Sound Positioning Using a Small-scale Linear Microphone Array

Ling Pei[1, 2], Liang Chen[2], Robert Guinness[2], Jingbin Liu[3], Heidi Kuusniemi[2], Yuwei Chen[3], Ruizhi Chen[4], Stefan Söderholm[2]

[1]School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University, China
[2]Department of Navigation and Positioning, Finnish Geodetic Institute, Masala, Finland,
[3]Department of Remote Sensing and Photogrammetry, Finnish Geodetic Institute, Masala, Finland,
[4]Conrad Blucher Institute for Surveying & Science, Texas A&M University Corpus Christi, Texas, USA
Email: ling.pei@sjtu.edu.cn

Abstract— Microphone arrays, also known as acoustic antennas, have been extensively used for sound localization. Small-scale microphone arrays have especially been used in teleconferences and game consoles due to their small dimension and easy deployment. In this article, we present an approach to locating a sound source using a small linear microphone array. We describe the fundamentals of linear microphone arrays and analyze the impact of geometry in terms of positioning accuracy using the dilution of precision (DOP) concept. The generalized cross-correlation (GCC) based on the phase transform (PHAT) weighting function is used to estimate the time difference of arrivals in a microphone array. Given the time differences, we use both closed-form and iterative optimization solutions to calculate the coordinates of the sound source. In order to evaluate the performances of the solutions applied in this paper, simulations and field tests were conducted. Simulation results show that the closed-form algorithm gives a positioning error of less than 5 cm in a 10-by-10 meter room when the geometry of a microphone array is good and the signal to noise ratio (SNR) is high. Linear small microphone arrays have lower performances compared to a non-linear distributed array. When the scale of a linear array is reduced, the positioning accuracy decreases dramatically. With a small linear array, the iterative optimization algorithm gives much better performance compared to the closed-form algorithm. Field tests were conducted in an 11-by-5.6 meter room using a linear array with a length of 0.23 meters. Positioning results show an average error of 0.25 meters along the axis parallel to the linear array and 0.53 meters error along the axis which is perpendicular to the linear array.

Keywords- Positioning, Microphone array, TDOA, Linear, Small-scale.

I. INTRODUCTION

Location technologies and applications have been developed and researched in a pervasive way in recent years. GNSS (global navigation satellites systems) have been widely adopted for positioning and navigation outdoors. Indoor environments, on the other hand, bring about increasing LBS (location based services) opportunities meanwhile more challenges for locating because of the lower availability of navigation satellite signals and complicated built infrastructures hindering signal propagation.

Alternatives to GNSS, short range RF signals such as WiFi, Bluetooth [1]−[6], RFID (radio-frequency identification) [7], UWB (ultra wideband) [8], long range signals of such as, GSM (global system for mobile communications) [9], DTV (digital television) [10] have been introduced in location applications using methods such as proximity, fingerprinting, or triangulation.

MEMS (microelectromechanical systems) motion sensors offer the opportunity of continuous relative navigation when localization with the positioning infrastructures is unavailable. Typically, MEMS motion sensors such as accelerometers, magnetometers, and gyroscopes can be utilized to calculate the speed, heading, orientation, or motion mode of a moving object. The above mentioned outputs can then be applied in a DR (dead reckoning) algorithm to assist positioning in challenging environments where the GPS performance is poor or WLAN positioning is unavailable [12]−[20]. Taking advantage of the magnetometer in modern smartphones, IndoorAtlas Ltd. (Oulu, Finland) pioneers magnetic anomaly-based indoor positioning [21][11]. In addition, the low cost camera phone is also a potential positioning sensor. Ruotsalainen et al. [22] used a camera on a Nokia N8 smartphone to detect the heading change of a mobile phone user. Lastly, hybrid solutions [23]−[26] are adopted to improve the availability and reliability of positioning by integrating multiple signals of opportunities and data of motion sensors.

A microphone array, known as an acoustic antenna, consists of multiple microphones placed at different spatial locations. The microphones are combined together to act as a single device. The array can be defined as a small-scale array if the microphones are physically located in one single device and the dimension of the device is smaller than 1 meter in any direction. A large-scale acoustic antenna is defined so that each microphone is a separate physical device and the dimension of the system exceeds 1 meter in every direction.

978-1-4799-4043-1/13/$31.00

Microphone arrays have promising potentials in practical applications. For instance, the microphone array has been used for scream and gun shooting detection and localization [27], monitoring vehicles, e.g. airplane, vessel, auto, etc., noise diagnosing [28]−[30], noise reducing, speech reorganization, teleconference, speaker tracking [31], game console, etc. Large-scale microphone arrays have also been widely applied for localization [32]−[37],while the small-scale linear microphone where the maximum dimension is normally less than one meter is mainly used for discriminating sounds based on direction, locating sound sources in term of direction, and enhancing the microphone performance by reducing noises.

In this paper, we propose an indoor positioning solution which uses a small-scale linear microphone array to locate a sound source. The rest of the paper is organized as follows: we describe the fundamentals of a small-scale linear microphone array in Section II. Section III introduces the TDOA (time difference of arrival) approach applied in this paper. Then, in Section IV, we explain the positioning algorithms in details. Conducted experiments and results are presented in Section V. Finally, we draw our conclusions of this research in Section VI.

II. SMALL-SCALE LINEAR MICROPHONE ARRAY

A. *Speed of Sound*

The sound is a compression wave which travels around 342 meters per second in dry air with temperature at 20 ˚C. The speed of sound in various temperature conditions can be calculated from [36]:

$$C = 331.3 \times \sqrt{1 + T / 273.15} \tag{1}$$

where C is the speed of sound in m/s, and T is the temperature in ˚C.

The propagation speed of sound is significantly slower compared to electromagnetic waves, which lowers the demand for precise clock in a sound based positioning system. In a GPS receiver a clock error of one nanosecond introduces a range measurement error of 0.3 meter. In a sound based localization system the same error will be introduced with a clock error of one millisecond.

B. *Far-field and Near-field cases*

The wave front from a sound source is curved and not flat and this introduces much more complexity in the actual position calculations using TDOA. However, this wave front can be approximated as flat which introduces an error in the calculated position. The relative size of this error compared to other error sources is dependent on the distance between the sound source and the microphone array and the size of the array. If the distance between the sound source and the array is large compared to the dimensions of the array the approximation does not introduce much additional error in the position. This is often referred to as a far-field case [38].

If the distance between the sound source and the array is small compared to the dimensions of the array itself we cannot do this approximation. This is referred to as a near-field case.

In both far-field and near-field cases the position of the microphone can be calculated using the time difference of arrival between all combinations of two microphones in an array.

C. *Dilution Of Precision*

The geometry of a microphone array with respect to the sound source heavily affects the positioning accuracy. In order to evaluate how this geometry affects the positioning quality, we borrow the PDOP (position dilution of precision) concept from GPS [39]. Higher PDOP value indicates poorer geometry of a microphone array with respect to the sound source. Therefore, the higher the PDOP value is, the lower accuracy can be achieved. Because of its better geometry, a nonlinear microphone array distributed around the sound source has a lower PDOP value compared to a linear array. Similarly, a microphone array within a small dimension has a higher PDOP value than a large scale array.

III. TIME DELAY ESTIMATION

A. *TDOA*

The range from a sound source to the *i*th microphone of a microphone array can be calculated as

$$r_i = C \cdot t_i \tag{2}$$

where C is the speed of sound and t_i is the traveling time from a sound source to the *i*th microphone. In order to measure the traveling time accurately, the clock of the microphone array need to be synchronized with the clock of the sound source, which increases the system complexity and is usually impractical. Instead, we use a TDOA solution which does not require that the clock of the sound source is synchronized to the clock of the microphone array.

Given a sound source with a location $\mathbf{s}$, the true TDOA relative to the *i*th microphone pair ($\mathbf{m_{i1}}, \mathbf{m_{i2}}$) is denoted as [46]

$$\Delta t_i = \frac{|\mathbf{s} - \mathbf{m_{i1}}| - |\mathbf{s} - \mathbf{m_{i2}}|}{C} \tag{3}$$

where $|\bullet|$ denotes the Euclidian distance between two points in a space. Eq. (3) indicates that we can solve the unknown location of a sound source given at least three independent TDOA measurements of microphone pairs.

B. *Time delay estimation*

TDE (time delay estimate) is an extensively studied topic with rich literature [40]−[42]. The delay estimate between two microphones is obtained as the time-lag that maximizes the cross-correlation between the received signals. The most popular and useful algorithms in practice are based on the generalized cross-correlation (GCC) method proposed by Knapp and Carter [42], as shown in Eq. (4), which implements a frequency domain cross correlation by introducing a weighting function $W(\omega)$ to mitigate noise disturbances.

$$R_{x_1x_2}(\tau) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} W(\omega)\, X_1(\omega) X_2^*(\omega) e^{j\omega\tau} \mathrm{d}\omega \tag{4}$$

where $X_1(\omega)$, $X_2(\omega)$ are the Fourier transforms of two microphone signals. $X_2^*(\omega)$ is a conjugate transposed matrix of $X_2(\omega)$. The Maximum likelihood (ML) approach and the Phase Transform (PHAT) weighting functions are two widely used and well studied functions. The ML approach performs well in moderately noisy and non-reverberant environments [43][44]. In environments with high reverberations, the PHAT weighting function [42] shown in Eq. (5), has more robust performance.

$$W_{PHAT}(\omega) = \frac{1}{|\, X_i(\omega) X_j^*(\omega)\,|}\,. \tag{5}$$

IV. POSITIONING ALGORITHMS

In this section, typical linear and non-linear approaches are applied for solving the problem of positioning.

A. Linear solution

Linear solutions, namely closed-form solutions, have been widely applied for triangulation because of the low computation requirements. These closed-form algorithms for the sound localization based on a microphone array have been validated by extensive research such as [49]–[51]. According to the definition of TDOA in Eq. (3), suppose the first microphone of a microphone array is selected as a reference. All the possible TDOAs respect to the first microphone can be defined as:

$$\Delta r_{i,1} = r_i - r_1 \equiv C \cdot \Delta t_{i1}\ . \tag{6}$$

Then, we have:

$$r_i^2 - r_1^2 = \Delta r_{i,1}^2 + 2 r_1 \Delta r_{i,1} \tag{7}$$

$$\begin{aligned} r_i^2 - r_1^2 = {} & (m_{x_i} - s_x)^2 + (m_{y_i} - s_y)^2 + (m_{z_i} - s_z)^2 \\ & - (m_{x_1} - s_x)^2 - (m_{y_1} - s_y)^2 - (m_{z_1} - s_z)^2 \end{aligned} \tag{8}$$

where $\mathbf{M}_i = [m_{x_i}, m_{y_i}, m_{z_i}]^T$ is the 3D coordinates of the ith microphone and $\mathbf{X_s} = [s_x, s_y, s_z]^T$ is the location of a sound source. In the solution, the nonlinear Eq. (8) has to be solved by linearization techniques. In this paper, we applied the methods proposed in [45].

B. Nonlinear solution

As opposed to the closed-form solutions which linearize the nonlinear problem, a nonlinear approach attempts to solve the problem by iteratively minimizing a (weighted) least squares estimate function such as:

$$F = \arg \min_{\mathbf{X}_S} (\sum_i \epsilon_i^2) \tag{9}$$

where

$$\begin{aligned} \epsilon_i = \Delta r_{i,1} & - \left[(m_{x_i} - \hat{s}_x)^2 + (m_{y_i} - \hat{s}_y)^2 + (m_{z_i} - \hat{s}_z)^2\right]^{\frac{1}{2}} \\ & + \left[(m_{x_1} - \hat{s}_x)^2 - (m_{y_1} - \hat{s}_y)^2 - (m_{z_1} - \hat{s}_z)^2\right]^{\frac{1}{2}} \end{aligned} \tag{10}$$

and $\hat{\mathbf{X}}_\mathbf{s} = [\hat{s}_x, \hat{s}_y, \hat{s}_z]^T$ is the optimal location estimation of a sound source based on the measured range difference $\Delta r_{i,1}$ with respect to the reference microphone.

In this paper, we applied the Levenberg-Marquardt [47][48] method which is a typical solution for solving nonlinear least squares problem. Then, according to [48], we have

$$(\lambda^{(k)}\mathbf{D} - \mathbf{A}^{(k)})\Delta \mathbf{X_s}^{(k)} = \upsilon^{(k)} \tag{11}$$

where $\upsilon^{(k)} = \mathbf{J}^{(k)^T} \boldsymbol{\epsilon}^{(k)}$, $\mathbf{J}$ is the Jacobian matrix, and $\boldsymbol{\epsilon}$ is a vector of residuals. $\mathbf{A}^{(k)} = \mathbf{J}^{(k)^T}\mathbf{J}^{(k)}$. $\mathbf{D}$ is a suitable diagonal matrix of scales. It is often chosen to be a unity matrix $\mathbf{I}$ or a diagonal of the matrix $\mathbf{A}^{(k)}$. $\lambda^{(k)}$ is a scale parameter. $\Delta\mathbf{X_s}^{(k)}$ is the increment of $\mathbf{X_s}^{(k)}$, the source location estimate at the kth step.

The Levenberg-Marquardt method actually combines two minimization methods: the gradient descent method and the Gauss-Newton method. If $\mathbf{D} = \mathrm{diag}(\mathbf{A}^{(k)})$, Eq. (11) is strongly influenced by a scale parameter $\lambda^{(k)}$. The higher it is, the closer the result is to the stable solution of steepest descent. When the scale parameter $\lambda^{(k)}$ equals to zero, Eq. (11) simplifies to the Gauss-Newton method. This strategy is based on the comparison of a forecast of the solution for the next iteration. If the forecast is close to the true result, the $\lambda^{(k)}$ will be lowered. If it is bad, $\lambda^{(k)}$ should become higher in order to stabilize the process. Therefore, the Levenberg-Marquardt method behaves like a gradient descent method when the parameters are far from their optimal value, and acts more like the Gauss-Newton method when the parameters are close to their optimal value.

The Levenberg-Marquardt algorithm terminates when at least one of the following conditions are met [48]:

1. υ, the magnitude of the gradient of F, drops below a threshold ε_1.
2. $\Delta\mathbf{X_s}^{(k)}$, the increment of $\mathbf{X_s}^{(k)}$, drops below a threshold ε_2.

3. The value of cost function F drops below a threshold ε_3.
4. The maximum number of iterations is reached.

V. EXPERIMENTAL RESULTS

In order to evaluate the performance of the algorithms and methods described in this paper, simulations and field tests were conducted.

A. *Positioning algorithms comparison*

In order to compare linear and nonlinear algorithms under different geometries of a microphone array, we utilized the simulation method adopted from [45]. Three geometries of a microphone array listed in Table I have been used in this work. The ground truth of the test sound source location was (3, 4) in the local coordinate system. Both linear and nonlinear algorithms are applied under the same conditions. Various noise levels are generated using the method proposed in [45] and 1000 ensample runs are simulated. The root mean square errors (RMSE) for the positioning solution obtained with two different positioning algorithms and for three different geometries are calculated. The results are shown in Fig 1, 2 and 3.

In Geo 1 shown in Fig 1 the closed-form solution performs better than the iterative optimization only in a low noise environment. In Geo 2 and 3 the nonlinear optimization solution has better performance than the closed-form solution at all noise levels.

TABLE I. GOMETRY DISTRIBUTIO.NS OF A MICROPHONE ARRAY

Geo ID	Mic1	Mic2	Mic3	Mic4	HDOP
Geo. 1	(5,1)	(10,5)	(5,10)	(1,5)	1.03
Geo. 2	(1,1)	(4,1)	(7,1)	(10,1)	1.07
Geo. 3	(5.35,1)	(5.45,1)	(5.55,1)	(5.65,1)	22.73

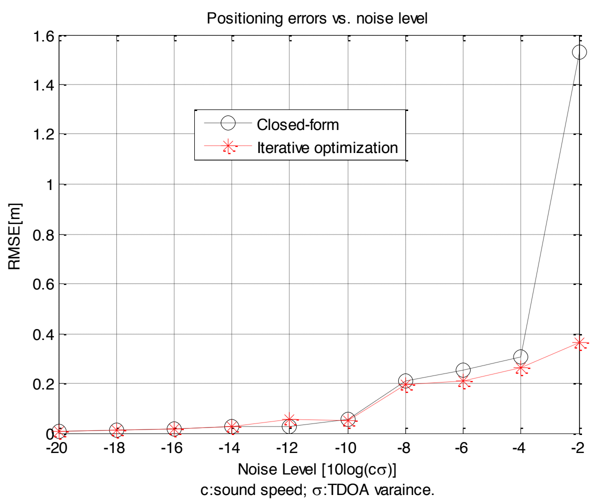

Figure 1. Source (4,3), Geo. 1.

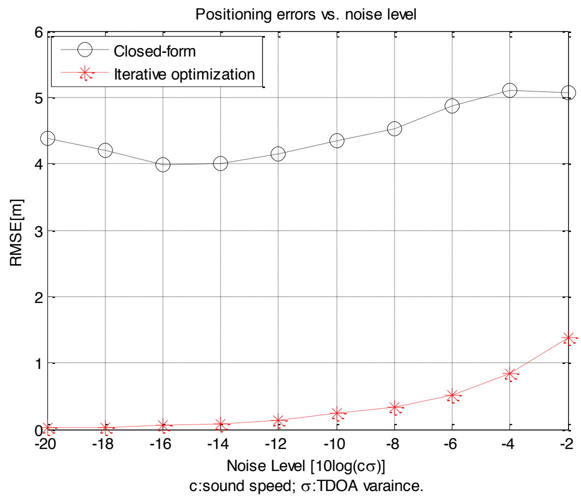

Figure 2. Source (3,4), Geo. 2.

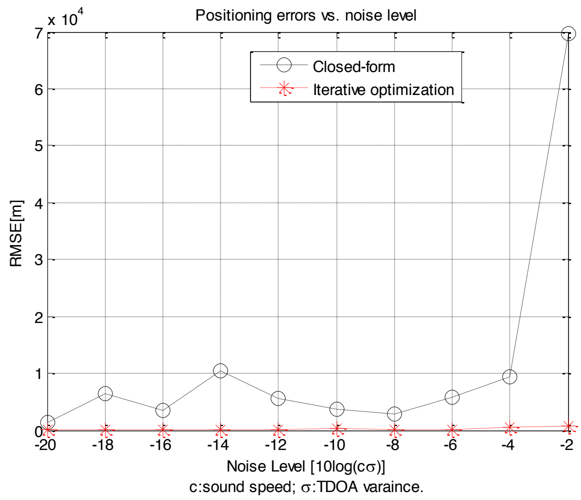

Figure 3. Source (3,4), Geo. 3.

B. *Field tests*

According to the results of simulations, the iterative optimization is a more suitable algorithm for sound localization using a small-scale linear microphone array. We also tested the iterative optimization solution in a field test environment. Kinect [52], a motion sensing input device by Microsoft for the Xbox 360 game console and also Windows PCs, has been used for tests. The Kinect system has a microphone array which enables the device to conduct ambient noise suppression. Therefore, the player can interact with the game via voice recognition. A Kinect microphone array features four microphone capsules with each channel processing 16-bit audio at a sampling rate of 16 kHz. Kinect has a similar geometry of a microphone array as Geo. 3 described in Section V. The dimensions of the Kinect microphone array are shown in Fig. 4. The distance between the leftmost and the rightmost microphones is 0.226 m. Only one microphone is placed on the left-hand side and the three other microphones are located on the right-hand side of the device. The Kinect microphone array was originally designed for beamforming [53] a technique used to amplify sound from one direction and suppressing sound coming from other directions. Previous research on the Kinect microphone array mainly focuses on the direction detection of a sound source instead of the exact location. In this paper, we used a Kinect for locating a sound source and estimating its coordinates.

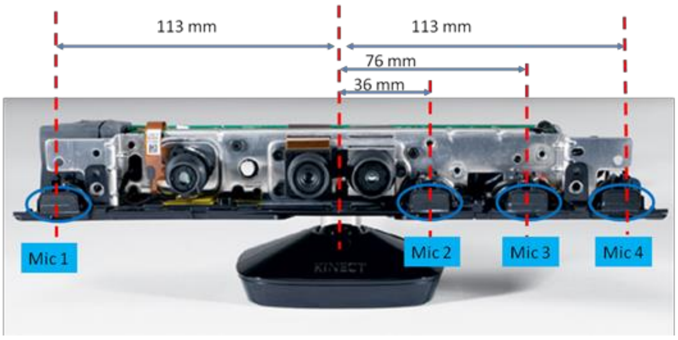

Figure 4. The microphone array in a Kinect

In the test environment, shown in Fig. 5, a Kinect was placed on a table and connected to a laptop via a USB port. A sound source was placed on another table with the same height as the one the Kinect was standing on.

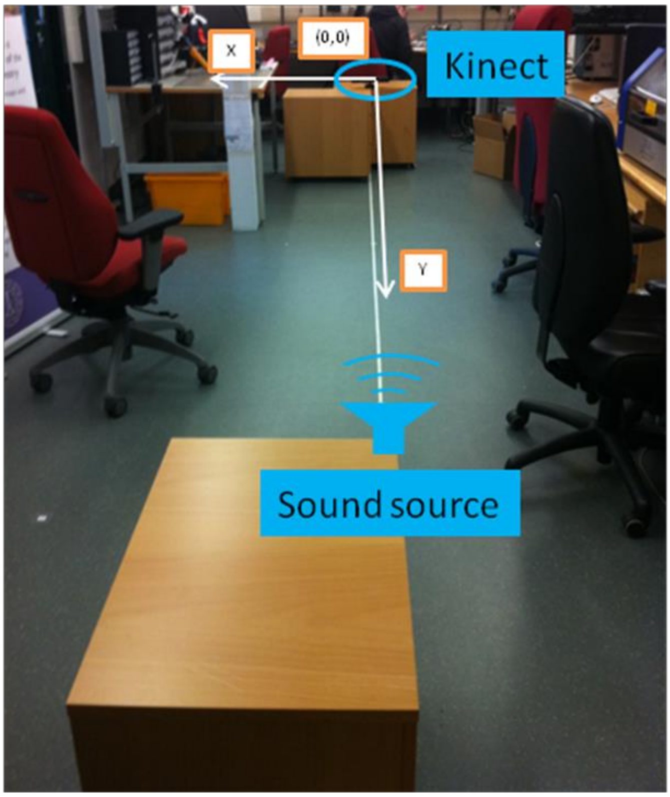

Figure 5. Test environment

Fig. 6 shows a test example in which a sound source emits a signal four times, and the signal is recorded by the four channels of a Kinect microphone array. Each channel handles the signals of one microphone. The signals are utilized for the time delay estimation using the PHAT-GCC algorithms described in Section III.

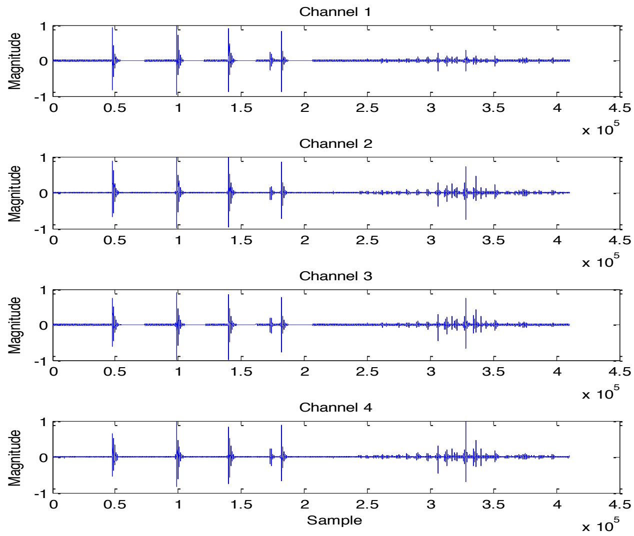

Figure 6. Signals of Kinect's four channels

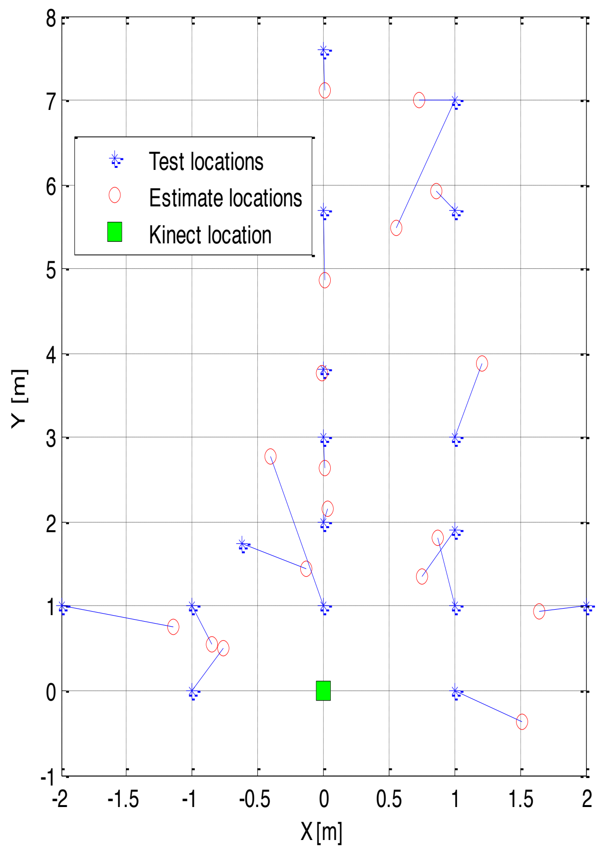

Figure 7. Test results

The positioning tests were carried out in a 10-by-5.6 meters room. As shown in Fig. 5, a Kinect was placed on one side of the test room close to a wall. The local coordinate system defined for these tests is depicted in Fig. 5. Eighteen points were selected from the target area and the sound source was placed at these points. The X and Y components for each location of the sound source were calculated based on the measurements from the microphone array.

Ground truth and estimate positions are shown in Fig. 7. Compared to the ground truth, the mean error of the estimate X component is 0.25 meters and Y component is 0.53 meters. The maximum error of the observed X component is 0.86 meters and Y component is 1.77 meters. The number of iterations used for each point was different. The maximum number of iterations needed was 71 and the minimum was 5.

VI. CONCLUSION

Microphone arrays have been widely used for sound localization. Small-scale microphone arrays have been deployed in consumer electronic devices such as Kinect and smartphones. Most smartphones nowadays have at least two microphones. Some emerging phones, e.g. Lumia 925, even have three microphones, which will enable more microphone array based applications. This paper studies the possibility of using a small-scale linear microphone array for the purpose of locating a sound source. The generalized cross-correlation (GCC) approach based on the phase transform weighting function is applied to estimate the time difference of arrival in a microphone pair. Given the time difference estimations, we use both the closed-form and iterative optimization solutions to calculate the coordinates of the sound source. In order to evaluate the performances of the applied solutions, simulations and field tests are conducted. Simulation results show that the closed-form algorithm gives positioning errors of less than 5 cm in a 10-by-10 meters room when the geometry of the microphone array is good and signal to noise ratio is high. Linear microphone arrays have lower performances than a non-linear large scale distributed array. When the scale of a linear array is reduced, the positioning accuracy decreases significantly. With a small-scale linear array, the iterative optimization algorithm has much better performance compared to the closed-form algorithm. Field tests were executed in a 10-by-5.6 meters room using a linear array with a length of 0.226 meters. Positioning results give an average error of 0.25 meters along the axis parallel with the linear array and 0.53 meters error along the axis which is perpendicular to the linear microphone array.

The scope of this paper is aimed at positioning solutions. Therefore, the efficiency of involved algorithms is ignored. In the future, we will further study these topics, for instance, using the geometry information to constrain the iterative optimization and restrict the search space. Iterative optimization methods are locally optimized methods and are very sensitive to the parameters of the algorithms, for instance, the initial point of iteration or the thresholds of terminating conditions, which are described at the end of section IV. Proper parameter selection will accelerate the convergence of the solution. Therefore, we will also work on the parameter selection for iterative optimization methods in the next steps of research. Besides the closed-form and iterative optimization, we will investigate other microphone array based algorithms, for instance, stochastic region contraction (SRC) base sound localization. This paper only presents static test results. In the future, the effect of obstacles e.g. people, furniture, equipment etc., .and moving sound target tracking will be studied.

ACKNOWLEDGMENT

The authors are grateful to Dr. Hongmei Hu from the Medical Physics Section of Carl von Ossietzky Universität Oldenburg, Germany for her suggestions on acoustic signal processing. The authors also would like to thank Anttoni Jaakkola from the Finnish Geodetic Institute for providing the test device.

REFERENCES

[1] Pei, L.; Chen, R.; Liu, J.; Kuusniemi, H.; Tenhunen, T.; Chen, Y. Using inquiry-based Bluetooth RSSI probability distributions for indoor positioning. *J. Glob. Position. Syst*. **2010**, *9*, 122–130.

[2] Priyantha, N.B.; Chakraborty, A.; Balakrishnan, H. The Cricket Location-Support System. In *Proceedings of the 6th Annual International Conference on Mobile Computing and Networking*, Boston, MA, USA, 6–11 August 2000; pp. 32–43.

[3] Bahl, P.; Padmanabhan, V.N. Radar: An In-Building RF Based User Location and Tracking System. In *Proceedings of Infocom—Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, Tel-Aviv, Israel, 26–30 March 2000; pp. 775–784.

[4] Pei, L.; Chen, R.; Liu, J.; Tenhunen, T.; Kuusniemi, H.; Chen, Y. Inquiry-Based Bluetooth Indoor Positioning via RSSI Probability Distributions. In *Proceedings of the Second International Conference on Advances in Satellite and Space Communications (SPACOMM 2010)*, Athens, Greece, 13–19 June 2010; pp. 151–156.

[5] Gomes, G.; Sarmento, H. Indoor Location System Using ZigBee Technology. In *Proceedings of Third International Conference on Sensor Technologies and Applications*, Athens/Glyfada, Greece, 18–23 June 2009; pp. 152–157.

[6] Chen, L.; Pei, L.; Kuusniemi, H.; Chen, Y.; Kröger, T.; Chen, R. Bayesian fusion for indoor positioning using bluetooth fingerprints. *Wirel. Pers. Commun*. **2012**, *67*, 1–11.

[7] Ni, L.M.; Liu, Y.; Lau, Y.C.; Patil, A.P. Landmarc: Indoor location sensing using active RFID. *Wirel. Netw*. **2004**, *10*, 701–710.

[8] Pahlavan, K.; Akgul, O.F.; Heidari, M.; Hatami, A.; Elwell, M.J.; Tingley, D. R. Indoor Geolocation in the Absence of Direct Path. *IEEE Trans. Wireless Comm*. **2006**, *13*, 50-58.

[9] Syrjärinne, J. Studies on Modern Techniques for Personal Positioning. Ph.D. Thesis, Tampere University of Technology, Tampere, Finland, 2001.

[10] Chen, L., Yang, L. and Chen, R., Time delay tracking for positioning in DTV networks, In *Proceedings of Ubiquitous Positioning, Indoor Navigation, and Location Based Service (UPINLBS*), 2012 , Helsinki, pp.1-4, 3-4 Oct. 2012.

[11] Storms, W.; Shockley, J.; Raquet, J. Magnetic field navigation in an indoor environment. In *Proceedings of Ubiquitous Positioning Indoor Navigation and Location Based Service (UPINLBS)*, Kirkkonummi, Finland, 3-4 October 2010.

[12] Pei, L.; Chen, R.; Liu, J.; Chen, W.; Kuusniemi, H.; Tenhunen, T.; Kröger, T.; Chen, Y.; Leppäkoski, H.; Takala, J. Motion Recognition Assisted Indoor Wireless Navigation on a Mobile Phone. In *Proceedings of the 23rd International Technical Meeting of The Satellite Division of the Institute of Navigation*, Portland, OR, USA, 21–24 September 2010; pp. 3366–3375.

[13] Pei, L.; Chen, R.; Liu, J.; Kuusniemi, H.; Chen, Y.; Tenhunen, T. Using Motion-Awareness for the 3d Indoor Personal Navigation on a Smartphone. In Proceedings of the 24rd International Technical Meeting of the Satellite Division of the Institute of Navigation, Portland, OR, USA, September 19-23, 2011. pp. 2906-2912.

[14] Pei, L.; Liu, J.; Guinness, R.; Chen, Y.; Kuusniemi, H.; Chen, R. (2012) Using LS-SVM based motion recognition for smartphone indoor wireless positioning. *Sensors*, *12 (5)*, 6155–6175.

[15] Pei, L., Guinness, R., Chen, R., Liu, J., Kuusniemi, H., Chen, Y., Chen, L., & Kaistinen, J. (2013). Human Behavior Cognition Using Smartphone Sensors. Sensors, 13(2), 1402-1424.

[16] R. Chen, L. Pei, and Y. Chen, "A Smart Phone Based PDR Solution for Indoor Navigation," in ION GNSS 2011, Portland, OR, 2011, p. 1404-1408.

[17] Foxlin, E. Pedestrian Tracking with Shoe-Mounted Inertial Sensors. *IEEE Comput. Graph. Appl.* **2005**, *25*, 38-46.

[18] Chen, W.; Chen, R.; Chen, Y.; Kuusniemi, H.; Wang, J.; Fu, Z. An Effective Pedestrian Dead Reckoning Algorithm Using a Unified Heading Error Model. In *Proceedings of IEEE/ION PLANS 2010 Conference*, Indian Wells: Palm Springs, CA, USA, 4–6 May 2010.

[19] Mathews, M.B.; Macdoran, P.F.; Gold, K.L. SCP Enabled Navigation Using Signals of Opportunity in GPS Obstructed Environments. *J. Navig.* **2011**, *58*, 91-110.

[20] Chen, R.; Chen; W.; Chen, X.; Zhang, X.; Chen, Y. Sensing Strides Using EMG Signal for Pedestrian Navigation. *GPS Solutions* **2011**, *15*, 161-170.

[21] IndoorAtlas Ltd. Oulu, Finland. Available online: http://www.indooratlas.com (accessed on 22 October 2012).

[22] Ruotsalainen, L.; Kuusniemi, H.; Chen, R. Visual-aided Two-dimensional pedestrian indoor navigation with a smartphone. J. Glob. Position. Syst. **2011**, 10, 11–18.

[23] Liu, J.; Chen, R.; Pei, L.; Chen, W.; Tenhunen, T.; Kuusniem, H.; Kröger, T.; Chen, Y. Accelerometer Assisted Wireless Signals Robust Positioning Based on Hidden Markov Model. In *Proceedings of the IEEE/ION Position, Location and Navigation Symposium (PLANS) 2010*, Indian Wells, CA, USA, 3–6 May 2010; pp. 488–497.

[24] Liu, J.; Chen, R.; Pei, L.; Guinness, R.; Kuusniemi, H. Hybrid smartphone indoor positioning solution for mobile LBS. *Sensors* **2012**, *12*, 17208–17233.

[25] Liu, J.; Chen, R.; Chen, Y.; Pei, L.; Chen, L. iParking: An Intelligent Indoor Location-Based Smartphone Parking Service. *Sensors*. 2012; 12(11):14612-14629.

[26] Kuusniemi, H.; Liu, J.; Pei, L.; Chen, Y.; Chen, L.; Chen, R. Reliability considerations of multi-sensor multi-network pedestrian navigation. *Radar Sonar Navig. IET* **2012**, *6*, 157–164.

[27] Valenzise, G., Gerosa, L., Tagliasacchi, M., Antonacci, F., & Sarti, A. "Scream and gunshot detection and localization for audio-surveillance systems". In *Advanced Video and Signal Based Surveillance, 2007. AVSS 2007. IEEE Conference on*. pp. 21-26. , September ,2007.

[28] Michel, U.; Barsikow, B.. (2003): Localisation of sound sources on moving vehicles with microphone arrays. In: Proceedings 5th European Conference on Noise Control Euronoise 2003, Napels, Italy, 19. - 21. May 2003

[29] B. Schulte-Werning, K. Jäger, R. Strube, L. Willenbrink, "Recent developments in noise research at Deutsche Bahn (noise asessment, noise source localization and specially monitored track)." Journal of Sound and Vibration, 267, (2003), pp. 689-699.

[30] J. Piet, U. Michel and P. Böhning, "Localization of the acoustic sources of the A340 with a large microphone array during flight tests," 8th AIAA/CEAS Aeroacoustics Conference, Breckenridge, Colorado (USA), 17-19 June 2002, AIAA-2002-2506.

[31] Khalil, F., Jullien, J. P., & Gilloire, A. (1994). Microphone array for sound pickup in teleconference systems. *Journal of the Audio Engineering Society*, *42*(9), 691-700.

[32] Filonenko, V., Cullen, C., & Carswell, J. (2010, September). nvestigating ultrasonic positioning on mobile phones. In *Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on* (pp. 1-8). IEEE.

[33] Pertila, P., Mieskolainen, M., & Hamalainen, M. S. (2012, August). Passive self-localization of microphones using ambient sounds. In *Signal Processing Conference (EUSIPCO), 2012 Proceedings of the 20th European* (pp. 1314-1318). IEEE.

[34] Janson, T., Schindelhauer, C., & Wendeberg, J. (2010, September). Self-localization application for iPhone using only ambient sound signals. In *Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on* (pp. 1-10). IEEE.

[35] Filonenko, V., Cullen, C., & Carswell, J. D. (2012). Asynchronous ultrasonic trilateration for indoor positioning of mobile phones. In *Web and Wireless Geographical Information Systems* (pp. 33-46). Springer Berlin Heidelberg.

[36] Hoflinger, F., Zhang, R., Hoppe, J., Bannoura, A., Reindl, L. M., Wendeberg, J., ... & Schindelhauer, C. (2012, November). Acoustic Self-calibrating System for Indoor Smartphone Tracking (ASSIST). In *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on* (pp. 1-9). IEEE.

[37] Holm, S. (2012, November). Ultrasound positioning based on time-of-flight and signal strength. In *Indoor Positioning and Indoor Navigation (IPIN), 2012 International Conference on* (pp. 1-6). IEEE.

[38] B. D. Steinberg. Principles of aperture and array system design: Including random and adaptive arrays. *New York, Wiley-Interscience, 1976*.

[39] Langley, R. B. (1999). Dilution of precision. *GPS world*, *10*(5), 52-59.

[40] M. Brandstein and D. Ward, Microphone Arrays. Berlin, Germany: Springer, 2001

[41] Carter, G. C. (1987). Coherence and time delay estimation. Proceedings of the IEEE, 75(2), 236-255.

[42] C. H. Knapp and G. C. Carter, "The generalized correlation method for estimation of time delay," IEEE Trans. Acoust., Speech, Signal Process., vol. ASSP-24, no. 4, pp. 320–332, Aug. 1976.

[43] Ianniello, J. (1982). Time delay estimation via cross-correlation in the presence of large estimation errors. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, *30*(6), 998-1003.

[44] DiBiase, J. H., Silverman, H. F., & Brandstein, M. S. (2001). Robust localization in reverberant rooms. In *Microphone Arrays* (pp. 157-180). Springer Berlin Heidelberg.

[45] Y. T. Chan and K. C. Ho, "A Simple and Efficient Estimator for Hyperbolic Location," IEEE Trans. Signal Processing, vol. 42, no. 8, pp. 1905-1915, Aug. 1994.

[46] Brandstein, M. S. (1995). A framework for speech source localization using sensor arrays.

[47] Moré, J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis* (pp. 105-116). Springer Berlin Heidelberg.

[48] Lourakis, M. I. (2005). A brief description of the Levenberg-Marquardt algorithm implemented by levmar. *Institute of Computer Science, Foundation for Research and Technology*, *11*.

[49] Smith, J., & Abel, J. (1987). Closed-form least-squares source location estimation from range-difference measurements. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, *35*(12), 1661-1669.

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

[51] Smith, J., & Abel, J. (1987). Closed-form least-squares source location estimation from range-difference measurements. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, *35*(12), 1661-1669.

[52] Webb, J., & Ashley, J. (2012). *Beginning Kinect Programming with the Microsoft Kinect SDK*. Apress

[53] Thomas, M. R., Ahrens, J., & Tashev, I. (2012, September). Optimal 3D beamforming using measured microphone directivity patterns. In Acoustic Signal Enhancement; Proceedings of IWAENC 2012; International Workshop on (pp. 1-4).