ROBUST SELF-CALIBRATION OF CONSTANT OFFSET TIME-DIFFERENCE-OF-ARRIVAL

K. Batstone[1], G. Flood[1], T. Beleyur[3], V. Larsson[2], H. R. Goerlitz[3], M. Oskarsson[1], K. Åström[1]

Centre for Mathematical Sciences
Lund University, Sweden[1]

Dept. of Computer Science
ETH Zurich, Switzerland[2]

Acoustic and Functional Ecology
Max Planck Inst. for Ornithology
Seewiesen, Germany[3]

ABSTRACT

In this paper we study the problem of estimating receiver and sender positions from time-difference-of-arrival measurements, assuming an unknown constant time-difference-of-arrival offset. This problem is relevant for example for repetitive sound events. In this paper it is shown that there are three minimal cases to the problem. One of these (the five receiver, five sender problem) is of particular importance. A fast solver (with run-time under 4 μs) is given. We show how this solver can be used in robust estimation algorithms, based on RANSAC, for obtaining an initial estimate followed by local optimization using a robust error norm. The system is verified on both real and synthetic data.

Index Terms— Time-difference-of-arrival, Constant Offset, RANSAC, Minimal Problem

1. INTRODUCTION

The problem of estimating receiver-sender node positions from measured arrival times of radio or sound signals is a key issue in different applications such as microphone array calibration, radio antenna array calibration, mapping and positioning. This field is well researched but in this paper we will focus on the anchor-free sensor network calibration both in terms of time-of-arrival measurements (TOA) and time-difference-of-arrival measurements (TDOA). For time-of-arrival the planar case of three receivers and three senders (3R/3S) was solved in [1]. For the full 3D case the over-determined problem (10R/4S) was studied in [2], where a solver for this non-minimal case was provided. There are actually three minimal cases for the 3D case, namely (4R/6S), (5R/5S) and (6R/5S). A practical solver was presented in [3]. There are in general 38, 42 and 38 solutions respectively for the three different set ups. Faster solvers for these minimal cases were provided in [4].

This work was partially supported by the following funding bodies: strategic research projects ELLIIT and eSSENCE, The Sten K. Johnson foundation, Swedish Foundation for Strategic Research project - Semantic Mapping and Visual Navigation for Smart Robots - (grant no. RIT15-0038), Wallenberg Artificial Intelligence, Autonomous Systems and Software Program (WASP) funded by Knut and Alice Wallenberg Foundation, Emmy Noether research (grant no. GO2029/2-1) of the Deutsche Forschungsgemeinschaft to HRG.

In this paper we study the constant offset TDOA self-calibration problem. It is a problem that naturally arises e.g. when signals are emitted with a known period. As an estimation problem it lies between TOA and full TDOA. In the paper we study the minimal (5R/5S) problem and provide a fast (few μs) solver. Robust parameter estimation often use the hypothesize and test paradigm, e.g. using random sampling consensus, [5] or one of its many variants [6, 7, 8]. In these frameworks minimal solvers are important building blocks for generating model hypotheses, and we show in the paper how a minimal solver can be used for robust parameter estimation of sender positions, receiver positions and unknown offset. The system is capable of handling missing data, outliers and noise. The algorithms are tested on synthetic data as well as real data, in an office environment and in a cave. The methods are straightforward to generalize for degenerate configurations which arise if senders or receivers are restricted to a plane or to a line.

2. TIME-DIFFERENCE-OF-ARRIVAL SELF CALIBRATION

The problem we are considering involves m receiver positions $\mathbf{r}_i \in \mathbb{R}^3$, $i = 1, \ldots, m$, and n sender positions $\mathbf{s}_j \in \mathbb{R}^3$, $j = 1, \ldots, n$. This could for example represent the microphone positions and locations of sound emissions, respectively. Assume that the arrival time of a sound j to receiver i is t_{ij} and that the time that sound j is emitted is T_j. Multiplying the travel time $t_{ij} - T_j$ with the speed v of the signal we obtain the distance between senders and receiver,

$$v(t_{ij} - T_j) = \|\mathbf{r}_i - \mathbf{s}_j\|_2, \quad (1)$$

where $\|.\|_2$ is the l^2-norm. The speed v is throughout the paper assumed to be known and constant.

In many settings the times of emissions T_j are unknown, but regular, *e.g.*

$$T_j = k_1 j + k_0, \quad (2)$$

where the interval k_1 is known. Inserting (2) into (1) we obtain

$$v(t_{ij} - k_1 j - k_0) = \|\mathbf{r}_i - \mathbf{s}_j\|_2. \quad (3)$$

Assuming an erroneous (but regular) emission time $\tilde{T}_j = k_1 j + \tilde{k}_0$ and introducing (the measured) $z_{ij} = v(t_{ij} - \tilde{T}_j)$ and (the unknown) $o = v(k_0 - \tilde{k}_0)$ yields the following expression

$$z_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o. \tag{4}$$

Note that this is a simplified variant of the general time-difference-of-arrival problem (see *e.g.* [9]), which allows for a different offset o for every j,

$$z_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o_j. \tag{5}$$

Problem 1 *(Constant Offset Time-Difference-of-Arrival Self-Calibration) Given measurements* $\tilde{z}_{ij}$

$$\tilde{z}_{ij} = \|\mathbf{r}_i - \mathbf{s}_j\|_2 + o + \epsilon_{ij}, \tag{6}$$

for a subset $W \subset I$ *of all the receiver-sender index pairs* $I = \{(i,j) | i = 1, \dots m, j = 1, \dots, n\}$ *determine receiver positions* $\mathbf{r}_i$, $i = 1, \dots, m$ *and sender positions* $\mathbf{s}_j$, $j = 1, \dots, n$ *and offset* o*. Here the errors* ϵ_{ij} *are assumed to be either* **inliers***, in which case the errors are small (*$\epsilon_{ij} \in N(0, \sigma)$*) or* **outliers***, in which case the measurements are way off.*

Here we will use the set W_{in} for the indices (i, j) corresponding to the inlier measurements and W_{out} for the indices corresponding to the outlier set.

3. LOCAL OPTIMIZATION AND THE LOW RANK RELAXATION

If an initial estimate of the parameters $\theta_1 = \{R, S, o\}$ is given and if the set of inliers is known, then refinement of the estimate can be found by optimization methods, *e.g.* Levenberg-Marquardt (LM) [10, 11],

$$\min_{\theta_1} f(\theta_1) = \sum_{(i,j) \in W_{\text{in}}} (z_{ij} - (\|\mathbf{r}_i - \mathbf{s}_j\|_2 + o))^2. \tag{7}$$

There is an interesting relaxation to the problem, that exploits the fact that the matrix with elements $(z_{ij} - o)^2$ is rank 5, [2]. Further simplifications use the double compaction method [9]. The double compaction matrix M is defined as the matrix with elements

$$M_{ij} = (z_{ij} - o)^2 - a_i - b_j, \tag{8}$$

and it can be shown to have rank 3, i.e. $M = U^T V$, where U is of size $3 \times m$ and V is of size $3 \times n$. The relaxed problem involves a set of parameters $\theta_2 = \{U, V, b, a, o\}$. Here the constraints can be written as

$$z_{ij} = \sqrt{u_i^T v_j + a_i + b_j} + o, \tag{9}$$

where u_i denotes column i of U and v_j denotes column j of V. Refinement of parameters can be done by performing local optimization on

$$\min_{\theta_2} f(\theta_2) = \sum_{(i,j) \in W_{\text{in}}} \left(z_{ij} - (\sqrt{u_i^T v_j + a_i + b_j} + o) \right)^2. \tag{10}$$

4. MINIMAL PROBLEMS AND SOLVERS

By counting equations and unknowns, one finds that there are three minimal problems. The first two are the symmetric case when $m = 4, n = 7$ or $m = 7, n = 4$. This case is not addressed in this paper, but we believe it to be difficult to solve. The other case is $m = n = 5$. Here, we first present a solver for the constant offset and then discuss how to solve for sender and receiver positions.

Given a 5×5 matrix, Z, with time-difference-of-arrival measurements z_{ij}, the rank 3 constraint on the double compaction matrix in (8) can be written as

$$f(o) = \det(C^T (Z - o)^{\circ 2} C) = 0, \tag{11}$$

where

$$C = \begin{pmatrix} -1 & -1 & -1 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{12}$$

and $^{\circ 2}$ denotes element-wise squaring (Hadamard power). Although the elements of $(Z - o)^{\circ 2}$ are of degree 2 in o, the quadratic terms cancel out after multiplication with C^T and C. Thus the elements of $C^T (Z - o)^{\circ 2} C$ are linear in o. Since the determinant is linear in each column, the determinant $f(o)$ is a polynomial of degree four in the offset o. This can be summarized as

Theorem 1 *Given time-difference-of-arrival measurements from five receivers to five senders, there are four possible offsets* o*, given as the roots to the fourth degree polynomial* $f(o)$*, counting complex roots and multiplicity of roots.*

For each solution o it is possible to generate a solution θ_2 to the relaxed problem, according to

$$b = \left((z_{11} - o)^2 \;\; (z_{12} - o)^2 \;\; (z_{13} - o)^2 \;\; (z_{14} - o)^2 \;\; (z_{15} - o)^2 \right),$$

$$a = \begin{pmatrix} 0 \\ (z_{21} - o)^2 - (z_{11} - o)^2 \\ (z_{31} - o)^2 - (z_{11} - o)^2 \\ (z_{41} - o)^2 - (z_{11} - o)^2 \\ (z_{51} - o)^2 - (z_{11} - o)^2 \end{pmatrix}, \tag{13}$$

$$U = \begin{pmatrix} 0 & u_2 & u_3 & u_4 & u_5 \end{pmatrix}, \tag{14}$$

$$V = \begin{pmatrix} 0 & v_2 & v_3 & v_4 & v_5 \end{pmatrix}, \tag{15}$$

where $\begin{pmatrix} u_2 & u_3 & u_4 & u_5 \end{pmatrix}^T \begin{pmatrix} v_2 & v_3 & v_4 & v_5 \end{pmatrix}$ is any rank 3 factorization of the matrix $C^T (Z - o)^{\circ 2} C$.

From a solution θ_2 to the relaxed problem it is possible to upgrade to a solution θ_1 to the original problem. This involves solving a system of polynomial equations. The procedure was first described in [3], where an algorithm for solving this was presented. Recently, a faster algorithm was presented in [4].

Table 1. Execution times for 5×5 minimal solvers steps. Notice that the steps of calculating o and the relaxed solution is significantly faster than upgrading to the full solution

Implementation	Matlab	C++
Calculation of o	$38\,\mu s$	$3.7\,\mu s$
Calculation of $\theta_2 = \{U, V, a, b, o\}$	$100\,\mu s$	N/A
Calculation of $\theta_1 = \{R, S, o\}$	$600\,ms$	$22\,ms$

An efficient implementation for calculating the four solutions of the offset o given the measurements z takes $4\ \mu s$ for a C++-implementation. Generating the solution θ_2 to the relaxed problem adds a few μs. However, calculating a solution θ_1 to the original problem takes another $22\ ms$. Thus, it is advantageous to estimate the parameters of the relaxed problem and postpone the upgrade from θ_2 to θ_1 as a final step, see Table 1.

5. USING RANSAC FOR FIVE ROWS

We propose the use of the fast minimal solver in an hypothesize and test framework to obtain (i) a initial estimate on the offset o and (ii) an initial inlier set. The steps are described in Algorithm 1

Algorithm 1 Offset RANSAC

1: Randomly select 5 rows and columns. Find the four solutions on o given the time-difference-of-arrival measurements.
2: For each solution o, calculate the relaxed solution $\theta_2 = \{U, V, a, b, o\}$.
3: For selected rows and for each remaining column, check for inliers according to the residuals in (10).

6. ROBUST ESTIMATION OF PARAMETERS

We use these minimal solvers with RANSAC as described in the previous section to find one or several initial estimates of the parameters θ_2 for a subset of five receivers and k senders. The solution is extended to additional rows and/or columns using robust techniques as described in [12]. During this process it is useful to keep the errors down by occasionally refining the solutions using local optimization. This has shown to reduce failures, see e.g. [13, 14]. In the proposed estimation algorithm we postpone the upgrade from θ_2 to θ_1 until we have found a good solution involving a large portion of the receiver and sender positions.

7. EXPERIMENTAL VALIDATION

7.1. Minimal Solver

To test the numerical accuracy and robustness of our minimal solver we conducted an experiment using simulated data

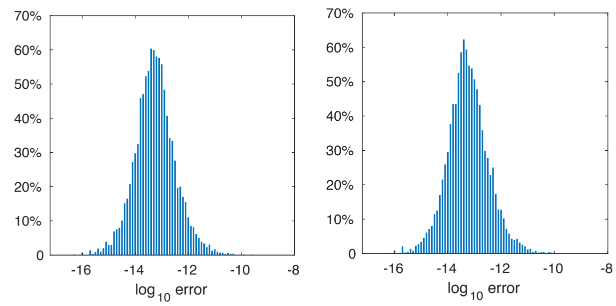

Fig. 1. Left shows the histogram of the logarithm of the absolute errors, for the Matlab implementation of our minimal solver. To the right the corresponding histogram for the C++ implementation.

without noise. We generated a large number of instance problems (10,000) with known offsets. We then ran our solvers and compared the returned solutions with the ground truth solution. For each instance problem we recorded the distance to the closest solution. In Figure 1 the resulting histogram of the logarithm of the absolute errors are shown. As can be seen, both implementations get close to machine precision.

7.2. Experimental Setup for Real Data

We have tested our system on (i) experiments made in an office environment and (ii) experiments made at the Orlova Chuka cave, Bulgaria.

For the office experiments, 12 microphones (8x t.bone MM-1, 4x Shure SV100) were positioned around a room ($\sim 3 \times 5\ m^2$) and measured using a laser to obtain ground truth positions of the microphones with an error of $\pm 2\ mm$. The space was cleared of most the furniture to create an open space to conduct the experiment in. The sound recordings were captured using a Roland UA-1610 Sound Capture audio interface and automatically amplified. The recordings were made using the open source software Audacity 2.3.0 with a sampling frequency of 96 kHz on a laptop. A synthetically generated chirp was then played using a simple loudspeaker every half second for 30 s while moving the speaker around in the room.

For the cave experiments, 12 microphones (4x Sanken CO-100K, 8x Knowles SPU0410) were positioned in a section of the cave, four microphones were placed on an inverted T array near one wall, while the other eight microphones were placed on the adjacent wall. The sound recordings were captured using pre-amplifiers (Quadmic, RME) and two synchronised Fireface 800 (RME) audio interfaces running at a sampling frequency of 192 kHz. Recording and playback were controlled via a custom written script based on the sound device library [15] in Python 2.7.12 [16]. Ultrasonic chirps ($8\ ms$, $16 - 96\ kHz$ upward hyperbolic sweep) were played every second via one of the audio interfaces, amplified (Basetech AP-2100) and presented through a Peer-

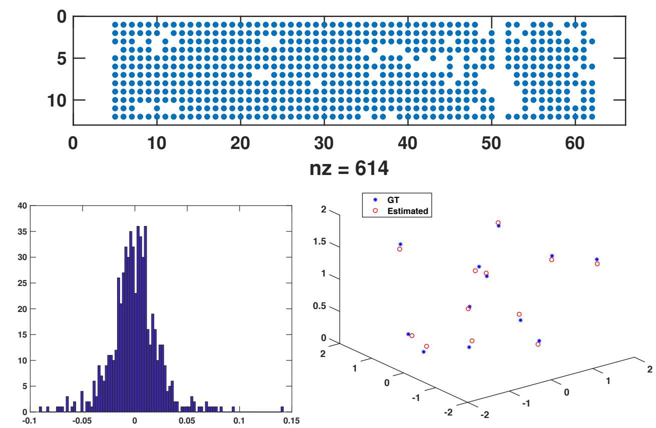

Fig. 2. For the office experiment the figure shows detected inliers W_{in} (top), inlier residual histogram (bottom left), and estimated and ground truth microphone positions (bottom right).

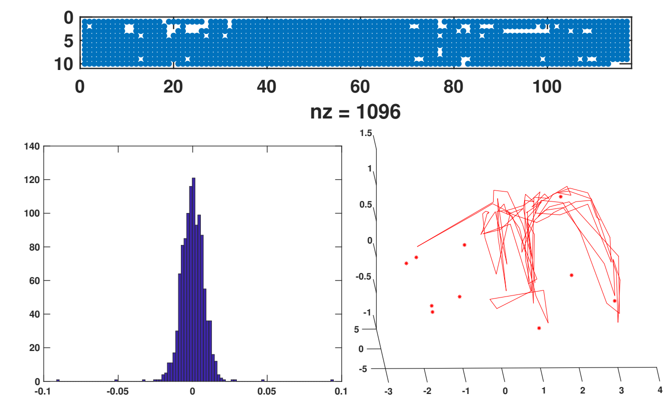

Fig. 3. For the cave experiment the figure shows detected inliers W_{in} (top), inlier residual histogram (bottom left) and estimated microphone and sound source positions, red dots and line respectively (bottom right).

less XT25SC90-04 loudspeaker. The speaker was attached to a 3-m-long pole and slowly waved in the approximately $5 \times 9 \times 3\ m^3$ recording volume. Playbacks were done past 6:00 am to prevent disturbing the resident bat population.

7.3. Experimental Evaluation for Real Data

Once the office recordings were taken, an algorithm was used to find the chirps in the captured sound recordings and the algorithm then outputs the z_{ij} matrix. This can then be used in our RANSAC scheme, Algorithm 1. For this experiment we used the (5R/5S) minimal solver. A fixed number of iterations was used; 100 iterations for the initial selection of 5 receivers and senders, then the extension to more columns and rows was allowed until there was no better solution. The tolerance was set to $T = 0.01$ for the initial selection and extension of rows and column.

Once the initial values have been estimated, it underwent l^2 optimization on the inlier set. The results of the estimated microphone positions after the optimization are shown in Figure 2.

This produced an Euclidean distance error between each of the microphones calculated position and its ground truth position as $(0.2016, 0.0587, 0.1444, 0.1153, 0.2017, 0.1326, 0.1407, 0.1198, 0.2041, 0.2010, 0.1908, 0.2110)\ m$.

For graphical purposes, a Procrustes fitting was used on the microphone positions to spread the total error over all 12 microphones. In the Procrustes fitting only rotation and translation were allowed.

For the cave experiment a similar scheme was devised and the results are shown in Figure 3.

8. CONCLUSIONS

In this paper, a novel method has been constructed to efficiently solve a TDOA problem with a constant offset. This has been verified using simulated data to test the solver and real experimental data to test our algorithms in realistic scenarios.

Looking at Figure 1 and Table 1, it can be seen that the calculation of the offsets and the calculation of the relaxed form θ_2 are very fast solvers without loss in numerical accuracy. The advantage of this is that when using a RANSAC approach, the iterations are performed quickly, giving a good initial estimate in which to optimize over, which is important in highly non-linear systems such as this.

Looking at the results from the office experiment, Figure 2, we can see that the calculated microphone positions are accurate and the residuals are small, mostly in the range $\pm 0.04\ m$. Further to this our inlier set appears to be accurate. The first and last few columns (corresponding to sound emissions) are not used in our initialisation. This is correct because the recording started before the chirps were sounded and ended after, so the chirp detection algorithm falsely determined that they were also chirps but our method decided that the data in those regions do not fit the model. A comparison of the calculated microphone positions were made to a solution from a Full TDOA system, [9], which produced similar results and very similar residuals. This provided a sanity check that the chirp detection was working correctly and that from this dataset a better solution could not be found.

For the cave experiment, similar conclusions can be made, since the residuals are very low, we can conclude that we have an accurate model. This gives a real life example of how algorithms such as the one proposed can be used.

For future work, the study of the number of inliers could be of use. At the moment our algorithm may not extend to more rows and columns if the initial solution is poor, perturbing our final solution. Perhaps a method which could adapt the initial selection in order to give a required amount of inliers could be more advantageous.

9. REFERENCES

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

[2] 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.

[3] Yubin Kuang, Simon Burgess, Anna Torstensson, and Kalle Åström, "A complete characterization and solution to the microphone position self-calibration problem," in *The 38th International Conference on Acoustics, Speech, and Signal Processing*, 2013.

[4] Viktor Larsson, Kalle Åström, and Magnus Oskarsson, "Polynomial solvers for saturated ideals," in *International Conference on Computer Vision (ICCV)*. IEEE, 2017.

[5] 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.

[6] Ondřej Chum, Jiří Matas, and Josef Kittler, "Locally optimized ransac," in *Joint Pattern Recognition Symposium*. Springer, 2003, pp. 236–243.

[7] Rahul Raguram, Ondrej Chum, Marc Pollefeys, Jiri Matas, and Jan-Michael Frahm, "Usac: a universal framework for random sample consensus.," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 35, no. 8, pp. 2022–2038, 2013.

[8] Simon Korman and Roee Litman, "Latent ransac," in *Proc. Conf. Computer Vision and Pattern Recognition*, 2018, pp. 6693–6702.

[9] Yubin Kuang and Kalle Astrom, "Stratified sensor network self-calibration from tdoa measurements," in *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21st European*. IEEE, 2013, pp. 1–5.

[10] Kenneth Levenberg, "A method for the solution of certain non-linear problems in least squares," *Quarterly of applied mathematics*, vol. 2, no. 2, pp. 164–168, 1944.

[11] Donald W Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *Journal of the society for Industrial and Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[12] Kenneth Batstone, Magnus Oskarsson, and Kalle Åström, "Robust time-of-arrival self calibration with missing data and outliers," in *Signal Processing Conference (EUSIPCO), 2016 24th European*. IEEE, 2016, pp. 2370–2374.

[13] C. Engels, H. Stewenius, and D. Nister, "Bundle adjustment rules," in *PCV06*, 2006.

[14] Georg Klein and David Murray, "Parallel tracking and mapping for small ar workspaces," in *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*. IEEE, 2007, pp. 225–234.

[15] M Geier, "sounddevice 0.3.5," `https://python-sounddevice.readthedocs.io/en/0.3.5/_modules/sounddevice.html`, 2015.

[16] Guido Van Rossum and Fred L Drake Jr, *Python reference manual*, Centrum voor Wiskunde en Informatica Amsterdam, 1995.