AUDIO GUNSHOT DETECTION AND LOCALIZATION SYSTEMS:

HISTORY, BASIC DESIGN, AND FUTURE POSSIBILITIES

by

JORDAN R. GRAVES

B.S., Northern Michigan University, 2010

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Master of Science

Recording Arts, with emphasis in Media Forensics Program

2012

This thesis for the Master of Science degree by

Jordan R. Graves

has been approved for the

Recording Arts, with emphasis in Media Forensics Program

by

Catalin Grigoras, Chair and Advisor

Jeff Smith

Lorne Bregitzer

Charles Ganzert

Date: 11/16/2012

Graves, Jordan, R (MS, Recording Arts, with emphasis in Media Forensics)

Audio Gunshot Detection and Localization Systems: History, Basic Design, and Future Possibilities

Thesis directed by Dr. Catalin Grigoras.

ABSTRACT

For decades, law enforcement organizations have increasingly utilized audio detection and localization systems to identify potential gunshot incidents and to respond accordingly. These systems have grown from simple microphone configurations used to estimate location into complex arrays that seem to pinpoint gunfire to within mere feet of its actual occurrence.

Such technology comes from a long and dynamic history of developing equipment dating back to the First World War. Additionally, though basic designs require little in terms of programming or engineering experience, the mere presence of this tool invokes a firestorm of debate amongst economists, law enforcement groups, and the general public, which leads to questions about future possibilities for its use.

The following pages will retell the history of these systems from theoretical conception to current capabilities. This work will also dissect these systems to reveal fundamental elements of their inner workings, in order to build a basic demonstrative system. Finally, this work will discuss some legal and moral points of dissension, and will explore these systems' roles in society now and in the future, in additional applications as well.

The form and content of this abstract are approved. I recommend its publication.

Approved: Catalin Grigroas

DEDICATION

I dedicate my work to my parents and my brother, Rick, Beth, and Connor, because their support and love have directly translated to my success. They taught me how to be kind, focused, and passionate with all that I do, and I cannot thank them enough for that.

I would also like to dedicate my work to my friends and family, who are one in the same. Knowing that what I do may impact all the wonderful people in my life keeps me ever-steady and ever-motivated to accept only the best of myself.

Finally, I would like to dedicate my work to my soulmate, Wendy. She is the driver of my ambitions and pushes me to achieve my dreams, both of which I gratefully do with her by my side.

> Nothing about me is truly perfect, but the people in my life make living it seem pretty spectacular.

ACKNOWLEDGMENTS

I would like to thank my classmates in the MSRA-MF program, my professors and teachers at UCD and at NMU, the faculty and staff of the NCMF, my coworkers, supervisors, and mentors at USACIL, and the countless others who have offered guidance for this project.

I could not have even begun this on my own, and I appreciate every contribution you all have made along the way.

TABLE OF CONTENTS

CHAPTER

I. INTRODUCTION	1
II. HISTORY	
III. BASIC DESIGN	
Planning	
Equipment and Configuration	14
Programming	17
Testing and Results	
Discussion	
IV. FUTURE POSSIBILITIES	
The Current State of Gunshot Detection and Localization Systems	
Prospects for the Basic System	
V. CLOSING REMARKS	
REFERENCES	

LIST OF TABLES

Table	
1.1 Test Trial 1 Configuration and Results	27
1.2 Test Trial 2 Configuration and Results	27
1.3 Test Trial 3 Configuration and Results	27
2.0 Test Trial Potentially Useful Evidence Results	28

LIST OF FIGURES

Figure
1.1 Sound Ranging Diagram
1.2 'Boomerang' Gunshot Detection System, Outfitted to US Army Humvee 6
2.1 Waveform of .22 Caliber Rifle Gunshot with Reflection
2.2 Angle Determination from Sound Delay between Two Microphones 11
2.3 Angle Determination from Sound Delay between Two Microphones (II) 12
2.4 Wide Angle Determination from Sound Delay between Two Microphones 12
2.5 Distanced Angle Determination from Two-Microphone Sound Delay
3.1 Gunshot Detection and Localization System, Basic Design – Overall
3.2 Basic System Workflow
3.3 Gunshot Detection and Localization System, Arduino and Servo 17
3.4 Audio Detection and Image Response Script 20
3.5 Thresholds in Audio Detection and Image Response Script 20
3.6 Discriminatory Thresholds for Audio Event Exclusion
3.7 Channel Delay and Angle Calculation in Detection and Response Script 22
3.8 Delay and Angle Calculation Example 1
3.9 Delay and Angle Calculation Example 2
3.10 FFT of .22 Caliber Rifle Gunshot with Reflection
4.1 Source Determination from Three Microphones in Three Dimensions
4.2 Source Determination from Three Microphones in Two Dimensions

CHAPTER I

INTRODUCTION

Until now, no known measure exists outside of science fiction that can successfully predict and prevent crime before it takes place. Therefore, law enforcement agencies must maximize the efficiency and accuracy of the response effort to criminal activity. The sooner law enforcement personnel can have a presence at the scene where a crime took place, the sooner the area may return to an orderly state of safety. Agents can defuse dangerous situations, gather evidence, and build a case used to apprehend those responsible for the crime.

Many factors limit the amount of influence law enforcement agencies have on the general public, and some of these factors introduce obstacles when striving for timely and effective crime response. Budgets and available funding dictate the amount of staff and equipment an agency can use, and legislation regulates agency power. Law enforcement groups may intervene in a given situation, but only when logistically capable and legally permitted. While positive intended results of this control include safe and cost-effective law enforcement, unfortunately negative side-effects also may arise. Agencies may be understaffed or ill-equipped, and may be restricted from responding as quickly or as soundly as desired. It is then vital for law enforcement groups and the people they protect to reach a compromise--keep the enforcers of the law within clear legal and financial boundaries while providing them with enough tools to help maintain safe communities.

With that compromise in mind, law enforcement groups have begun utilizing the gunshot detection and localization system. This technology provides information for law

enforcers in two regards: it identifies possible gunshot events based on audio information acquired by microphones and interpreted by algorithmic processing, and it also provides the perceived location of the sound source. The system is semi-automatic, which is to say it operates largely by automated computer programming but still requires human interface to complete its task as designed. Installed systems passively "listen" for specific audio characteristics and alert operators of potential detected gunshot events, but the decision to include or exclude an audio event as a gunshot (and requiring response at the scene) still belongs to a human at the controls.

The following sections include a history of audio gunshot detection systems, a simple design plan for a basic system, and a discussion of the potential problems facing the implementation of these systems, with some speculation on their future use in law enforcement and for other applications.

CHAPTER II

HISTORY

The origins of many technological advances are often traced back to innovations in different fields, later made applicable through simple redesign. Modern-day gunshot detection systems share similar roots.

The onset of World War I brought about a technique known as "sound ranging," which provides information regarding the coordinates of artillery weaponry. Developed by William Lawrence Bragg, a British military officer and physicist, initial sound ranging techniques involved arrays of microphones carefully placed in the field of battle to detect sound events from the fired weapons and report back to a monitor at an operating base, as depicted in Figure 1.1. At times, the resulting information contained valuable clues about the sound events' origins. Though the technique's success was less-than-desirable in the early years, nations from each of the opposing sides made adjustments to the process to find increasingly useful results.¹

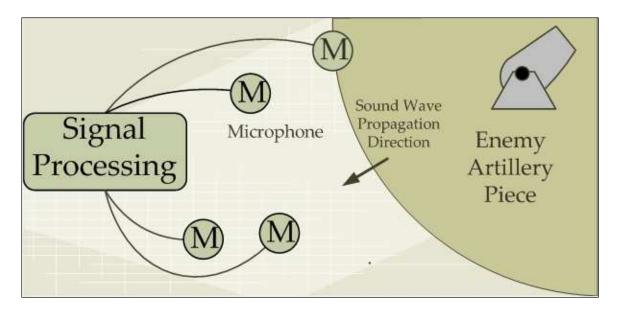


Figure 1.1 Sound Ranging Diagram

By World War II, most major military players used sound ranging for mortar detection and counter-artillery measures. In particular, British forces and United States Marines made good use of sound ranging in defensive operations.² Although sound ranging equipment had been growing more sophisticated and less costly over the years, radar systems and aerial surveillance took over as the primarily-used gun locating methods in military operations.³ Radar operators were capable of locating large weaponry faster, derived more conclusive data in settings of extreme terrain or overgrown vegetation, equipment could be outfitted on more mobile units for determining the location of airplanes and vehicles as well, and most importantly, radar could operate without waiting for shots fired.⁴ Sound ranging still held a place in combat, but acted mostly as a backup to rapidly expanding radar capabilities.

Techniques involving sound ranging for gunfire location receded in popularity until as recently as the 1980s and 1990s. Researchers borrowed sound-ranging techniques from seismologists studying earthquakes and began testing capabilities of detecting small arms

activity in urban areas.⁵ Organizations such as ShotSpotter Incorporated, now SST Inc., tested detection and localization systems in areas with high crime rates, and US police departments along the Californian Pacific coast began working with the technology to improve incident response time and subsequently help deter future crimes.⁶

Meanwhile, the military returned to using gunfire detection and location in combat zones, mainly to assist in evading and countering enemy sniper attacks.⁷ Technology is now mountable both to vehicles and personnel, and war fighting units currently rely on these tools in the Middle East and other theaters worldwide.⁸ ⁹ Figure 1.2 consists of a Boomerang system outfitted to a US Army Humvee.



Figure 1.2 'Boomerang' Gunshot Detection System, Outfitted to US Army Humvee *Photo courtesy Marine Corps Warfighting Lab via Office of Naval Research*

Back in the United States, agencies nationwide have deployed gunshot detection and localization systems in cities and other urban areas that are prone to gunfire-related crimes and random gunfire incidents. These systems are receiving more consideration as a large contributor to community safety and law enforcement success, and offer not only increased response capabilities but potential video evidence as well, incorporating video capture components in the system designs.¹⁰ While some critics raise concerns including costs, privacy issues, and accuracy, gunshot detection and localization systems used in American cities have had a significant impact in the way authorities identify and respond to criminal activity.

CHAPTER III

BASIC DESIGN

Planning

Before assembling and testing a basic audio detection system, a general strategy must be outlined. The system to be designed in the following steps will detect sound of a certain set of characteristics, will derive a direction of origin of the sound source, and will rotate a camera to point in the derived direction. In a real-world setting, such a system would "activate" when triggered and would automatically aim a camera towards the determined sound source, in hopes of capturing potentially valuable video evidence to aid investigators.

With these expectations in mind, the system should include microphones to capture audio, a computer to process the incoming audio and send commands, a microcontroller to receive the commands and send corresponding voltages, a servo to receive those voltages and rotate a platform, and a camera affixed to the platform to quickly capture the scene on video. The camera may then be wired back to the computer to display or record the incoming video information. To keep things simple, the servo will only rotate the camera along the horizontal x-axis, and will have a range of 180 degrees of rotation. Along with the equipment planning, a strategy should be made for the system programming. The two main questions to answer are: how will the system discriminate gunshot-like sounds from other sounds? And how will the system determine the direction of the sound source?

To discriminate gunshot-like sounds from others, the sounds of interest must be characterized in terms of measurable traits. To the human ear, the most obvious of these traits are the perceived loudness and short duration of the event. According to Michael and Lucien Haag, a gunshot sound measured from 1 meter away often reports louder in dB than chain saws, jackhammers, and even a jet taking off 100 feet away.¹¹ Additionally, the "rise time," or time from the start of the event to the first peak, is nearly instantaneous. One study in particular found that the "muzzle blast," or explosive shock wave and sound energy emanating from the weapon's barrel, often lasts for less than 3 milliseconds.¹² This means the shape and relative intensity of a gunshot's "waveform," or visual representation of an audio signal or recording (used to show changes in amplitude over time), can separate a gunshot sound from others. Though costs and timeframes limit the materials used in this project, these audio characteristics can still be harnessed using readily available components and intuitive programming.

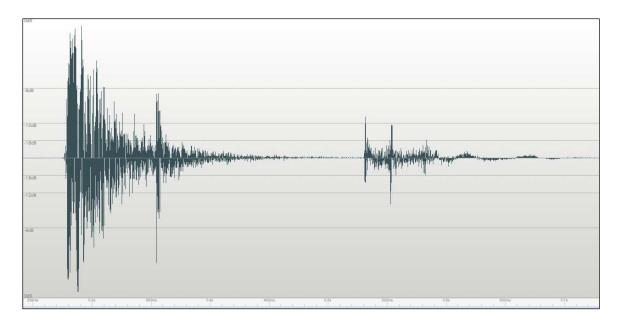


Figure 2.1 Waveform of .22 Caliber Rifle Gunshot with Reflection

Audio courtesy of user gezortenplotz via FreeSound.org, recorded with Nady wireless microphone to minidisc

The waveform in Figure 2.1 demonstrates the primary characteristics of a gunshot sound, the high signal power and the near-instantaneous first peak from relative silence. This recording in particular was purported to have taken place at an outdoor firing range. Notice a pronounced reflection recorded very quickly after the originating event takes place, most likely the "response" of the original sound event bouncing off the rearward retaining wall or barrier used to stop incoming bullets.

Finally, the means of determining the direction of the sound source should be addressed. In a plane, an object's velocity can be derived from the time elapsed over a known distance, assuming the object's speed is constant. This is represented by v = d/t. When a traveling wave maintains a constant speed over a known distance, the elapsed time will be constant as well. However, when the wave begins at a third point and travels at a constant speed along any trajectory other than perpendicular to the midpoint between two microphones, the velocity and distance can be constant and the "arrival times" to each point can vary. The wave will reach the closer point first and the further point second. Then, using the delay between the signals arriving at each channel, a source bearing can be derived, with the source originating from a point along the bearing. These are the working principles behind sound ranging, past and present, and are demonstrated in Figures 2.2 through 2.5.

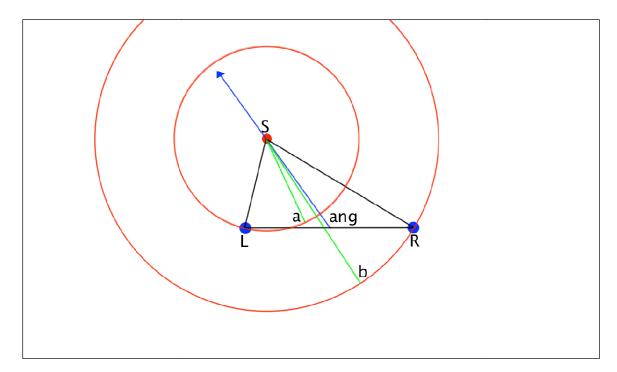


Figure 2.2 Angle Determination from Sound Delay between Two Microphones

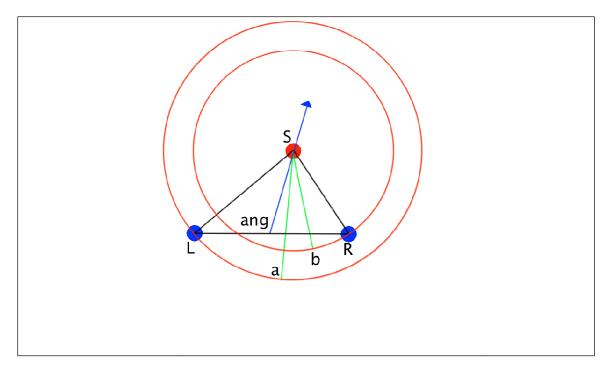


Figure 2.3 Angle Determination from Sound Delay between Two Microphones (II)

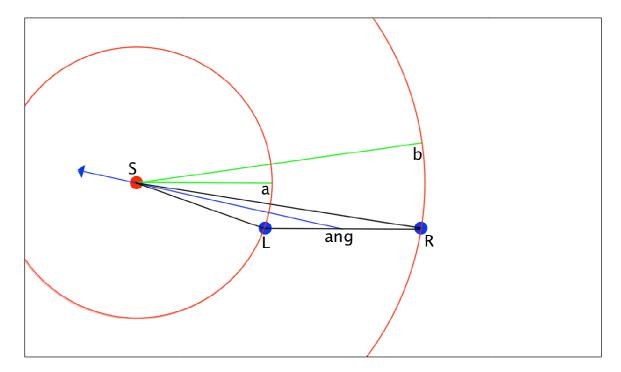


Figure 2.4 Wide Angle Determination from Sound Delay between Two Microphones

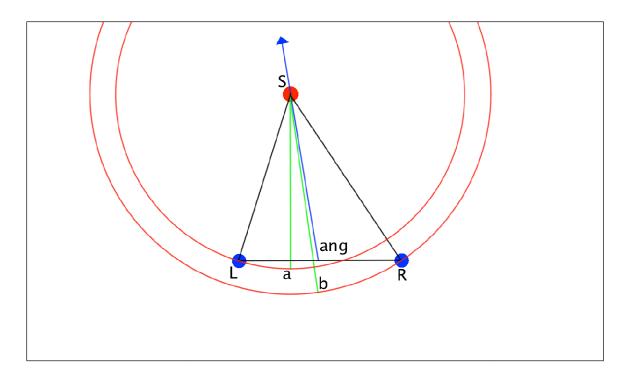


Figure 2.5 Distanced Angle Determination from Two-Microphone Sound Delay

With sound source *S*, the produced sound propagates outward at a constant speed. In Figure 2.2, because *S* is closer to *L* than *R* (SL < SR) the sound from *S* reaches *L* first, then *R*. This also means radius *a* is shorter in length than radius *b* (ra < rb). The delay of the arrival of the sound to each channel determines the calculated direction from which the sound came (*ang*), which is assessed from the midpoint between *L* and *R*.

Like in Figure 2.2, the scenario in Figure 2.3 demonstrates how *ang* can be derived from the difference between *SL* and *SR*. This time, since *S* is closer to *R* (or *rb* is shorter in length than *ra*), the corresponding angle is in the direction from the midpoint to the *R*-side.

Example 2.4 shows that even extreme angles can be determined using the difference between *SL* and *SR*, or *ra* and *r b*.

The resulting angles in all of these examples are independent of the distance from the sound source to the microphones, since the derived angle is a bearing, not an absolute

point. Even though it involves the furthest distance from the microphones of all the given examples, Figure 2.5 results in the same calculation process for delay and subsequent angle.

Equipment and Configuration

The system begins with a pair of microphones. Microphones with high tolerance to loud impulses would be ideal for a fully functioning system used in a real operation, but an inexpensive pair of smaller microphones is suitable for this design. The microphones used in this test are a pair of Olympus ME-15 microphones. These are considered a stereo pair and both capture audio simultaneously. Next, the microphones are connected to a laptop computer via a stereo input cable into the stereo mic-in port. This computer is equipped with MatLab, which is a versatile computation and programming software. MatLab handles both the audio input and the command output to the microcontroller. The actual programming scripts used in MatLab and with the microcontroller will be discussed later.

A microcontroller is then attached to the computer via serial interface. In this case, the connection is via USB cable. The microcontroller of choice is the Arduino Uno, due to its versatility and extensive open-source support. The Arduino accepts commands from the computer and sends a corresponding voltage to a servo motor which rotates a mounted webcam. The servo motor is a standard HS-422 servo, and the webcam is a 5-megapixel USB webcam that is connected back to the laptop computer for display purposes.

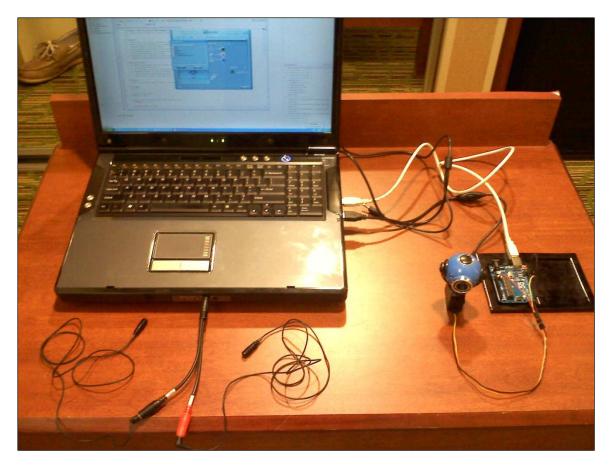


Figure 3.1 Gunshot Detection and Localization System, Basic Design – Overall

Figure 3.1 is a photo of the overall view of the system designed, assembled, and utilized for the testing outlined in this project. The microphones acquire incoming sound and send to MatLab for processing. Should the incoming signal meet the threshold requirements, MatLab would process the signal delay and compute the angle. The Arduino receives the angle rotation command via serial connection (the white cord on the right-hand side of the laptop), and communicates to the servo motor with the camera mounted atop. Then, the image information from the camera is sent to the laptop through another USB connection for display and potential recording purposes. This workflow is outlined in Figure 3.2 below.

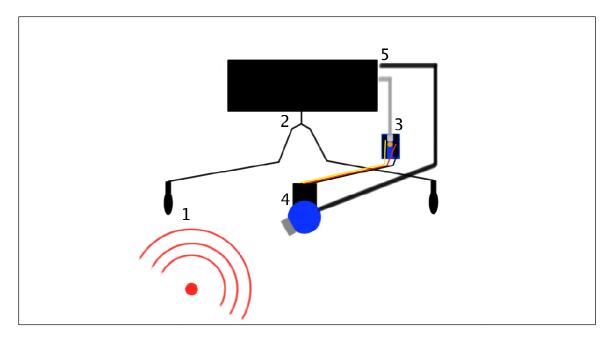


Figure 3.2 Basic System Workflow

 sound emanates from source, microphones capture upon arrival; 2) laptop receives and processes incoming audio information, determines possible target signal confirmation, delay, and resultant angle; 3) laptop sends command to Arduino;
 Arduino sends voltage to servo motor with camera mount; 5) laptop receives resultant image information from camera



Figure 3.3 Gunshot Detection and Localization System, Arduino and Servo Close-up image of Arduino and servo assembly

Programming

With the system configured properly, the Arduino and MatLab need to be programmed properly. The Arduino platform works fluently with MatLab, enough so that the microcontroller can be programmed to work continuously, waiting for MatLab serial commands, executing the commands, and returning results if asked. To prepare the Arduino for this setup, the *MATLAB Support Package for Arduino (aka ArduinoIO Package)* must be downloaded to the laptop computer. From this package, the motorsrv.pde file must be uploaded to the Arduino IDE, and the appropriate AFMotor.h

and AFMotor.cpp files must be allocated properly. For instructions on how to perform

these steps, refer to the corresponding forum at MatLab's home page.

Next, a script must be written for MatLab to automatically process the incoming audio and send commands accordingly (Figure 3.4).

```
%% Automatic Audio Detection/Image Response (AADIR) system
% Mr. Jordan Graves, BS and Dr. Catalin Grigoras, PhD
8 2012
%% Purpose:
% 1. This system will acquire live audio signals.
% 2. Based on defined thresholds, this system will discriminate
particular acquired audio events from others.
% 3. Using the perceived delay of incoming audio signals between the
pair of recording channels, this system will estimate the directional
source of the discriminated audio signal.
% 4. This system will command the servo motor to rotate the cameara
array towards the perceived source of the discriminated audio signal.
%% Materials Used:
% (1) Arduino microcontroller with serial connection to computer and
signal connection to servo motor
% (1) 5V rotational servo motor (180-degree range) connected to Arduino
% (1) webcam attached to rotating mechanism of servo motor
% (2) omnidirectional microphones arranged to acquire stereo audio
signal, connected to computer via stereo microphone input, through Y-
adapter
%% Notes
% works with motorsrv and add AFMotor.cpp and AFMotor.h to path:
...Arduino\libraries\Servo
%% Script
%delete(a)
%connect to the board
a=arduino('COM13')
% define Pin#9 as output and attach the motor to it
a.pinMode(9,'output');
% Attach servo#2 to Pin#9
a.servoAttach(2);
a.servoWrite(2,90); %reset servo to center
% define the audio settings
% sampling frequency
fs=48000;
% resolution (bits)
```

```
nbits=16;
% no. of channels
ch=2;
% each "extraction" length in sec
t=0.5;
% signal power threshold
th=1200;
% window threshold size
win=200;
% define the audio object
recObj=audiorecorder(fs,nbits,ch); %begin recording
get(recObj) %collect/display values as they are recorded
disp '***BEGINNING ACQUISITION***' %status message
for k=1:2000
% aquire the audio signal
recordblocking(recObj,t); %record without on-the-fly control until
recording is stopped
% Store data in double-precision array.
x=getaudiodata(recObj,'int16'); %signed integers mapped to set
parameters (anything outside will be "rounded")
% find absolute value of incoming signal
xa=abs(x);
% extract L and R channels
L=double(x(:,1));
R=double(x(:,2));
[k max(L) max(R)] %query for maximum values during sampling "window"
if max(L)>th && max(R)>th %set power threshold
   if xa(k:k+win) < win % set rise time threshold
    % Plot the waveform (grid on, tight to L/R)
    subplot(211),plot(L,'r'), grid on
   axis([0 length(L) -2^15 2^15])
    subplot(212),plot(R,'g'), grid on
    axis([0 length(R) -2^15 2^15])
   disp '***SYSTEM ARMED, DATA COLLECTED***' %status message
    [c,lags]=xcorr(L,R); %cross-correlation between vectors
(automatically adjusts for length differences), returns a "lag vector"
    [a1,b1]=max(L); %fs/time of max values
    [a2,b2] = max(R);
```

```
[a3,b3]=max(c); %define c's maximum values as a3,b3
   delay2=fs/2-b3 %delay is half of sampling frequency minus b3
(maximum value for c), in samples
    s=delay2;
   if s<-127 %round values outside degree parameters to furthest
degree value left or right (to maintain 180 degree range)
        s = -127;
    elseif s>127
        s=127;
    end
    % convert the delay s into degrees ang
    ang=round((s+128).*179/256)
    % rotate angle ang
   a.servoWrite(2,ang);pause(0.01); c;
    end
end
end
delete(a)
```

Figure 3.4 Audio Detection and Image Response Script

As is standard for MatLab scripts, green lines of text with a percentage sign preceding the

content are considered notes and are not executed with the actual programming language.

The comments provide guidelines for each portion of the script.

Primary points of interest in the script include the "win" and "th" thresholds and the

delay calculation and angle conversion elements.

```
% signal power threshold
th=1200;
% window threshold size
win=200;
```

Figure 3.5 Thresholds in Audio Detection and Image Response Script

According to Figure 3.5, which is a reference to the threshold element of the script, the "th," threshold is 1200 quantization levels of relative signal power. This is a setting dependent on multiple factors, including microphone gain settings, expected distance from sound source to microphones, and expected background noise. Due to these many

factors, the "th" setting requires careful calibration for each deployment. "Win" corresponds to a threshold of 200 samples. Note, the sampling frequency of the incoming audio is 48kHz.

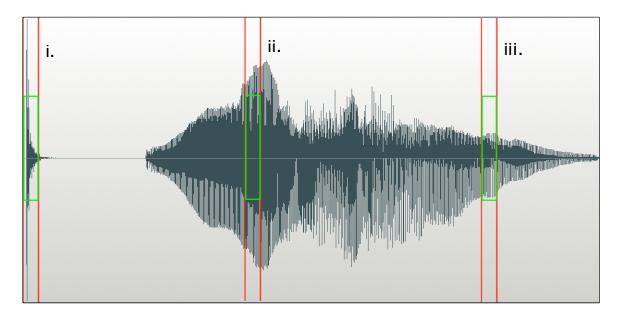


Figure 3.6 Discriminatory Thresholds for Audio Event Exclusion

Figure 3.6 illustrates how a set of thresholds might work in discriminating audio signals by power and duration. At window *i*, the signal meets both the minimum power and maximum duration thresholds. At *ii*, the signal meets the minimum power threshold but is too long in duration. At *iii*, the signal does not meet the power threshold. In the script used in this project, the signal must meet the power threshold, then the duration threshold, in order to provoke a system response.

The other significant portion of the script used in the project pertains to the delay calculation and angle conversion portions (Figure 3.7).

```
[c,lags]=xcorr(L,R); %cross-correlation between vectors (automatically
adjusts for length differences), returns a "lag vector"
    [a1,b1]=max(L); %fs/time of max values
    [a2,b2] = max(R);
    [a3,b3]=max(c); %define c's maximum values as a3,b3
   delay2=fs/2-b3 %delay is half of sampling frequency minus b3
(maximum value for c), in samples
   s=delay2;
   if s<-127 %round values outside degree parameters to furthest
degree value left or right (to maintain 180 degree range)
        s = -127;
    elseif s>127
        s=127;
    end
    % convert the delay s into degrees ang
    ang=round((s+128).*179/256)
```

Figure 3.7 Channel Delay and Angle Calculation in Detection and Response Script

The lower portion of the script determines the delay of the incoming audio event between channels, and then produces the corresponding angle for the sample delay. The Arduino uses 0 degrees as a valid degree integer, so the 180-degree range actually includes 0 and spans from 0-179.

Some further explanation is necessary for the delay and angle calculation portions of the script. The *xcorr*, max(L), and max(R) portions of the script mark the initial peak values of the incoming signal in each "sampling window" (defined earlier in *t* as .5 seconds in length). Each initial peak is marked in the numerical sample it was measured to take place. The delay is then determined from the difference in those sample values; if the value fell outside the allotted range, it would be rounded to the high or low extreme, depending on whether it was above or below those extremes, respectively.

Assuming the speed of sound is approximately 350 meters per second, a sound wave would travel 6 feet (or approximately 1.829 meters) in around .0053 seconds. 6 feet is the

prescribed distance between microphones used in the system, .0053 seconds is the maximum delay between channels. Since the sampling frequency defined above is 48kHz, or 48000 samples per second, the maximum delay between channels can also be measured as approximately 256 samples. The delay is then added to 128 to account for the delay reference to the R channel instead of L, and then compared to the ratio of samples to angles. The resultant value is rounded to the nearest whole integer and is the calculated angle for rotation.

So, if the delay were 98 samples (signal reaching R 98 samples faster than L):

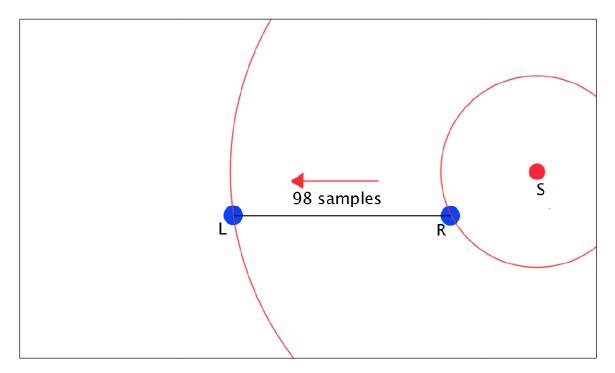


Figure 3.8 Delay and Angle Calculation Example 1

ang = (98+128)*179/256 ang = (226)*179/256

ang = 158 degrees

On the other hand, a delay of -110 samples (signal reaching L 110 samples faster than R):

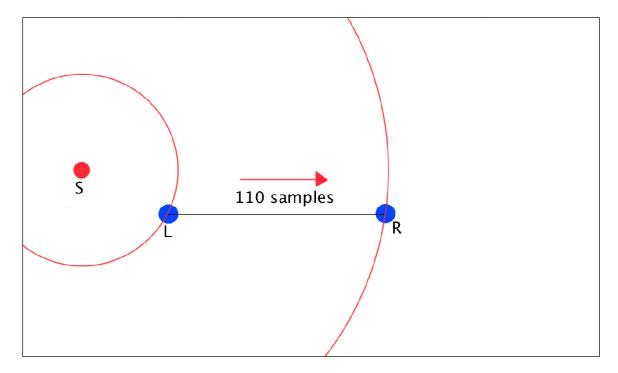


Figure 3.9 Delay and Angle Calculation Example 2

ang = (-110+128)*179/256 ang = (18)*179/256 ang = 13 degrees

The angle calculated from this script is measured from left to right of the midpoint between microphones. This means angles between 0 and 88 degrees will rotate the camera counterclockwise from neutral, while angles between 90 and 179 degrees will rotate the camera clockwise from neutral. Of course, the user may reverse which side, L or R, to assign reference to, among many other customizable features (thresholds, microphone distance calibration, etc).

Testing and Results

To evaluate the system's functionality, a simple test was formulated and executed. The system was assembled as described above, in a series of open, outdoor tennis courts. This

place was chosen with the intent of minimizing potential interfering reverberations, as well as other variables introduced in more crowded areas. The test was executed at night, to reduce the chance of external noise interference from wind or passerby. The air temperature was approximately 37 degrees Fahrenheit. This is significant because, although relatively small changes in temperature would not affect the speed of sound in a drastic way, it is well known that larger temperature variations could introduce complications in calculating the speed of sound.¹³

As mentioned earlier, the microphones used were elevated and spaced at approximately 72 inches apart, with the notion that spacing should be towards the wider end of the spectrum to emphasize the delay between incoming audio channels.

Markers were placed at 5, 10, and 15-foot distances from the center point between microphones, all distances at 10, 30, 50, and 70-degree angles from that same center point in either direction. In total, 24 markers were made. These markers indicated the intended positions from where the test sound would originate.

At the time of the test, an actual firearm was not an available sound source. Instead, a loud, sharp clap of the hands was utilized at each marker. The overall waveform shape of a hand clap could properly simulate a gunshot because both events can be characterized with high intensity and short duration. Though the claps were kept at consistent volume, some variation in signal intensity must be acknowledged. However, the variations were considered acceptable because of the multiple factors that introduce variations in sound in a real-life situation. The test itself was designed to be controlled in most reasonable aspects, yet allowed for some semblance of a realistic environment.

After the system was assembled and initiated, the testing began. After each instance of a hand clap at each marker, the system was reviewed for a response and possible camera movement. At each marker, the possible responses for the system were:

1. Rotation of the camera towards the sound source, stopping with the marker in the center of the camera frame (represented by \downarrow in the tables below)

2. Rotation of the camera towards the sound source, stopping with the marker in the frame but not in the center (represented by O in the tables below)

3. Rotation of the camera, stopping without the sound source in frame (represented by X in the tables below)

4. No camera movement in the response (represented by - in the tables below)

The responses were determined after some camera movement in response to the claps, or after a maximum 5 clap attempts at the marker.

Each marker was tested in a trial, with three total trials making up the test. The order for marker tests varied by trial; the first two trials were in order of each degree at one distance followed by the remaining two distances, while the last trial proceeded in a more staggered pattern. Tables 1.1 through 1.3 illustrate each trial and set of results.

Trial 1 – Results	5 feet	10 feet	15 feet
70° Stage Left (SL)	\downarrow	\downarrow	\rightarrow
50° SL	-	-	\rightarrow
30° SL	-	-	0
10° SL	0	Х	0
10° Stage Right (SR)	0	Х	0
30° SR	Х	Х	Х
50° SR	-	-	0
70° SR	\downarrow	-	0

Table 1.1 Test Trial 1 Configuration and Results

Trial 2 – Results	5 feet	10 feet	15 feet
70° Stage Right (SR)	\downarrow	\downarrow	\rightarrow
50° SR	-	-	0
30° SR	Х	Х	Х
10° SR	0	0	-
10° Stage Left (SL)	\downarrow	0	0
30° SL	Х	Х	Х
50° SL	Х	-	\downarrow
70° SL	-	-	\downarrow

Table 1.2 Test Trial 2 Configuration and Results

Trial 3 – Results	5 feet	10 feet	15 feet
70° Stage Right (SR)	\rightarrow	\rightarrow	\downarrow
10° Stage Left (SL)	0	\downarrow	0
50° SR	0	Х	Х
30° SL	\downarrow	Х	Х
30° SR	\rightarrow	Х	Х
50° SL	\rightarrow	0	0
10° SR	0	0	0
70° SL	\downarrow	\downarrow	\downarrow

Table 1.3 Test Trial 3 Configuration and Results

Each test proceeded through the first column, then the middle column, then the last

column, each column from top to bottom ('Stage' directions refer to the direction from

the viewpoint of the camera outward towards the markers).

Discussion

At first glance, the results from the test as a whole seem mixed and inconsistent, with only Trial 3 producing responses to each marker with camera movement. To make the results more relevant and truly evaluate the test results for inconsistent system responses, each trial was evaluated in terms of how many "Potentially Useful Evidences" (or "PUEs") were created. These are defined within the realm of the test as camera responses that end with the location of the sound source somewhere in frame, either centered or not centered. This is to simulate a real-world scenario, where a video recording of an incident would be submitted for evidence. Potentially useful evidence in such a scenario would require the event itself or the immediate aftermath to be captured somewhere in frame, either centered or not. Within those parameters, the test resulted with Table 2.0.

PUE Results	5 feet	10 feet	15 feet
Trial 1	4 of possible 8	1 of possible 8	7 of possible 8
Trial 2	3 of possible 8	3 of possible 8	5 of possible 8
Trial 3	8 of possible 8	5 of possible 8	5 of possible 8

Table 2.0 Test Trial Potentially Useful Evidence Results

The 8 possible instances for PUE response refer to each marker at the given distance.

This table of PUE results helps clarify exactly what about the system's responses are inconsistent. The inconsistency does not necessarily exist on a trial-by-trial basis, but more within the changes between distances. While the potential for useful evidence increases in Trial 3 from the others, Trial 1 and 2 exhibit increases in PUEs as the distance grows. The opposite results occur in Trial 3, where the PUE decreases and stays the same at each respective distance. This is a counterintuitive result, since one would assume that events at further distances are more likely to take place within the frame of view of a camera.

The overall inconsistency in the system can be attributed to multiple factors. First of all, the microphones used may not be most suitable for the task of discrimination. Due to the very short duration of most gunshot sounds, microphones used should be sensitive enough to accurately define the incoming audio information, to the point in which a series of loud impulses spaced closely together in rapid succession would be recorded as such, instead of one long, loud impulse. The microphones used in this test were not specifically designed with that task in mind.

Second, the surrounding environment plays a large role in the effectiveness of these systems. Even though the testing location and time of day were chosen with the intent that uncontrollable variables would be minimized, not everything could possibly be accounted for, and minor changes in the test conditions could introduce fluctuations in data.

Finally, the programming language used in MatLab itself could use some review and potential upgrading. In order to derive the correct thresholds at a particular setting, testing must be done to determine a combination that works best. A level of automation for thresholds might be worthy of some attention, where the system could be designed to automatically adjust for changes in the noise floor, etc. Real-time adaptive filters might also work in terms of limiting the amount of extraneous and useless sound information that would only hinder progress, especially sounds of frequencies below around 400Hz and above around 2.5kHz, which are the primary frequencies exhibited by gunfire.¹⁴

These instructions, tests, and results are useful in describing the process for basic detection and response, but the design does carry an underlying flaw. Any impulse that is short and loud enough would trigger a camera movement, not necessarily a gunshot. This

is because further discrimination of gunshot sounds from other noises of the same shape involves higher-level filtering and analysis of the sounds, via the programming. Unfortunately, gunshot audio usually exhibits peaks around 630Hz.¹⁵ These ranges are common for a wide variety of other sounds, making discrimination much more complicated.

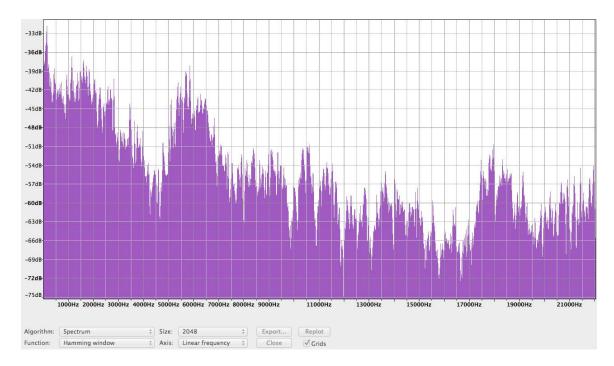


Figure 3.10 FFT of .22 Caliber Rifle Gunshot with Reflection

Using Fast Fourier Transform to bring a signal into the frequency domain, Figure 3.10 allows for some frequency analysis of the example recording shown earlier as a waveform in Figure 2.1. The frequencies of highest intensity range from the lowest up to 2kHz, with maximum values below 500Hz. These results do incorporate some "environmental acoustics" due to the reflection in the recording, but that would be expected for most audio events in realistic scenarios.

To properly discriminate from other sounds, the best approach at this point is an algorithmic learning strategy, such as those proposed by Morton and Collins¹⁶ or

Valenzise et al.¹⁷ Algorithmic learning strategies are not used in this project, but will be discussed later.

To properly simulate the products used in the field and attempt to replicate their functionality, vendors such as ShotSpotter should provide scientists with some data in regards to their product specifications and schematics, simply for research purposes. This conflicts with the vendors' rights to withhold proprietary information, but more importantly encourages unbiased review and testing in a scientific and peer-reviewed forum. Those kinds of procedures would help alleviate doubts about these systems' capabilities, and may improve their general public image, as observed below. At the time of this test, representatives from both ShotSpotter and Boomerang elected not to field questions about specific elements of their respective systems' designs, functionalities, and test data.

CHAPTER IV

FUTURE POSSIBILITIES

The Current State of Gunshot Detection and Localization Systems

Gunshot detection and localization systems, ShotSpotter in particular, have been used to aid law enforcement for some time now, and have been met with mixed reactions. In some situations, law enforcement agencies report dramatic improvements in gunfire incident response time. In others however, the system is reported as inaccurate, costly, and overly demanding of valuable resources, including taxpayer money.

In 2009, the New York Times followed ShotSpotter use in the greater New York, New Jersey, and Connecticut areas. Officials praised the system for helping reduce shooting response time, which led to faster aid for victims and better suspect apprehension. According to one Sergeant, the system proves especially helpful in areas where citizens are so accustomed to gunfire that 9-1-1 calls simply do not take place any longer.¹⁸

In 2011, city officials of Wilmington, North Carolina approved a 2-year contract with ShotSpotter, with the mindset that personnel is the greatest expense, and the technology could be used to better manage staff in the field. The contract expires in mid-December of 2012, and will be reevaluated by the city to determine whether or not to extend the service.¹⁹

Other agencies have seen less success with ShotSpotter. In March of 2012, the New Haven Police Department adopted a new protocol of sending audio data to ShotSpotter headquarters prior to redirection to the New Haven 9-1-1 center. Though the change in routing only delays information from reaching dispatch by mere seconds, the system has reported enough false positives to cause concern and necessitate the change.²⁰ More

worrisome are the gunshot events that take place but are not detected, which has happened in the area.

That very situation factored into a decision made by city council members of Trenton, New Jersey to reject a ShotSpotter expansion in January 2012. The previous Christmas, a man was fatally shot and left to die on a sidewalk of a major avenue. Although a ShotSpotter sensor was set up only blocks away, no alarms were triggered. The South Ward Councilman, a former police officer in Trenton, concluded that "[ShotSpotter] does not work, at least not for Trenton."²¹

To further cloud judgment on system value, the city of Detroit rejected ShotSpotter installation because it was too taxing on available personnel. The Broward County Sheriff in Florida previously used ShotSpotter but decided to remove it, citing the system was not cost effective and came with too many false alarms. However, the Rochester Police Department in New York swears by ShotSpotter, asserting "...its value is always a relative question. We think it's valuable or we wouldn't have done it." Further, a Criminal Justice professor and 26-year RPD veteran asks, "If it gets police to one more victim sooner, how do you put a price on that? If it adds to the evidence to convict someone, how would you add value to that?"²²

To better decide whether or not to employ a gunshot detection and localization system like ShotSpotter, authorities need reliable, unbiased information in the form of extensive testing and reporting. While efficacy studies and evaluations are available, most do not publish complete statistical results.^{23 24} Some studies are conducted with tangible and measurable results, but are not current within the past decade.²⁵ A long-term study with complete transparency of test materials and results will be the best means of assessing the effectiveness of gunshot detection and localization systems, determining strengths and weaknesses, and choosing whether to utilize the system or not. Until the day comes when leading manufacturers reveal their design plans and schematics, researchers must continue to build replicas as similar as possible to those used in the field, so results gathered will be relevant and useful.

Prospects for the Basic System

Within the scope of this project, the first future developments available should be improvements and further testing. As mentioned earlier, additional testing using superior equipment such as high-quality microphones, cameras, etc. would logically introduce different results than those found with the current configuration. In particular, microphones more sensitive to changing dynamics in audio signals should theoretically offer more precise measurements and findings. Not only could additional upgrades offer improvements on current functions, new additions to the current microphone arrangement could literally add another dimension to the system response capabilities. A third microphone could allow for three-dimensional localization, either in the form of triangulation in a single plane or in space.

In much the same way that two-microphone systems would give a direction in a single plane, three microphones would accomplish the same feat—both vertically and horizontally. If the corresponding microphones are arranged as described in Figure 4.1, such a system could theoretically provide azimuth and elevation information about a signal source.

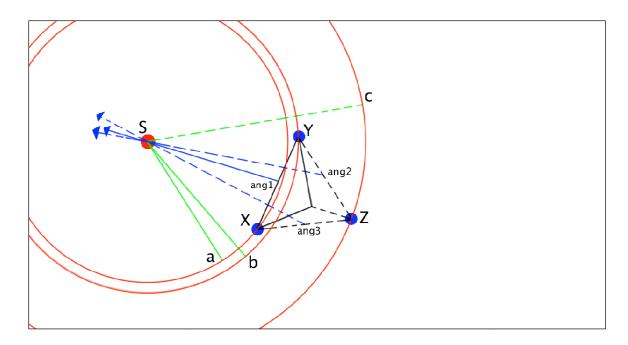


Figure 4.1 Source Determination from Three Microphones in Three Dimensions

With two microphones, such as in LR earlier, two-dimensional hemispherical sensing is feasible. Three microphones could allow for three-dimensional sensing, along an X, Y, and Z axis accordingly. Each resulting angle produced from each pair of microphones would converge on a point in space and thereby provide a location for the source of sound. This is shown in Figure 4.1.

Triangulation, often used outdoors by navigators or cartographers with compasses, involves the use of geometry and known relationships to estimate an object's location within two-dimensional space. By spreading the three microphones in such a way that the sound event is "contained" within the triangle formed by the sensors, the resulting angles derived from each pair would localize and pinpoint the perceived source to a specific location. Technically, two angles could produce a location for the sound source. However, a third angle would not only be available by default after using three different microphones, the third angle would provide additional correction to minimize error in the results. Triangulation in this context is described in Figure 4.2.

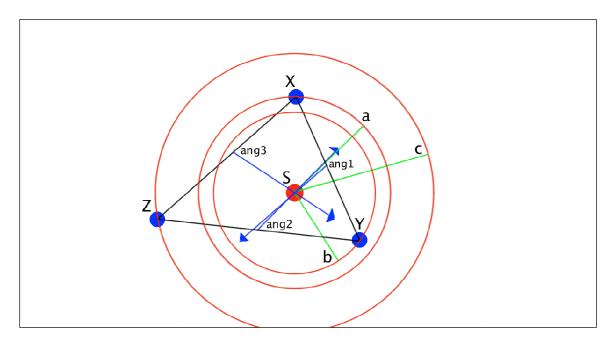


Figure 4.2 Source Determination from Three Microphones in Two Dimensions

Using the *a-b*, *b-c*, and *a-c* differences and resultant angles for the *XY*, *YZ*, and *XZ* pairs respectively, the relative location for the sound source can be calculated. Therefore, assuming each sensor is fixed in a known coordinate, the true coordinate of the sound source can be derived from its relationship to the sensors. This particular example shows how the source can be derived in a plane.

Additional microphones would not only increase the capabilities of the system, but would minimize potential error while maximizing sensing capabilities. Sound events, such as muzzle blasts emanating from the firearm may be highly directional and vary significantly at different angles.²⁶ The greater the microphone coverage available, the greater the potential for capturing an adequate amount of audio information.

Once these upgrades have been made, further testing would undoubtedly take place. To better estimate accuracy, the camera may be replaced with a laser pointer, which will

provide a more quantitative means of evaluating system response. The laser would produce a beam, terminating in a "point." The distance of this point from the originating sound source could offer measurable results for analysis and future calibrations.

On the topic of calibrations, the programming language could also be adjusted to accommodate for different microphone spacing. This would be used to determine a minimum or perhaps even an optimum distance between microphones, which would become more important as more microphones are added and the system complexity grows.

Since this system in particular was adjusted to react to gunshot-like audio events, it would not be recommended for other applications as-is. However, the signal power and window thresholds allow for a variety of potential "target sounds." For instance, researchers might consider using a system described above to monitor animal activity in their natural habitats. This could even entail the building and installation of systems in extreme environments, such as seafloors or mountaintops, so scientists may carry out their studies from the safety of a remote laboratory. Research using these sorts of systems in national parks has already commenced, in hopes of combating outbreaks of poaching on protected territory.²⁷

A system such as this might also be used by an instructor teaching an on-line course, so that he or she might feel free to move about the setting without worrying about whether the camera will remain fixated upon him or her while talking. While the teacher shifts to draw attention to a demonstrative object at his or her side, the detection system, attuned to his or her voice, would compensate accordingly with a camera adjustment. Of course, such possibilities would depend heavily on far superior programming to prove successful in a long-term scenario. These kinds of uses would almost certainly necessitate the use of learning algorithms to allow the system to more accurately suggest whether to accept or reject incoming signals, perhaps to the point where it may make the decision for itself and allow the user to review the decisions made and correct any errors.

Though some might suspect a complicated set of equations and commands must steer a discriminatory algorithm, one study in particular²⁸ actually listed a series of relatively simple correlations as one of the more robust methods tested. Templates of gunshot signals taken from 30 metros (Spanish for "meters" in English) and 90 metros are averaged and stored in vectors 1000 samples in length. At 39 samples per iteration, a correlation between the incoming signal and averaged templates is calculated for each iteration and stored in a pair of vectors for comparison against given thresholds. Testing revealed a True Positive Rate (true positives detected divided by total number of positives) of .91, with a False Positive Rate (false positives detected divided by total number of additional programming skill or hardware resources, and could be easily customizable for a user's need.

Similar to the time-based correlation method is another adapted from a document released by ShotSpotter to inform about gunshot location systems.²⁹ With sensitive equipment and comprehensive programming, an accurate frequency envelope can be developed for the gunshot event. This envelope could be stored as a series of data points, along with a large volume of multiple series of data points generated in similar fashion, all normalized to the same length. Storing such data would require much less in terms of

relative capacity when compared to storing numerous uncompressed recordings. An incoming signal could be broken down into the same sort of envelope, and then compared at target frequencies to the stored "templates" for correlation values. If the incoming signal shared a high enough correlation in frequency to known templates, an automated decision could be made to alert the system and direct a response. On the other hand, data points could be stored for false positives (heavy machinery, backfiring automobiles, and the like), so that if the incoming signal shared high enough correlation with false positives, the system could notify the operator of a potential false alarm needing review and confirmation. As each event takes place, the incoming signal values could be designated as true or false positive, and then the system would subsequently "learn." This sort of strategy would function best using highly sensitive microphones and large databases for comparison, which could restrict processing speeds and inflate costs.

While some might question the effectiveness of this type of programming against false positives as deceptive as other small explosions like fireworks, the opposing viewpoint is that a system prone to more false positives is much more acceptable than one prone to error in false negatives. Nevertheless, researchers have shown that even firecrackers differ from gunshots in terms of frequency domain bandwidth,³⁰ meaning this proposal in particular requires further development but shows some promise.

CHAPTER V

CLOSING REMARKS

Gunshot detection and localization systems have attracted attention of nearly every sort the past few years. While some praise the systems' implementation in law enforcement scenarios, others remain skeptical of their effectiveness versus costs. At this time, it is unclear whether or not gunshot detection and localization systems deserve widespread installation or banishment, but the most reasonable course of action is to continue testing and development, with hopes of constant improvement and simultaneous research transparency. Fortunately, basic designs such as the one outlined in this thesis, combined with the enormous potential for their use in gunshot response or other related applications, allow for progressive scientific development and advancements. This in turn may give rise to a universally reliable system to use in law enforcement. No matter the means, the bottom line is this: authorities and law enforcement personnel must continue to pursue a healthy partnership with technology to combat the ever-changing threat of crime and violence. In the educated words of Criminal Justice professor Dennis Kenney, "Guns are more ubiquitous than they used to be. There's inevitably going to be more gunplay on the streets, and it's inevitable that police will want to begin to use technology to help address that."³¹

REFERENCES

- 1. Van der Kloot W. Lawrence Bragg's role in the development of sound-ranging in World War I. Notes and Records of the Royal Society. 2005;(59):273-284.
- 2. Melson CD. Organization and Equipment for the Defense Battalion. In: Condition Red: Marine Defense Battalions in World War II. United States Marine Corps-Editing and Design Section, History and Museums Division. 1996.
- Locating Surveillance and Target Acquisition Association. Locating Artillery Overview. http://www.locatingartillery.org/overview.htm
- 4. Varshney L. Ground Surveillance Radars and Military Intelligence. North Syracuse (NY): Syracuse Research Corporation; 2002 Dec.
- 5. Lahr JC, Ward PL, Stauffer PH, Hendley JW. Earthquake Technology Fights Crime. United States Geological Survey-Earthquake Hazard Reduction Program. http://jclahr.com/science/psn/gunshots/factsheet/index.html
- 6. ShotSpotter, History. http://shotspotter.com/company/history
- 7. Hsu J. Tracing a Bullet Back to Its Gun. ScienceLine.org. 2007 Jan. http://scienceline.org/2007/01/tech-hsu-guns/
- Defense Industry Daily. Sniping at US Forces Beginning to Boomerang. DefenseIndustryDaily.com. 2011 Feb. http://www.defenseindustrydaily.com/sniping-at-us-forces-beginning-toboomerang-01128/
- 9. Raytheon BBN Technologies. Boomerang Warrior-X. 2010 Jul. http://bbn.com/resources/pdf/Boomerang-Warrior-X-072210.pdf
- Al-Muslim A. Gunfire detection coming to Hempstead Village. Newsday.com. 2012 Jul. http://www.newsday.com/long-island/nassau/gunfire-detection-coming-tohempstead-village-1.3828324
- 11. Haag MG, Haag LC. Shooting Incident Reconstruction. 2nd rev. ed. San Diego: Academic Press, 2011.
- 12. Maher RC. Acoustical Characterization of Gunshots. Institute of Electronics and Electrical Engineers. 2007 Apr. 109-113.

- 13. Dean EA. Atmospheric Effects on the Speed of Sound. El Paso (TX): United States Army Electronics Research and Development Command; 1979. Report No.: ASL-CR-79-0100-4.
- 14. Haag MG, Haag LC.
- 15. Ibid.
- Morton KD, Collins L. Bayesian Detection of Acoustic Muzzle Blasts. Sensors and Command, Control, Communications, and Intelligence Technologies for Homeland Security and Homeland Defense VIII (proceedings of SPIE). 2009; 7305.
- 17. Valenzise G, Gerosa L, Tagliasacchi M, Antonacci F, Sarti A. Scream and Gunshot Detection and Localization for Audio-Surveillance Systems. Institute of Electronics and Electrical Engineers. 2007.
- 18. Buckley C. High Tech 'Ears' Listen for Shots. NYTimes.com. 2009 Nov. http://www.nytimes.com/2009/11/22/nyregion/22shot.html?pagewanted=1&_r=2 &emc=eta1
- Feskos B. Wilmington police hope gunfire sensor will help response times. StarNewsOnline.com. 2011 Jun. http://www.starnewsonline.com/article/20110609/ARTICLES/110609693?p=1&t c=pg
- 20. Kaempffer W. New Haven's ShotSpotter reliability in question. NHRegister.com. 2012 Mar. http://www.nhregister.com/articles/2012/03/12/news/new_haven/doc4f5d6e0426b cf101484344.txt?viewmode=fullstory
- 21. Fair M. Trenton council rejects expansion of 'ShotSpotter' gunshot detection system. NJ.com. 2012 Jan. http://www.nj.com/mercer/index.ssf/2012/01/trenton_council_rejects_expans.htm l
- 22. Hong A. RPD: ShotSpotter Valuable to Crime Solving. 13WHAM.com. 2012 Jun. http://www.13wham.com/news/local/story/rochester-shotspotter-police-alertssensors/uSDQnaKgw0i-CRyzn4HQ_w.cspx
- 23. Selby N, Henderson D, Tayyabkahn T. ShotSpotter Gunshot Location System Efficacy Study. CSG Analysis, endorsed by the National Organization of Black Law Enforcement Executives. 2011.

- 24. Platt J. ShotSpotter Stat Update. TheHuntingtonian.com. 2012 May. http://thehuntingtonian.com/2012/05/25/shotspotter-stat-update/
- 25. Mazerolle LG, Watkins C, Rogan D, Frank J. Random Gunfire Problems and Gunshot Detection Systems. Washington (DC): National Institute of Justice, United States Department of Justice; 1999 Dec. NCJ No.: 179274.
- 26. Beck SD, Nakasone H, Marr KW. An Introduction to Forensic Gunshot Acoustics. Proceedings of the 162nd Meeting of the Acoustical Society of America; 2011 Oct 31-Nov 4; San Diego (CA).
- 27. González-Castaño FJ, Alonso JV, Costa-Montenegro E, López-Matencio P, Vicente-Carrasco F, Parrado-García F, et al. Acoustic Sensor Planning for Gunshot Location in National Parks: A Pareto Front Approach. Sensors 2009 Nov;(9):9493-9512.
- 28. Chacón-Rodríguez A, Julián P. Evaluation of Gunshot Detection Algorithms. Proceedings of the Argentine School of Micro-Nanoelectronics, Technology and Applications; 2008:49-54.
- 29. Showen RL, Calhoun RB, Chu WC, Dunham JW. Gunshot Location in Complex Environments-Concepts and Results. In: Carapezza, EM, editor. Acoustic Sensors, and Command, Control, Communications, and Intelligence Technologies for Homeland Security and Homeland Defense VII (proceedings of SPIE). 2008; 6943.

30. Ibid.

31. Buckley C.