since ellipse and hyperbola based methods are equivalent duals [15], hyperbola-based methods are also expected to suffer from this last problem.

In this work we propose to estimate a reflector by fitting it to a set of reflection points. These points correspond to the locations on the walls where sound waves from the different loudspeakers are reflected, and are computed with knowledge of the array geometry and estimated positions of the real and image microphones. In [4]–[8], candidate reflection points are given by ellipses for each microphone-loudspeaker pair and a TOA. The reflector is then given by their common tangent. The proposed method replaces the search for a common tangent by computing the actual points of tangency, i.e., the reflection points. Furthermore, it is shown in this work that the aforementioned ill-conditioning can be mitigated by assuming that the loudspeaker array is positioned near one of the reflectors.

This paper is organized as follows: Section II formulates the problem mathematically, Section III outlines the algorithm in the case of a single reflector, Section IV expands the previous reasoning to multiple reflectors, Section V presents an objective performance evaluation of the proposed approach, and Section VI concludes the paper.

II. PROBLEM FORMULATION

In this work, a linear array with L loudspeakers and a single microphone is assumed. Considering only the direct-path and

[†]A joint institution of the Friedrich-Alexander-University Erlangen-Nürnberg (FAU) and Fraunhofer IIS, Germany.

first-order reflections, the RIR between the microphone and the j^{th} loudspeaker can be modeled by

$$h_j(t) = \alpha_{0j} \delta(t - \tau_{0j}) + \sum_{r=1}^R \alpha_{rj} \delta(t - \tau_{rj}) + \eta_j(t), \quad (1)$$

where the index r designates one of R reflectors, and α_{0j} and α_{rj} the attenuation coefficients of the direct and reflected paths, respectively. The function $\delta(t)$ represents the delta function and $\eta_j(t)$ represents the noise in the case of measured RIRs, as opposed to simulations. The time shifts τ_{0j} and τ_{rj} represent the TOAs of the direct and reflected wave fronts, respectively.

These TOAs in addition to the known relative loudspeaker positions in the array (i.e. the array geometry) constitute the input data. TOAs are retrieved through automatic peak picking, and are disambiguated into separate sets $\{\{\tau_{rj} : \forall j \in \{1..L\}\} : \forall r = 1..R\}$, each belonging to a single reflector, in a supervised manner¹.

The aim is to obtain from these TOAs the desired line equations characterizing the different reflectors in the form of

$$f_r x + g_r y + k_r = 0, \quad (2)$$

where $\mathbf{n}_r = [f_r, g_r]^T$ is the corresponding normal vector and k_r the line's offset. These lines lie on the intersections between the planes containing the real or image microphones and the loudspeakers, and the different physical reflectors.

The proposed solution presented below uses the input data to obtain the equations in (2) using a three-step procedure. First, SSL is performed to estimate the positions of real and image microphones. Second, using these position estimates, the positions of reflection points are computed. These then determine unequivocally the reflector equations.

In the following formulation, 2D position vectors are denoted by $\mathbf{v} = [v_x \ v_y]^T$, where v_x and v_y denote the x and y coordinates. The triplet (M, I_r, S_j) denotes the real microphone, the image microphone mirrored w.r.t the r^{th} reflector, and the j^{th} loudspeaker, respectively, and \mathbf{m} , \mathbf{i}_r and \mathbf{s}_j denote their respective positions. The positions of points are given as vectors in the following coordinate systems: \mathcal{C}_{S_j} denotes the loudspeaker-dependent system with origin S_j (Fig. 1); and $\mathcal{C}_{E_{rj}}$ the reflector- and loudspeaker-dependent system conveniently centered on the midpoint between S_j and M (the dependence on the reflector is made clear in Section III). The orientations of the systems are shown by the arrows in Fig. 1. All vectors in $\mathcal{C}_{E_{rj}}$ are indicated by a left-subscript E , such as ${}_E \mathbf{v}$, while other vectors are in \mathcal{C}_{S_j} .

III. SINGLE REFLECTOR LOCALIZATION

The main interim step between TOA grouping and obtaining the reflector equations in (2) is the computation of sets of reflection points

$$\{\mathbf{P}_{rj} : \forall j \in \{1..\tilde{L}\}\} \quad (3)$$

¹Other authors assuming provided TOA disambiguation include [10].

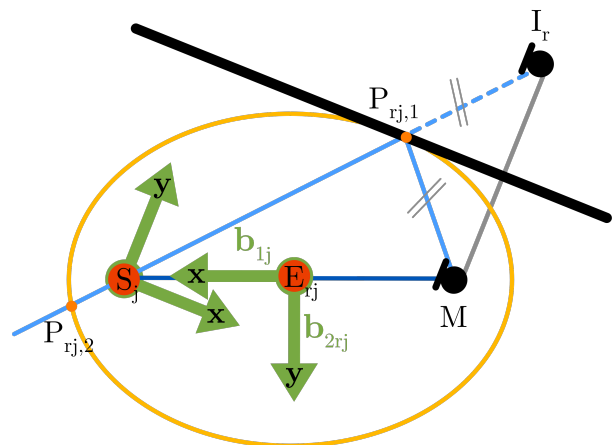


Fig. 1. Vector definitions, points and coordinate systems for a specific triplet (M, I_r, S_j) . The ellipse represents the constraint. The thick black line represents the reflector under consideration. The blue lines represent the sound propagation paths, and the red dots denote the coordinate systems.

where $2 \leq \tilde{L} \leq L$ denotes the number of reflection points. For RL in 2D it is sufficient to compute the locations of $\tilde{L} = 2$ points to obtain the equation of their corresponding reflector unequivocally.

For a fixed arrangement of a loudspeaker array and one microphone, a reflector r induces a corresponding image microphone I_r , and L triplets (M, I_r, S_j) for $j = 1..L$. Since in this section the only wave fronts available are those of the direct sound, and the sound bouncing off the available reflector, only two TOA sets $\{\tau_{rj} : \forall j \in \{1..L\}\}$ for $r \in \{0, 1\}$ exist for these two wave fronts. These two sets are relatively straightforward to separate by an iterative scheme: namely running the peak picking scheme a first time, replacing the earliest detected TOAs in the RIRs with zeros and then running the peak picking scheme a second time. Since the reflected sound TOAs do not interfere with the direct sound TOAs², this makes it possible to distinguish the two sets. The two TOA sets are then used with the known loudspeaker array geometry to estimate through SSL the positions \mathbf{m} and \mathbf{i}_r of the real and image microphones, for $r = 0$ and $r = 1$ respectively.

Each triplet (M, I_r, S_j) of the reflector r under consideration induces a reflection point localization (RPL) *sub*-problem involving three vectors \mathbf{m} , \mathbf{i}_r and \mathbf{s}_j and one reflection point \mathbf{P}_{rj} . For such a triplet, the candidate reflection point is computed with the help of an elliptical constraint: the locus of all points whose summed distances to S_j and M (Fig. 1) add up to τ_{rj} . Finding the point is easier when working in a convenient coordinate system such as $\mathcal{C}_{E_{rj}}$, as it is not only centered on — but also rotated to fit — the ellipse with its y axis oriented away from I_r ; thereby simplifying the formulation considerably. To this end a projection³ scheme is used to change to this coordinate system. In the following the r and

²Thanks to the direct line of sight between the loudspeakers and the real microphone.

³This projection is perfectly reversible.

j indices are omitted for brevity whenever possible since a specific triplet and its reflector are considered.

For a given triplet, the plane containing M , I_r and S is spanned⁴ by the vectors $\mathbf{s} - \mathbf{m}$ and $\mathbf{m} - \mathbf{i}_r$. A convenient orthogonal basis for its coordinate system \mathcal{C}_E is obtained by applying the Gram-Schmidt procedure on $(\mathbf{s} - \mathbf{m})$ and $(\mathbf{m} - \mathbf{i}_r)$:

$$\mathbf{b}_1 = (\mathbf{s} - \mathbf{m}) / \|\mathbf{s} - \mathbf{m}\| \quad (4)$$

$$\begin{aligned} \tilde{\mathbf{b}}_2 &= (\mathbf{m} - \mathbf{i}_r) - \langle \mathbf{m} - \mathbf{i}_r, \mathbf{b}_1 \rangle \mathbf{b}_1 \\ \mathbf{b}_2 &= \tilde{\mathbf{b}}_2 / \|\tilde{\mathbf{b}}_2\|, \end{aligned} \quad (5)$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product between vectors. Vectors like $\mathbf{i}_r - \mathbf{s}$ are projected from \mathcal{C}_S onto \mathcal{C}_E using \mathbf{b}_1 and \mathbf{b}_2 :

$${}_E(\mathbf{i}_r - \mathbf{s}) = \begin{bmatrix} {}_E(\mathbf{i}_r - \mathbf{s})_x \\ {}_E(\mathbf{i}_r - \mathbf{s})_y \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1^T \\ \mathbf{b}_2^T \end{bmatrix} (\mathbf{i}_r - \mathbf{s}). \quad (6)$$

The valid reflection point P lies at the intersection between the aforementioned ellipse and the *segment* $[S, I_r]$ (Fig. 1). First, two candidate intersections between the ellipse and the *line* (S, I_r) are computed, then the valid point is chosen between them according to the logic outlined below.

The line (S, I_r) is characterized in \mathcal{C}_E by $y = \lambda x + o$ where λ and o are respectively its slope and offset, with $\lambda = {}_E(\mathbf{i}_r - \mathbf{s})_y / {}_E(\mathbf{i}_r - \mathbf{s})_x$ and $o = -\lambda \|\mathbf{s} - \mathbf{m}\|/2$. The ellipse is unequivocally characterized by its foci corresponding to S and M , its semi-major axis $a = \|\mathbf{i}_r - \mathbf{s}\|/2$, and its semi-minor axis $b = (\sqrt{2a^2 - \|\mathbf{s} - \mathbf{m}\|^2})/2$. These expressions are formulated as such thanks to the alignment of the \mathcal{C}_E system with the ellipse's axes (Fig. 1). The positions of the two candidate intersections in \mathcal{C}_E are computed with

$$E\mathbf{P}_{1/2} = \frac{\begin{bmatrix} -a^2\lambda o \pm ab\sqrt{a^2\lambda^2 + b^2 - o^2} \\ b^2o \pm ab\lambda\sqrt{a^2\lambda^2 + b^2 - o^2} \end{bmatrix}}{(a^2\lambda^2 + b^2)}. \quad (7)$$

Then the choice between the two points is based on the relative distance between $P_{1/2}$ and S on (S, I_r) , which is given by

$$D_{1/2} = \langle {}_E(\mathbf{P}_{1/2} - \mathbf{s}), {}_E(\mathbf{i}_r - \mathbf{s}) \rangle / \|{}_E(\mathbf{i}_r - \mathbf{s})\|,$$

and the valid point P is chosen according to

$$E\mathbf{P} = \begin{cases} E\mathbf{P}_1 & \text{if } 0 \leq D_1 \leq 1 \\ E\mathbf{P}_2 & \text{otherwise} \end{cases}. \quad (8)$$

This procedure is performed for all $j = 1 \dots \tilde{L}$ loudspeakers. Since the array's geometry is known it is possible to map the results to a single coordinate system, say, that of the first loudspeaker (\mathcal{C}_{S_1}). Together they form the desired set $\{\mathbf{P}_{r,j} : \forall j \in \{1 \dots \tilde{L}\}\}$ in that common system. The reflector is then computed as the line passing through these points.

⁴In degenerate cases where (M, I_r, S_j) are collinear, this does not hold, however a trivial solution is available, i.e. the midpoint between M and I_r .

IV. EXTENSION TO MULTIPLE REFLECTORS & TOA DISAMBIGUATION

The presence of multiple reflectors produces multiple and potentially overlapping first-order reflections, and possibly also second and higher-order reflections. This makes the grouping of TOAs into sets more involved. While the above deals only with first-order reflections, in practice TOA sets from higher-order reflections produce phantom reflectors that can be discarded at a later step.

The process of grouping TOAs into sets belonging to a single reflector, i.e. disambiguation, alleviates the ambiguity in terms of which TOAs belong together, and ideally to which reflector they correspond in the case of reflected wave fronts. This process is needed because RL algorithms give phantom results if provided with sets consisting of TOAs belonging to different reflectors [9]. Similarly, SSL estimates phantom image transducers when provided with such incorrect sets. This occurs for both incorrectly disambiguated first-order and higher-order TOA sets. Several approaches have been proposed to disambiguate TOAs: in [9] a graph-based algorithm is outlined, in [14] the authors employ a combinatorial search scheme with validity checks, and in [5] the authors perform TOA disambiguation and RL jointly in a scheme based on the generalized hough transform.

For a fixed arrangement of loudspeakers and one real microphone, the scheme described in Section III is applicable to multiple image microphones $(I_r, \forall r)$, each corresponding to a specific reflector and its L triplets (M, I_r, S_j) . The positions of $(I_r, \forall r)$ are obtained from those *disambiguated* sets of TOAs $\{\tau_{r,j} : \forall j \in \{1 \dots L\}\}$ by means of SSL. This transforms the RL problem involving multiple reflectors into several smaller *sub-problems* involving only a single reflector.

Based on a coupling of SSL with analytical reflector computation, the proposed approach effectively circumvents the ill-conditioning discussed previously. The latter is limited to SSL where it can be more practically addressed. Namely, if the loudspeaker array is placed near a reflector, a reasonable assumption would be that the image microphones mirrored w.r.t. all the other reflectors lie in the region in front of the array and its nearby reflector. Hence, localizing real and image microphones in this frontal region is practically sufficient, giving an extra constraint to the SSL scheme and alleviating the front/back ambiguity causing the ill-conditioning in SSL. With the proposed approach, the ill-conditioning in RL is also alleviated.

V. PERFORMANCE EVALUATION

As a first step, RL was investigated using the standard ellipse algorithm family [4]–[8] on the given limited setup for multiple reflectors in a shoebox room. It was found that the time difference of arrival (TDOA)-based SSL [16] resulted in large errors w.r.t. the room dimensions, which in turn resulted in erroneous RL estimates. Moreover, when the real microphone's location was provided to the RL algorithm, other problems due to ill-conditioning still appeared: because of the missing dimension in the loudspeaker array, multiple reflector

candidates according to the logic in [4]–[8] were plausible but erroneous solutions. When a loudspeaker was added in the missing dimension orthogonal to that of the linear array, the algorithm provided better RL estimates.

As to the approach proposed in this paper, experiments were conducted using both simulated and real RIRs measured in shoebox rooms. The TOAs were estimated from the RIRs assuming that the system was temporally synchronized (known hardware latency for real measurements). TOA disambiguation was performed in a supervised manner. Next, real and image microphone positions were estimated with SSL using those pre-grouped TOA sets and RPL was applied on all available triplets (M, I_r, S_j) and then reflector lines fitted to the obtained sets of reflection points $\{P_{rj} : \forall j \in \{1..L\}\}$. To limit the processing to 2D space, TOAs corresponding to sound waves reflecting off the ceiling and floor were discarded manually. In addition, the TOA set corresponding to the reflector near the loudspeaker array was discarded, due to the frontal region assumption mentioned in Section IV.

To assess the accuracy of the real and image microphone position estimates, the Euclidean distance between the true and estimated microphone positions $E_{ML} = \|\mathbf{m} - \hat{\mathbf{m}}\|$ as in [7] was used, where \mathbf{m} is a real or image microphone position and $\hat{\mathbf{m}}$ its estimate. To assess the accuracy of RL and therefore the accuracy of RPL, the orientation error $O_{RL} = |\arccos(\langle \mathbf{n}, \hat{\mathbf{n}} \rangle)|$ as in [17] between the true and estimated reflectors' normal vectors was used; additionally the offset $D_{RL} = ||\langle \mathbf{n}, (\mathbf{m} - \mathbf{x}) \rangle| - |\langle \hat{\mathbf{n}}, (\mathbf{m} - \hat{\mathbf{x}}) \rangle||$ as in [17] in terms of distance to the real microphone location \mathbf{m} was used, where \mathbf{x} and $\hat{\mathbf{x}}$ represent points on the true and estimated reflectors, respectively. These three metrics are represented graphically in Fig. 2. Lower metric values indicate better performance.

In the following, each setup (one room and one loudspeaker array arrangement) with a specific microphone position formed a separate scenario with a set of RIRs on which the above procedure was applied. The presence of multiple microphone positions in each setup enabled more testing scenarios, as for each setup the metrics were averaged across all such available positions (scenarios) and the reflectors. The end result was that RL was performed on a total of 157 simulated and 23 real reflectors, respectively. Some reflectors were excluded because they could not be resolved due to extraneous constraints.

Additionally, the SSL step in this approach assumes temporally synchronized measurements (known latency) for good accuracy as in that case it is possible to use a TOA-based multilateration scheme instead of one using TDOAs as in [16], since the latter are subject to higher errors. The scheme employed finds the spatial position of the microphone at which the TOAs from multiple loudspeakers, translated into ranges in meters, intersect. The search is performed in the least-squares sense.

A. Simulated RIRs

Six different setups with different configurations were prepared for evaluation of RL using RPL, with RIRs simulated

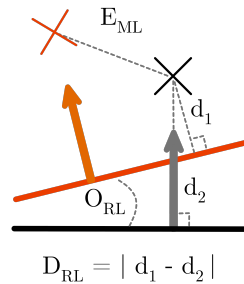


Fig. 2. Visual representation of RL error metrics. The black line, arrow and cross indicate the true reflector, its normal vector and (real or image) microphone position, while the red line, arrow and cross indicate their estimated counterparts.

using the image method of first order [18]. Each setup consisted of a shoebox room of specific dimensions, one 1.5 m long uniform linear loudspeaker array with 16 drivers, and included 9 microphones. The 9 microphones were placed on a cross pattern centered in the middle of each room, in a range of 1 – 6 m from the loudspeaker array. The attenuation coefficients α_{rj} for all reflectors were set to 0.7. While 5 setups featured a loudspeaker array parallel to one of the reflectors, one setup had the array rotated 15° counter-clockwise. This rotated setup tested the generalizability of the proposed approach. A total of 54 scenarios were evaluated.

One reflector was discarded from the evaluation in the rotated setup because its corresponding image microphone was perfectly collinear with the loudspeaker array. In that case all the reflections are located on a single point because of the reduced array geometry, leaving the reflector line fitting with an infinite number of solutions. To avoid this problem in practice, it is useful to restrict the array's rotation to small intervals.

The results of these evaluations for the 6 different setups are shown in rows 1–6 in Table I, in addition to the respective sizes of the rooms. The proposed approach performs similarly to what is reported in the literature, in spite of the ill-conditioned setups.

B. Measured RIRs

RIRs were measured for each (loudspeaker, microphone) pair in a setup involving a 1.56 m long linear loudspeaker array with 26 drivers, where 8 omnidirectional microphones were placed at specific positions — on a grid in one quadrant, 4.6 – 7.8 m away from the loudspeaker array — across a shoebox room, forming therefore 8 scenarios. The hardware latency was provided to the algorithm based on a comparison with the simulated replica of the measurements. The results of this evaluation are shown in the 7th row in Table I, with the size of the room. Reverberation time (RT60) was estimated for all RIRs and had a value of 0.432s, averaged across all microphone/loudspeaker position pairs. One scenario with its SSL and RL estimates is shown as an example in Fig. 3.

VI. CONCLUSION

A new approach for RL using RPL has been presented, having the advantage of being almost exclusively analytical and compatible with ill-conditioned setups. In spite of these setups, the performance of the proposed approach has been

TABLE I

PERFORMANCE METRICS FOR 6 SIMULATED SETUPS (ROWS 1-6) AND 1 REAL SETUP (ROW 7). SHOWN ARE THE METRICS AVERAGED ACROSS ALL MICROPHONE POSITIONS AND REFLECTORS, WITH A MEAN VALUE FOLLOWED BY \pm THE STANDARD DEVIATION (SEE SECTION V).

Setup, room dimensions (m)	E_{ML} (cm)	D_{RL} (cm)	O_{RL} (degrees)
4.5x5	1.105 \pm 0.404	0.546 \pm 0.379	0.106 \pm 0.091
6x4	0.976 \pm 0.288	0.527 \pm 0.294	0.073 \pm 0.076
6x8.5	4.131 \pm 8.876	1.061 \pm 1.461	0.320 \pm 0.656
9x7.5	1.226 \pm 0.656	0.562 \pm 0.295	0.073 \pm 0.063
6x12	1.840 \pm 1.391	0.425 \pm 0.310	0.080 \pm 0.071
4.5x5, rotated by 15°	4.700 \pm 9.251	1.992 \pm 2.324	0.870 \pm 1.201
12.66x10.42, measured RIRs	9.472 \pm 9.660	4.631 \pm 4.206	0.485 \pm 0.377

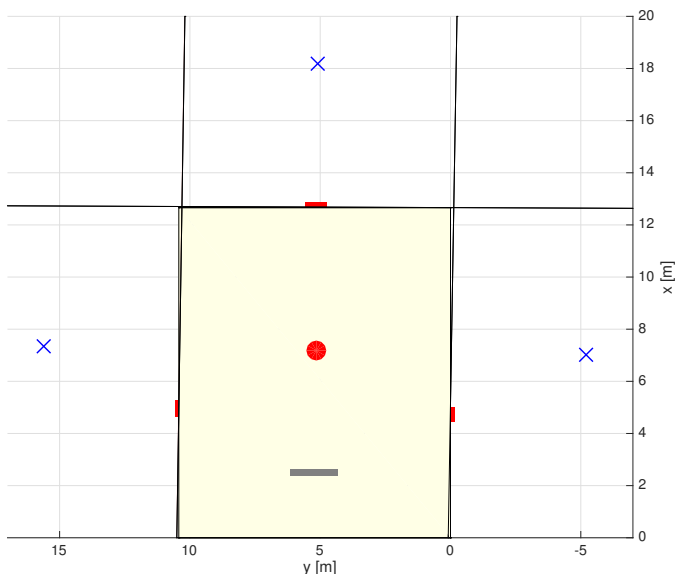


Fig. 3. Example of RL estimates (slightly non-vertical and non-horizontal lines) for a specific real set of RIRs. Real microphone is indicated by red dot, image microphones by smaller blue crosses; all are estimated with SSL. The reflection points are indicated by little red lines, and the linear array by a gray line. The true reflectors are perfectly horizontal and vertical on the graph, with their estimates showing the slight estimation offset.

found to be similar to what is reported in the current literature. As a side benefit, the proposed approach can, for a large part, be efficiently implemented using vector operations.

Future research will focus on extending the proposed approach to 3D, in addition to developing a robust TOA disambiguation scheme.

REFERENCES

- [1] D. de Vries and M. Boone, "Wave field synthesis and analysis using array technology," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 1999, pp. 15–18.
- [2] F. Ribeiro, D. Ba, C. Zhang, and D. Florêncio, "Turning enemies into friends: Using reflections to improve source localization," in *Proc. Intl. Conf. Multimedia and Expo (ICME)*, Singapore, Jul. 2010, pp. 731–736.
- [3] Y. Peled and B. Rafaely, "Method for dereverberation and noise reduction using spherical microphone arrays," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2010, pp. 113–116.
- [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)*, Mar. 2010, pp. 2822–2825.
- [5] F. Antonacci, J. Filos, M. Thomas, E. Habets, A. Sarti, P. Naylor, and S. Tubaro, "Inference of room geometry from acoustic impulse responses," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 20, no. 10, pp. 2683–2695, Dec. 2012.
- [6] A. Canclini, F. Antonacci, M. R. P. Thomas, J. Filos, A. Sarti, P. Naylor, and S. Tubaro, "Exact localization of acoustic reflectors from quadratic constraints," in *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*, 2011.
- [7] 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.
- [8] E. Nastasia, F. Antonacci, A. Sarti, and S. Tubaro, "Localization of planar acoustic reflectors through emission of controlled stimuli," in *Proc. European Signal Processing Conf. (EUSIPCO)*, 2011, pp. 156–160.
- [9] J. Scheuing and B. Yang, "Disambiguation of TDOA estimation for multiple sources in reverberant environments," *IEEE Trans. Audio, Speech, Lang. Process.*, vol. 16, no. 8, pp. 1479–1489, Nov. 2008.
- [10] A. Moore, M. Brookes, and P. Naylor, "Room geometry estimation from a single channel acoustic impulse response," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Sep. 2013, pp. 1–5.
- [11] A. O'Donovan, R. Duraiswami, and D. Zotkin, "Imaging concert hall acoustics using visual and audio cameras," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2008, pp. 5284–5287.
- [12] E. Mabande, H. Sun, K. Kowalczyk, and W. Kellermann, "On 2D localization of reflectors using robust beamforming techniques," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 2011, pp. 153–156.
- [13] S. Tervo and T. Tossavainen, "3-D room geometry estimation from room impulse responses," in *Proc. IEEE Intl. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2012, pp. 513–516.
- [14] I. Dokmanic, R. Parhizkar, A. Walther, Y. M. Lu, and M. Vetterli, "Acoustic echoes reveal room shape," *Proceedings of the National Academy of Sciences*, vol. 110, no. 30, pp. 12 186–12 191, 2013.
- [15] R. Schmidt, "A new approach to geometry of range difference location," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 8, no. 6, pp. 821–835, Nov. 1972.
- [16] Y. A. Huang, J. Benesty, G. W. Elko, and R. M. Mersereau, "Real-time passive source localization: a practical linear-correction least-squares approach," *IEEE Trans. Speech Audio Process.*, vol. 9, no. 8, pp. 943–956, Nov. 2001.
- [17] J. Filos, A. Canclini, F. Antonacci, A. Sarti, and P. Naylor, "Localization of planar acoustic reflectors from the combination of linear estimates," in *Proc. European Signal Processing Conf. (EUSIPCO)*, Aug. 2012, pp. 1019–1023.
- [18] J. B. Allen and D. A. Berkley, "Image method for efficiently simulating small-room acoustics," *J. Acoust. Soc. Am.*, vol. 65, no. 4, pp. 943–950, Apr. 1979.