

# **VIDEO TRACKING USING ACOUSTIC TRIANGULATION**

An Honors Fellow Thesis

by

ALEXANDER IVANOV IVANOV, ALEXANDRU RADUCANU

Submitted to Honors and Undergraduate Research  
Texas A&M University  
in partial fulfillment of the requirements for the designation as

**HONORS UNDERGRADUATE RESEARCH FELLOW**

May 2012

Major: Electrical Engineering

# VIDEO TRACKING USING ACOUSTIC TRANGULATION

An Honors Fellow Thesis

by

ALEXANDER IVANOV IVANOV, ALEXANDRU RADUCANU

Submitted to Honors and Undergraduate Research  
Texas A&M University  
in partial fulfillment of the requirements for the designation as

HONORS UNDERGRADUATE RESEARCH FELLOW

Approved by:

Research Advisor:

Associate Director, Honors and Undergraduate Research:

Jean-Francois  
Chamberland  
Duncan MacKenzie

May 2012

Major: Electrical Engineering

## ABSTRACT

Video Tracking Using Acoustic Triangulation. (May 2012)

Alexander Ivanov Ivanov, Alexandru Raducanu  
Department of Electrical Engineering  
Texas A&M University

Research Advisor: Dr. Jean-François Chamberland  
Department of Electrical and Computer Engineering

This study focuses on the detection and triangulation of sound sources. Specifically, we focus on the detection of sound in order to track a person's position with a video camera. Acoustic tracking, an alternative to visual tracking, is relatively inexpensive, passive (does not emit energy), and effective in low lighting environments. Our project is broken into two major aspects: accurately discerning input as opposed to background noise and the localization of the sound source. In order to focus on the input signal, we analyze two methods: time averaging and impulse culling. After the sound is analyzed and filtered we focus on the triangulation of the source in 2-D space using direct and estimation techniques requiring three microphones. This process is geared towards finding a compromise between performance and complexity which allows implementation on a standard micro-controller. Results show that this method of tracking is feasible with modern micro-controller technology. In addition, preliminary data shows algorithm optimization is required to reduce computational intensity.

## ACKNOWLEDGMENTS

I would like to thank everyone who has contributed to this project. Although I no longer remember all the conversations I've had, many friendly fellow engineers had a contribution to this work. More specifically, I would like to thank: Andrew Targetta for helping us debug microcontroller issues despite have no time or reason to do so, my father Ivan Ivanov for giving ideas on mathematical approaches to our problem and methods, Dr. Chamberland for advising us and making this project possible, and my partner Alexandru Raducanu for putting in serious work and putting up with me along the way.

## NOMENCLATURE

TDOA	Time Difference of Arrival
FPGA	Field Programmable Gate Array
AOA	Angle of Arrival
IID	Independent Identically Distributed
BLUE	Best Linear Unbiased Estimator
ROC	Region of Convergence

## TABLE OF CONTENTS

	Page
ABSTRACT .....	iii
ACKNOWLEDGMENTS.....	iv
NOMENCLATURE.....	v
TABLE OF CONTENTS .....	vi
LIST OF FIGURES.....	vii
LIST OF TABLES .....	viii
 CHAPTER	
I INTRODUCTION.....	1
II METHODS.....	3
Assumptions .....	4
Explanation of System Setup .....	4
Sampling Means .....	5
Explanation of Cross Correlation .....	5
Explanation of Geometric System of Equations .....	7
Hyperbolic Localization Using Maximum Likelihood .....	8
Gauss-Markov Theorem: Linearization and Tracking .....	12
Analysis Techniques .....	14
III RESULTS.....	16
Data Acquisition Setup.....	16
Initialization .....	18
Convergence of Estimation Methods .....	19
Computational Speed .....	22
IV CONCLUSIONS.....	24
REFERENCES.....	25
CONTACT INFORMATION .....	26

## LIST OF FIGURES

FIGURE	Page
1 Graph of Shifted Input Signals .....	6
2 Cross Correlation of Input Signals .....	7
3 Pictorial Representation of Physical Setup .....	8
4 Hyperbolic Equations with Noisy Measurements .....	9
5 Hyperbolic Sheath .....	10
6 Overall Probability Distribution .....	10
7 Region Restriction .....	12
8 Gauss-Markov Parameters .....	13
9 Representation of ‘Impulse Jump’ .....	15
10 Noise Induced Position Differences Near Coordinate Axis .....	17
11 Calculated Initial Points vs. Source Location .....	18
12 ROC with Perfect Fidelity Data .....	20
13 Empirical ROC .....	20
14 Effects of Impulse Culling on ROC .....	21
15 Effects of Time Averaging on ROC .....	21

## LIST OF TABLES

TABLE	Page
1 Comparison of Computational Times vs. Search Coarseness .....	23



# CHAPTER I

## INTRODUCTION

In the past 10 years, the use and implementation of cameras has increased dramatically. Cameras are now part of almost all cell phones, present in many buildings, and will soon survey every intersection. Still, camera placement is not effective if the camera does not focus on desired objects. Recently, the concept of autonomous visual tracking was considered by Ling and Mei [1]. However, this approach is very expensive, maladjusted to drastic illumination changes [2]. In addition, it is ineffective if the subject does not display key visual elements (e.g. face) in the camera's view [1].

A more subtle approach used to focus on a desired subject is acoustic triangulation. By accurately focusing in on human speech, a person's position can be determined and homed in on without the use of visual elements. Figueroa and Mahajan, 1994 [3] showed triangulation of wave sources with self-calibration for the speed of propagation of the wave is possible using only four sensors (microphones).

Elsewhere, Furguson et al. [4] showed that, even with long range sensors spanning several kilometers, a high degree of accuracy in a two-dimensional (2D) plane (15-25

---

This thesis follows the style of *IEEE Transactions on Consumer Electronics*.

meters) can be achieved using only two sensors. Furguson also showed that a large majority of this error is due to instabilities in sound speed and noise. In another 2D experiment, Canistraro and Jordan [5] showed that the location of an impacting object can be acoustically triangulated within 0.05 meters in a 5.25 square meter area.

We predict that, by reducing noise through voice isolation and applying some of the methodologies above, a more accurate solution for video acoustic tracking may be possible. Acoustic triangulation can make autonomous video tracking of individuals more affordable, accurate, and effective in relatively short range homogeneous environments such as conference rooms.

## CHAPTER II

### METHODS

Determining the location of a sound source in space can be done by calculating the Time Difference of Arrival (TDOA) between pairs of microphones. That is, determining the difference between when one microphone detects a sound and a second microphone detects the same sound. By using this information and the known location of these microphones, the location of a sound source can be estimated. Locating a source using TDOA can be broken down into two steps:

- 1) The first step consists in computing the TDOA. After this, the difference in length from the source to each microphone can be estimated.
- 2) The second step is to obtain a solution for the location of source by utilizing the TDOA of pairs of microphones in an algorithm.

To compute the TDOA, several methods were studied: cross correlation, source intensity, and crosspower-spectrum phase. Of these three methods, the cross-correlation was used in our implementation. Several factors influenced this decision including low computational intensity, ease of implementation by a microcontroller, and robustness to noisy measurements.

Several different methods were taken into consideration for performing the second stage: hyperbolic localization, a geometric system of equations, and estimation using the Gauss-Markov theorem. All three methods were tested in the calculation of our source location. Either the hyperbolic localization method or the geometric system of equations can be used to obtain the location of the source on a plane, after which the Gauss-Markov theorem is used to converge more closely to the sound source as well as provide tracking of the source.

### **Assumptions**

In order to use these methodologies we have made several assumptions about the nature of our system concerning hardware and sound propagation. The most important assumption we have made is that the source is a single point which emits sound radially. In addition, we assume the sound which reaches the microphones will not be heavily degenerated or disturbed by noise. Furthermore, we assume that the relative location of the microphones to each other is known. In addition, in order to constrain our problem, we assume that the source will be located within the first quadrant. The speed of sound was assumed to be 340 m/s. Finally, we assume that any noise inherent in our system can be modeled as a zero-mean, Gaussian, stationary random process.

### **Explanation of system setup**

The physical setup consists of an array of three microphones placed in an L shape. The array forms a right isosceles triangle. The center microphone acts as the origin of the

plane while the other two are separated by a distance of 20 cm along the x and y axis.

The source can take any location in the first quadrant.

### **Sampling means**

One of the most important factors in achieving a high fidelity TDOA is accurate sampling. The sound wave must be accurately sampled simultaneously at the three microphones at a frequency which achieves a high resolution. Although the Nyquist sampling criterion only dictates a frequency of twice the maximal sound frequency, we decided to utilize raw data instead of reconstructing our received signal. This was practical due to the large computational overhead associated with accurately reconstructing a signal. Several devices such as FPGAs, microcontrollers, and sampling software were considered. However, to accurately sample our data a dsPIC33 microcontroller was chosen. The effectiveness of this microcontroller in simultaneously sampling functionality and computational power made it a viable candidate for our procedure.

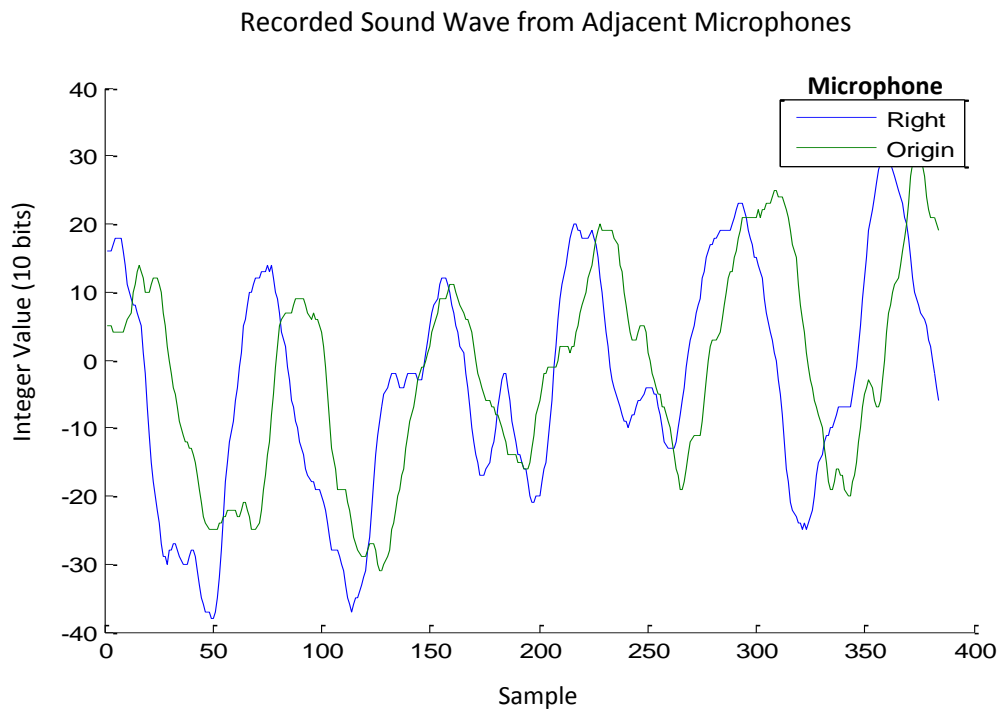
### **Explanation of cross correlation**

The cross correlation of two signals is a measurement of similarity between the signals. By overlapping two signals and multiplying them, one can see how 'similar' the two signals are. An example is shown in Fig. 1. By then shifting one signal in the time domain, similarity can be shown as a function of time. By finding the maximum of the cross correlation of two identical, but delayed signals, the time delay between the two

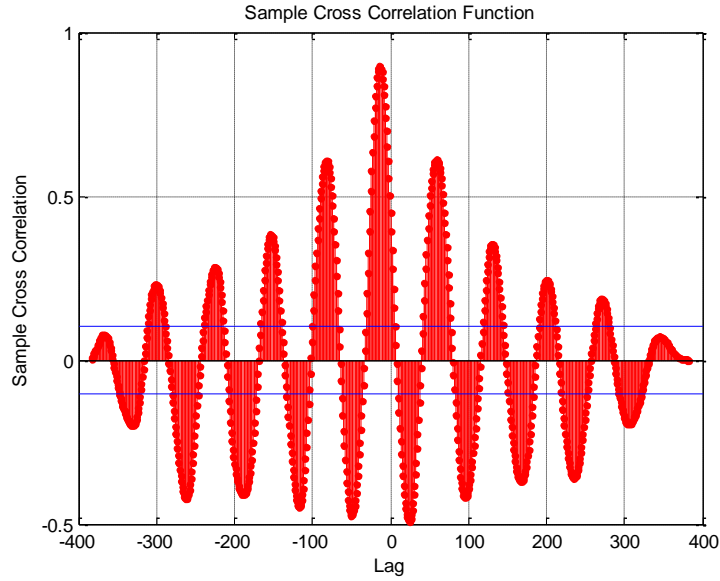
can be easily computed. The cross correlation is used to obtain the TDOA between sound sources in our setup. The cross correlation of signals in Fig. 1 is shown in Fig. 2.

The cross-correlation between two signals  $e$  and  $u$  is described by:

$$R_{eu}^N(\tau) = \sum_{k=1}^N e(k)u(k - \tau) \quad (1)$$



**Fig. 1. Graph of Shifted Input Signals**



**Fig. 2. Cross Correlation of Input Signals**

### Explanation of geometric system of equations

Traditionally, the location of a source was computed through a system of equations, which focused on the geometrical properties of the system. In this way, if the distance between microphones and the TDOA is perfectly known, then the location of the source can be obtained. Taking the physical setup of our system into consideration the system of equations we used is:

$$d_1^2 = l_1^2 + l_2^2 - 2l_1^2 l_2^2 \cos(\theta) \quad (2)$$

$$d_2^2 = l_3^2 + l_1^2 - 2d_1^2 l_1^2 \cos(\phi) \quad (5)$$

$$l_1^2 = d_1^2 + l_2^2 - 2d_1^2 l_2^2 \cos(\beta) \quad (3)$$

$$l_1^2 = d_2^2 + l_3^2 - 2d_2^2 l_3^2 \cos(\rho) \quad (6)$$

$$l_2^2 = d_1^2 + l_1^2 - 2d_1^2 l_1^2 \cos(\alpha) \quad (4)$$

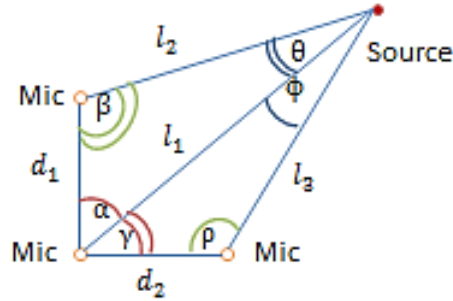
$$l_3^2 = d_2^2 + l_1^2 - 2d_2^2 l_1^2 \cos(\gamma) \quad (7)$$

$$l_2 - l_1 = \Delta L_{21} \quad (8)$$

$$l_3 - l_1 = \Delta L_{31} \quad (9)$$

$$\cos(\gamma)^2 \cos(\alpha)^2 = (1 - \cos(\gamma)^2)(1 - \cos(\alpha)^2) \quad (10)$$

The variables are implicitly defined in Fig. 3.



**Fig. 3. Pictorial Representation of Physical Setup**

A pictorial representation of our setup can be observed in Fig. 3. The first six equations were derived from laws of cosine: while the latter equation utilizes the fact that the microphones are separated by an angle which is an integer multiple of  $90^\circ$ .

Unfortunately, this method has computational overhead associated with the nonlinear nature of the system used. A linearization of this system is much more suited to low power applications such as a microcontroller.

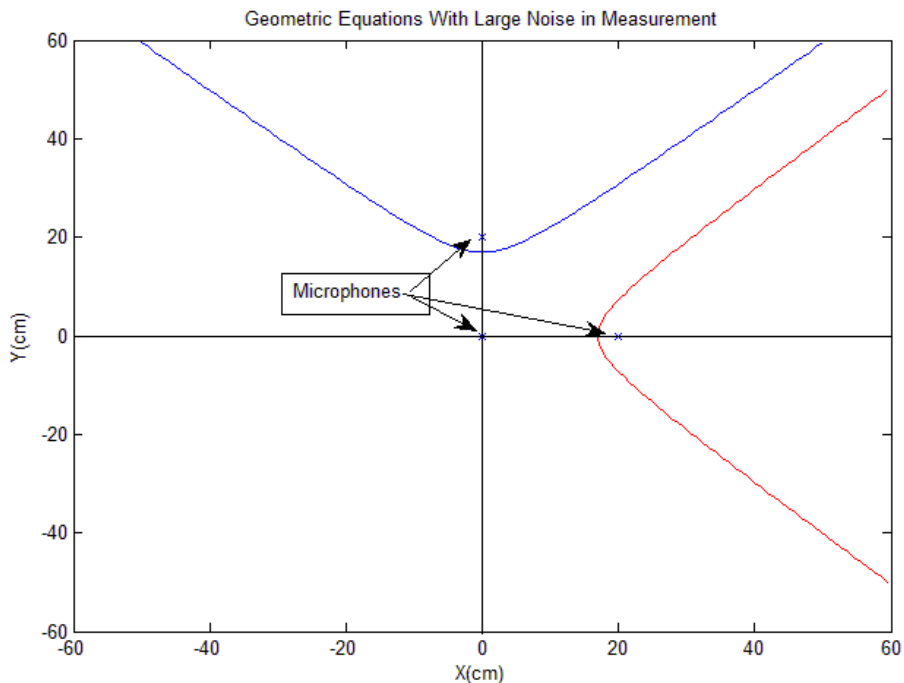
### **Hyperbolic localization using maximum likelihood**

#### *Hyperbolic localization*

A second method we consider for direct triangulation utilizes hyperbolic localization. This technique is well known and has been studied extensively. The concept behind this style of triangulation relies on the intersection of two hyperboloids. In the case of 2D triangulation, these are simple hyperbolas. In the 3D case, these are hyperbolic surfaces.



Using TDOA between two sets of microphones, we can create a hyperbola which is defined by the constraint: at all points on this curve, the difference in distance between the two microphones is constant. Fig. 4 shows an example of this.



**Fig. 4. Hyperbolic Equations with Noisy Measurements**

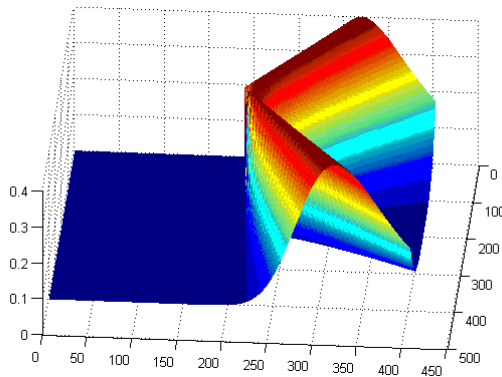
The benefit to using this method stems from non-ideal data acquisition. In the case of a non-ideal system, there is error associated with TDOA measurements. This means that the two hyperbolic surfaces may or may not intersect given the data. When no intersection occurs, traditional triangulation fails and there is no solution to our system of equations (see Fig. 4). In order to avoid this situation we couple this method with Maximum Likelihood analysis.

### *Maximum likelihood analysis*

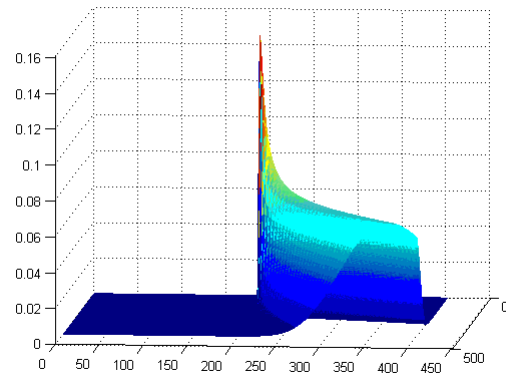
When taking a single hyperboloid into consideration, the effect of a corrupted TDOA measurement can be easily characterized. The equation below shows this,

$$[(x - x_{m1})^2 + (y - y_{m1})^2] - [(x - x_{m2})^2 + (y - y_{m2})^2] = D_s + \Delta D \quad (11).$$

Above,  $(x,y)$  is the location of the source and  $(x_{m1},y_{m1})$ ,  $(x_{m2},y_{m2})$  are the locations of microphones 1 and 2, respectively. The distance between source-microphone1 and source-microphone2 is modeled as the true difference with the addition of an error term  $\Delta D$ . This error term is directly proportional to an error in a TDOA measurement.



**Fig. 5. Hyperbolic Sheath**



**Fig. 6. Overall Probability Distribution**

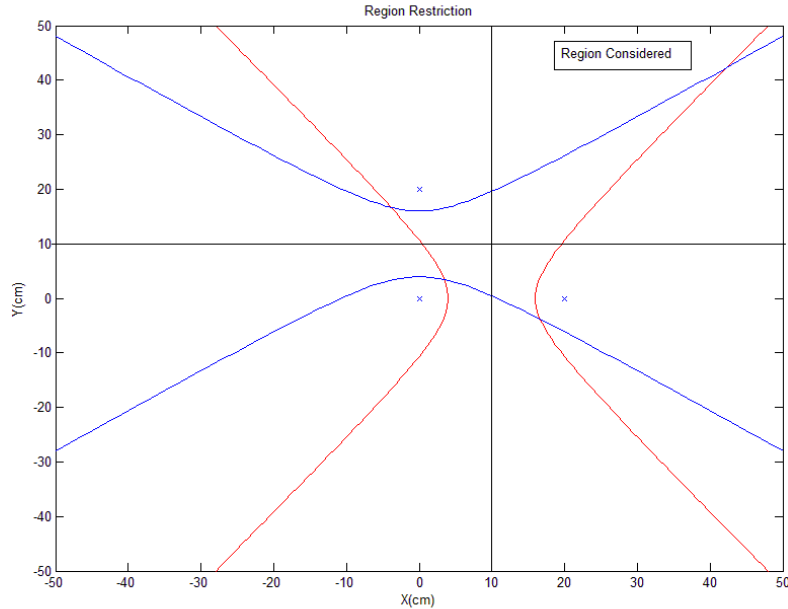
If we assume that errors in TDOA measurements can be modeled by Gaussian random variables with zero mean and finite variance, we can utilize this result. By solving for  $\Delta D$ , we can use the resultant equation as a parameter to a normal distribution. The resultant surface is a three dimensional hyperbolic manifold. An example is shown in Fig. 5.

$$P = \left[ \frac{1}{\sqrt{2\pi\sigma}} * e^{-\frac{\Delta D_1^2}{2\sigma^2}} \right] * \left[ \frac{1}{\sqrt{2\pi\sigma}} * e^{-\frac{\Delta D_2^2}{2\sigma^2}} \right] \quad (12)$$

Now, if we assume that the error using one set of microphones is Independent Identically Distributed (IID) with respect to that of the other set, we can multiply two resultant sheaths using two sets of microphones. This creates a probability surface with a sharp maximum, as seen in Fig. 6. This maximum corresponds to the most likely location of our source. Therefore, by finding the maximum of this function we can locate our source with high fidelity while being robust to error in the system.

### *Considerations*

Some considerations are necessary when utilizing this method. One of the most evident issues is that the equation of a hyperbola creates two distinct curves. In our case, one of the resultant curves does not correspond to a real measurement. In order to avoid superfluous results, we restrict our system to a subset of the first quadrant. By discarding points in other quadrants as well as those which correspond to negative TDOA measurements, we ensure the existence of a unique, real solution. An example is shown in Fig. 7.



**Fig. 7. Region Restriction**

### **Gauss-Markov theorem: linearization and tracking**

Once we have identified an initial or nominal location for our intended target, we no longer have to use computationally intensive methods to continue locating the source. Instead, the problem can be linearized and a Best Linear Unbiased Estimator (BLUE), used to iteratively converge on the location of the source. This is beneficial in two ways. First it reduces computational intensity and involves only solving a few linear equations. In addition, this method can realistically be used for tracking without any further complications. In other words, as the source moves the BLUE will simply begin to converge to its new location. Finally, it accounts for error in measurements and effectively eliminates its effect under certain conditions. A pictorial representation is shown in Fig. 8. The inputs to the system are again the time differences.

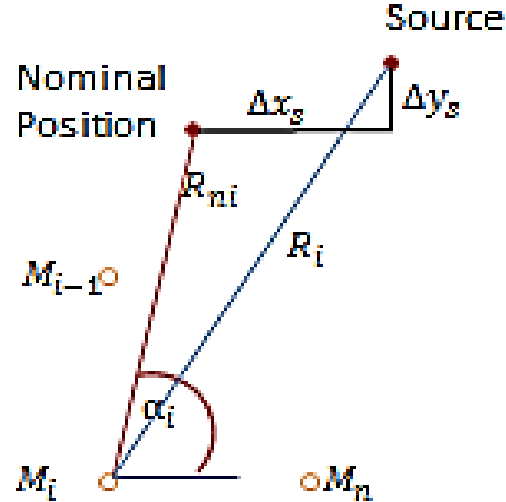


Fig. 8. Gauss-Markov Parameters.

The BLUE for our system utilizes the Gauss-Markov theorem [6]. Due to the fact that this is random modeling, it is important to note that the convergence to the source location is in fact a random process dependent on incoming data and the initial location. This estimation technique has been well studied and utilized in this type of problem. For a detailed description of system linearization, see [6].

A key feature of this method is that, due to linearization, it is only effective within the proximity of the actual location of the sound source. This means that this iterative process is limited to a Region of Convergence (ROC). This ROC is the area around the actual location of the sound source, which can be utilized as the initial guess for the

source's location. If this process is used with a point outside the ROC, it may diverge and lead to a solution which places the estimated source location at infinity.

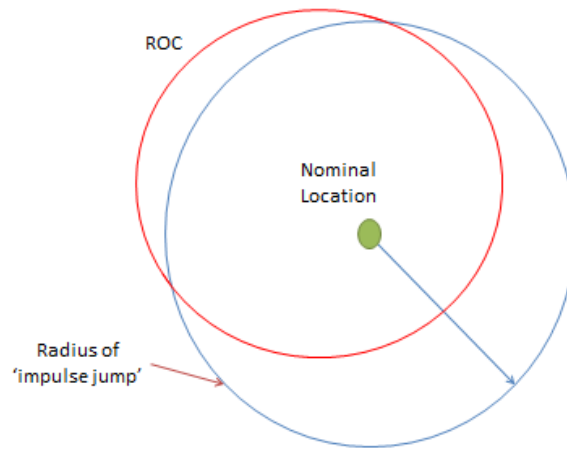
### **Analysis techniques**

Several methods are used to reduce the effects of noise and improve the ROC of our algorithm. These methods are time averaging and impulse culling.

#### *Impulse culling*

Impulse culling is a method of filtering out large amplitude, high frequency noise. This method is applied after the computation of time differences and is only applicable to the estimation and tracking stage of our methods. Impulse culling consists of discarding any input data which is considered an 'impulse'. In this case, an impulse is any input which creates a large change in estimated location with respect to Gauss-Markov estimation. This technique is especially relevant given the problem parameters.

In this case, it is known that the estimation technique only converges within a finite ROC. If data is obtained that would cause a change in estimated location with magnitude greater than the radius of the ROC, this data has a large likelihood of being erroneous and causing the estimated location to 'jump' outside the ROC and consequently diverge. By utilizing this type of filtering technique the stability region of the system can be improved.



**Fig. 9. Representation of 'Impulse Jump'**

### *Time averaging*

Time Averaging is a well-known filtering strategy which acts as a low pass filter. This technique is both applicable directly to incoming data as well as the results of estimation. In the estimation stage, time averaging can be used to give a more stable source location. By averaging over several computed source locations, this method effectively smooths the resulting location and filters 'jumpy' high frequency changes. This jump is seen in Fig. 9. This again is relevant due to problem specifications. In this experimental setup, human tracking is the main objective. As such, it is important to notice that human motion is much slower than the modeled white Gaussian noise. Therefore this noise would contribute to fast high frequency position changes and should be filtered. In this work, a simple two-point un-weighted time average is used. The previous location  $N_{i-1}$  is averaged with the new calculated location  $C_i$  to produce a new nominal location,

$$\frac{(N_{i-1} + C_i)}{2} = N_i \quad (13) .$$

## CHAPTER III

### RESULTS

The results section will be organized in chronological order in terms of procedures required from data acquisition to motor actuation. Each subsection will describe specific, related results.

#### **Data acquisition setup**

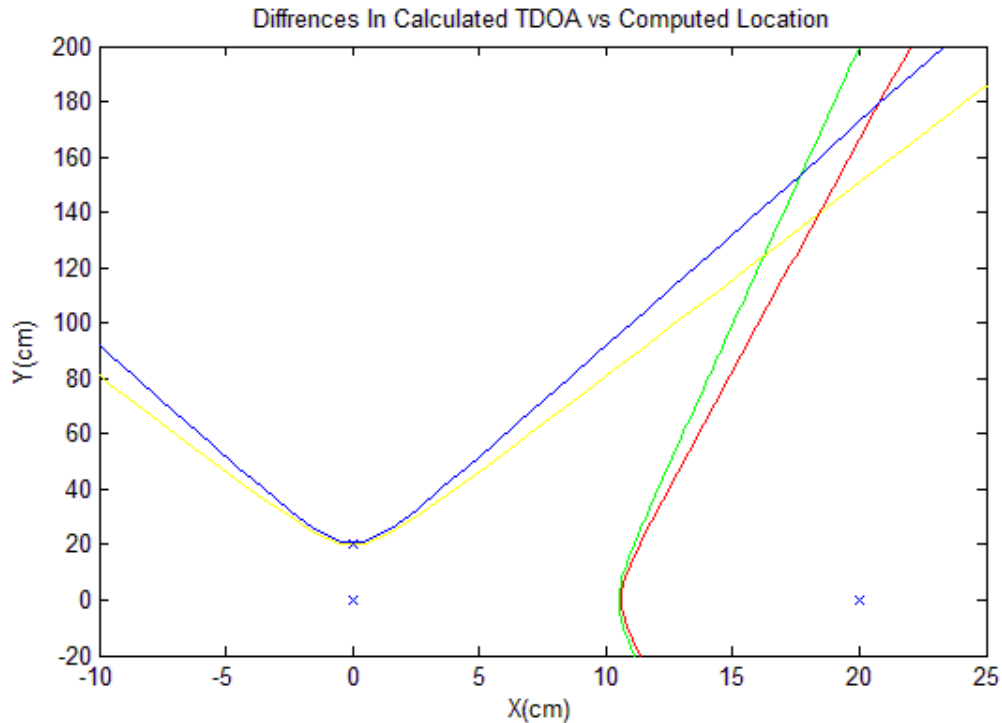
As described in Chapter II, our two dimensional data acquisition set up consisted of three microphones set in an isosceles right triangle with 20cm long legs. The results of this setup can be summarized by a tradeoff between compactness and error susceptibility.

The small distance between microphones proved desirable in terms of total area used. This is beneficial in terms of producing a viable, plug-and-play style device. Having this limited distance would allow for a product that could be stand alone and would not require intricate tuning, but could rather simply be brought into a desired location and utilized directly.

Unfortunately this small setup proved highly susceptible to error. In a case where the object being tracked is a significant distance away (5-10 meters), small changes in time differences acquired can lead to a large variation in the calculated location of the source.



An example of this phenomenon can be seen in Fig.10. Notice that the slight difference in time differentials corresponds to a large shift in location.

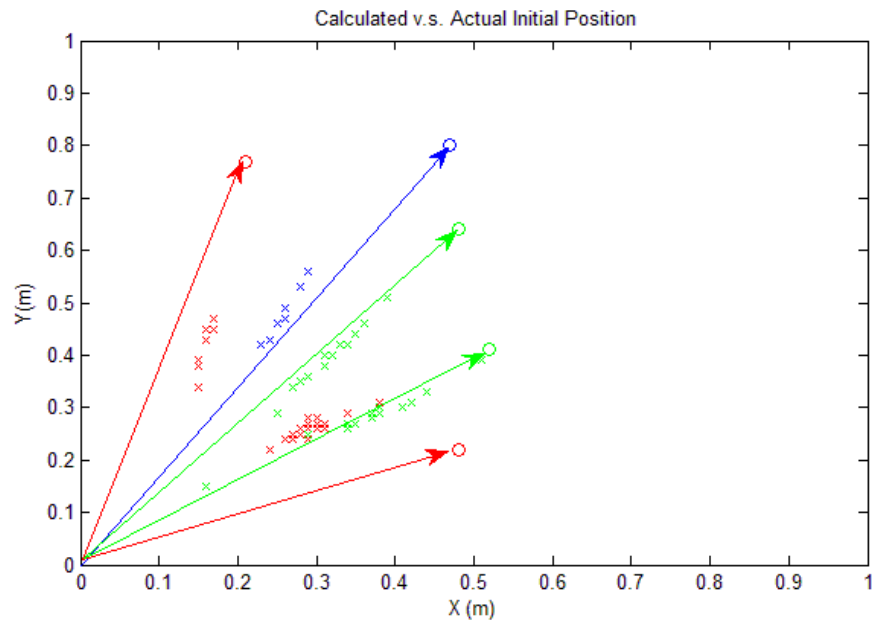


**Fig. 10. Noise Induced Position Differences Near Coordinate Axis**

Finally, it is important to note that due to the nature of our acquisition techniques, large errors are incurred when source locations are near each of the coordinate axes. It is difficult to quantify exactly where our methods break down because this location is extremely susceptible to small variations in error which can be caused by a variety of latent variables. These include, acoustic reflections caused by the physical environment, type of sound sampled, and any co-dependence of the error between microphone pairs.

## Initialization

As stated previously, we utilized the approach of greatest likelihood to calculate an initial or nominal position. This method proved highly accurate and efficient. By subdividing our search area into discrete intervals a high search speed could be achieved while not sacrificing much in terms of accuracy. There is a fundamental compromise which must be made in terms of search speed and accuracy. By more coarsely subdividing our search field, it is possible to reduce search time drastically. Unfortunately, an artifact of our small physical setup made the calculation of initial nominal positions highly variable due to the large impact of noise.



**Fig. 11. Calculated Initial Points vs. Source Location**

Despite this, the experimental setup managed to retain the directionality of the angle of arrival (AOA). Some typical results are shown in Fig. 11. For the problem being studied, this is what is essentially required. Even if the actual location of a source is incorrectly computed, as long as the AOA of the system is maintained, standard autofocus algorithms which are commercially available can be used to focus a camera on the intended target.

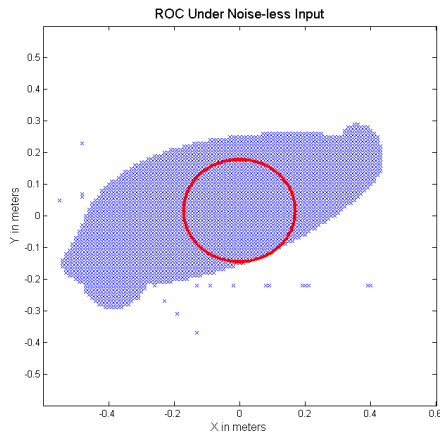
It is important to note that this method begins to have large error when a source is located radially near the axes. Again, the problem dictates that in practical application a source can be expected to stay within the region which maintains high accuracy. In addition, if microphones are located on the camera itself, they move as the camera tracks. As such the source will necessarily remain within the accurate region.

Finally, observations show that this technique was successful in computing initial points, which were within the ROC of our estimation methods.

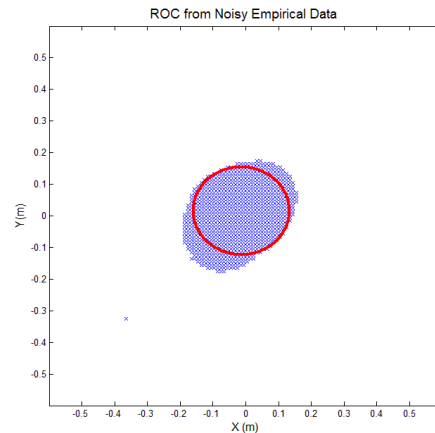
### **Convergence of estimation methods**

An important limitation of our linear estimation approach is the size of the Region Of Convergence (ROC) associated with the linearization of this nonlinear system. Two relevant measurements were taken to determine the ROC of our system. First, sample nominal points were utilized and coupled with perfect fidelity input data (simulated) to determine the ideal ROC. Second, the same nominal points were utilized with empirically acquired data to determine the empirical ROC. It is important to note that the

sample nominal points used were systematically chosen and not empirically calculated by our initialization phase. In addition, note that the decreased fidelity in acquired data corresponds to a decreased ROC. Fig. 12 and Fig. 13 show the ideal versus empirical ROCs.



**Fig. 12. ROC with Perfect Fidelity Data**



**Fig. 13. Empirical ROC**

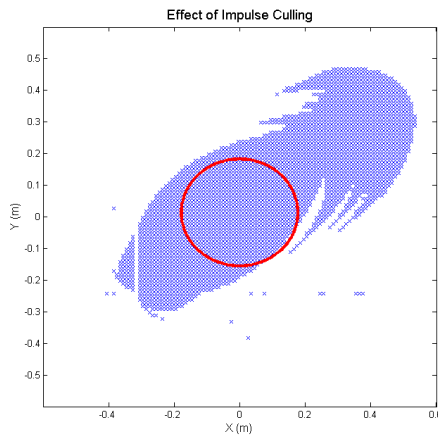
As noted in Chapter II, several well-known techniques for acoustic analysis were proposed to isolate the desired source from other potential erroneous sound sources. Due to time limitations, we were only able to utilize the two simplest forms of analysis proposed: time averaging, impulse culling.

#### *Time averaging and impulse culling*

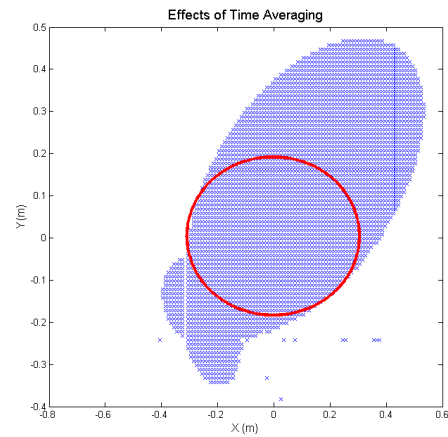
Time averaging was used in addition to 'impulse culling' to increase the ROC of our system and reduce susceptibility to high frequency position changes. These two methods were vital to the performance of our system because they allowed negation of data which

could potentially throw our tracking algorithm out of its ROC a cause divergence.

The results of impulse culling are somewhat qualitative and further experimentation is required to quantize this form of analysis. During implementation we noted that, due to the inherent error in initialization and incoming data, utilizing this culling technique was necessary in order to assure convergence in a majority of cases. In almost all tests where this technique was not used, our tracking algorithm failed to converge. In addition, an observed increase in the ROC of empirical data sets was observed. The maximum observed change was approximately 2.5 cm (Fig. 14). A more systematic analysis needs to be performed in order to accurately quantify the average ROC improvement.



**Fig. 14. Effects of Impulse Culling on ROC**



**Fig. 15. Effects of Time Averaging on ROC**

The results of time averaging were far more noticeable. By utilizing a simple, two point un-weighted average we were able to enlarge our ROC. The observed change in ROC was substantial in comparison to the change observed with impulse culling. The

maximal change observed was approximately 12.5cm (Fig. 15). In addition, it is important to notice the effects of time averaging on the Region of Convergence. The region not only expands, but takes on a more expected, symmetrical, ovular shape. This is also somewhat noticeable with impulse culling but the ovular shape is not as well formed and distinct. The ovular shape is intuitively correct when taking into account that sensors were located near the origin while the source was located in the first quadrant. Overall, time averaging was observed to be a more effective method of expanding ROC than impulse culling. Further testing should consider widening the averaging window, as well as the effects of performing a weighted average.

### **Computational speed**

There were several major areas in the proposed procedure that led to large computational overhead. The two most significant were: scanning for an initial point of maximum likelihood, and performing the convolution between signals.

In these experiments, 384 point sample data sets are used. This data is symmetrically padded with 768 zeroes before any computation takes place. This implies that there were  $3N$  multiplications and  $3N$  additions for each point in a convolution with a total of  $36N^2$  computations for each pair of signals. This is unacceptable for a microcontroller at 60MHz and would result in a computational time of approximately 180ms. To avoid this, a sampling frequency of 108kHz was used resulting in a total of 192 data points for the same sampling time. This results in a delay of only 44ms, which is acceptable for human motion.

The most computationally intensive section of the used algorithm was by far scanning for the maximum likelihood point. Due to the highly nonlinear equation used, the computational complexity of this function is high. Typical performance on a system using an Intel I7-740 processor with 6GB of ram is given in Table 1.

Table I  
Comparison of Computation Times vs. Search Coarseness

Scan Interval	Search Distance	Typical Time Taken	Delay Increase
1cm	9m <sup>2</sup>	~3.32s	
1cm	64m <sup>2</sup>	~23s	6.9 times
5cm	9m <sup>2</sup>	~.132s	
5cm	64m <sup>2</sup>	~.956s	7.24times
10cm	9m <sup>2</sup>	~.037s	
10cm	64m <sup>2</sup>	~.238s	6.46times
15cm	9m <sup>2</sup>	~.018s	
15cm	64m <sup>2</sup>	~.108s	6times

Although performance on a powerful machine is more than adequate at intervals of 5cm or more, these results are unacceptable for the much slower microcontroller. It is important to note that this code was run using MATLAB functions and was not optimized. Performance could be increased by hand optimizing code.

The brute force approach used in this experiment was inadequate for real time tracking on a microcontroller. One way to improve search speed is by successive refinement of search areas. For example, we can first search in 100 cm intervals then search with 15 cm intervals around the maximum located using the 100 cm search and so forth.

## **CHAPTER IV**

### **CONCLUSIONS**

Results indicate that this implementation is feasible and easily instantiated on a current microcontroller. Regardless of the error in accuracy of initial position finding, tracking was achievable with acceptable errors in angle of arrival. Modern auto focus techniques would allow a system to focus on the subject once the AOA is found. Observations show that this experimental setup was able to achieve an initialization value within the ROC of our system and converge to within an acceptable range of the source location.

Future work is essential to solidify the feasibility of this system. From a theoretical perspective, upper bounds on the ROC of our system should be found. In addition the ROC should be formulated as a function of acoustic noise strength and physical setup parameters (distance between microphones etc.) Another area of concentration should be the expansion of functionality by utilization of signal processing techniques to isolate human voice. Computational intensity should also be addressed by the creation of an optimized search algorithm. Finally, from an implementation stand point, a complete working prototype must be constructed and this system should then be implemented in 3D space.



## REFERENCES

- [1] X. Mei and H. Ling, "Robust Visual Tracking using l1 Minimization," in *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, Kyoto, Japan, 2009.
- [2] R. E. Bland, *Acoustic and seismic signal processing for footstep detection*, Cambridge, Massachusetts: Massachusetts Institute of Technology, 2006.
- [3] F. Figueroa and A. Mahajan, "A robust method to determine the coordinates of a wave source for 3-D position sensing," *J. Dynamic Systems, Meas. Control, Trans.*, vol. 116, no.3, pp. 505-511, 1994.
- [4] B. Ferguson, L. Criswick, K. Lo, "Locating far-field impulsive sound sources in air by triangulation," *J. Acoustical Society of America*, vol. 111, no.1, pp. 104-116, 2002.
- [5] H. Canistraro, E. Jordan, "Projectile-impact-location determination: an acoustic triangulation method," *Meas. Sci. Technol.*, vol. 7, no. 12, pp. 1755-1760, 1996.
- [6] S.M. Kay, *Fundamentals of statistical signal processing*, 2nd ed., Upper Saddle River, New Jersey: Prentice Hall, 1998, pp. 141-146.

## CONTACT INFORMATION

Name: Alexander Ivanov Ivanov,

Professional Address: c/o Dr. Jean-Francois Chamberland  
Department of Electrical Engineering  
MS 4227  
Texas A&M University  
College Station, TX 77843

Email Address: a.i.ivanov08@gmail.com

Education: B.S., Electrical Engineering, Texas A&M University,  
May 2012  
Summa Cum Laude  
Honors Undergraduate Research Fellow

Name: Alexandru Raducanu

Professional Address: c/o Dr. Jean-Francois Chamberland  
Department of Electrical Engineering  
MS 4227  
Texas A&M University  
College Station, TX 77843

Email Address: alexraducanu89@gmail.com

Education: B.S., Electrical Engineering, Texas A&M University,  
May 2012  
Cum Laude  
Undergraduate Research Fellow