ARMY RESEARCH LABORATORY

# Acoustic Transient Localization: A Comparative Analysis of the Conventional Time Difference of Arrival Versus Biomimetics

**by Latasha Solomon, Yirong Pu, and Allyn Hubbard**

**ARL-TR-5039** **November 2009**

# NOTICES

## Disclaimers

# Army Research Laboratory

Adelphi, MD 20783-1197

**ARL-TR-5039** **November 2009**

# Acoustic Transient Localization: A Comparative Analysis of the Conventional Time Difference of Arrival Versus Biomimetics

**Latasha Solomon**
**Sensors and Electron Devices Directorate, ARL**

**and**

**Yirong Pu and Allyn Hubbard**
**Boston University**
**Dept of Electrical & Computer Engineering**
**Boston, MA 02215**

| **REPORT DOCUMENTATION PAGE** | *Form Approved* *OMB No. 0704-0188* |
|---|---|



| **1. REPORT DATE** *(DD-MM-YYYY)* | **2. REPORT TYPE** | **3. DATES COVERED (From - To)** |
|---|---|---|
| November 2009 | Summary | January to September 2009 |

| | |
|---|---|
| **4. TITLE AND SUBTITLE**<br>Acoustic Transient Localization: A Comparative Analysis of the Conventional Time Difference of Arrival Versus Biomimetics | **5a. CONTRACT NUMBER** |
| | **5b. GRANT NUMBER** |
| | **5c. PROGRAM ELEMENT NUMBER** |
| **6. AUTHOR(S)**<br>Latasha Solomon, Yirong Pu, and Allyn Hubbard | **5d. PROJECT NUMBER** |
| | **5e. TASK NUMBER** |
| | **5f. WORK UNIT NUMBER** |
| **7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**<br>U.S. Army Research Laboratory<br>ATTN: RDRL-SES-P<br>2800 Powder Mill Road<br>Adelphi, MD 20783-1197 | **8. PERFORMING ORGANIZATION REPORT NUMBER**<br>ARL-TR-5039 |
| **9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)** | **10. SPONSOR/MONITOR'S ACRONYM(S)** |
| | **11. SPONSOR/MONITOR'S REPORT NUMBER(S)** |

**12. DISTRIBUTION/AVAILABILITY STATEMENT**
Approved for public release; distribution unlimited.

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT**
Accurately localizing threats of interest remains a high priority for today's military. The conventional least-squares time difference of arrival algorithm and a novel biomimetic algorithm will be applied to acoustic transient events to localize points of origin. This report will compare the accuracy of the output of these two signal processing algorithms as it relates to that of truth data.

**15. SUBJECT TERMS**
Acoustics, biomimetics, localization, least-squares

| **16. SECURITY CLASSIFICATION OF:** | | | **17. LIMITATION OF ABSTRACT** | **18. NUMBER OF PAGES** | **19a. NAME OF RESPONSIBLE PERSON**<br>Latasha Solomon |
|---|---|---|---|---|---|
| **a. REPORT** | **b. ABSTRACT** | **c. THIS PAGE** | | | **19b. TELEPHONE NUMBER** (*Include area code*) |
| Unclassified | Unclassified | Unclassified | UU | 20 | (301) 394-2180 |

**Standard Form 298 (Rev. 8/98)**
**Prescribed by ANSI Std. Z39.18**

## Contents

## List of Figures



## List of Tables

## Summary

There is a critical need to develop enhanced acoustic direction finding sensors and algorithms to provide the individual Soldier additional situational awareness useful for reduced casualties and possible counter-insurgency. With the onset of unconventional warfare, it is crucial that these sensors perform accurately in urban and mountainous terrains. Current localization systems that address these needs are satisfactory at best, often performing poorly in highly reverberant environments. This research compares the output of the conventional least squares (L-S) time difference of arrival (TDOA) algorithm with that of a novel biomimetic approach. Preliminary analysis indicates that the biomimetic algorithm is superior to that of L-S TDOA with a detection rate of 93%, outperforming the L-S by 9%.

INTENTIONALLY LEFT BLANK.

# 1. Introduction

Mortar rounds, roadside bombings, and sniper fire are all viable threats to Soldiers fighting the current war on terrorism. Thousands of Soldiers are injured and/or killed every year from the abovementioned threats. Providing a two-dimensional (2-D) grid location enables quicker response for first responders and possible return fire.

The U.S. Army Research Laboratory (ARL) and Boston University have long worked in the area of acoustic direction finding (*1, 2*). Both have successfully detected, localized, and tracked various military targets to varying degrees of certainty. These algorithms are critical for survivability and provide actionable intelligence to our military personnel. Such algorithms should be robust, highly reliable, and adaptable to a range of environments. In this report, previously developed time difference of arrival (TDOA) and biomimetic algorithms are applied to acoustic transients. This research compares the accuracy of conventional signal processing techniques with that of a novel biomimetic approach.

# 2. Signal Processing

A least-squares (L-S) estimator using TDOA was initially applied to the acoustic data to determine direction of arrival. The L-S approach chooses the value of θ that best minimizes the squared difference between the given data and the assumed signal. The process is described in the following equation

$$\theta_{L\text{-}S} = P^{+}\hat{\tau}\,, \tag{1}$$

where $P^{+}$ represents the difference in microphone locations and $\hat{\tau}$ are the estimated time delays between corresponding microphone locations (*3*). Triangulation of lines of bearing from each individual sensor is then used to calculate a 2-D grid solution. Tracking acoustic transients via a 2-D grid coordinate is often a complex data association problem. The tracker must be sophisticated enough to update older tracks as necessary and detect additional targets as new reports are acquired. The initial tracker applied to the transient data uses a genetic algorithm (GA) to search for the best solution over a sliding window of time. This technique has been simulated as part of a simple tracker that uses an alpha/beta filter for track prediction given a predetermined interval of time (*4*).

This method is ideal when trying to solve a problem for which little information is known a priori. GAs use the principles of selection and evolution to produce several solutions to a given problem (*5*). This algorithm inputs lines of bearings from a distributed network of sensors to form tracks related to transient targets of interest. The tracking algorithm evaluates the intersection of associated lines of bearings to determine the likelihood of an acoustic target of interest.  Next, the algorithm attempts to estimate the number of targets and their expected positions.

The algorithm then uses a L-S estimate to minimize the angular distance (i.e., how much does the line of bearing generated from the sensor miss the cross point) given 1 s of data.  Finally, the time cost is computed by subtracting the estimated time of arrival from the actual time of arrival. The tracks that satisfy both criteria based on predetermined constraints are reported and all others are discarded.

Next, a biomimetic approach, consisting of front-end hardware and back-end algorithms, mimicking the human auditory system was applied to the same set of data.  The acoustic direction finding system is a symmetric system that has one left audio channel and a separate right audio channel, representing the left and right ear, respectively.  Because each sensor site contains four microphones mounted in a tetrahedral configuration, there are six two-eared pairings possible, although not all pairings are necessarily used.  The acoustic signals received by the microphones first pass through a Gamma-tone filter bank, which mimics the inner ear's filtering functionality.  The characteristic frequency components are extracted and processed through different auditory nerve channels.  Spike trains are produced at the output of the acoustic direction finding (ADF) system and are then processed with the back-end localization algorithms (*6*).

The back-end algorithm consists of three stages:  detection, direction finding, and localization. In the detection stage, which also acts as a classification stage, the onset of a weapon sound is detected and classified as either a targeted event or not, and the onset time is recorded as the event time.  The radial basis function (RBF) neural network is applied during this stage.  The center locations of the RBF network are decided through a supervised learning procedure.  The spike trains are mapped to 2-D data arrays as the input of the RBF network.  The data arrays use the spiking neuron firing time as the X coordinate and the frequency channel as the Y coordinate.  Figure 1 illustrates an example in which the RBF network is applied to the spiking neuron firings.

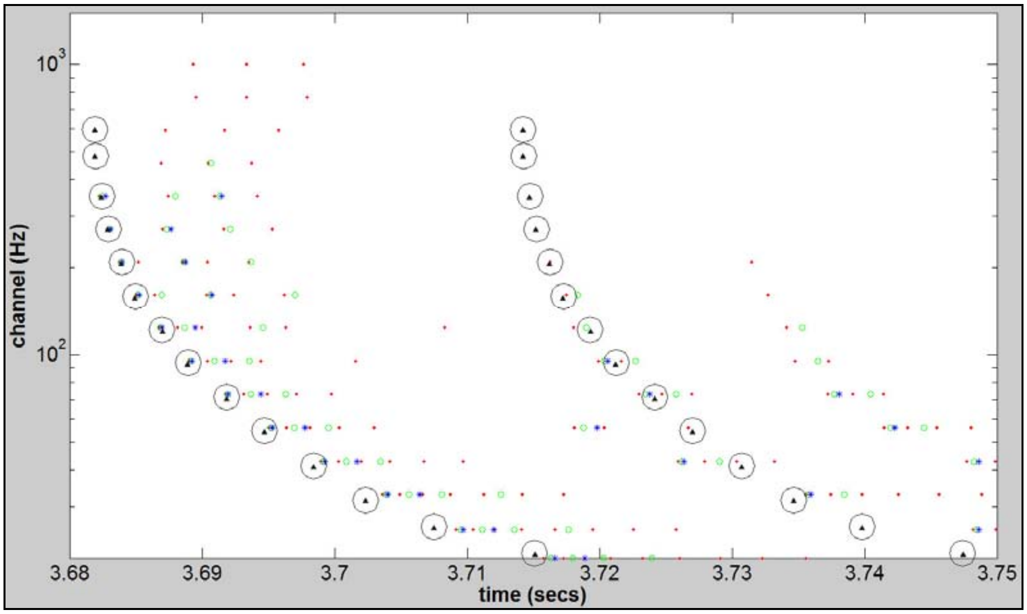



Figure 1. The spiking neuron firing generated from a 60-mm mortar launch record.

In figure 1, the *x*-axis is time and the *y*-axis is the 16 frequency channels distributed exponentially from 20 Hz to 1 kHz. The three populations of spiking neuron firings are represented by red dots, green circles, and blue stars, respectively. The black dot inside the circle is the RBF center. The spiking neuron firings inside the circle are classified as valid firings. The classified results of all the frequency channels are grouped for weapon sound type detection and classification decision. The first fired spiking neuron inside the circle is considered as the event time of this frequency channel. For each separate event, there is one event time in each frequency channel. There are two possible events in figure 1. However, there are no valid spiking neuron firings in about half of the frequency channels in the event on the right. Therefore, this event is not classified as the expected weaponry sound.

The direction finding stage uses the interaural time difference (ITD) results from up to six microphone pairs on the same site and generates an overall azimuth result from this site. The ITD results are calculated by the subtraction of the event times and the output of a Jeffress model. The Jeffress model is a cross-correlation-like model to calculate ITD based on delay lines. It was proposed by L. A. Jeffress that the neurons on the delay lines act as coincidence detectors by firing maximally when receiving simultaneous inputs from both ears. Two signals from the cochlea of each ear converge synchronously on a coincidence detector, or a neuron, in the auditory cortex based on the magnitude of the ITD (*7*). The minimum resolution of ITD result is the input signal's time period.

A temporal difference value for each frequency channel is calculated from the subtraction result of the detected event times in the microphone pair. The temporal differences from all the frequency channels are averaged to get a value $t_{sub}$. Then, as shown in figure 2, a short time window, which is 10 ms in the present algorithm, is applied to the detected event time.

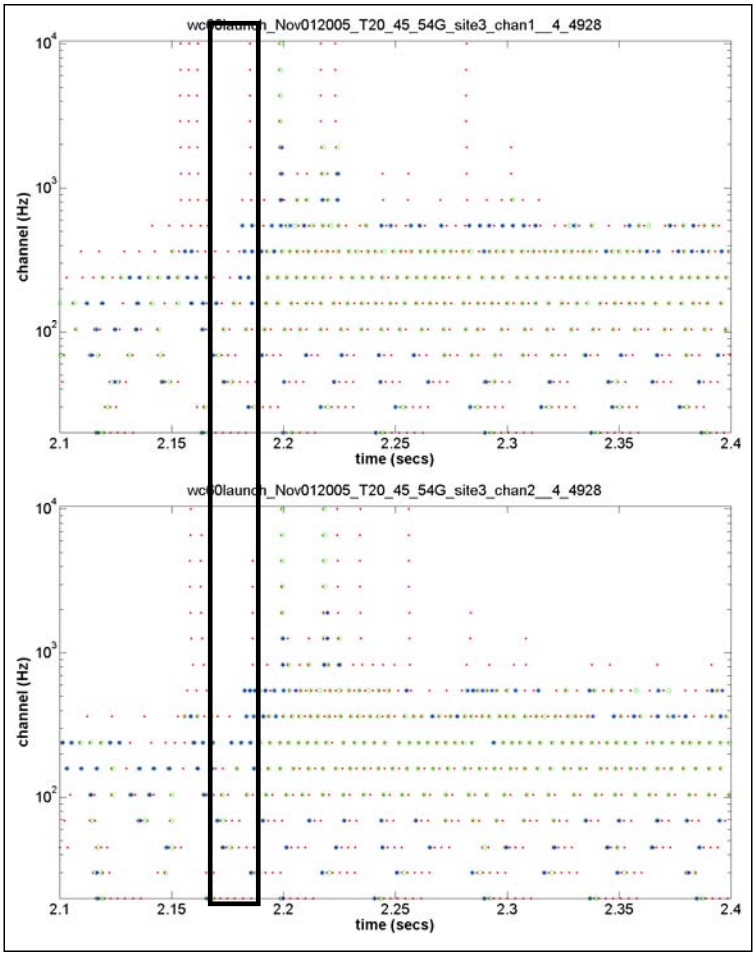



Figure 2.  ITD calculation, showing the spiking neuron firings from two microphones.

The rectangle illustrates the short time window wrapped around the detected event, which happens at approximately 2.2 s in the record.  The Jeffress model takes the spiking pulse trains inside the rectangle and gives out an ITD output for each frequency channel.  Note that the width of the rectangular shown is larger than the actual time window.  It is for the purpose of illustration.

Weighted average value $t_{jeffress}$ is calculated from the 16 results of all the frequency channels, because in the higher frequency channels where there are very few spiking neuron firings the Jeffress model does not work very well.  This result $t_{jeffress}$ is averaged with the value $t_{sub}$ to estimate a final ITD value of this microphone pair.  The reliability of $t_{sub}$ is reduced when the signal is noisy, because it is calculated only from the first spiking neuron firings in the RBF circle.  The background noise might stimulate the spiking neuron earlier than the real targeted weaponry sound.  However, the reliability of $t_{jeffress}$ does not reduce a lot when there is noise, because the calculation looks at all the spiking neuron firings within the time window.  With the introduction of the Jeffress model, the algorithm can still get a reasonably accurate direction finding result despite the fact that the input sound file was recorded in a noisy environment. Even though the signal-to-noise ratio (SNR) value is not given for the provided weaponry sound records, by rough estimation, the SNR value is less than 5 dB for some files.

The Duplex theory (*8*) explains the ability of humans to localize a sound by using arrival time difference. Figure 3 shows how this theory is implemented. The azimuth result is calculated by using the ITD value as described next. This theory has the problem of deciding the front or back ambiguity of the sound source location; however, this ambiguity is solved when there are more than two microphones available. $x_1$ and $x_2$ are the distance from the sound source to the human left and right ear, respectively, which can be simulated by two microphones; $t_1$ and $t_2$ are the time sound travels in the air before it reaches the microphones; φ is the azimuth result, which can be negative or positive depending on whether the sound source is nearer to the left microphone or the right microphone; and $\Delta t_{max}$ is the maximum possible time difference, which equals the distance between the two microphones divided by the speed of sound.

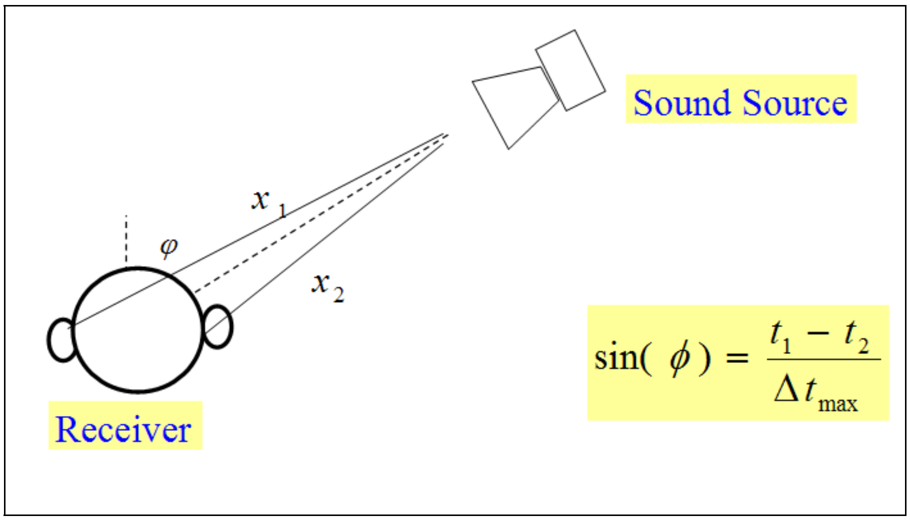



Figure 3. Duplex theory.

The time difference, $t_1 - t_2$, between the two microphones must be equal to or smaller than $\Delta t_{max}$. Occasionally, there are spiking neuron firings generated from the background noise that are mistaken for the spiking neuron firings from the weapon sound. If this happens, it is possible that the azimuth result of the microphone pair is invalid because the calculated ITD value is larger than $\Delta t_{max}$. In this case, the microphone pair is called an invalid pair and its results are excluded from further calculations.

The direction finding stage aims at finding an azimuth result of the microphone site, which contains six microphone pairs in the current hardware platform. The azimuth result from the microphone site is calculated from the azimuth results of all the valid microphone pairs based on a standard deviation value-checking criterion. The algorithm deletes the maximum or the minimum value or both of the azimuth results from the valid microphone pairs until the standard deviation value is less than a threshold value, which is set to be 20° at present.  A site might not have a valid overall azimuth result if fewer than two microphone pairs agree on the similar direction.  A site without a valid azimuth result does not necessarily mean that it does not have valid sound detection.  The final direction finding result is the average value of the azimuth results from the valid microphone pairs from one site.

In the localization stage, a Gaussian function (equation 2) is used to calculate the probability of where the sound source is located.

$$P(x, y) = \sum_{i=1}^{N} w_i \cdot e^{-\frac{\Delta\theta_i^2}{2\sigma^2}} \tag{2}$$

where $P(x, y)$ is the probability that the sound source is located at position $(x, y)$ and $i$ is the microphone site number.  $N$ is the number of sites, which is four in the given testing dataset.  A weight value $w_i$ is assigned to each site.  It is usually 1 for all the sites, but this value can be changed if one or more sites have better direction finding confidence.  This happens when one or more sites have malfunctioning microphones or one site is much further away from the sound source when comparing this distance to the sensor sites.  Usually the microphone site that is nearer to the sound source or has all microphones working properly has a higher weight value that the others.  $\Delta\theta_i$ is the angle difference between the calculated direction finding result and the angle from the position $(x, y)$ to site $i$. The variance, $\sigma^2$, is set to 25 square degrees in the current algorithm.

The final estimated location is defined by the point that has the largest probability, $P(x, y)$.  The advantage of the introduction of the Gaussian function in the localization algorithm is that it is error tolerant.  It is useful especially when one out of four microphone sites has an azimuth result that largely differs from that of the other microphones.  This single microphone site does not affect the final localization result greatly since the more accurate azimuth result that is agreed by the other three microphone sites takes over the less accurate result that is only supported by one microphone site.  Additionally, by applying the Gaussian function, the nearer a point is within a preset range of the direction finding result from a microphone site, the larger its probability is. The output of the algorithm can be a specific point that has the highest probability in the space or it can be a small range that includes all the points whose probability is higher than a preset confidence value.  The decision of which output to take can be made according to the specific applications and requirements.

## 3. Experimental Results

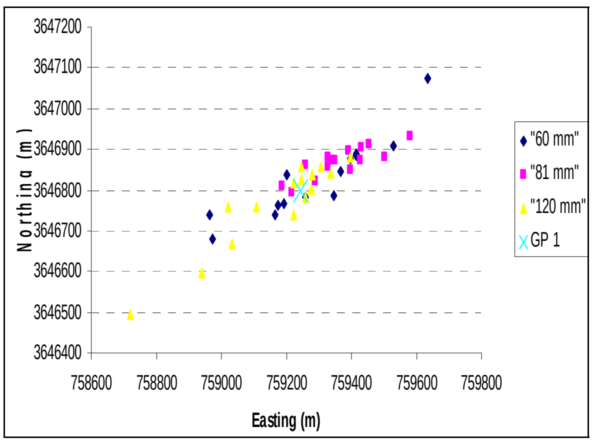
Experimental data was collected during a field experiment at Yuma Proving Ground, AZ, in November 2005.  Data analyzed consists of mortar launches (60, 81, and 120 mm) of varying charge launched at two separate gun positions (GP) with a maximum distance of 5 km from the farthest acoustic array. Four tetrahedral arrays spaced approximately 2 km apart were used to collect acoustic data. Figure 4a–d contains the projected location of the respective GPs calculated by each of the algorithms. Statistics relating to the L-S and biomimetic approach are listed in table 1. A successful launch detection is one where the distance error for an individual event is below 3 km; values above 3 km are considered outliers and discarded. The mean error for easting and northing is calculated using the following formulas

$$ME_x = \frac{1}{n}\sum_{i=1}^{n}\left(x - \hat{x}_i\right) \tag{3}$$

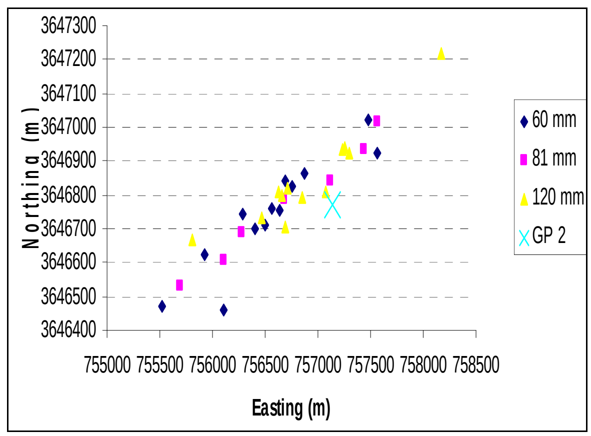
and

$$ME_y = \frac{1}{n}\sum_{i=1}^{n}\left(y - \hat{y}_i\right) \tag{4}$$

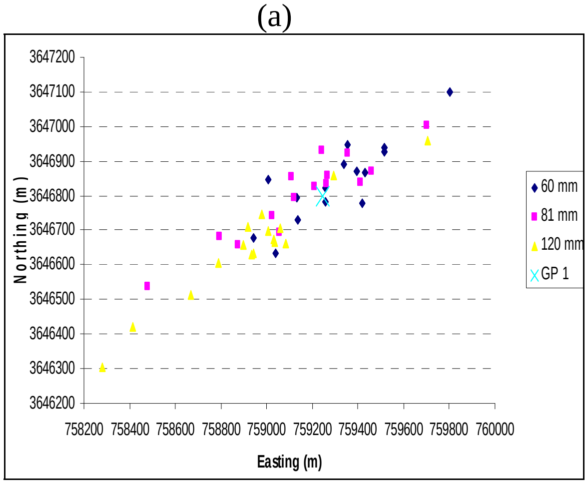
where *n* is the total number of rounds fired, *i* is the current round at a specific time; *x* and *y* correspond to the easting and northing gun locations; and $\hat{x}_i$ and $\hat{y}_i$ are the estimated easting and northing gun locations.



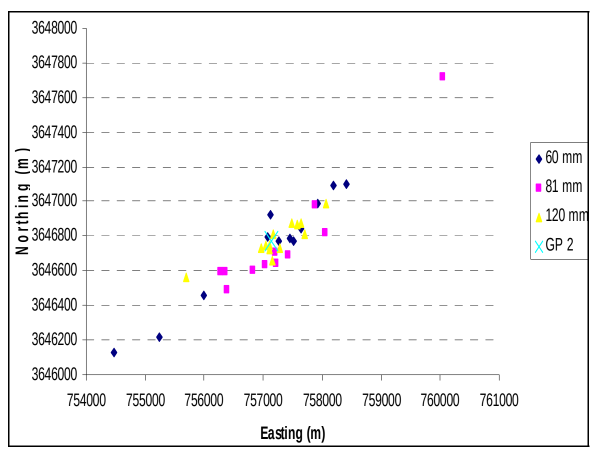

(a)




(b)




(c)




(d)

Figure 4. Estimated launch location for (a) GP 1 calculated via L-S TDOA, (b) GP 2 calculated via L-S TDOA, (c) GP 1 calculated via biomimetics, and (d) GP 2 calculated via biomimetics.

Table 1. Comparison of the number of rounds detected and the true mean error for L-S and biomimetic approaches. GP 1 and GP 2 are approximately 2.5 and 4.5 km from the center of the acoustic sensors, respectively.

| | | | Least Squares - TDOA | | | Biomimetic | | |
|---|---|---|---|---|---|---|---|---|
| **GP** | **Caliber (mm)** | **Rounds Fired** | **Detected** | **$ME_x$ (m)** | **$ME_y$ (m)** | **Detected** | **$ME_x$ (m)** | **$ME_y$ (m)** |
| 1 | 60 | 16 | 14 | 43.1 | 29.2 | 16 | 60.9 | 42.6 |
| 1 | 81 | 16 | 15 | 120 | 70.6 | 15 | –89.7 | 6.4 |
| 1 | 120 | 16 | 15 | –70.3 | –29.5 | 16 | –304.4 | –144.6 |
| 2 | 60 | 16 | 13 | –580.5 | –25.8 | 12 | –116.3 | –31.3 |
| 2 | 81 | 11 | 7 | –445.9 | 1.1429 | 11 | –193.3 | –1.2 |
| 2 | 120 | 16 | 12 | –239.8 | 73.3 | 15 | –29.4 | –7.4 |

Preliminary results indicate the biomimetic algorithm outperforms the conventional L-S TDOA approach when comparing the percentage of shots detected and their accuracy. The L-S approach has a detection rate of 84%, while the biomimetic approach has a detection rate of 93%. In general, the easting and northing mean error in the L-S approach is also lower than that of the biomimetic approach. One would have expected the mean error to increase as the distance between sensor and launch site is increased; however, this is not the case. This could be a direct result of estimating lines of bearings (LOBs) given varying atmospheric conditions. Further investigation of this phenomenon is necessary. The biomimetic approach appears to be most promising given the current results; however, the data set is relatively small and other factors such as cost and processing time should be considered when ultimately deciding upon which algorithm and associated hardware is most desirable.

Figure 5a–b illustrates the sound localization results from the biomimetic algorithm where the *x-y* axis is denoted in meters. The color describes the probability of where the sound source is located. Deep red indicates the most probable location and deep blue indicates the lowest probability regions. The region to the right of the sites is all deep blue because there is no front and back ambiguity of the sound source location. The four-microphone implementation on each site eliminates the ambiguity in the direction finding stage. Figure 5b only shows two direction finding results because there were sound files originally recorded from only two of the four sites. It proves that the algorithm is still able to localize the sound source even when not all the microphone sites are working properly. As the number of microphone sites that provide valid sound direction finding results increase, the error associated with the sound localization result decreases. However, it is noted that the localization accuracy does not necessarily increase with the number of valid direction finding results. The localization accuracy depends more on the accuracy of direction finding results.

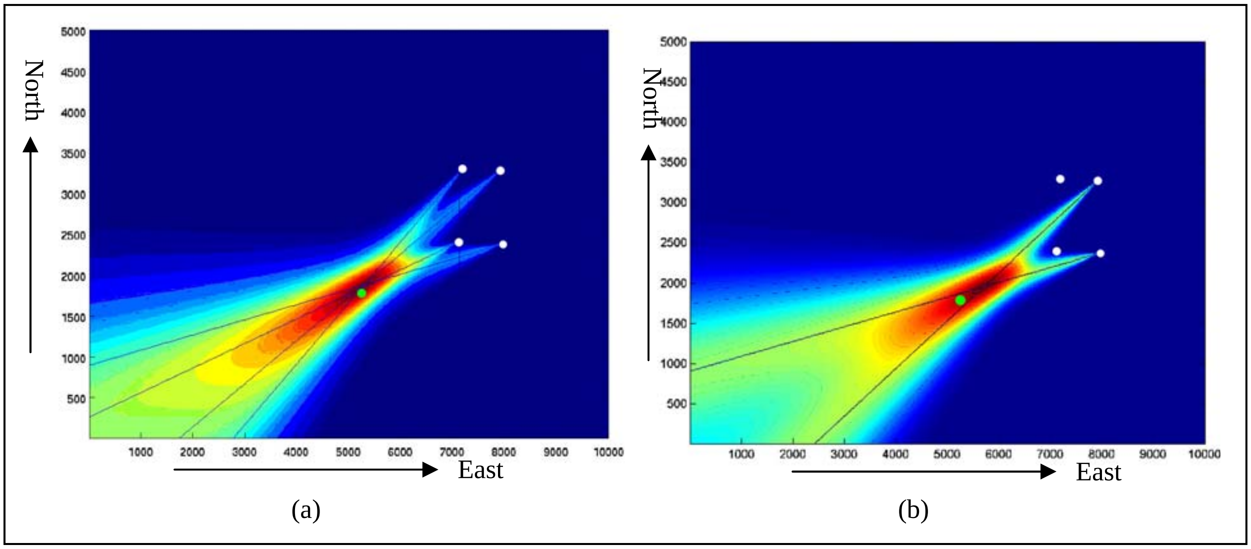



Figure 5. Valid sound direction finding results from (a) all of the four microphone sites and (b) two out of the four microphone sites.

Note: The lines show the direction finding results from the four sites, respectively. The four white dots represent the locations of the microphone sites. The green dot is the location of the actual launch sound source.

## 4. Conclusions

Acoustic sensors continue to provide the individual Soldier improved situational awareness during times of conflict. These sensors and associated algorithms should be capable of accurately detecting threats of interest with a high degree of certainty. This research compares the L-S TDOA approach with an approach using a biomimetic algorithm. Mean error and percentage of detection were computed for the estimated 2-D localization results for acoustic mortar launch signatures. Analysis of the results indicates that the biomimetic approach outperforms the L-S approach with respect to number of detections and overall accuracy of the launches detected.

The data set should be expanded to include additional mortar rounds as well as other transients such as rocket propelled grenades, C4, and small arms fire. Future comparisons should also consider processing time and cost for associated equipment. Other factors that should be considered include atmospheric conditions, environmental terrain, and outliers detected from nearby testing not associated with the test.

## 5. References


1. Scanlon, M. V.; Young, S. H.; Hillis, D. B. Networked Acoustic Sensor Array's Performance During 2004 Horizontal Fusion – Warrior's Edge Demonstration. *Proceedings of 24th Army Science Conference*, Orlando, FL, 2005.

2. Deligeorges, Socrates; et al. The Development of a Biomimetic Acoustic Direction Finding System for use on Multiple Platforms. *Proceedings of 2008 SPIE*, Orlando, FL, April 2008.

3. Tran-Luu, D. Fractional Time Delay for Direction-of-Arrival Estimation. *Proceedings of MSS-BAMS*, Laurel, MD, August 2007.

4. Hillis, D. Using Genetic Algorithms for Multitarget Tracking. *IEEE*, January 1998.

5. Genetic Algorithms Overview.http://geneticalgorithms.ai-depot.com/Tutorial/Overview.html (accessed March 4, 2007).

6. Pu, Y. et al. Back-End Algorithms that Enhance the Functionality of a Biomimetic Acoustic Gunfire Direction Finding System. *Proceedings of 2009 SPIE*, Orlando, FL, April 2009.

7. Jefferess, L. A. A Place Theory of Sound Localization. *J. of Comparative and Physiological Psychology* **1948**, *41*, 35–39.

8. Rayleigh, L. On Our Perception of Sound Direction. *Phil. Mag.* **1907**, *13*, 214–232.

## List of Symbols, Abbreviations, and Acronyms

| | |
|---|---|
| 2-D | two-dimensional |
| ADF | acoustic direction finding |
| ARL | U.S. Army Research Laboratory |
| GA | genetic algorithm |
| GPs | gun positions |
| ITD | interaural time difference |
| LOBs | lines of bearings |
| L-S | least-squares |
| RBF | radial basis function |
| SNR | signal-to-noise ratio |
| TDOA | time difference of arrival |

| NO. OF COPIES | ORGANIZATION |
|---|---|
| 1 ELEC | ADMNSTR<br>DEFNS TECHL INFO CTR<br>ATTN DTIC OCP<br>8725 JOHN J KINGMAN RD STE 0944<br>FT BELVOIR VA 22060-6218 |
| 1 | DARPA<br>ATTN IXO S WELBY<br>3701 N FAIRFAX DR<br>ARLINGTON VA 22203-1714 |
| 1 CD | OFC OF THE SECY OF DEFNS<br>ATTN ODDRE (R&AT)<br>THE PENTAGON<br>WASHINGTON DC 20301-3080 |
| 1 | US ARMY RSRCH DEV AND ENGRG CMND<br>ARMAMENT RSRCH DEV AND ENGRG CTR<br>ARMAMENT ENGRG AND TECHNLGY CTR<br>ATTN AMSRD AAR AEF T J MATTS<br>BLDG 305<br>ABERDEEN PROVING GROUND MD 21005-5001 |
| 1 | PM TIMS, PROFILER (MMS-P)<br>AN/TMQ-52<br>ATTN B GRIFFIES<br>BUILDING 563<br>FT MONMOUTH NJ 07703 |
| 1 | US ARMY INFO SYS ENGRG CMND<br>ATTN AMSEL IE TD A RIVERA<br>FT HUACHUCA AZ 85613-5300 |
| 1 | COMMANDER<br>US ARMY RDECOM<br>ATTN AMSRD AMR<br>W C MCCORKLE<br>5400 FOWLER RD<br>REDSTONE ARSENAL AL 35898-5000 |
| 1 | US GOVERNMENT PRINT OFF<br>DEPOSITORY RECEIVING SECTION<br>ATTN MAIL STOP IDAD J TATE<br>732 NORTH CAPITOL ST NW<br>WASHINGTON DC 20402 |

| NO. OF COPIES | ORGANIZATION |
|---|---|
| 1 | US ARMY RSRCH LAB<br>ATTN RDRL CIM G T LANDFRIED<br>BLDG 4600<br>ABERDEEN PROVING GROUND MD 21005-5066 |
| 8 | US ARMY RSRCH LAB<br>ATTN IMNE ALC HRR<br>MAIL & RECORDS MGMT<br>ATTN RDRL CIM L TECHL LIB<br>ATTN RDRL CIM P TECHL PUB<br>ATTN RDRL SES P L SOLOMON<br>(5 COPIES)<br>ADELPHI MD 20783-1197 |

TOTAL: 17 (1 ELEC, 1 CD, 15 HCS)