

Minimal Structure and Motion Problems for TOA and TDOA Measurements with Collinearity Constraints

Erik Ask, Simon Burgess and Kalle Åström

Centre for Mathematical Sciences, Lund University, Lund, Sweden

Keywords: Structure from Sound, TOA, TDOA, Minimal Solvers.

Abstract: Structure from sound can be phrased as the problem of determining the position of a number of microphones and a number of sound sources given only the recorded sounds. In this paper we study minimal structure from sound problems in both TOA (time of arrival) and TDOA (time difference of arrival) settings with collinear constraints on e.g. the microphone positions. Three such minimal cases are analyzed and solved with efficient and numerically stable techniques. An experimental validation of the solvers are performed on both simulated and real data. In the paper we also show how such solvers can be utilized in a RANSAC framework to perform robust matching of sound features and then used as initial estimates in a robust non-linear least-squares optimization.

1 INTRODUCTION

Sound ranging or sound localization has been used since world war I, to determine the sound source using a number of microphones at known locations and measuring the time-difference of arrival of sounds. The same mathematical model is today used both for applications based on acoustics and radio and both for signal strength or time-based information such as time of arrival (TOA) or time differences of arrival (TDOA), or a combination thereof. 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, i.e. the initialization problem for sensor network calibration, has received much less attention. One technique used for sensor network calibration is to manually measure the inter-distance between pairs of microphones and use multi-dimensional scaling to compute microphone locations, (Birchfield and Subramanya, 2005). Another option is to use GPS, (Niculescu and Nath, 2001), or to use additional transmitters (radio or audio), close to each receiver, (Elnahrawy et al., 2004; Raykar et al., 2005; Sallai et al., 2004). Sensor network calibration is treated in (Biswas and Thrun, 2004). In (Chen et al., 2002) it is shown how to estimate additional microphones, once an initial estimate of the position of some microphones are known. In (Thrun, 2005) the far field approximation is used to

initialize the calibration of sensor networks. Initialization of TOA networks has been studied in (Stewénius, 2005), where solutions to the minimal case of three transmitters and three receivers in the plane is given. The minimal case in 3D is determined to be four receivers and six transmitters for TOA, but this is not solved. Initialization of TDOA networks is studied in (Pollefeys and Nister, 2008), where solutions were given to two non-minimal cases of ten transmitters and five receivers, whereas the minimal solution for far field approximation in this paper are six transmitters and four receivers. In (Wendeberg et al., 2011) a TDOA setup is used for indoor navigation based on non-linear optimization, but the method can get stuck in local minima and is dependent on initialization.

In this paper we will study the effects of restricting one set of synchronized sensors to a line (we will assume receivers). For TOA measurements applications could be to determine all positions by travelling along a line and measuring distances to fixed positions. In TDOA it could be used to calibrate linear sensor-arrays, easily setting up scenarios for indoor navigation by placing sensors along a wall. A more complicated setting could be if the line synchronization could be emulated, by for instance using known periodic signals from the transmitters, to again estimate positions of both a receiver and known transmitters by a linear motion. For example a moving car in range of cellular antennas.

2 PROBLEM FORMULATION

We will denote the position of transmitter i as \mathbf{z}_i and position of receiver j as \mathbf{m}_j . As is shown in the next section one may without loss of generality assume planar configurations, i.e. $\mathbf{z}_i = (x_i, y_i)$ and $\mathbf{m}_j = (u_j, v_j)$. Throughout the paper $\mathbf{D} = [d_{ij}]$ is a matrix with time measurements between transmitters and receivers. Since we are only interested in measuring distances or relative distances it is indifferent what type of sensor is placed collinearly, assuming synchronization is on line, but for consistency we shall assume the receivers are placed collinearly for our arguments.

2.1 Linear Restriction

The main purpose of the linear restriction is to reduce the number of unknowns, and hence the number of necessary equations. This will reduce the size of the minimal case and has the dual advantage of more stable numerical performance and a reduced requirement in the number of transmitters and receivers needed. The cost being the reduced usability in that either transmitters or receivers need to be placed in linear constellations. Since we only measure distances from a common line, the results presented hold in all dimensions, with the dimension of the solution set being of size 2 smaller than the dimension of the problem. To see this consider any measurement $d_{ij} = \|\mathbf{z}_i - \mathbf{m}_j\|$ and let the first dimension, denoted u be defined by the collinear receivers. Then \mathbf{z}_i has an orthogonal decomposition $\mathbf{z}_i = \tilde{\mathbf{z}}_i + \tilde{\mathbf{z}}_i^\perp$ with $\tilde{\mathbf{z}}_i = u_{zi}$ a unique point on the u axis. Then u_{zi} can be uniquely determined from any two known measurement points on the line but only the length of $\tilde{\mathbf{z}}_i^\perp$. The ambiguity in direction gives us several equally feasible solutions. In 2D there are 2 directions that with known distance fixates two points (0D). In 3D the perpendicular vector can be rotated around the line 360 degrees tracing out a circle (1D). We summarize the above in a theorem.

Theorem 1. *The Structure and motion problem for linear motion based on measurements to reference points is equivalent in all dimensions ND up to the dimension of the solution set.*

Another important aspect is that the linear configuration most likely is a degenerate case of any full minimal solver. In (Stewénius, 2005) It is shown that the solver for the planar unrestricted case for TOA the linear setup causes the algorithm to become unstable and it is impossible to solve for the non-line placed sensors without adapting the method for handling null spaces. There are no minimal polynomial solvers for TDOA in any dimension, but theorem 1 essentially

states that the linear case results in similar considerations on null spaces.

2.2 TOA

The TOA case occurs when time synchronization is possible between transmitters i and receivers j . By our assumptions this implies that all distances d_{ij} are known.

With k receivers are placed on a line and all n transmitters are unrestricted we get kn measurements and $2n+k-1$ unknowns, where one receiver is placed in origo and the remaining receivers placed on the first axis. The minimal case is then given by the smallest possible integer solution to

$$kn = 2n + k - 1 \quad (1)$$

easily confirmed to be $k = 3$ and $n = 2$. Using what we know from Theorem 1 we now have

Lemma 1. *The minimal case for linear TOA in $N \geq 2$ dimensions is 3 receivers and 2 transmitters, and has a $N - 2$ dimensional solution set.*

To derive the solution we have, since by assumption $v_j = 0$ for each measurement d_{ij} that

$$E_{ij} : d_{ij}^2 = x_i^2 - 2x_i u_j + u_j^2 + y_i^2. \quad (2)$$

This gives a total of 6 equations and 6 unknowns since we can set $u_1 = 0$. Forming the two combinations $(E_{21} - E_{22}) - (E_{11} - E_{12})$ and $(E_{21} - E_{23}) - (E_{11} - E_{13})$ gives

$$\begin{aligned} d_{12}^2 - d_{22}^2 - d_{11}^2 + d_{12}^2 &= u_2(2x_2 - 2x_1) \\ d_{12}^2 - d_{23}^2 - d_{11}^2 + d_{13}^2 &= u_3(2x_2 - 2x_1) \end{aligned}, \quad (3)$$

and hence that

$$\frac{d_{12}^2 - d_{22}^2 - d_{11}^2 + d_{12}^2}{d_{12}^2 - d_{23}^2 - d_{11}^2 + d_{13}^2} = \frac{u_2}{u_3}, \quad (4)$$

giving us the possibility to exchange u_2 for a constant times u_3 . A second order equation containing only u_3 can then be obtained by

$$\frac{d_{12}^2 - d_{23}^2 - d_{11}^2 + d_{13}^2}{d_{12}^2 - d_{22}^2 - d_{11}^2 + d_{12}^2} (E_{12} - E_{11}) - (E_{13} - E_{11})$$

with u_2 substituted in $(E_{12} - E_{11})$. This polynomial is trivial to solve and the remaining variables can be obtained by back substitution in intermediate results.

2.3 TDOA

The motivation for TDOA is to avoid the restriction of synchronization between transmitters and receivers. In this setting it is not possible to directly transform measured times to distances as it is unknown at what

point in time the signal was originally transmitted. By instead imposing a restriction that all collinear receiver are synchronized we can instead look at the difference in time of arrival. First the relation in equation 2 is modified as to account for the ambiguity by for each transmitter i introducing an unknown offset o_i as

$$d_{ij} = \sqrt{(x_i - u_j)^2 + (y_i - v_j)^2} + o_i. \quad (5)$$

The following lemma gives the minimal cases under these settings

Lemma 2. *The minimal case in $N \geq 2$ dimensions with no synchronization between transmitters and receivers, but synchronized receivers is either 2 transmitters and 5 receivers or 3 transmitters and 4 receivers, and has a $N - 2$ dimensional solution set.*

Proof: We place the receivers on the first axis and one receiver in origo. As all k receivers are assumed synchronized we have no new parameters and we still get $k - 1$ unknowns. For the n transmitters we now have both the unknown spatial coordinates as well as the offset giving us $3n$ unknowns. As before we get kn equations. It is simple to verify that $k = 5, n = 2$ and $k = 4, n = 3$ are the minimal integer solutions to $k - 1 + 3n = kn$. The independence of dimension follows directly from theorem 1 ■

In the subsequent discussions we will refer to the above situations as (5,2) and (4,3) respectively. The relation between measurements and positions as given in equation 5 are not on polynomial form and hence can not be solved directly by polynomial solvers. By first eliminating the square root one obtains

$$(d_{ij} - o_i)^2 = (x_i - u_j)^2 + (y_i - v_j)^2,$$

which with $v_j = 0$ can be written as

$$d_{ij}^2 - 2d_{ij}o_i + o_i^2 = x_i^2 - 2x_iu_j + u_j^2 + y_i^2. \quad (6)$$

By subtracting any two such relations for any fixed i effectively eliminates both the o_i^2, x_i^2 and y_i^2 terms. We choose again to set $u_1 = 0$ and will use the corresponding equations to subtract obtaining

$$d_{ij}^2 - d_{i1}^2 - 2o_i(d_{ij} - d_{i1}) + 2x_iu_j - u_j^2 = 0. \quad (7)$$

If we interpret this as a linear system in the monomial o_i, x_i and 1 we get for each transmitter j the system

$$\begin{pmatrix} d_{i2} - d_{i1} & u_2 & u_2^2 - d_{i2}^2 + d_{i1}^2 \\ d_{i3} - d_{i1} & u_3 & u_3^2 - d_{i3}^2 + d_{i1}^2 \\ d_{i4} - d_{i1} & u_4 & u_4^2 - d_{i4}^2 + d_{i1}^2 \\ d_{i5} - d_{i1} & u_5 & u_5^2 - d_{i5}^2 + d_{i1}^2 \end{pmatrix} \begin{pmatrix} -2o_i \\ 2x_i \\ -1 \end{pmatrix} = 0, \quad (8)$$

for the (5,2) case and an equivalent system with the last line in the matrix removed for the (4,3) case. By

basic linear algebra such systems have non-trivial solutions exactly when the determinant of the matrix is zero. For the over determined (5,2) case this must hold for all 3×3 sub matrices. In the (4,3) case we have a total of 3 square matrices and hence as many determinants. The determinants form polynomial equations in the unknowns u_2, u_3 and u_4 . This means we have reduced our problem from 12 equations in 12 unknowns to just solving 3 equations in 3 unknowns. For the (5,2) case we get 2 rectangular matrices with a total of 8 sub determinants for the unknowns u_2, u_3, u_4 and u_5 . A subset would be sufficient, but as more equations will be generated later in the solution algorithm in practice all are used. Again the number of unknowns and equations are reduced. From 10 unknowns and 10 equations to 4 unknowns and 4 to 8 equations. Both reductions are important for keeping the size of the problems manageable when solving them.

3 SOLVING POLYNOMIAL SYSTEMS

For the (3,2) case solving the system is a matter of solving a series of 1 variable 2nd degree polynomials as described above. For the (4,3) and (5,2) cases solvers based on (Byröd et al., 2009) were implemented. The technique is based on forming an expanded set of equations, by multiplying the original equations with a number of monomials, typically low order monomials up to a certain degree. All expanded equations are then expressed as a sparse coefficient matrix C times a monomial vector m , i.e. the equations are $Cm = 0$. Using numerical linear algebra it is possible to calculate the action matrix M of the linear mapping $T_{m_0} : p \mapsto pm_0$ for some monomial m_0 . The solutions to the original equations can then be calculated from the eigenvectors and eigenvalues of the action matrix M .

4 EXPERIMENTAL VALIDATION

4.1 Numerical Stability

Receivers were placed randomly in the interval $[0, 1]$ and transmitters randomly in the square $([0, 1], [0, 1])$. Figure 1, 2 and 3 shows histograms of the error residuals of recovered positions for the (3,2), (4,3) and (5,2) cases respectively. The residual is the l_2 norm between the true receiver positions and the reconstructed positions given by the minimal solvers. All

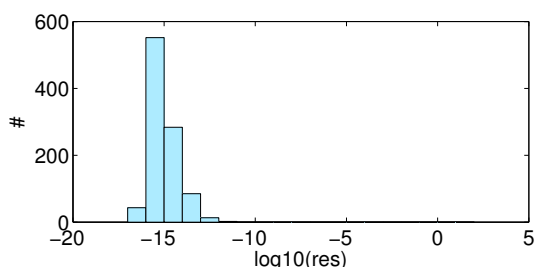


Figure 1: Residuals for the (3,2) solver.

solvers have excellent numerical performance, in particular the (3,2) solver, which is expected as it is just a series of one variable 2nd degree solvers. In a few instances the (5,2) solver gives high residuals or outright fail. This is related to a higher sensitive to both proximity of degenerate cases and due to a larger problem size being more prone to cancellation errors. A total of 1000 experiments per solver were run.

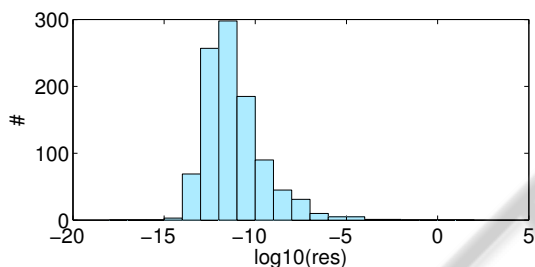


Figure 2: Residuals for the (4,3) solver.

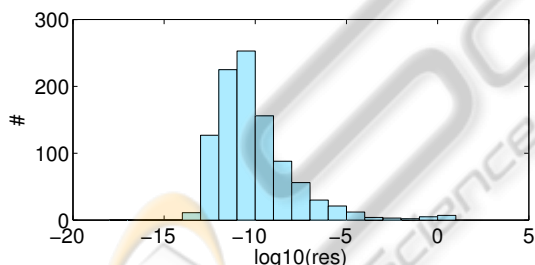


Figure 3: Residuals for the (5,2) solver.

4.2 Real Data

For the experiments with real data 8 microphones (Shure SV100) were placed on a line along wall in an office and connected to an audio interface (M-Audio Fast Track Ultra 8R), which is then connected to a computer. Relatively distinct sounds were generated by moving around in the room and clapping. The 8 synchronized channels were recorded at 44.1kHz. Signal processing was performed by a crude interest point detector on each of the eight signals. Interest points were defined as edges between periods

Table 1: Reconstructed microphone array (top) compared with ground truth (bottom) with origo omitted.

Microphone positions (m)						
0.34	0.66	1.00	1.40	1.80	2.14	5.05
0.42	0.69	1.02	1.37	1.77	2.09	4.95

with low energy and periods with high energy. Each interest point was then matched to the other seven signals using normalized cross-correlation. Thus approximately 180 hypothetical matches were found in the dataset. Among the several error sources in the setup were reflections in hard surfaces (walls, books, shelves, computer monitor), receivers not placed perfectly collinearly and non-exact estimate of the speed of sound. A RANSAC procedure using 50 iterations randomly selecting points and solver (5-2,4-3) saving the best hypothesis. Scoring here are how many additional audio signals were consistent within 2dm other than the 2 or 3 randomly selected. The final result is obtained using a bundle adjustment (non-linear least squares) on the found inlier set. Table 1 shows ground truth and reconstructed coordinates for the points on the line. Given the error sources and the fact that the microphones had a diameter of 3 cm the results are very satisfactory. No ground truth were taken for the sound sources But the spatial layout (not shown) is reasonable in regards to the proportions of the office.

5 CONCLUSIONS

In this paper we have studied, modelled and solved three important minimal cases for structure from sound assuming that e.g. the microphones are positioned on a line. For each of the case we present and publish efficient and numerically stable solvers. Such solvers could be used in RANSAC schemes to weed out the outliers in real data or be integrated in the low-level audio or radio matching schemes. In the paper we demonstrate the efficiency and numerical stability on simulated data and demonstrate a small system using low-level feature detection, matching, RANSAC and bundling, to enable automatic microphone sensor array calibration using only synchronized audio as input.

REFERENCES

Birchfield, S. T. and Subramanya, A. (2005). Microphone array position calibration by basis-point classical multidimensional scaling. *IEEE transactions on Speech and Audio Processing*, 13(5).

- Biswas, R. and Thrun, S. (2004). A passive approach to sensor network localization. In *IROS 2004*.
- Byröd, M., Josephson, K., and Åström, K. (2009). Fast and stable polynomial equation solving and its application to computer vision. *Int. Journal of Computer Vision*, 84(3):237–255.
- Chen, J. C., Hudson, R. E., and Yao, K. (2002). Maximum likelihood source localization and unknown sensor location estimation for wideband signals in the near-field. *IEEE transactions on Signal Processing*, 50.
- Elnahrawy, E., Li, X., and Martin, R. (2004). The limits of localization using signal strength. In *SECON-04*.
- Niculescu, D. and Nath, B. (2001). Ad hoc positioning system (aps). In *GLOBECOM-01*.
- Pollefeys, M. and Nister, D. (2008). Direct computation of sound and microphone locations from time-difference-of-arrival data. In *Proc. of International Conference on Acoustics, Speech and Signal Processing*.
- Raykar, V. C., Kozintsev, I. V., and Lienhart, R. (2005). Position calibration of microphones and loudspeakers in distributed computing platforms. *IEEE transactions on Speech and Audio Processing*, 13(1).
- Sallai, J., Balogh, G., Maroti, M., and Ledeczi, A. (2004). Acoustic ranging in resource-constrained sensor networks. In *eCOTS-04*.
- Stewénius, H. (2005). *Gröbner Basis Methods for Minimal Problems in Computer Vision*. PhD thesis, Lund University.
- Thrun, S. (2005). Affine structure from sound. In *Proceedings of Conference on Neural Information Processing Systems (NIPS)*, Cambridge, MA. MIT Press.
- Wendeberg, J., Hoflinger, F., Schindelbauer, C., and Reindl, L. (2011). Anchor-free tdoa self-localization. In *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on*, pages 1–10.

