LOCALIZING MULTIPLE AUDIO SOURCES FROM DOA ESTIMATES IN A WIRELESS ACOUSTIC SENSOR NETWORK

Anthony Griffin[1] *Athanasios Mouchtaris*[1,2*]

[1] Foundation for Research & Technology – Hellas (FORTH-ICS), Heraklion, Crete, Greece, GR-70013
[2] University of Crete, Department of Computer Science, Heraklion, Crete, Greece, GR-70013
agriffin@ics.forth.gr, mouchtar@ics.forth.gr

ABSTRACT

In this work we propose a method to estimate the position of multiple sources in a wireless acoustic sensor network, where each sensor node only transmits direction-of-arrival (DOA) estimates each time interval, minimizing the transmissions to the processing node. Our method is based on the intersection of DOA estimates with outlier removal, and as such is very computationally efficient. We explore the performance of our method through extensive simulations and real measurements.

Index Terms— Wireless sensor networks, source localization, acoustic sensors, microphone arrays, position estimation

1. INTRODUCTION

The recent advent of small, low-cost microphones has created a surge in interest of localization of entities from their acoustic emissions (voice, cries, movement noise, footsteps, etc). Application areas such as security, wildlife monitoring, and health care for the elderly exploit the fact that monitoring acoustic emissions is a passive, non-intrusive localization method.

To do this accurately in a manner that doesn't require a very dense distribution of microphones throughout the area to be monitored, a popular method is to treat the microphones as multiple arrays and jointly process the acoustic signals [1–4]. Although these methods can achieve accurate performance, they require the acoustic signals to be very well synchronized. This is not an issue if the area to be monitored is relatively small, but if the area is larger, then many hundreds of meters of microphone cables may be required.

An obvious way to avoid cables is to use a wireless acoustic sensor network (WASN), which also offers extreme flexibility in sensor (microphone) placement. Unfortunately, the independent sample clocks in the nodes of a WASN are not synchronized, and this dramatically affects the performance of standard acoustic localization algorithms [5,6]. The work in [7] circumvented this problem by using special nodes that used their internal GPS chips to resample the audio samples with a network-common timestamp. Obviously this requires expensive, resource-hungry nodes, and transmitting the full audio signals to the central processing node reduces the network's life significantly.

There are a few methods that are more suitable for WASNs due to the fact that they consider multiple unsynchronized microphone arrays. By allowing increased computational ability in the nodes, reductions in transmission are achieved in the work of [8] which only transmits time difference of arrival (TDOA) information between microphone pairs at each node along with associated reliability information to the central node. The absolute minimum transmission possible is attained in the work of [9, 10], where each sensor node only transmits a direction of arrival (DOA) estimate to the central processing node.

We follow a similar model to [9, 10], but we extend their work by removing less reliable data, and particularly by considering multiple sources in the region of interest. We present an extensive investigation of our algorithm's performance through both simulations and real measurements.

2. THE FRAMEWORK

Our framework is a wireless sensor network whose M nodes are each equipped with a microphone array—which we will also refer to as a sensor. This enables each node to generate a direction-of-arrival (DOA) estimate for any source that it can "hear" (any source whose SNR at the node is high enough to be detected). It is important to note that each node estimates a direction only, and no range information, thus any one node's DOA estimate is not sufficient to obtain an absolute position for a source.

In our modeling, we assume that the signal of a source radiates as a spherical wave, and the attenuation experienced by the source signal in travelling from r_1 meters from the source to r_2 meters from the source is given by [11]

$$a = 20 \log_{10} \frac{r_2}{r_1} \quad \text{dB}. \tag{1}$$

Thus by specifying a signal-to-noise ratio (SNR) at the centre of a cell, we can determine the SNR at each sensor. Note that although we assume that the noise power at each sensor is the same, the SNR at each sensor will be different as the signal energy depends on the differing distances between the source and the sensors.

We also assume that the accuracy of the each sensor's DOA estimate of a source is determined only by the SNR of that source's signal at that sensor. Our previous work has shown this to be a valid assumption [12, 13].

3. LOCALIZING A SINGLE SOURCE FROM MULTIPLE DOA ESTIMATES

In the ideal case—i.e., perfect DOA estimates—the DOA vectors from each node will all intersect at the same point and a source could be localized by simply finding the intersection point. In practice—or any realistic simulation—the DOA estimates will not

*This research has been co-financed by the European Union (European Social Fund ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: Thales-DISFER.

978-1-4799-0972-8/13/$31.00 ©2013IEEE

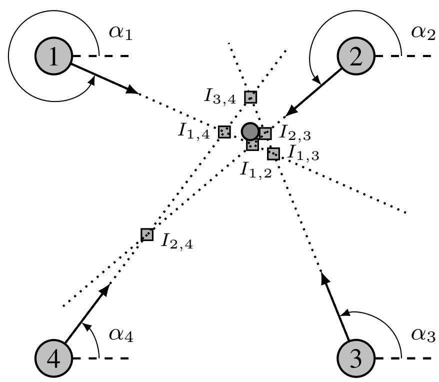


Figure 1: Example square cell with four sensor nodes (blue circles, numbered 1 to 4), the DOAs (α_1–α_4) to a source (the red circle), and the intersection points (grey squares, labeled $I_{1,2}$–$I_{3,4}$) of DOA vector pairs.

be perfect, and we must mitigate the effect of the DOA estimation errors. As we show in this paper, we can still accurately estimate the location of a source by taking the centroid of the intersections of *pairs* of DOA vectors. The centroid is simply the mean of the set of intersection points, and minimizes the sum of squared Euclidean distances between itself and each point in the set. Note that this is similar to the work in [9, 10] but here we extend their work by removing less reliable intersection points.

Fig. 1 illustrates this idea with an example, where the DOA estimates have an error of up to $\pm 5°$. The locations of sensors 1 to 4 are: (0, 4), (4, 4), (4, 0), (0, 0), respectively, and the source is at (2.6, 3.0). The estimated location from the centroid of the intersection points is (2.40, 2.77), which is a distance error of 0.43, or 11% of the inter-sensor spacing, S. Further inspection of Fig. 1 reveals that the effect of $I_{2,4}$ is significant. By excluding this point from the centroid, the estimated location becomes (2.64, 2.99) and the error drops to 0.03, or 1% of S.

A question that then naturally arises is: how can we detect and exclude outliers such as $I_{2,4}$? It can be shown that these outliers are caused by DOA vectors that are almost parallel. A small change in the slope of either of these lines—due to DOA estimation error—can move their point of intersection significantly. Thus excluding the intersection points of pairs of DOA vectors that are almost parallel improves the accuracy of the location estimation. This is further demonstrated in the results of Section 5.

Now let $\gamma_\parallel$ be a "parallelness" threshold, and let $A(X,Y)$ be a function that calculates the *angular distance* between two angles X and Y and returns a value in the range $[0°, 180°]$, such as

$$A(X,Y) = 2\sin^{-1}\frac{|\exp(jX) - \exp(jY)|}{2}. \tag{2}$$

Note that there are other more efficient implementations of $A(X,Y)$ than (2), but this is one of the most compact.

Source localization using our intersection point method can then be summarized as:

1. Collect the M DOA estimates.
2. Take each of the pairs of DOA estimates $\alpha_{m_1}, \alpha_{m_2}$ and discard it if either of the two conditions are met:

$$A(\alpha_{m_1}, \alpha_{m_2}) \ < \ \gamma_\parallel, \tag{3}$$
$$A(\alpha_{m_1}, \alpha_{m_2}) \ < \ 180° - \gamma_\parallel. \tag{4}$$

3. Calculate the points of intersection of the remaining pairs.
4. The estimate of the source location is then given by the centroid of the points of intersection.

Note that this method's resolution has no inherent limitations, and is affected only by the accuracy of the DOA estimates.

4. LOCALIZING MULTIPLE SOURCES FROM MULTIPLE DOA ESTIMATES

We now consider the case where there are two sources in the cell, and each sensor provides two DOA estimates. Unfortunately, the processing node receiving these estimates cannot know to which source they belong, and the localization algorithm must take this into account. An additional complication is that some sensor nodes may only detect one source, as the sources' DOAs may be too close together for that node to discriminiate between them. We call this the *minimum angular source separation* (MASS) of a sensor node, i.e., if the angular distance between two sources is less than the MASS, then the sensor node will only detect one source. Thus our localization algorithm must deal with the ambiguity that each DOA estimate may originate from either source, and that some (or even all) of the sensor nodes may underestimate the number of sources.

Let us first describe the concept of our geometric approach to solving this problem. Let X_2 be the set of sensors surrounding a cell detecting two sources in that cell, and let C_2 be the size of that set, i.e., $C_2 = |X_2|$. We take advantage of the fact that each DOA estimate—from a sensor in X_2—can only belong to one source. By dividing the possible locations for sources into the 2^{C_2} unique combinations of DOA estimates, we obtain up to 2^{C_2} regions, (we say "up to" as some of these regions may be null, depending on the orientation of the DOA estimates). By counting the number of intersection points in each region, and choosing the one that contains the most intersection points, we obtain the one that is most likely to contain one of the sources. Once we have chosen a region—and thus one of the combinations of DOA estimates—there will only be one other possible combination of DOA estimates pointing to the second source. Our proposed algorithm to localize two sources can be more formally stated as:

1. Find the intersection points of all of the pairs of DOA vectors, removing any pair whose vectors are too parallel, as in step 2 of the single-source algorithm of Section 3.
2. Count the number of sensor nodes detecting two sources in the current time frame, C_2. The next step then depends on the value of C_2.
3. If $C_2 = 0$ then the locations of the sources are given by the centroid of the intersection points. This is the best that can be done for this case, but this case only occurs if the sources are very close together.
4. Otherwise, find the C_2 circular means of the DOAs from the sensors detecting two sources.
5. The vectors of these circular means form $2C_2$ half-planes, find the regions defined by all the intersection of all the possible combinations of pairs of half-planes from different sensors. There will be 2^{C_2} of them.
6. Find the region with the most intersection points. If there is a tie, choose the region whose intersection points have the minimum variance. The location of the first source is given by the centroid of the intersection points in this region.
7. There will be only one other possible combination of half-planes, and the centroid of the intersection points in this region will give the location of the second source.

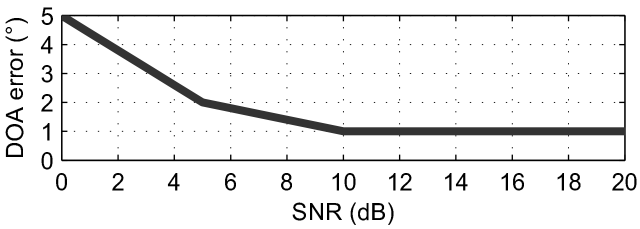


Figure 2: Modeling the effect of SNR on mean absolute DOA estimation error for an 8-element circular microphone array.

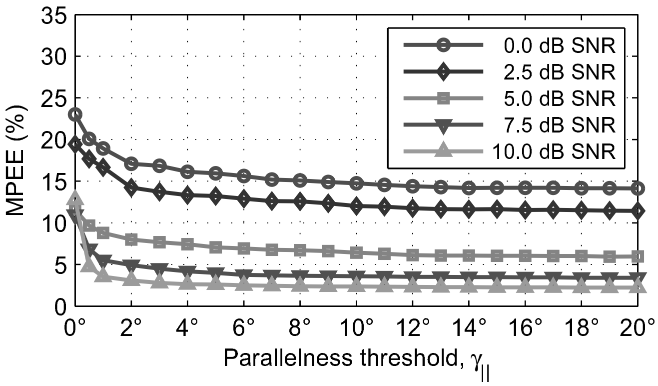


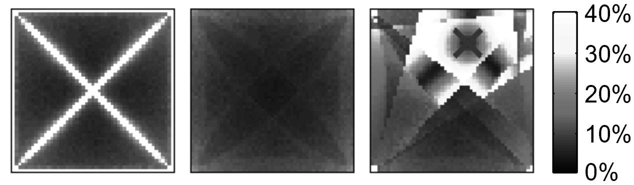
Figure 3: Effect of the parallelness threshold $\gamma_{||}$ on mean position estimation error (MPEE) as a percentage of cell size S using the intersection point method in a square 4-node cell for various values of SNR measured at the center of the cell.

Note that we have described this algorithm conceptually, but it can be implemented very efficiently by using line tests—testing whether a point is above, below, or on a line—and binary masks.

5. SIMULATION RESULTS AND DISCUSSION

In order to investigate the performance of our proposed localization method, we performed some simulations of a square 4-node cell of a WASN, similar to that of Fig. 1, with the nodes separated by S. The error in the DOA estimation at each sensor was assumed to be dependent only upon the SNR at each sensor, which is in turn determined by the length of the path from the source to the sensor, as discussed in Section 2. By specifying a reference SNR at the center of the cell, the SNR at each sensor can be calculated through geometry and the use of (1). We used the results of our previous work on DOA estimation [12, 13] to model the magnitude of this SNR-dependent error as the curve in Fig. 2, and this was added or subtracted—with equal likelihood—from the true DOA. We define the mean position error estimate (MPEE) as the mean distance between the sources' estimated and true locations, with the sources being located anywhere within the cell with independent, uniform probability densities.

In Fig. 3 we present simulation results for a single source, following the method of Section 3, which shows the effect of discarding pairs of DOA vectors that are almost parallel. A value of 20° for $\gamma_{||}$ is clearly sufficient to reduce the error, and this is further illustrated in Fig. 4 (a) & (b), where the high error zones (in white) between sensors are dramatically improved. An example of the effect of a second source is illustrated in Fig. 4 (c), following the method of Section 4. The presence of the second source clearly degrades the accuracy of localization when the other source is near it. Note though, that the majority of the cell's error is relatively unaffected.



(a) $\gamma_{||} = 0^\circ$ (b) $\gamma_{||} = 20^\circ$ (c) $\gamma_{||} = 20^\circ$

Figure 4: Position errors throughout a square 4-node cell as a percentage of cell size S, with a center reference SNR of 10 dB, and MASS of 20°, for the cases of: (a) & (b) one source, and (c) two sources, with one fixed at the blue X.

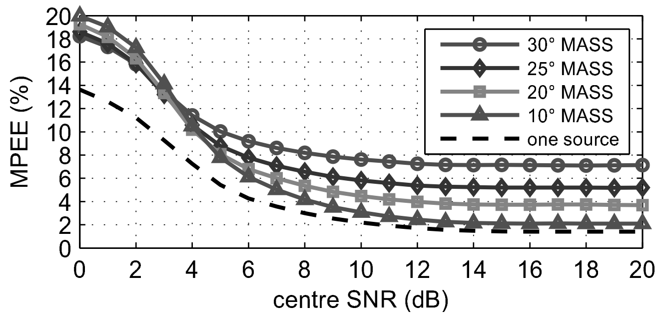


Figure 5: Effect of the minimum angular source separation (MASS) on the mean position estimation error (MPEE) as a percentage of cell size S for two sources in a square 4-node cell, for various values of SNR measured at the center of the cell.

The localization error of the two source case is further explored in Fig. 5. The importance of having a high angular resolution—and therefore a low MASS—here is evident. As the MASS decreases, the two-source performance approaches that of one source at high SNR, which is intuitively satisfying. Although not explicitly shown in these figures, we found that the higher C_2—the number of sensors detecting two sources—the better the localization performance, as more sensors have two DOA estimates, and the intersection points are therefore more accurate. A higher MASS results in more likelihood of a lower C_2, or equivalently a lower localization accuracy, clearly illustrated in Fig. 5.

All the previous results have considered the DOA estimation error at the sensors to be modeled as in Fig. 2. In Fig. 6 we consider the position error for two sources with increased DOA estimation error when the centre SNR is 20 dB. This is modeled by adding 0°–9° to the magnitude of the DOA estimation error at each sensor node. The curves in Fig. 6 show that the effect of more DOA estimation error is more important than an increasing MASS in determining the position error. These results suggest that more design effort should be put in to reducing a sensor's DOA estimation error, than decreasing its MASS.

Finally, we performed some real recordings[1] of acoustic sources in a 4-node square cell with sides 4 meters long. The sensors on the nodes were circular 4-element microphone arrays, and the DOA estimation was performed by our real-time system of [12, 13]. The sources were recorded speech played back simulta-

[1]The authors would like to thank Anastasios Alexandridis for his assistance in performing the experiments.

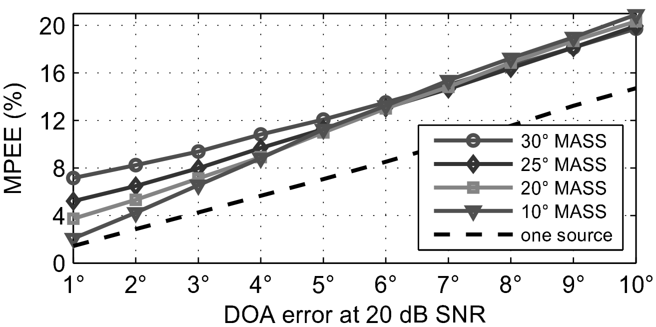
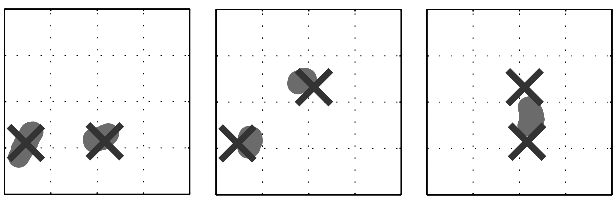


Figure 6: Effect of MASS on the mean position estimation error (MPEE) as a percentage of cell size S for two sources in a square 4-node cell, for various values of DOA error at 20 dB SNR, and signals having 20 dB SNR at the center of the cell.



(a) MPEE = 3.3% (b) MPEE = 7.1% (c) MPEE = 15 %

Figure 7: Position errors (the red clouds of estimates) throughout a square 4-node cell as a percentage of cell size S, for real recordings of two simultaneous sources (the blue X's).

neously through loudspeakers at different locations, and their SNR at the center of the cell was measured to be about 13 dB. Fig. 7 shows the results for three different positions of the sources. For the cases in Fig. 7 (a) & (b), the proximity and orientations of the sources meant that only two sensors detected two sources, yet the MPEE is still very good in both of those cases. In Fig. 7 (c) the sources were so close together that none of the sensors detected two sources, and the position error is larger. As discussed in Section 4 however, this only occurs when the sources are quite close together, which has the effect of bounding the error in these cases. It should be noted that these recordings took place outdoors, and as such did not have many reflections, but there was a significant level of distant noise sources, such as cars and dogs barking. Furthermore, the orientations of the sensors were not finely calibrated, and the DOA estimates likely have unintended offsets of a few degrees. Thus the conditions were far from ideal, making the results of our proposed localization method even more encouraging.

6. CONCLUSIONS

In this work we have considered the framework for estimating the position of a source in a wireless acoustic sensor network where each sensor node only transmits direction-of-arrival estimates each time interval, thus minimizing the transmissions to the central processing node. We presented a low-complexity method to perform the position estimation of multiple sources, along with an improvement to standard single source position estimation. Through extensive simulations and experiments we showed the effectiveness of our proposed method, and investigated some of its important parameters. Note that we have only considered localizing up to two sources here due to space considerations, but the method may be extended to localize a greater number of sources.

7. REFERENCES

[1] P. Aarabi, "The fusion of distributed microphone arrays for sound localization," *EURASIP Journal of Applied Signal Processing*, January 2003. [Online]. Available: http://dx.doi.org/10.1155/S1110865703212014

[2] T. Ajdler, I. Kozintsev, R. Lienhart, and M. Vetterli, "Acoustic source localization in distributed sensor networks," in *Conf. Rec. of Asilomar Conf. on Signals, Systems and Computers*, vol. 2, November 2004.

[3] I. Potamitis, H. Chen, and G. Tremoulis, "Tracking of multiple moving speakers with multiple microphone arrays," *IEEE Trans. on Speech and Audio Processing*, vol. 12, no. 5, 2004.

[4] A. Brutti, M. Omologo, and P. Svaizer, "Comparison between different sound source localization techniques based on a real data collection," in *Hands-Free Speech Communication and Microphone Arrays (HSCMA)*, May 2008.

[5] T. Schmid, R. Shea, Z. Charbiwala *et al.*, "On the interaction of clocks, power, and synchronization in duty-cycled embedded sensor nodes," *ACM Trans. on Sensor Networks*, vol. 7, no. 3, 2010. [Online]. Available: http://doi.acm.org/10.1145/1807048.1807053

[6] D. T. Blumstein, D. J. Mennill, P. Clemins *et al.*, "Acoustic monitoring in terrestrial environments using microphone arrays: applications, technological considerations and prospectus," *Journal of Applied Ecology*, vol. 48, no. 3, 2011. [Online]. Available: http://dx.doi.org/10.1111/j.1365-2664.2011.01993.x

[7] D. J. Mennill, M. Battiston, D. R. Wilson *et al.*, "Field test of an affordable, portable, wireless microphone array for spatial monitoring of animal ecology and behaviour," *Methods in Ecology and Evolution*, vol. 3, no. 4, 2012.

[8] A. Canclini, F. Antonacci, A. Sarti, and S. Tubaro, "Acoustic source localization with distributed asynchronous microphone networks," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 21, no. 2, 2013.

[9] H. Wang, C. E. Chen, A. Ali *et al.*, "Acoustic sensor networks for woodpecker localization," in *Proc. of SPIE Conference on Advanced Signal Processing Algorithms, Architectures, and Implementations*, vol. 5910, August 2005.

[10] A. Ledeczi, G. Kiss, B. Feher *et al.*, "Acoustic source localization fusing sparse direction of arrival estimates," in *Proc. of Int. Workshop on Intelligent Solutions in Embedded Systems*, June 2006.

[11] M. J. Crocker, *Handbook of Acoustics*. Wiley, 1998. [Online]. Available: http://books.google.com/books?id=YoxQAAAAMAAJ

[12] D. Pavlidi, M. Puigt, A. Griffin, and A. Mouchtaris, "Real-time multiple sound source localization using a circular microphone array based on single-source confidence measures," in *Proc. of IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, 2012.

[13] D. Pavlidi, A. Griffin, M. Puigt, and A. Mouchtaris, "Source counting in real-time sound source localization using a circular microphone array," in *Proc. of IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2012.