

---

## LONG RANGE SOUND SOURCE LOCALIZATION EXPERIMENTS

Flaviu Ilie BOB

*Faculty of Electronics, Telecommunications and Information Technology  
Technical University of Cluj-Napoca  
26-28 George Barițiu Street, 400027 Cluj-Napoca, Romania  
ilie.bob@bel.utcluj.ro*

**Abstract:** Sound Source Localization (SSL) can be used in various domains, ranging from surveillance and security to medicine or multimedia. Most SSL systems rely on different knowledge about the emitted sound or the environment in order to determine the position of the source. In this paper the author presents the experimental results obtained for an unknown sound source. The distance towards the sound source is determined without prior knowledge about the emitted sound, and the environment is assumed to be noise free, including reverberations. Using a system that has only three microphones placed at 1 meter apart, experiments were conducted at ranges of up to 50 meters, at various angle orientations. Results show that it is possible to measure within an acceptable degree of error the distance towards a sound source situated up to 35 meters away.

**Keywords:** *Sound Source Localization, Time Delays of Arrival, three microphones localization, unknown acoustic source localization, General Cross-Correlation, Sound Source Ranging.*

### I. INTRODUCTION

Unknown single acoustic emission localization in terms of distance is a complex problem. Simple solutions that do not require prior training are not able to estimate the distance at high ranges. Most solutions available are limited in range (up to 2 meters) for systems that have the same size (of 2 meters). Slightly bigger ranges can be determined using complex systems. These are expensive and difficult to use considering size and cost [1]. For artificial intelligence to make assumptions about the environment properties or retain great amount of data for sound recognition is challenging and implies unjustified costs.

Sound localization has been first reported during the World War I, in the form of a system that captured the sounds emitted by the enemy artillery as it fired and sent them to a head quarter where the calculations were done. Microphones were spread over the battle field, thus the system was not easily portable. It was also easily affected in case of communication lines interruption. The army has not lost its interest into this topic, continuing even nowadays to develop different passive sound localization systems in order to fulfill anti aircraft missions or to quickly locate snipers [1].

Teleconferencing systems are the most known sound localization devices to the general public. The speaker's location is detected and the camera is automatically pointed towards him.

The greatest efforts in the sound localization problem have been done in robotics, trying to teach robots having a maximal manlike behavior or to do special tasks concerning sound localization. Also another important use for these methods is the interaction between robots and humans. This is done by using computational voice detection and computational linguistics.

Another very important field is the development of artificial hearing aids for persons with a hearing disorder.

Human prosthesis will use sound localization consequent with further development of the technology. Next step for the sound localization domain is to be extended in surveillance and security domain. Sound localization is less invasive than video surveillance, and is effective even in low or no visibility conditions.

The current state of research provides mostly systems that are able to localize only the direction of the sound source. The few systems that are able to determine also the distance of the source require complex computational power or complex structures of statistic data for prediction [2]. Databases containing pre-recorded samples or environment geometric attributes, models or propagation characteristics are the ones that load the memory, and the processors are kept busy by the advanced calculations that have to be done. Portability and simplicity is not their strong point. Along with complexity comes the size, as the main used processing unit is the PC. Also wires, advanced sensors, or other related equipment increase the volume of the systems.

We aim to design a system that is small, portable and capable to locate a single unknown acoustic emission. This must be a low cost and easy to use system. Having met all this conditions, a solution for various fields would be obtained. Some uses of this solution would be in surveillance and security, medicine, multimedia and robotics.

### II. PRINCIPLE, METHOD AND ALGORITHM

In order to describe the specific problem of localizing a sound source, three main components have to be considered: the source, the environment and the acquisition system. This brings a lot of uncertainties that a passive localization system has to cope with.

A sound source has some characteristics that influence the localization [3]. Directivity and constancy are two of them, but its own noise should not be neglected either. Also

the sound emitted from the source has parameters like frequency and power that can bring prejudice to the localization.

Presuming the source and sound are reliable, the environment is the next factor to be considered. It can be quiet, noisy, reverberant or in specific labor applications even acoustically dead. The environment also has humidity, temperature and texture that influence attributes like the speed of sound. Since a signal has to propagate through a medium it can also be easily influenced by noise, which can be other sounds or reverberations of the source itself.

Sound transducers are also perturbing the signal, as there are no two identical transducers. Components that amplify or repeat the signal in order to get acquired by an analog to digital converter also insert noise in the system. In digital domain conversion errors and mathematical rounding errors may appear.

A sound emitted from a source travels a certain distance to the sound receptor. Different positioned microphones means different distances for the sound to travel and using the speed of sound, the distance to each microphone is covered in a different time interval. Not knowing the exact moment when the sound was emitted, we can only use the Time Delays Of Arrival (TDOA) between the microphones. Based on these delays, implicitly on the differences of distance, the position of the source can be calculated using at least two differences.

The used system is based on a linear configuration of three microphones, as shown in Figure 1. The distance between each pair of center and side microphone is 1 meter.

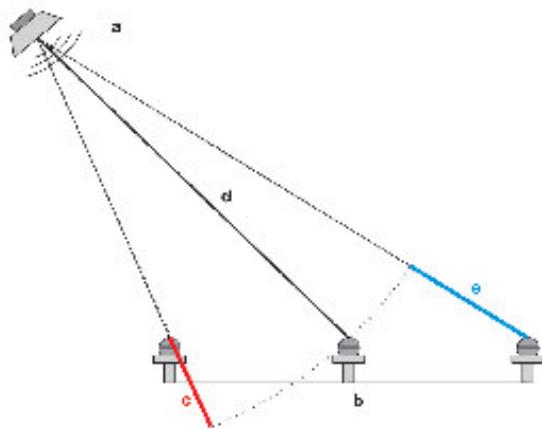


Figure 1. Microphone setup: a) Sound source b) Measuring device c) Difference between Left and Center microphones d) Distance form source to device e) Difference between Center and Right microphones.

The TDOAs differ depending on the position of the source regarding the microphone setup. If the distance between the source of sound and the microphone setup remains constant, and only the angle changes, the ratio of the delays changes. Similar, if the distance changes and the angle remains constant, the sum of the delays changes.

Using these observations, based on the ratio and sum of the delays, the exact position of the source can be calculated. The delays obtained from the two pairs of center and side microphone, annotated  $Samples_R$  and  $Samples_L$ , are measured in Analog to Digital Converters (ADC)

samples. These are multiplied with the Sampling Period ( $T_S$ ) and the Speed of Sound ( $v_s$ ) in order to obtain a meter measurable value. It is all divided by  $10^6$  to obtain a result measured in meters.

$$DR = \frac{Samples\_R \cdot T_S \cdot v_s}{10^6} \quad (1)$$

$$DL = \frac{Samples\_L \cdot T_S \cdot v_s}{10^6} \quad (2)$$

After having the two differences measurable in meters ( $DR$  and  $DL$ ), two intermediary variables are used: the sum ( $S$ ) and the ratio ( $R$ ) of the two differences. The sum is calculated in Equation (3).

$$S = DL + DR \quad (3)$$

To calculate the ratio, the next step is to establish the *orientation* of the sound source, meaning determining in which half-plane the sound source is located. This is realized by comparing the values of  $DL$  and  $DR$ . If they are equal, the source is located right ahead. If  $DR$  is greater, then the source is located in the left half-plane and Equation (4) is used, otherwise it is located in the right half-plane and Equation (5) is suitable.

$$R = \frac{DL}{DR} \quad (4)$$

$$R = \frac{DL}{DR} \quad (5)$$

In Equation (6) we introduce parameter  $a$ . It is equal to the distance between the microphones, how far are *apart* the left and the center microphones, respectively the center and the right ones.

$$Dt = \frac{-2 \cdot a^2 - 4 \cdot R \cdot a^2 - 2 \cdot R^2 \cdot a^2 + S^2 + S^2 \cdot R^2}{2 \cdot S \cdot (1+R)^2} \quad (6)$$

Having determined the distance, an angle determination is also possible, using Equation (7).

$$An = \frac{360}{2\pi} \cdot \arcsin \left( \frac{(2 \cdot Dt - S) \sqrt{-S^2 + 4 \cdot Dt \cdot S + 4 \cdot a^2}}{4 \cdot a \cdot Dt} \right) \quad (7)$$

The method is based on computing TDOAs. For determining the time differences, the General Cross-Correlation (GCC) is the most reliable method. Considering the above mentioned, we can state that the performance of the system is highly dependent on the capability of determining the TDOAs and the accuracy of the determination.

### III. SYSTEM AND LIMITATIONS

For the sound to be acquired, the electret microphones are used as transducers. The signal is then amplified and fed to the analog inputs of the Microchip® dsPIC33F256GP710A processor, equipped on an Explorer16 Development Board. The microprocessor determines the TDOAs, calculates the sound source parameters (angle and distance) and shows them on the display. The system's block schematic is presented in Figure 2:

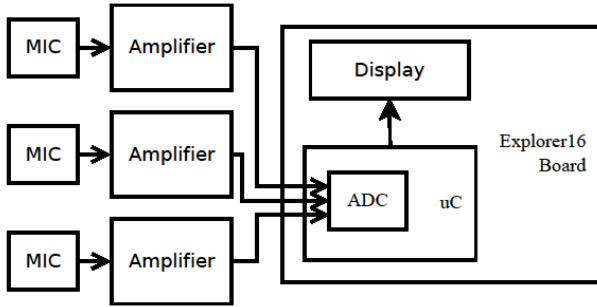


Figure 2. Hardware block schematic

The microprocessor's ADC uses a *Sampling Frequency* of 256.416 kHz, and the microphones are placed at a *Distance Apart* equal to 100 centimeters. The microprocessor determines the distance and the angle based on the delays between the received signals measured in number of ADC Sampling Periods ( $T_S$ ). To highlight the system limitations Matlab® simulations were performed. We simulated the position of the source for distances from 1 to 50 meters and an angle range of values from -45.0 to 45.0 degrees. Simulation steps of 0.01 meters for distances and 0.1 degrees for angles were used. A distribution of the possible calculated results was obtained, and using the *Comprehensive Polar Plots* [4] the simulation results are shown in Figure 3:

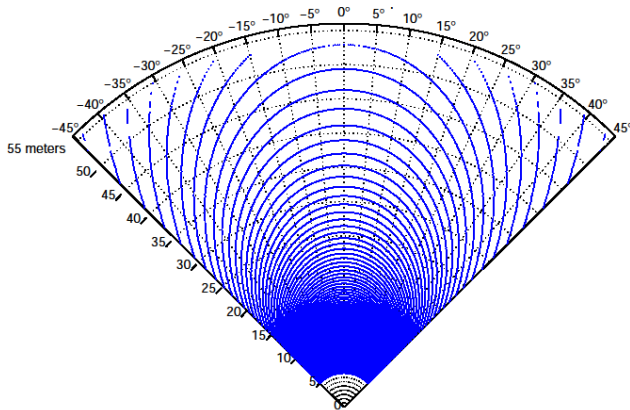


Figure 3. Distribution of possible results.

For each point of simulation, the corresponding TDOAs were calculated and rounded to the nearest integer value as the obtained delays by the microprocessor cannot be floating point results. Using the rounded value, the distance and angle towards the sound source were calculated in the same

manner as the microprocessor would have done.

Due to sampling rounding stated before, there are visible gaps between "lines" of distances. This means that for more points of simulation, the same distance and angle is obtained. Ideal would have been the results scattered uniformly.

The main cause for not obtaining a smoother distribution is the TDOAs determined between the channels. The delays are determined by the GCC algorithm in steps relative to *Sampling Frequency* ( $f_s$ ), as the delays can only be measured in integer values of *Sampling Period* ( $T_S$ ). Equation (6) stated in Section II emphasizes the fact that the calculated distance towards the sound emission is also directly proportional with the *Distance Apart*. This is another parameter that does not allow a smoother results distribution.

Detecting an ongoing emission of a periodic signal, like a sine of audio frequency is also an issue. To obtain a correct time delay between signals, using cross-correlation, the delay between them must be smaller than half the signal's period:

$$t_d < \frac{1}{2 \cdot f} \quad (8)$$

Considering the worst case scenario in which the source is co-linear with the three microphones, the maximum delay between two channels is relative to the *Distance Apart* ( $d_a$ ):

$$t_{d_{\max}} = \frac{d_a}{v_{\text{sound}}} \quad (9)$$

Combining Equations (8) and (9) we obtain the restrictions:

$$d_a < \frac{v_{\text{sound}}}{2 \cdot f} \quad (10)$$

$$f < \frac{v_{\text{sound}}}{2 \cdot d_a} \quad (11)$$

Calculating the parameters for our system, results that for 20 kHz sound frequency, at  $v_{\text{sound}} = 346.13 \text{ m/s}$  at a temperature of  $25^\circ \text{C}$ , the maximum *Distance Apart* between two microphones would be:

$$d_{a_{\max}} = \frac{346.13}{2 \cdot 20000} = 0.00865325 \text{ m} = 8.65 \text{ mm} \quad (12)$$

This would be great if we had a greater sampling frequency, as the portability of the system would be considerably improved, dare I say, beyond "pocket size".

This case however applies only for continuous emission signals. Also the presented worst case scenario is for a source co-linear with the microphones. In this study we consider that the sound source emission is not continuous and the frequency of the sound is not constant. We also imposed limits regarding the minimum distance towards the

sound source and the maximum deviation of the source from the straight ahead direction.

**IV. RESULTS**

Experiments were made in an open environment, a free field garden, with low echo and distortions. We did not consider multiple sources. Although, the test were subject of wind and other external conditions.

Experiments were done for distances between 1 to 50 meters, at different angles. Errors were introduced by approximating positioning of the sound source, as far as 0.1 meters, as the sound source was manually positioned at the desired location.

We used unknown test tones, such as hand claps or sounds produced by metal stroke with a hard object. Therefor, there are no two sound emissions alike. We also used other sound emissions, like yelling. All the sound emissions were recorded for later use. This allows a later sound processing in order to obtain the best results.

The configuration of the tests placement is shown in Figure 4. The distance shown is the ideal one, an the angle is the one determined by our system, using GCC.

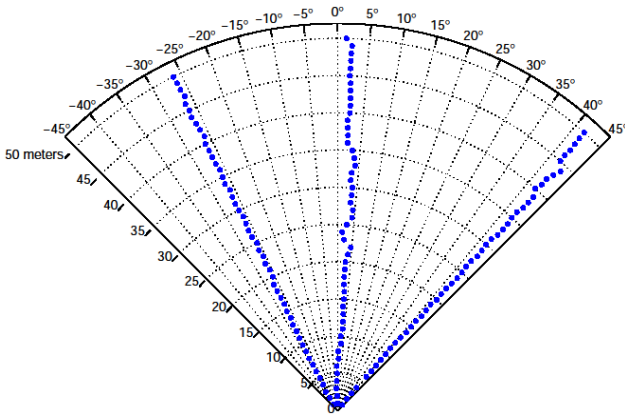


Figure 4. Distribution of experiment instances.

The measurements done and processed using GCC proved to be quite accurate, their representation being shown in Figures 5, 6 and 7:

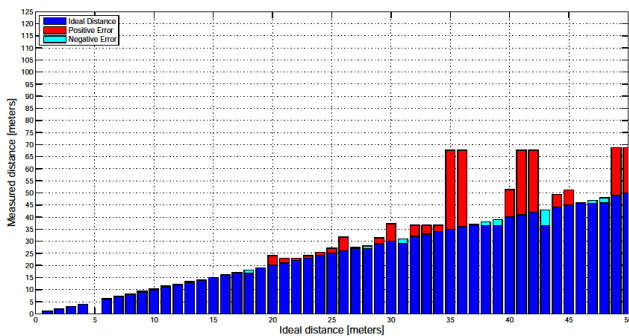


Figure 5. Performed measurements and corresponding errors for the left direction.

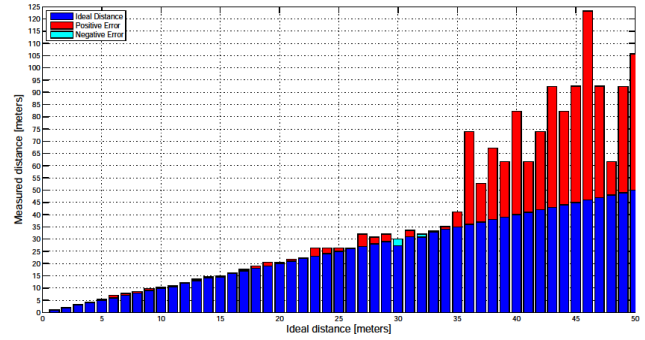


Figure 6. Performed measurements and corresponding errors for the ahead direction.

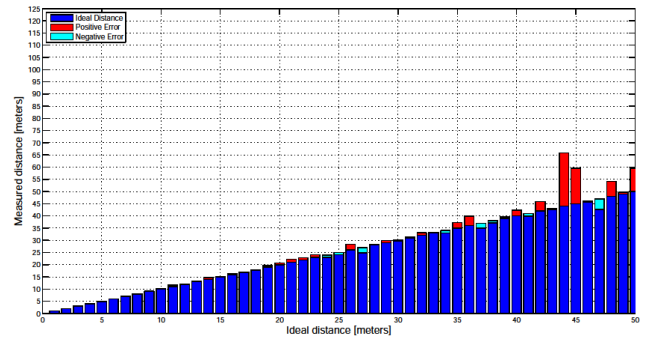


Figure 7. Performed measurements and corresponding errors for the right direction.

For the *Ahead* direction the measurements contain the biggest errors, but only for high distances. As the angle increases, which is the case of *Left* direction at about -25 degrees, the errors become smaller, and for the *Right* direction where the angle is around 40 degrees the errors are the smallest. The cause for this is that the delays between the microphones increase as the angle deviates from the 0 degrees direction. After centralizing all the test results, the system using GCC was proven quite reliable obtaining an overall *mean error* of just 16.75%. The *maximum error* although was as high as 168%:

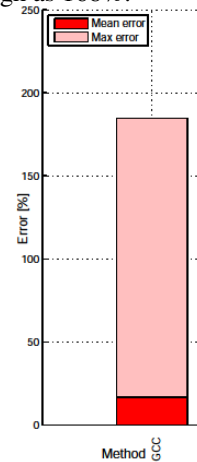


Figure 8. Overall error.

Because the experiments were done for 3 different directions, a statistic for every one of them was also calculated, and shown in Figure 9. Here is visible that the *Ahead* direction reported the worst results. This is caused by the small TDOAs available for high distances at straight forward direction. The main cause is the *Sampling Frequency* that could be bigger, even though the used value of 256.416 kHz is above the one used by most audio systems.

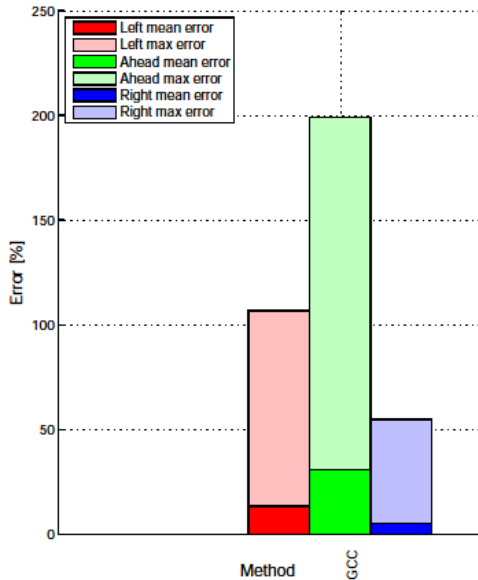


Figure 9. Overall direction error.

Observing the maximum errors in Figure 9, we can draw the conclusion that the maximum errors, that most likely appear at high ranges, influence the mean error in a negative way. A statistic was done for 3 intervals having as "critical points" the distances of 19 and 34 meters. Splitting the three intervals on directions, in Figures 10, 11 and 12 we can see that the *Ahead* direction has always reported the worst results. Also the tested system reports errors below 25% for distances up to 35 meters.

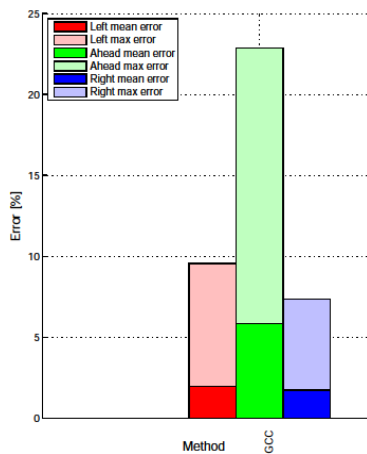


Figure 10. Direction error for the [1...19] meters interval.

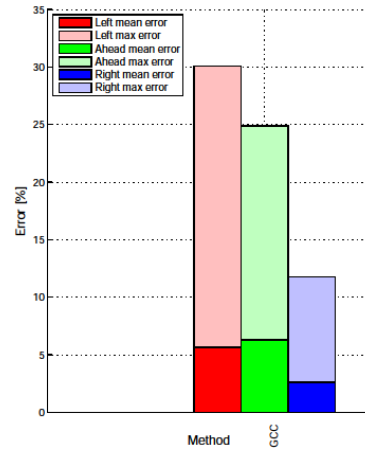


Figure 11. Direction error for the [1...34] meters interval.

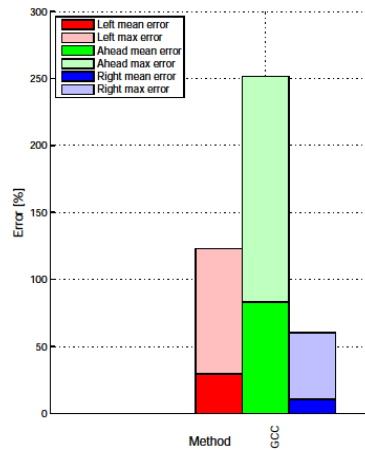


Figure 12. Direction error for the [35...50] meters interval.

In Figure 13 we can see that the *mean error* for distances up to 19 meters is 3.17%, with the *maximum error* not overcoming 17%. Distances from 1 to 34 meters also report a small value *mean error*, of 4.86%. Looking at the error for the 35 to 50 meters the *mean error* goes as high as 41.22% and reaching the maximum value of 168%.

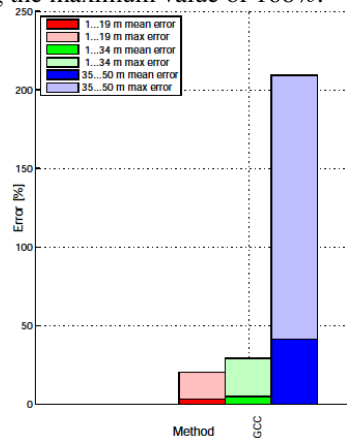


Figure 31. Overall interval error.

## V. CONCLUSIONS

We successfully implemented a system for determining the distance towards an unknown single sound emission in an homogeneous noise free and non-echoic environment. There are some restrictions imposed for our experiments, but we proved the feasibility of the system. It achieved the main goal of measuring the distance towards the sound source. The statistics provided show that for distances up to 35 meters the method and system proposed are viable, obtaining *mean errors* and *maximum errors* not exceeding 10% and respectively 30%.

The GCC has proven to be a simple, robust and precise method, mentioning that no other signal processing was done on the signals. This makes the method viable for use in an embedded system. Its applications would require it to be used at a great scale, either in combination with other devices like cameras and other sensors or robots and airborne drones for surveillance, rescue missions and other applications. There is also space for future improvements of the system, depending on the desired application. The dimensions of the system itself are reduced, considering that the microphones can be mounted on telescopic supports. Being small, the system is portable and can be deployed and used easily. The *Sampling frequency* is the most necessary feature that must be improved, considering the reduced dimensions that the system must preserve.

A calibration algorithm could improve the results, as the positioning of the microphones raises complex issues in the physical design. The calibration algorithm could also be used for measuring the sound intensity at the receiver. In use with other sensors like humidity and temperature, both accuracy and capabilities could be improved, as this would allow the system to use a more accurate speed of sound and also to determine the sound intensity at emission place. The addition of echo-cancellation algorithms to the design could lead to usage of the system in reverberant environments.

Our study acts as a milestone for future studies, easing access to starting point for further implementations for any sound source localization system. The system could eventually locate simultaneously multiple sound sources. Certain algorithms were already implemented, but experiments on large distances were not performed.

Depending on the necessities, the system could raise from a cheap embedded solution to a great processing factory, but still fulfilling the mission of unknown sound source localization, as proven in this study.

## ACKNOWLEDGEMENT

This paper was supported by the project "Improvement of the doctoral studies quality in engineering science for development of the knowledge based society-QDOC" contract no. POSDRU/107/1.5/S/78534, project co-funded by the European Social Fund through the Sectorial Operational Program Human Resources 2007-2013.

## REFERENCES

- [1] C. Lenz, "Localization of sound sources", ETH Zurich, Tech. Rep., 2009.
- [2] B. Choi, "Acoustic source localization in 3d complex urban environments", Master's thesis, Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, Virginia, April 2012.
- [3] A. M. Ali, K. Yao, T. C. Collier, C. E. Taylor, D. T. Blumstein, and L. Girod, "An empirical study of collaborative acoustic source localization," in *Proc. 6th International Conference on Information Processing in Sensor Networks (IPSN'07)*, Apr. 2007.
- [4] D. Hanselman, *Matlab Central File Exchange - Comprehensive Polar Plots*, University of Maine, October 2012. [Online]. Available: <http://www.mathworks.com/matlabcentral/fileexchange/38855-comprehensive-polar-plots/content/mmpolar.m>