MINIMAL SOLUTIONS FOR DUAL MICROPHONE RIG SELF-CALIBRATION

Simayijiang Zhayida, Simon Burgess, Yubin Kuang and Kalle Åström

Centre for Mathematical Sciences, Lund University
{zhayida,simonb,yubin,kalle}@maths.lth.se

ABSTRACT

In this paper, we study minimal problems related to dual microphone rig self-calibration using TOA measurements from sound sources with unknown positions. We consider the problems with varying setups as (i) if the internal distances between the microphone nodes are known a priori or not. (ii) if the microphone rigs lies in an affine space with different dimension than the sound sources. Solving these minimal problems is essential to robust estimation of microphone and sound source locations. We identify for each of these minimal problems the number of solutions in general and develop non-iterative solvers. We show that the proposed solvers are numerically stable in synthetic experiments. We also apply our method in a real indoor experiment and obtain accurate reconstruction using TOA measurements.

1. INTRODUCTION

Time-of-arrival (TOA) and time-difference-of-arrival (TDOA) measurements are used in applications ranging from radio based positioning to beamforming and audio sensing. Although such problems have been studied extensively in the literature in the form of localization of e.g. a sound source using a calibrated detector array, the problem of calibration of a sensor array using only measurement has received comparatively less attention.

Several contributions addressing the self-calibration problem rely on prior knowledge or extra assumptions of locations of the sensors [1–6]. Iterative methods exist for TOA or TDOA based self-calibration [7, 8]. However, such methods are dependent on initialization and can get stuck in local minima. Initialization of TOA sensor networks using only measurements has been studied in [9, 10], where solutions to the minimal cases of three senders and three receivers in the plane, or six senders and four receivers in 3D are given. Initialization of TDOA networks is studied in [11] and refined in [12] where a solution to non-minimal case of 9 receivers and 4 speakers in 3D was derived. In [13] a far field approximation was utilized to calibrate both TOA and TDOA networks. [9–13] attempt to solve the self-calibration problem with either minimal or close to minimal data. Studying minimal cases is both of theoretical importance and essential to develop fast stable algorithms suitable in random sample consensus (RANSAC) [14] schemes.

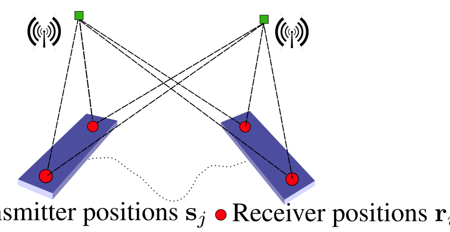

Fig. 1: Sketch of setting. Rigid pair of receivers with unknown positions get TOA signals from unknown sender positions

In this paper we focus on the previously unsolved problem of finding positions of a set of receivers and speakers, where pairs of receivers are set on a rigid rig, using only time-of-arrival (TOA) measurements between receivers and speakers with unknown positions (Fig.1). We show in what constellations of receivers and speakers the self-calibration problem has a solution, and present numerically stable closed form solvers for these minimal cases. Applications can be in robotics and SLAM, where a robot is equipped with stereo receivers in a rigid constellation, moving through a room with unknown transmitter positions. Recently mobile devices e.g. iPhone 5s also come equipped with dual microphones. Furthermore, solving the corresponding time-difference-of-arrival (TDOA) calibration problem often involves a two step process: First, figuring out the offsets and then solving the TOA calibration problem, see e.g. [12, 15]. The dual receiver rig self-calibration, we study here, also has the advantage of needing fewer measurements than the corresponding self-calibration problems for unconstrained receivers, and is thus better suited in RANSAC schemes where the setting applies.

2. THE TOA-BASED MIC-RIG CALIBRATION PROBLEM

We study the TOA-based mic-rig calibration problem for dual microphone rigs. A dual microphone rig is a rigid array with two receivers and we set all rigs to have same length c between receivers. The problem setting can be seen in Fig.1. We assume that receivers can distinguish which TOA signal comes from which sender. This can be done in practice by e.g.

m/n	2	3	4	5
2	-	$G_{2,3}$	$(G_{3,3}, G^u_{2,3}, G^u_{3,2})$	$G^u_{3,3}$
3	$G_{2,3}$	$G_{3,2}$	-	-
4	$G^u_{2,3}$	$G^u_{3,2}$	-	-

Table 1: Here m and n represents number of rigs and speakers respectively. G_{k_m,k_n} denotes the corresponding case in the table is minimal for calibrated dual microphone rigs in dimension k_m and sounds in dimension k_n. $G^u_{k_m,k_n}$ denotes the uncalibrated case of G_{k_m,k_n}.

separating the signals temporally or by frequency. The TOA dual rig calibration problem can then be defined as follows.

Problem 1 *(Calibrated) Given absolute distance measurements $d_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2$ and length between receiver pairs on the same rig c, determine receiver positions $\mathbf{r}_i$ and transmitter positions $\mathbf{s}_j$.*

Problem 2 *(Uncalibrated) Given absolute distance measurements $d_{ij} = ||\mathbf{r}_i - \mathbf{s}_j||_2$, determine receiver positions $\mathbf{r}_i$ and transmitter positions $\mathbf{s}_j$ as well the constant length of the receiver rigs c.*

One can only hope to recover a solution up to rotation/mirroring and translation of coordinate system, as any such transformation applied to a solution $\mathbf{r}_i, \mathbf{s}_j$ result in the same measurements d_{ij}.

2.1. Identifying Minimal Problems

Depending on the number of receiver rigs and speakers, we first characterize when such problems are well-defined i.e. when there are finite number of solutions to the problem. We are particularly interested in the minimal problems where minimal number of receiver rigs and speakers are required to solve the problem. One way to identify such minimal problems is to study the degrees of freedom of the problems F and the number of measurements M. The necessary condition for a problem to be minimal is that $F = M$. For instance, for the case where the m receiver rigs and the n speakers both span a 3D affine space, we have $F = 6m + 3n - 6$ (here -6 takes care of the *gauge freedom* i.e. rotation and translation ambiguity in the reconstruction). We have $M = 2mn + m$ and $M = 2mn + m - 1$ for calibrated and uncalibrated cases, respectively. By finding m, n such that $F = M$, we can identify potential minimal cases. With this type of analysis, we have identified a set of potential minimal problems. These cases are summarized in Table 1.

2.2. Problem Formulation

We start by deriving a set of new equations. Since the distance measurements are assumed to be real and positive one does not lose any information by squaring them, i.e.

$$d_{ij}^2 = (\mathbf{r}_i - \mathbf{s}_j)^T(\mathbf{r}_i - \mathbf{s}_j) = \mathbf{r}_i^T\mathbf{r}_i + \mathbf{s}_j^T\mathbf{s}_j - 2\mathbf{r}_i^T\mathbf{s}_j.$$

The problem is significantly easier to analyze and solve by forming new equations according to the following linear combinations of d_{ij}^2. Given the TOA measurements between receivers and the speakers, we have $2mn$ equations in the following four types:

(A) 1 equation $d_{11}^2 = (\mathbf{r}_1 - \mathbf{s}_1)^T(\mathbf{r}_1 - \mathbf{s}_1)$.

(B) $n-1$ equations of type $d_{1j}^2 - d_{11}^2 = \mathbf{s}_j^T\mathbf{s}_j - \mathbf{s}_1^T\mathbf{s}_1 - 2\mathbf{r}_1^T(\mathbf{s}_j - \mathbf{s}_1)$, for $j = 2, \ldots, n$.

(C) $2m-1$ equations of type $d_{i1}^2 - d_{11}^2 = \mathbf{r}_i^T\mathbf{r}_i - \mathbf{r}_1^T\mathbf{r}_1 - 2(\mathbf{r}_i - \mathbf{r}_1)^T\mathbf{r}_1$, for $i = 2, \ldots, 2m$.

(D) $(2m-1)(n-1)$ equations of type $d_{ij}^2 - d_{i1}^2 - d_{1j}^2 + d_{11}^2 = -2(\mathbf{r}_i - \mathbf{r}_1)^T(\mathbf{s}_j - \mathbf{s}_1)$, for $i = 2, \ldots, 2m, j = 2, \ldots, n$.

And we also have following equations for known or unknown rig distance, respectively:

(E) m equations of type $c_{2i-1,2i}^2 = (\mathbf{r}_{2i-1} - \mathbf{r}_{2i})^T(\mathbf{r}_{2i-1} - \mathbf{r}_{2i})$, for $i = 1, \ldots, m$.

(E′) $m-1$ equations of type $||\mathbf{r}_1 - \mathbf{r}_2|| - ||\mathbf{r}_{2i-1} - \mathbf{r}_{2i}|| = 0$, for $i = 2, \ldots, m$.

Using these equations, we now describe methods to solve the polynomial systems for different minimal problems.

3. SOLVING MINIMAL PROBLEMS

The solution strategy is to use a factorization step first to reconstruct the positions up to an unknown affine transformation $\mathbf{L}$ and $\mathbf{b}$. By collecting terms, the equations of type (D) can be written in matrix form $\tilde{\mathbf{D}} = \mathbf{r}^T\mathbf{s}$ with $\mathbf{r}_i$ as columns of $\mathbf{r}$ and $\mathbf{s}_j$ as columns of $\mathbf{s}$. The rank of $\tilde{\mathbf{D}}$ depends on the dimensionality of the affine span of the receivers and the speakers respectively. For instance, if we assume that both of the rigs and the speakers are in 3D, then the matrix $\tilde{\mathbf{D}}$ also has rank 3. By factorizing $\tilde{\mathbf{D}}$ which is of rank 3 using e.g. singular value decomposition, we can compute the vectors to all receivers and speakers from unknown initial/reference positions ($\mathbf{r}_1$ and $\mathbf{s}_1$) up to an unknown full-rank 3×3 transformation $\mathbf{L}$ such that $\tilde{\mathbf{D}} = \tilde{\mathbf{r}}^T\mathbf{L}^{-1}\mathbf{L}\tilde{\mathbf{s}} = \mathbf{r}^T\mathbf{s}$. Depending on how one fixes the gauge freedom, the unknown $\mathbf{b}$ enters the equations in different ways. By a good choice of parametrization of the problems it can be shown that the equations of Types C, E (or E′) are linear in the unknowns and the equations of Types A and B can be used to form polynomial equations.

3.1. 3D-Rigs and 3D-Speakers

In this section, we solve the minimal problems for cases where both the rigs and speakers are in 3D.

To solve for the unknown transformation and reference positions, we now utilize the nonlinear constraints in equations of Type A, B and C. First we can fix the gauge freedom by choosing the location $\mathbf{r}_1$ at the origin. Given that $\mathbf{r} = \mathbf{L}^{-T}\tilde{\mathbf{r}}$ and $\mathbf{s} = \mathbf{L}\tilde{\mathbf{s}}$, we can parameterize $\mathbf{s}_1$ as $\mathbf{L}\mathbf{b}$ where $\mathbf{b}$ is a 3×1 vector. This gives

$$\begin{aligned} \mathbf{r}_i &= \mathbf{L}^{-T}\tilde{\mathbf{r}}_i,\ i = 2\dots 2m, \\ \mathbf{s}_j &= \mathbf{L}(\tilde{\mathbf{s}}_j^* + \mathbf{b}),\ j = 2\dots n, \end{aligned} \tag{1}$$

where $\tilde{\mathbf{s}}_j^* = \tilde{\mathbf{s}}_j/(-2)$. Using the parametrization above and also letting $\mathbf{H} = (\mathbf{L}^T\mathbf{L})^{-1}$ the equations of type (A), (B), (C) and (E) become

$$d_{11}^2 = \mathbf{b}^T\mathbf{H}^{-1}\mathbf{b}, \tag{2}$$

$$d_{1j}^2 - d_{11}^2 = \tilde{\mathbf{s}}_j^{*T}\mathbf{H}^{-1}\tilde{\mathbf{s}}_j^* + 2\mathbf{b}^T\mathbf{H}^{-1}\tilde{\mathbf{s}}_j^*, \tag{3}$$

$$d_{i1}^2 - d_{11}^2 = \tilde{\mathbf{r}}_i^T\mathbf{H}\tilde{\mathbf{r}}_i - 2\mathbf{b}^T\tilde{\mathbf{r}}_i, \tag{4}$$

$$c_{2i-1,2i}^2 = \begin{cases} \tilde{\mathbf{r}}_2^T\mathbf{H}\tilde{\mathbf{r}}_2 \\ \tilde{\mathbf{r}}_{2i-1}^T\mathbf{H}\tilde{\mathbf{r}}_{2i-1} - 2\tilde{\mathbf{r}}_{2i-1}^T\mathbf{H}\tilde{\mathbf{r}}_{2i} + \tilde{\mathbf{r}}_{2i}^T\mathbf{H}\tilde{\mathbf{r}}_{2i}. \end{cases} \tag{5}$$

3.1.1. *Case of 2 Rigs and 4 Speakers Calibrated*

There are in total 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively). By utilizing $\mathbf{H}^{-1} = \mathrm{adj}(\mathbf{H})/\det(\mathbf{H})$, where $\mathrm{adj}(\mathbf{H})$ is the adjoint of $\mathbf{H}$, we can multiply equations in (2)-(5) by $\det(\mathbf{H})$ to rewrite them as polynomials equations. There are three linear equations of type C and two linear equation of type E. By linear elimination we can parameterize $\mathbf{H}$ and $\mathbf{b}$ in terms of $9 - 5 = 4$ unknowns $\mathbf{x} = (x_1, x_2, x_3, x_4)$. We now obtain four equations:

$$\det(\mathbf{H})d_{11}^2 = \mathbf{b}^T \mathrm{adj}(\mathbf{H})\mathbf{b} \tag{6}$$

$$\det(\mathbf{H})(d_{12}^2 - d_{11}^2) = \tilde{\mathbf{s}}_2^{*T} \mathrm{adj}(\mathbf{H})\tilde{\mathbf{s}}_2^* + 2\mathbf{b}^T \mathrm{adj}(\mathbf{H})\tilde{\mathbf{s}}_2^* \tag{7}$$

$$\det(\mathbf{H})(d_{13}^2 - d_{11}^2) = \tilde{\mathbf{s}}_3^{*T} \mathrm{adj}(\mathbf{H})\tilde{\mathbf{s}}_3^* + 2\mathbf{b}^T \mathrm{adj}(\mathbf{H})\tilde{\mathbf{s}}_3^* \tag{8}$$

$$\det(\mathbf{H})(d_{14}^2 - d_{11}^2) = \tilde{\mathbf{s}}_4^{*T} \mathrm{adj}(\mathbf{H})\tilde{\mathbf{s}}_4^* + 2\mathbf{b}^T \mathrm{adj}(\mathbf{H})\tilde{\mathbf{s}}_4^* \tag{9}$$

in the four unknowns. While multiplying $\det(\mathbf{H})$ introduces false solutions, we utilize the same saturation technique as in [10] to remove such solutions. Using algebraic geometric tools, we verify that this system has in general 12 solutions. We then solve this system with efficient polynomials solvers based on [16]. After solving for $\mathbf{H}$, L can be calculated with Cholesky factorization which is done for the other cases where the affine span is larger than one.

3.1.2. *Case of 2 Rigs and 5 Speakers Uncalibrated*

For this case, there are 9 equations (1 of Type A, 4 of Type B, 3 of Type C and 1 of Type E′) and 9 unknowns (6 and 3 unknowns for $\mathbf{H}$ and $\mathbf{b}$, respectively). We follow a similar solution scheme as for the case of 2 rigs and 4 speakers case. By linear elimination using the 4 linear constraints of type C and E, we can express $\mathbf{H}$ and $\mathbf{b}$ in terms of $9 - 4 = 5$ unknowns $\mathbf{x} = (x_1, x_2, x_3, x_4, x_5)$. The remaining five constraints (1 of Type A, and 4 of Type B) gives a polynomial system with 28 solutions after a saturation procedure similar to the previous case. Again we use the scheme in [16] to produce a numerically stable and efficient solution.

3.2. 2D-Rigs and 3D-Speaker

In this section, we solve the minimal problems for cases where the rigs in 2D and speakers are in 3D.

3.2.1. *Case of 2 Rigs and 3 Speakers Calibrated*

We can here parameterize the two rigs to be on the z-plane. Then we know that the matrix $\tilde{\mathbf{D}}$ is of rank-2, and $\mathbf{H}$ is a symmetric 2 by 2 matrix, and $\mathbf{b}$ is a 2 by 1 matrix. There are three linear equations of type C and two linear equations of type E and 5 unknowns (3 for $\mathbf{H}$ and 2 for $\mathbf{b}$). Thus we can solve for $\mathbf{H}$ and $\mathbf{b}$ linearly and resolve for the positions of $\mathbf{r}$ and the projection of $\mathbf{s}$ onto the z plane. Then we solve the z coordinates for the speakers simply using the equation of type B.

3.2.2. *Case of 3 Rigs and 2 Speakers Calibrated*

By factorizing $\tilde{\mathbf{D}}$ which is of rank 1 in this case, we can compute all receivers and speakers from unknown initial/reference positions up to the unknown transformation $L = l$, which is now only an unknown constant. We now utilize the nonlinear constraints in equations of Type A, B, C and E.

We know that one can reconstruct for $\mathbf{r}$ and $\mathbf{s}$ with factorization up to unknown transformation l such that $\tilde{\mathbf{D}} = \tilde{\mathbf{r}}^T \frac{1}{l} l\tilde{\mathbf{s}} = \mathbf{r}^T\mathbf{s}$, we put $\tilde{\mathbf{r}} = [0 \quad \tilde{\mathbf{D}}^T]$ and $\tilde{\mathbf{s}} = [0 \quad \mathbf{1}]$. We fix the gauge freedom by choosing the location $\mathbf{r}_1 = [r_{11}, 0, 0]^T$, $\mathbf{r}_i = [r_{1i}, r_{2i}, 0]^T, i = 2, \dots 6$ and $\mathbf{s}_j = [0, s_{2j}, s_{3j}]^T, j = 1, 2$, and also denote second row of $\mathbf{r}$ and $\mathbf{s}$ as $\bar{\mathbf{r}}$ and $\bar{\mathbf{s}}$, respectively. This gives

$$\begin{aligned} \bar{\mathbf{r}}_i &= \frac{1}{l}\tilde{\mathbf{r}}_i,\ i = 2, \dots, 2m, \\ \bar{\mathbf{s}}_j &= l(\tilde{\mathbf{s}}_j + b)/(-2),\ j = 1, \dots, n. \end{aligned} \tag{10}$$

The equations of type (A), (B), (C) and (E) then become

$$d_{11}^2 = r_{11}^2 + \frac{l^2b^2}{4} + s_{31}^2, \tag{11}$$

$$d_{1j}^2 - d_{11}^2 = \frac{l^2}{4}\tilde{\mathbf{s}}_j^2 + \frac{l^2 b}{2}\tilde{\mathbf{s}}_j + s_{3j}^2 - s_{31}^2, \tag{12}$$

$$d_{i1}^2 - d_{11}^2 = r_{1i}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_i + b\tilde{\mathbf{r}}_i - r_{11}^2, \tag{13}$$

$$c_{2i-1,2i}^2 = \begin{cases} r_{11}^2 + r_{12}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_2^2 - 2r_{11}r_{12}, \quad i = 1 \\ \\ r_{1,2i-1}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_{2i-1} + r_{1,2i}^2 + \frac{1}{l^2}\tilde{\mathbf{r}}_{2i} - \\ \quad -2r_{1,2i-1}r_{1,2i} - \frac{2}{l^2}\tilde{\mathbf{r}}_{2i-1}\tilde{\mathbf{r}}_{2i}, \quad i > 1. \end{cases} \tag{14}$$

There are in total 10 unknowns (1 for l, 1 for b, 6 for r_{1i} and 2 for s_{3j}). We have 10 equations (1 of Type A, 1 of Type B, 5 of Type C and 3 of Type E). By using the parametrization $\mathbf{x} = [h = \frac{1}{l^2}, b, v_1, v_2, v_3, v_4, v_5, v_6, u_{12}, u_{34}, u_{56}]^T$, where $v_i = r_{1i}^2, i = 1, \ldots, 6$ and $u_{2i-1,2i} = r_{1,2i-1}r_{1,2i}, i = 1, 2, 3$, we have 8 linear equations from (13) and (14). Thus we can express all the unknowns in $\mathbf{x}$ linearly in terms of $[h, b, v_1]$. We then proceed to solve the three equations

$$u_{2i-1,2i}^2 = r_{1,2i-1}^2 r_{1,2i}^2, \quad i = 1, 2, 3 \tag{15}$$

in the three unknowns $[h, b, v_1]$, using the techniques in [16]. Resubstitution gives us the coordinated of $\mathbf{r}$ and $\mathbf{s}$. In general there are 2 solutions.

3.2.3. *Case of 4 Rigs and 2 Speakers Uncalibrated*

Similar to the previous calibrated case, we now have 12 equations (1 of Type A, 1 of Type B, 7 of Type C and 3 of Type E′). We solve the problem in the same manner as the calibrated 3-rigs and 2-speakers case. For the corresponding polynomial system, there is in general 6 solutions.

3.2.4. *Cases of 2 Rigs and 4 Speakers Uncalibrated*

This case is actually underdetermined though the it satisfies $F = M$. One way to explain this is the following. Adding one transmitter seems to give 4 measurements and 3 unknowns (unknown speaker positions in 3D), thus one obtains an additional constraint which could indicate that we can use that to solve the rig length c. But in fact we only get 3 (linearly independent) measurements due to the rank constraints on $\tilde{\mathbf{D}}$. Thus it is unsolvable.

3.3. 3D-Rigs and 2D-Speaker

In this section, we solve the minimal problems for cases where the rigs are in 3D and speakers are in 2D.

3.3.1. *Case of 2 Rigs and 4 Speakers Uncalibrated*

We have 8 equations (1 of Type A, 3 of Type B, 3 of Type C and 1 of Type E′) and 9 unknowns (4 for z-coordinates of receivers and 5 unknowns for $\mathbf{H}$ and $\mathbf{b}$). We do not have enough information to solve the case.

3.3.2. *Case of 3 Rigs and 3 Speakers Calibrated*

Similar technique as 2D-3D case, but this time $\tilde{\mathbf{D}}$ has rank 2, and we have different characterization of $\tilde{\mathbf{r}}$ and $\tilde{\mathbf{s}}$ as $\tilde{\mathbf{r}} = [\mathbf{0}_{2\times 1} \;\; \tilde{\mathbf{D}}^T]$ and $\tilde{\mathbf{s}} = [\mathbf{0}_{2\times 1} \;\; \mathbf{1}]$. If we fix the translational part of the gauge freedom by choosing the location $\mathbf{r}_1 = [0, 0, r_{31}]^T$, $\mathbf{r}_2 = [r_{12}, 0, r_{32}]^T$ and also denote first two rows of $\mathbf{r}$ and $\mathbf{s}$ as $\bar{\mathbf{r}}$ and $\bar{\mathbf{s}}$, respectively, then we get our equation system in a similar way.

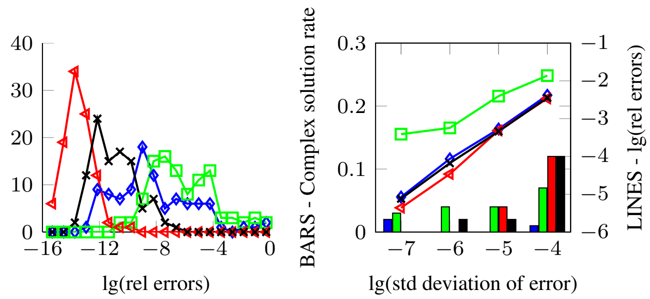

Fig. 2: Blue, Green, Red and Black correspond to the cases of Sec. 3.1.1, 3.1.2, 3.2.2, and 3.2.3. Left: Histogram of relative errors for 100 runs of solvers. Right: Mean of lg relative errors and complex solution rate for 100 runs, plotted against the standard deviation of the additive Gaussian noise.

For this case, there are 11 equations (1 of Type A, 2 of Type B, 5 of Type C and 3 of Type E) with 8 of them linear, and 11 unknowns (6 for z-coordinate of receivers and 5 unknowns for $\mathbf{H}$ and $\mathbf{b}$). We use the same technique which is used in Section 3.2.2. Using algebraic geometry tools, we verify that there are in general 16 solutions.

3.3.3. *Case of 4 Rigs and 3 Speakers Uncalibrated*

We have 13 equations (1 of Type A, 2 of Type B, 7 of Type C and 3 of Type E′) and 13 unknowns (8 for z-coordinate of receivers and 5 unknowns for $\mathbf{H}$ and $\mathbf{b}$). We can following the same strategy to parametrize the problem as the calibrated 3-rigs and 3-speakers case. This problem is much more difficult and has in general 29 solutions.

4. EXPERIMENTS

To be able to evaluate the quality of a solution, receivers and transmitter positions $\mathbf{r}_i$ and $\mathbf{s}_j$ are compared to ground truth positions $\mathbf{r}_{i,gt}$ and $\mathbf{s}_{j,gt}$. Positions are rotated, mirrored and translated so that the points are aligned [17]. For comparing with computer vision reconstruction, the alignment of points is also done over scale. Relative errors are defined as $\|[\mathbf{r} - \mathbf{r}_{gt} \;\; \mathbf{s} - \mathbf{s}_{gt}]\|_F / \|[\mathbf{r}_{gt} \;\; \mathbf{s}_{gt}]\|_F$ where $\|\cdot\|_F$ is the Frobenius norm. Simulations were run for 100 cases where ground truth receivers and transmitters where drawn uniformly over a unit cube around the origin. Half of the receivers were then fitted to a rig distance of 0.2 from their respective pairs, and measurements d_{ij} were created from ground truth. Relative errors for the solvers with and without additive Gaussian noise on the measurements can be seen in Fig. 2.

For testing the case of senders and receivers in 3D with a calibrated rig distance, an indoor experiment was carried out. A set of real data was obtained using four T-bone MM-1 microphones and four Roxcore portable speakers, connected to a Fast Track Ultra 8R sound card in an indoor environment,

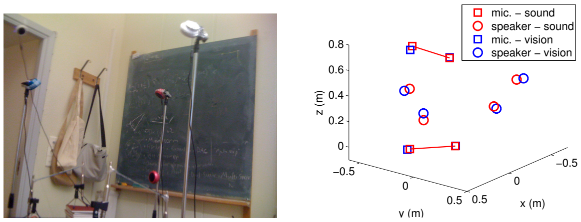

Fig. 3: Real microphone and speaker calibration setup, Left: the setup of microphones and speakers in an office, Right: the reconstructed sensor positions using TOA measurements (red) aligned with the positions estimated based on computer vision (blue).

with speakers and microphones placed in an approximate $1.5 \times 1.5 \times 1.5\ m^3$ volume (Fig.3, Left). TOA measurements were obtained by heuristically matching beginning of sounds from different speakers to beginning of sounds recorded from different microphones. A reconstruction of the scene was made using computer vision techniques to be used as ground truth. The reconstruction (Fig.3, Right) when compared to the vision-based reconstruction has an RMSE of $4.2\ cm$ and $5.6\ cm$ for microphones and speakers respectively.

5. CONCLUSION

In this paper, we have studied the TOA based self-calibration problem of dual microphone rigs for known and unknown rig distance, and also for affine space with different dimensions for receiver and sender spaces. We present extensive studies on the minimal problems including number of solutions, parameterizations as well as non-iterative numerically stable solvers. We also show good reconstruction using real data in an indoor environment. We see this problem as a building block for TDOA based self-calibration problem of dual microphone rigs, and we believe it can be used to further analyze problem within radio, Wi-Fi and ultrasound.

REFERENCES

[1] S. T. Birchfield and A. Subramanya, “Microphone array position calibration by basis-point classical multidimensional scaling,” *IEEE transactions on Speech and Audio Processing*, vol. 13, no. 5, 2005.

[2] D. Niculescu and B. Nath, “Ad hoc positioning system (aps),” in *GLOBECOM-01*, 2001.

[3] E. Elnahrawy, Xl. Li, and R. Martin, “The limits of localization using signal strength,” in *SECON-04*, 2004.

[4] V. C. Raykar, I. V. Kozintsev, and R. Lienhart, “Position calibration of microphones and loudspeakers in distributed computing platforms,” *IEEE transactions on Speech and Audio Processing*, vol. 13, no. 1, 2005.

[5] M. Crocco, A. Del Bue, M. Bustreo, and V. Murino, “A closed form solution to the microphone position self-calibration problem,” in *37th International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2012), Kyoto, Japan*, March 2012.

[6] J. C. Chen, R. E. Hudson, and K. Yao, “Maximum likelihood source localization and unknown sensor location estimation for wideband signals in the near-field,” *IEEE transactions on Signal Processing*, vol. 50, 2002.

[7] R. Biswas and S. Thrun, “A passive approach to sensor network localization,” in *IROS 2004*, 2004.

[8] J. Wendeberg, F. Hoflinger, C. Schindelhauer, and L. Reindl, “Calibration-free tdoa self-localization,” *Location Based Services*, vol. 7, pp. 121–144, May 2013.

[9] H. Stewénius, *Gröbner Basis Methods for Minimal Problems in Computer Vision*, Ph.D. thesis, Lund University, April 2005.

[10] Y. Kuang, S. Burgess, A. Torstensson, and K. Åström, “A complete characterization and solution to the microphone position self-calibration problem,” in *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Processing*, 2013.

[11] M. Pollefeys and D. Nister, “Direct computation of sound and microphone locations from time-difference-of-arrival data,” in *Proc. of International Conference on Acoustics, Speech and Signal Processing*, 2008.

[12] Yubin Kuang and Kalle Åström, “Stratified sensor network self-calibration from tdoa measurements,” in *21st European Signal Processing Conference*, 2013.

[13] Yubin Kuang, Erik Ask, Simon Burgess, and Karl Åström, “Understanding toa and tdoa network calibration using far field approximation as initial estimate,” in *ICPRAM*. 2012, pp. 590–596, SciTePress.

[14] M. A. Fischler and R. 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. 381–95, 1981.

[15] N. D. Gaubitch, W. B. Kleijn, and R. Heusdens, “Auto-localization in ad-hoc microphone arrays,” in *Proc. IEEE Int. Conf. on Acoust. Speech and Signal Processing*, 2013.

[16] Martin Byröd, Klas Josephson, and Kalle Åström, “Fast and stable polynomial equation solving and its application to computer vision,” *Int. Journal of Computer Vision*, vol. 84, no. 3, pp. 237–255, 2009.

[17] Arun K.S., Huang T.S., and Blostein S.D., “Least-squares fitting of two 3-d point sets,” *IEEE Transactions on pattern analysis and machine intelligence*, 1987.