Error analysis on source localization in ad-hoc wireless microphone networks

Wouter van Herpen  
Department of Electrical Engineering  
Eindhoven University of Technology  
Email: w.w.m.v.herpen@student.tue.nl

Sriram Srinivasan  
Digital Signal Processing Group  
Philips Research Eindhoven  
Email: sriram.srinivasan@philips.com

Piet Sommen  
Department of Electrical Engineering  
Eindhoven University of Technology  
Email: p.c.w.sommen@tue.nl

Abstract—As mobile devices with microphones (e.g. mobile phones, laptops) become increasingly common, it is possible to create an ad-hoc wireless network of such devices to make audio conferencing more mobile and practical. Such a network could then be used for efficient speech capture and source localization. However, the ad-hoc nature of the network introduces several errors, e.g., inaccurate clock synchronization, errors in estimating the sensor location, and constraints on the amount of bits available for wireless data transmission. In this paper, a framework is established to analyze the impact of such errors on source localization in the audio conferencing application.

I. INTRODUCTION

Microphone arrays have been used successfully for several applications such as speech capture in noisy environments, dereverberation, and source localization [1][2][3]. A majority of microphone array processing schemes assume that all elements in the array are wired to a single processing unit. A relevant trend is that mobile devices equipped with microphones and loudspeakers, e.g., mobile phones, PDA’s and laptops are becoming increasingly common. By forming an ad-hoc wireless sensor network using these devices, an array of microphones can be realized. A particularly relevant application is audio conferencing, using the available mobile devices of the participants instead of requiring a dedicated device.

In this paper, we focus on employing such an ad-hoc sensor network for the purpose of source localization in an audio conferencing application, to locate the position of the speaker who is currently active. Knowledge of this position could be used for instance to automatically focus a camera on the person who is speaking. While source localization using traditional wired microphone arrays has been extensively studied [1], new challenges arise when considering ad-hoc wireless sensor networks.

A commonly used method for source localization is based on Time-Differences-of-Arrival (TDOA) between the acoustic source and several microphones [4]. Energy based localization methods have also been proposed for ad-hoc arrays [5], but as the audio signals will anyway be available at the fusion center in the audio conferencing set-up, e.g., for beamforming, the more accurate TDOA based approaches are relevant here. The TDOA between pairs of microphones, together with knowledge of the location of the microphones, can be used to obtain an estimate of source location. While the microphone positions are known in a fixed array, these need to be estimated in the new set-up, which is prone to errors. TDOA estimation is also subject to errors arising from improper clock synchronization between the independent devices. Furthermore, TDOA estimation may also be affected by constraints on the number of bits that can be exchanged over the wireless link, due to the limited battery life of the mobile devices.

In this paper, for the particular application of source localization in audio conferencing, a framework is established to analyze the impact of the errors mentioned above. In a constrained resource setting, such a framework can be used to determine which error-handling requires more resources to obtain optimal performance. Furthermore, given a certain desired level of accuracy in source localization and a set of constraints on certain sensor parameters, the developed framework can be used to specify the minimum amount of needed precision in the remaining free parameters. In this way, different parameter precision trade-offs that obtain the desired level of accuracy can be realized, depending on application-specific constraints. We note that there are other sources of error, e.g., different sampling frequencies, unknown number of sensors, etc. While such errors can be handled through the exchange of control information, estimation of TDOA and sensor location is fundamental to source localization, and thus we restrict our attention in this paper to errors in these parameters.

This paper is organized as follows. In Section II, the source localization method used in the analysis is explained. Section III describes the methodology used to characterize and quantify the impact of errors on the performance of the source localization algorithm. In Section IV, the simulations and results are discussed. Finally, conclusions and directions for future work are given in Section V.

II. TDOA BASED SOURCE LOCALIZATION

Source localization can be performed using estimates of TDOA [4], which in turn can be estimated through correlation operations among sensors [6]. In this paper, the source localization problem is considered to be a 2-D problem, however, the method described in this section could be adapted to 3-D. Let \( r_i = [x_i, y_i]^T \) denote the position of sensor \( i \) in the Cartesian coordinate system, and let \( r_s = [x_s, y_s]^T \) denote the position of the source. Without loss of generality, \( i = 1 \).
In order to find the reference sensor for the computation of all differential time-delays, and will also serve as the origin of the coordinate system. Let \( v \) be the speed of sound in air. Then, the TDOA between the reference sensor and sensor \( i \) is given by

\[
t_{i1} = \frac{\| r_s - r_1 \|}{v} - \frac{\| r_s - r_i \|}{v}, \quad i = 2 \ldots N,
\]

where \( t_{i1} \) is the time taken by a signal traveling from the source to sensor \( i \), \( t_{11} = t_i - t_1 \), and \( \| \cdot \| \) is the \( l^2 \) norm.

All the equations from all \( N \) TDOA pairs can be combined into the following linear system [4]:

\[
A \theta = b,
\]

where

\[
A = \begin{bmatrix}
x_2 & y_2 & vt_{21} \\
x_3 & y_3 & vt_{31} \\
\vdots & \vdots & \vdots \\
x_N & y_N & vt_{N1}
\end{bmatrix},
\]

\[
b = \frac{1}{2} \begin{bmatrix}
x_2^2 + y_2^2 - (vt_{21})^2 \\
x_3^2 + y_3^2 - (vt_{31})^2 \\
\vdots \\
x_N^2 + y_N^2 - (vt_{N1})^2
\end{bmatrix},
\]

\[
\theta = \begin{bmatrix}
x_s \\
y_s \\
R_1
\end{bmatrix},
\]

with the constraint \( R_1 = \sqrt{x_s^2 + y_s^2} \). The constraint can be formulated as

\[
\theta^T \Sigma \theta = 0,
\]

where \( \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \). This constrained linear system can be solved through the method of Lagrange multipliers, which results in

\[
\theta = (A^T A + \lambda \Sigma)^{-1} A^T b. \tag{7}
\]

In order to find \( \lambda \), we can impose the quadratic constraint directly by substituting (7) into (6). This results in a fourth order equation where \( \lambda \) can be found by searching for a root around zero [4].

This approach depends on a system of three equations with three unknowns. Since a pair of sensors is needed to generate an equation, and one sensor is used as reference point for all time-delays, a minimum of four sensors is needed in this approach. The above framework allows us to introduce errors in estimates of sensor locations and TDOA pairs, and study the effect on the estimation of the source location.

### III. Error Analysis

Source localization is sensitive to errors in TDOA estimation, errors in sensor location estimation and errors in the estimation of the speed of propagation. This paper focuses on the first two types of errors. The speed of propagation, although dependent on temperature, is assumed to be known and set to \( v = 343 \text{ m/s} \). In this section, we describe the measure used to quantify performance, the array geometry considered, and finally the methodology adopted to analyze the impact of errors in TDOA and sensor location estimation on source localization.

#### A. Performance measure

To quantify the impact of these errors on source localization, the following measure is adopted:

\[
e_{sd} = ||r_s - r_a||,
\]

where \( r_a = [x_a, y_a]^T \) is the estimated source location, and \( e_{sd} \) is the error in meters.

#### B. Array geometry

The performance of TDOA-based source localization depends on the geometry of the sensor array, and the position of the speaker. We assume circular microphone arrays, which in the audio conferencing setup corresponds to a group of people sitting around a table. The diameter of the array is assumed to be 1 m. In this paper, the source location will be assumed to be in the same location as a sensor, corresponding to the physical situation of a speaker sitting close to a mobile audio-enabled device.

#### C. TDOA errors

In this paper, errors in TDOA are modeled by adding or subtracting a certain value from the true TDOA value in equations (3) and (4). We consider two sources of TDOA estimation errors. One is due to inaccurate clock synchronization between the different sensors. In this case, the error in source localization is studied as the error in TDOA estimation is varied. In the second case, the TDOA estimation error is due to the fact that a quantized version of the sensor signal is used to estimate the TDOA. Here, it is assumed that sensor \( i, 2 \leq i \leq N \) transmits its signal at a rate \( R \) to the reference sensor, where the TDOA \( t_{i1} \) is computed. Assuming flat signal and noise power spectral densities (PSD), the Cramer-Rao lower bound (CRLB) on TDOA estimation in this case is given as [7]:

\[
CRLB = \frac{1}{\epsilon} \left( \frac{\Delta \omega T}{2\pi} \right) \left[ \frac{1}{1 + \frac{\Delta \omega^2}{\omega_0^2}} \right],
\]

where

\[
\epsilon = \frac{\rho^2}{(1 + \rho^2)^2} - \rho^2, \tag{10}
\]

\[
\rho = \frac{\Phi_s}{\Phi_n} \text{ where } \Phi_s \text{ is the PSD of the desired acoustic signal and } \Phi_n \text{ is the noise PSD, } T \text{ is the observation duration, } \omega_0 \text{ is the center frequency of the acoustic signal used in the localization, and } \Delta \omega \text{ is its bandwidth. As the bit-rate } R \text{ is varied, the corresponding error in TDOA estimation is modeled by adding the square root of the CRLB at rate } R \text{ to the true TDOA values in equations (3) and (4).}
D. Sensor location errors

As the positions of the sensors in an ad-hoc network will not be known a priori, they have to be estimated, e.g., using acoustic localization schemes [8]. Modelling the errors in sensor location is not straightforward. An erroneous estimation of a sensor location in the x-direction will not necessarily have the same effect as the same amount of error in y-direction. Therefore, when studying an error in sensor location, the erroneous location is assumed to be on a random point on a circle around the actual location, as depicted in Figure 1. The actual sensor locations are depicted as squares. In one realization of the simulation, the erroneous location of the sensor, in this case the upper right sensor, is assumed to be one point on the circle with a radius \( r \), and the corresponding estimate of the source location is obtained using (3) and (4). Results are averaged over 10000 such realizations such that a good representation is obtained for the different positions on the circle of radius \( r \). This procedure is repeated as the value of \( r \) is varied.

IV. RESULTS

In Figure 2 results are shown for simulations of TDOA errors on an array containing 4 sensors, placed equidistantly on a circle of radius 1 m, thus forming a square as in Figure 1. Figure 2 shows the results when only one of the three TDOA estimates contains an error. Errors in TDOA are considered until a magnitude of 100 \( \mu s \), which covers the range of expected errors in TDOA when common time synchronization methods are applied [9],[10].

The source is assumed to be located at sensor 2, the right bottom sensor, and the error in TDOA is studied for two cases: an error on \( t_{21} \), and on \( t_{41} \). By comparing the curves corresponding to \( t_{21} \) and \( t_{41} \) in Figure 2, it can be seen that when the source is located on one of the sensors whose TDOA with respect to the reference sensor is erroneous (\( t_{21} \), source is assumed to be on sensor 2), then the impact of the error is significantly less than when the source is located elsewhere (\( t_{41} \)). To explain this behavior, we note that equation (1), which defines a particular TDOA estimate, describes a hyperbole in the x-y plane. The source is located somewhere on this hyperbole. When multiple TDOA estimates are available, the corresponding hyperbolic functions ideally intersect at exactly one point, which is the actual source location. However, when an error is introduced in the model, the shape of one of the hyperbolic functions is changed. The set of hyperbolic functions do not necessarily intersect in one point anymore, and a least squares solution to (2) is obtained. It can be verified that when the source is located on one of the sensors in the pair that results in an erroneous TDOA estimate, the corresponding hyperbolic function varies only slightly for the range of TDOA errors considered here, so that the influence on the least squares solution, and in turn the source location estimate is limited. When the TDOA error is on a sensor pair that does not include the source, the variation in the corresponding hyperbolic function is significant, which results in the behavior observed in Figure 2. For larger TDOA errors however, of the order of milliseconds, the corresponding source location estimation errors are large regardless of the location of the source.

A similar effect is seen in Figure 3, which analyzes errors in sensor location. The same sensor array is used, again with the source on sensor 2. Now, errors are introduced in location estimates of sensors 2 and 4. The influence of the error on sensor 2 is minor compared to the influence of the error on sensor 4. The reason is again that depending on the array geometry and source position, the influence of an error on the corresponding hyperbolic function is different.

Based on the behavior in Figures 2 and 3, one may conclude in the context of an audio conferencing application that it is beneficial for a sensor that is likely to introduce errors to be located close to the main speaker.

While Figure 2 corresponds to an error in only one TDOA estimate, Figure 4 considers the effect of errors in two TDOA estimates. Here, errors were inserted into both \( t_{31} \) and \( t_{41} \). The errors in TDOA could be both positive and negative. An interesting effect occurs; in the presence of one error, depending on the sign of the second error, the result in source localization can be worse but also better. Under certain conditions, as can be intuitively expected, the errors cancel each other. Also, the magnitude of the error in the source location estimate is different depending on which sensor pair contributes to the TDOA error.

To study the effect of multiple errors in sensor location, errors were added to the sensor location estimation of sensors.
V. CONCLUSIONS

In this paper, the behavior of a source localization method is studied when subjected to errors in the estimation of TDOA and sensor location, both of which are parameters required by the method. The analysis is performed in the context of an audio conferencing set-up using an ad-hoc wireless microphone array. Apart from serving to quantify the impact of these errors on source location, the analysis also leads to some interesting observations.

It is observed that the impact on source localization is less severe when the source is near the sensor or sensor pair experiencing the error. This observation also holds for a combination of errors in the estimation of both TDOA and sensor location. One may thus conclude in the context of an audio conferencing application that it is beneficial for a sensor that is likely to introduce errors to be located close to the main speaker.

Furthermore, it is observed that timing errors can become redundant when large errors occur in the estimation of sensor location. In such cases, the available resources may be devoted to improving the accuracy of sensor location estimation.

Errors in TDOA can also occur if the communication link between the sensors is rate-constrained, where the TDOA estimates need to be obtained using quantized signals. Here it was seen that the impact of a limited bit-rate on source localization can be reduced by increasing the observation duration.

In this paper, we have focused on source localization. Another important application of the wireless microphone networks considered here is the extraction of a desired signal in the presence of noise. Future work will analyze the impact of errors introduced by the wireless array on the performance of signal extraction algorithms.

REFERENCES


