Pose Estimation from Minimal Dual-Receiver Configurations

Simon Burgess, Yubin Kuang and Kalle Åström
Centre for Mathematical Sciences, Lund University, Lund, Sweden
{*simonb, yubin, kalle*}*@maths.lth.se*

Abstract

Using multiple receivers (microphones or antennas) in a rigid configuration, such as on a smartphone, it is possible to measure time difference of arrival to the receivers. This in turn can be used to determine the direction to the transmissions, if there are at least three receivers. When using two receivers it can be used to determine the angle to the transmissions relative to the line through the two receivers. In this paper we study three minimal problems for pose using such data: (i) determine position and orientation using five transmissions, (ii) determine position and orientation using four transmissions and known 'down' direction and (iii) determine position using three transmissions and known orientation. Numerically stable solvers are implemented. An experimental validation of the solvers are performed on simulated data.

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 technique has been used in warfare to determine hostile artillery battery, in crime fighting in cities to determine location of gun shots. Similar techniques is used today to with microphone arrays to enable beamforming and speaker tracking (smartrooms, surveillance and event recording). 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 a sound source using a calibrated detector array [2, 4, 5, 6], the problem of determining pose of the sensor array to known sources has received much less attention.

Using multiple receivers (microphones or antennas) in a rigid configuration, such as on a smartphone, it is possible to measure time difference of arrival to the receivers. This in turn can be used to determine the direction to the transmissions, if there are at least three receivers. When using two receivers it can be used to determine the angle to the transmissions relative to the line through the two receivers.

In this paper we study three minimal problems for pose using such data: (i) determine position and orientation using five transmissions, (ii) determine position and orientation using four transmissions and known 'down' direction and (iii) determine position using three transmissions and known orientation. Numerically stable solvers are implemented. Such solvers can 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. An experimental validation of the solvers are performed on simulated data.

2. The antenna array pose problem

Traditionally sound ranging was used in the world war I as a method for locating enemy artillery from time difference of arrival of sound to a number of receivers in known location. The inverse problem to locate the receivers using known position of the sound or radio sources is becoming increasingly useful for instance in mobile phone localization. When exploiting time difference of arrival data using just one receiver does not give any information. In this paper we investigate the limits of what can be done in the case of two receivers.

We assume that the two sound or radio receivers are placed relatively close to each other in relation to the sound or radio transmitters. Thus it is from the receiver end possible to use a far-field approach. We assume that for each transmission it is possible to measure the time difference of arrival. This is feasible by sound matching or phase difference for radio, if the antennas are less than half a wavelength from each other. The TDOA measurements are expressed in distance by multiplying with the speed of sound or light, depending on the type of transmission. If this distance difference of the signal is d and the distance between the two receivers are Δ, then a fair approximation is to say that the vector $U_i - C$ between the receiver-pair ,with centroid $C = [x\ y\ z]^T$, and the transmitter location at $U_i = [u_{x,i}\ u_{y,i}\ u_{z,i}]^T$, has an angle relative to the vector $n = [n_x\ n_y\ n_z]^T$ between the transmitter pair according to

$$n \cdot (U_i - C) = \frac{d}{\Delta} |U_i - C| \, |n| = cos(\Theta_i) |U_i - C| \, |n| \, .$$

See Figure 1 for an illustration. For all practical purposes we henceforward assume that the experimental setup gives us the angle measurements Θ_i. Note that since the relative distance d can be measured with sign it is possible to get the $\cos(\Theta_i)$ with sign. Notice also that for planar problems each measurement gives us the direction in the plane (up to a two-fold ambiguity), whereas in 3D the transmitter position is restricted to lie on a cone. Finally notice that for the 3D problem one could also obtain the direction in 3D to the transmitter using more than 2 receivers.

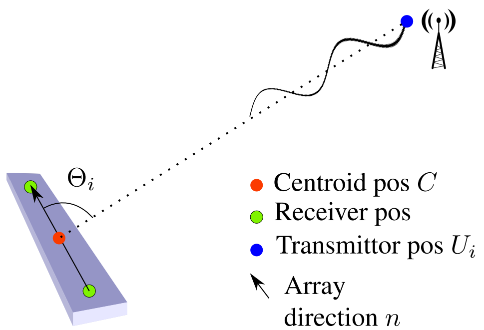

Figure 1. Two receivers with centroid C and direction n measuring angles Θ_i from transmitters U_i.

In a more general setting it is interesting to study the question: Given l transmissions from at known positions U_i and time-difference of arrival data to k receivers in a fixed pattern on a mobile unit, try to estimate the position and orientation of the unit.

For 3D problems when $k \geq 3$, one obtains directions in 3D. It is well known from computer vision that this problem is well posed when you have at $l \geq 3$ direction measurements. For the minimal case of 3 directions there are up to 4 solutions to the pose problem [8, 9].

For planar problems $k = 2$ suffice. It is also here well known that given 3 directions you can with linear techniques find the unique solution [1].

Thus the unsolved problems are 3D problems with only 2 receivers. In this paper we would like to characterize when it is possible to solve for position C and orientation n of the receiver pair using such angular measurements Θ_i

$$n \cdot (U_i - C) = cos(\Theta_i) |U_i - C| \, |n| \, . \tag{1}$$

In particular we will treat the following three problems.

1. Estimation of position and orientation from 5 transmitters
2. Estimation of position and orientation from 4 transmitters and known down direction
3. Estimation of position and orientation from 3 transmitters and known orientation

As will be shown all of these three problem formulations are minimal problems in the sense that there is precisely enough information to solve for the unknowns. In all of the cases there is typically a finite number of solutions.

3 Reformulation as polynomial equations

We here formally state the pose problems and reformulate them as solving sets of polynomial equations.

3.1 5 point pose polynomial problem

Problem 3.1 *Given a stationary antenna array of two antennas, base stations with known locations* $U_i = [u_{x,i}\ u_{y,i}\ u_{z,i}]^T, i = 1, \ldots, 5$ *and angle measurements* Θ_i *from the array direction* n *to each of the base stations, reconstruct the position* $C = [c_x\ c_y\ c_z]^T$ *and the direction* $n = [n_x\ n_y\ n_z]^T$ *of the antenna array so that* (1) *are fulfilled.*

Squaring (1) we get

$$\begin{aligned} (U_i - C)^T nn^T (U_i - C) = \\ \cos \Theta_i{}^2 (U_i - C)^T (U_i - C) n^T n. \end{aligned} \tag{2}$$

These are **fourth degree** polynomials in the six unknowns n and C. By either adding the equation

$$|n|^2 = 1 \tag{3}$$

or parametrizing the normal vector according to

$$n = [n_x\ n_y\ 1]^T \tag{4}$$

we get a problem that might have finite number of soultions. We choose to use representation (4) for n, as we have already lost the sign of n by squaring the equations, giving five unknowns.

In [10] it is shown that, under rather general settings, an upper bound of the number of solutions to a set of polynomial equations can be obtained by solving a single instance of the equations. Degenerate configurations with larger number of solutions can still exist, but with random coefficienst the probability of ending up in these degenerate cases are small. Relying on this, we run a solution over $\mathbb{Z}^p$, where p is a large prime number, in Macaulay2 [7]. We find that there are 24 solutions.

3.2 4 point pose with known 'down' direction

Problem 3.2 *Given a stationary antenna array of two antennas, base stations with known locations* $U_i = [u_{x,i}\ u_{y,i}\ u_{z,i}]^T, i = 1, \ldots, 4$ *and angle measurements* Θ_i *from the array direction* n *to each of the four base stations, and given the angle* Θ_d *from* n *to the 'down' direction* n_d*, e.g. from and accelerometer, reconstruct the position* $C = [c_x\ c_y\ c_z]^T$ *and the direction* $n = [n_x\ n_y\ n_z]^T$ *of the antenna array.*

By squaring (1) and taking the scalar product of n and n_d we get

$$\begin{gathered}(U_i - C)^T nn^T (U_i - C) = \\ \cos\Theta_i{}^2 (U_i - C)^T (U_i - C)\, n^T n,\end{gathered} \tag{5}$$

$$n^T n_d = \cos\Theta_d\, |n_d|\, |n|\,. \tag{6}$$

By assuming that $|n| = |n_d| = 1$ and rotating the coordinate system so that the downward direction $n_d = [0\ 0\ -1]^T$ we can simplify (5) and (6) to

$$\begin{gathered}(U_i - C)^T nn^T (U_i - C) = \\ \cos\Theta_i{}^2 (U_i - C)^T (U_i - C)\,,\end{gathered} \tag{7}$$

$$n_z = -\cos\Theta_d. \tag{8}$$

As we have now solved the problem for n_z using n_d and Θ_d, we are left with (7) four **fourth degree** polynomials in five unknowns in n_x, n_y and C. By adding (3) we get a polynomial problem that might have finitely many solutions. Again relying on the results in [10], running a solution over $\mathbb{Z}^p$, where p is a large prime number, in Macaulay2 [7] we find that there are 24 solutions.

3.3 3 point pose with known rotation

Problem 3.3 *Given a stationary antenna array of two antennas, base stations with known locations* $U_i = [u_{x,i}\ u_{y,i}\ u_{z,i}]^T, i = 1, \ldots, 3$ *and angle measurements* Θ_i *from the array direction* n *to each of the three base stations, and given known orientation of the unit* n*, e.g. from and accelerometer, reconstruct the position* $C = [c_x\ c_y\ c_z]^T$.

Squaring (1) we get

$$\begin{gathered}(U_i - C)^T nn^T (U_i - C) = \\ \cos\Theta_i{}^2 (U_i - C)^T (U_i - C)\, n^T n.\end{gathered} \tag{9}$$

These are **second degree** polynomial in three unknowns in C. Thus we obtain three quadratic equations in three unknown, a problem that typically has up to eight solutions. Running a solution over $\mathbb{Z}^p$, where p is a large prime number, in Macaulay2 [7] we find that there are 4 solutions.

4. Solving Polynomial Systems

For all of the three cases we implement numerically stable and efficient solvers based on the theory developed in [3], where a more extensive review ot the method can be found. 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 $Cm = 0$.

Let $\mathbb{C}[\mathbf{x}]$ be the set of all polynomials over $\mathbb{C}$ for the unknowns $\mathbf{x}$, and I is the polynomial ideal generated by the equations $Cm = 0$. For the set of equations $Cm = 0$ where the solutions are finite, the quotient space $\mathbb{C}[\mathbf{x}]/\mathbf{I}$ is finite dimensional. If $p \in \mathbb{C}[\mathbf{x}]$, let $[p]$ be the corresponding equivalence class in $\mathbb{C}[\mathbf{x}]/\mathbf{I}$.

Using numerical linear algebra on $Cm = 0$ 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.

The computational complexity of the algorithms can be characterized by the size of the sparse coefficient matrix C and the size of the action matrix M.

For the 3 point pose problem our implementation involves a coefficient matrix C of size 28×20 and an action matrix M of size 4×4.

For the 4 point pose problem our implementation involves a coefficient matrix C of size 1260×1436 and an action matrix M of size 24×24.

For the 5 point pose problem our implementation involves a coefficient matrix C of size 2625×2352 and an action matrix M of size 24×24.

5 Experimental validation

Simlulated experiements were run for the 3, 4 and 5 point polynomial pose problems. For all experiments, ground truth were constructed as follows: The centroid C was drawn from a multivariate normal distribution $\mathbf{N}([0\ 0\ 0]^T, 10I)$ where I is the identity matrix. Transmitters U_i were drawn from i.i.d. random variables uniformly distributed on spheres around C with radius $r_i \sim U(0, 100)$. The array direction n was drawn from a uniform distribution over the unit sphere. Angle measurements Θ_i are then calculated from ground truth.

The relative error for each experiment is defined as $\min_i ||x_{gt} - x_i||_2/||x_{gt}||_2$ where x_{gt} is a vector of the ground truth and x_i is a vector of the reconstructed solution i. For the 3 point case $x_{gt} = [C]$ and for the four and five point case $x_{gt} = [C\ ; n]$. In Figure 2 histograms of relative errors are shown for 500 experiemnts each for the 3, 4 and 5 point problem. That the 4 point problem tends to be more erroneous than the 5 point problem is due to the larger expanded set of equations used in the 5 point case, 2625 equations compared to 1260 equations in the 4 point case. The error for the 4-point case can be reduced by further expanding the set of equations, but at the cost of speed.

Mean execution time on a standard desktop computer 0.11s, 0.61s and 1.2s for the 3, 4 and 5 point solver respectively.

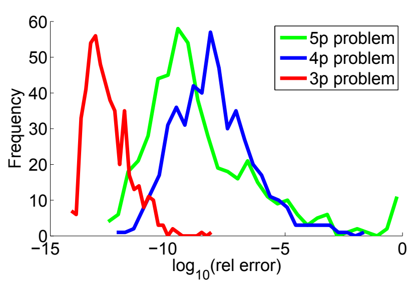

Figure 2. Histograms of relative errors

6 Conclusion

In this paper we study, model and solve three important minimal cases for pose determination of a configuration of two receivers using time difference of arrival or phase difference from transmitters in known location. For each of the minmial cases we present results from efficient and numerically stable solvers. Such solvers can 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.

7 Acknowledgements

The research leading to these results has received funding from the swedish strategic research project ELLIIT, the Swedish Strategic Foundation funded project ”Wearable Visual Systems” and the Swedish Research Council project ”Polynomial Equations and Geometry in Computer Vision”.

References

[1] K. Åström and M. Oskarsson. Solutions and ambiguities of the structure and motion problem for 1d retinal vision. *Journal of Mathematical Imaging and Vision*, 12(2):121–135, 2000.

[2] M. Brandstein, J. Adcock, and H. Silverman. A closed-form location estimator for use with room environment microphone arrays. *Speech and Audio Processing, IEEE Transactions on*, 5(1):45 –50, jan 1997.

[3] M. Byröd, K. Josephson, and K. Åström. Fast and stable polynomial equation solving and its application to computer vision. *Int. Journal of Computer Vision*, 84(3):237–255, 2009.

[4] A. Cirillo, R. Parisi, and A. Uncini. Sound mapping in reverberant rooms by a robust direct method. In *Acoustics, Speech and Signal Processing, 2008. ICASSP 2008. IEEE International Conference on*, pages 285 –288, 31 2008-april 4 2008.

[5] M. Cobos, A. Marti, and J. Lopez. A modified srp-phat functional for robust real-time sound source localization with scalable spatial sampling. *Signal Processing Letters, IEEE*, 18(1):71 –74, jan. 2011.

[6] H. Do, H. Silverman, and Y. Yu. A real-time srp-phat source location implementation using stochastic region contraction(src) on a large-aperture microphone array. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, volume 1, pages I–121 –I–124, april 2007.

[7] D. Grayson and M. Stillman. Macaulay 2. Available at http://www.math.uiuc.edu/Macaulay2/, 1993-2002. An open source computer algebra software.

[8] J. A. Grunert. Das pothenot'sche problem, in erweiterter gestalt; nebst bemerkungen über seine anwendung in der geodäsie. *Archiv der Mathematik und Physik*, 1:238–248, 1841.

[9] R. M. Haralick, C. Lee, K. Ottenberg, and M. Nölle. Analysis and solutions for the three point perspective pose estimation problem. In *Proc. Conf. Computer Vision and Pattern Recognition*, pages 592–598, 1991.

[10] R. Holt, R. Huang, and A. Netravali. Algebraic methods for image processing and computer vision. *IEEE Transactions on Image Processing*, 5:976–986, 1996.