Original Research Article

Evaluation of auto pilot situational awareness system using gunshot detection algorithm in a localized environment: case study federal polytechnic Offa, mini campus.

Abstract

Gunshot detection technologies are more applicable in many industries for the security enhancement of public places like the Federal Republic Territory FCT Abuja in Nigeria. Many factors affect the accuracy of the gun detection algorithm. This paper describes an audio-based video surveillance system in an auto pilot situational awareness to detect gunshots in Federal Polytechnic Offa. The video camera is steering regarding the initial position to localize the acoustic source's position. Implementing an auto pilot situational awareness system is an experimental procedure with a gunshot detection algorithm in a localized environment. In the direction of the weapon, the distance between firearms, types of ammunition, types of study environment, and diffraction of audio, the standard feature for gunshot recognition are Mel frequency cepstral coefficients in terms of uniform gamma-tone filters linearly spaced over the whole frequency range from 0KHZ to 16KHZ. Experiments show that our system can detect gunshots with a precision of 93% at a false rejection rate of 5% when the SNR is 10db while proving the estimate of the source direction of the gunshot with an accuracy of one degree. The research recommends a real-time system implementation for protecting the Federal Polytechnic Offa against any form of treats.

Keywords: Auto pilot awareness system, gunshot detection, algorithm, localized environment

1.0 Introduction

National security threat has been a significant issue for the nations in recent years and has become unbearable for the local environment. The level at which intruders interfere, even in the so-called sacred places, is alarming. Recently, Nigeria has been characterized by turmoil ranging from human abduction, political mayhem, terrorism and bomb attacks. Governments have tried several methods to curb these menaces. Still, all of them have been proved abortive because there are too many roads and insufficient security personnel to guard all the locations safely. Due to this, the deficiency of human power (security personnel) will lead to the adoption of the autonomous system of operations in public places to monitor events in a particular location.

In a diverse and active environment where there is poor situation awareness, decision-making is affected by many variables such as sudden sounds like gunshots or alarms. In such cases, the purpose of that or another decision is important and the ability to understand and analyze the current situation within a short period.

Situation awareness is vital in human information processing and pilots' decision-making processes. "Situational awareness" is formally defined as "a perception of the environment within a volume of time and space, the comprehension of their meaning and the projection of their status shortly" [1]. Situational awareness is a critical mental process affecting decision-making and performance [2]. Situational awareness is the leading paradigm in studying the human factor as the source of knowledge and investigating the effects it produces on the interaction with the environment. To improve situational awareness and secure the level of crime within Federal polytechnic Offa. Efforts have to autonomously monitor the regions using an autopilot gunshot detection system, especially the public location. A gunshot detection system is comprised of sensors to detect the sound of a gunshot, transmitters to send a message to the police dispatch centre or security point, and a computer to receive and display that message. When a signal arrives at the police station, the dispatcher decides whether or not to send a unit to respond to the signal.

The initial research into the effectiveness of gunshot detection systems is up-and-coming, particularly regarding the technology's usefulness in identifying and solving problems and detecting crime. As the technology develops and becomes more accurate and portable, these systems could prove to be highly effective tools for local police departments.

Gunshot detection systems use acoustic sensing technology to identify, discriminate, and report gunshots to the security personnel within seconds of the shot being fired. As a problem-solving tool, gunshot detection reports can be used with police or security personnel data (e.g., citizen reports of gunfire) and physical features of a neighbourhood (e.g., parks or liquor stores) to identify neighbourhood hot spots. Suppose demographics (e.g., income level or gun ownership) are considered. In that case, the data can be used to analyze various dimensions of the problem and to evaluate the effectiveness of responses to the problem.

1.1 Campus safety: gunshot detection systems benefit Measure

So far this year, the researcher have already matched the number of deaths from school shootings as conducted did in all of year 2021 [3]. While school shootings are unfortunately becoming a common stance in our country, implementing a gunshot detection solution could help prevent undue casualties by being the first link in the chain of events to dispatch police and warn staff of the imminent danger to our children.

Gunshot detection systems were initially developed for use by the military, but in recent years, private companies have begun urging school systems to adopt the technology. Tragedies that receive national attention have caused many parents to wonder if enough is being done to protect children in the classroom or school environment.

Multiple companies install and maintain gunshot detection systems, but generally, sensors are placed throughout a building that detects when and where a gun is fired. The information is relayed to police, administrators, and other first responders within milliseconds. Alerts are also sent out within a building to notify the people inside that there is an active shooter in the building.

Suppose a growing number of gunshot detection systems are being installed in schools and other buildings throughout the country. Although few schools currently have the technology, advocates say near-instant detection of gunshots in mass shooting situations could save lives.

Industry representatives said gunshot detection systems modernize the way organizations respond to active shooter situations. Instead of relying on human behaviour in a chaotic situation, such as a teacher or worker calling 112 of emergency code to notify the authorities, the system feeds law enforcement near real-time data about the threat and allows them to quickly develop a plan to neutralize it, representatives said.

Imagine in one second, boom, and one second, knowing it's a gunshot, and know where they are. In case of a mass shooting, every 18 seconds, considering a victim dies, so every second counts.

Earlier this year, one company, Shooter Detection Systems (SDS), received "Safety Act" certification from the Department of Homeland Security. Safety Act certification provides companies with protection against liability when deploying anti-terrorism technology.

SDS said their systems rely on both acoustic and infrared sensors, and the company said their systems have never registered a "false alarm."

Limitations

Extensive research of prior gunshot detection systems has proven that there are many approaches to existing gunshot detection systems with different algorithms; such as:

- Host of complicated power-consuming algorithms: These detection algorithms could range from the Mel Frequency Cepstral Coefficients to adaptive background noise cancellation through multiple layers of notch & band pass filtering.
- The triangulation (location) through TDOA (Time-Difference 4 on Arrival) of the gunshot must be calculated through the consistent speed of sound calculations, and generalized cross-correlation phase transforms. While these are computationally intensive tasks, this aspect may be handled by a computer receiving the data and not on the processors in the field, allowing them to be purely used for detection.

While some systems rely solely on acoustic information (and the false alarm rate quoted is likely about right for the worst of those), "the better systems utilize three sensor technologies and achieve false-positive rates less than 1/10 of 1%". In testing records, this is much better accuracy than the video analytics systems. Cost is always a factor; the dedicated systems can be as much as 70% less to install and maintain and do not require live monitoring from a central station. In addition, these systems provide real-time updates directly to law enforcement, including the firearm type and the shooter's location indicated on a floor plan of the building. Each facility is different, and requirements vary. However, let's not rule out the dedicated systems just yet.

2.0 Review on Evaluating Gunshot Detection System for Schools

According to[4], since the year 2010, the United States averages 15 lives lost to mass shootings annually in our school system. The average might have been higher if not for the COVID-19 pandemic, which kept our children and teachers out of school last year. According to the FBI, between the years 2000 and 2018, 74% of mass shootings in education occurred at or below the 12th-grade level.

A few months passed, and President Biden stated, "Gun violence in this country is an epidemic." It has become so bad that more and more schools are starting to implement gunshot detection technology within their halls.

The gunshot detectors will protect lives like a fire alarm system's manual pull station (pull stations do not prevent fires from starting, but hopefully, they are activated early enough to save as many lives as possible).

According to [5], the origins of many technological advances are often traced back to innovations in different fields, later made applicable through a 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 opposing side made adjustments to the process to find increasingly useful results.

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. Although sound-ranging equipment had grown 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. Radar operators were capable of locating large weaponry faster. This is derived from more conclusive data in settings of extreme terrain or overgrown vegetation. The equipment could be outfitted on more mobile units for determining the location of aeroplanes and vehicles, and most importantly, radar could operate without waiting for shots to be fired. Sound ranging still held a place in combat but acted mainly as a backup to rapidly expanding radar capabilities.

Techniques involving sound-ranging for gunfire locations 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. Organizations such as ShotSpotter Incorporated, now SST Inc., tested detection and localization systems in areas with high crime rates. US police departments along the Californian Pacific coast began working with the technology to improve incident response time and help deter future crimes.

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

Fig 1.2: Boomerang system outfitted to a US Army Humvee

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

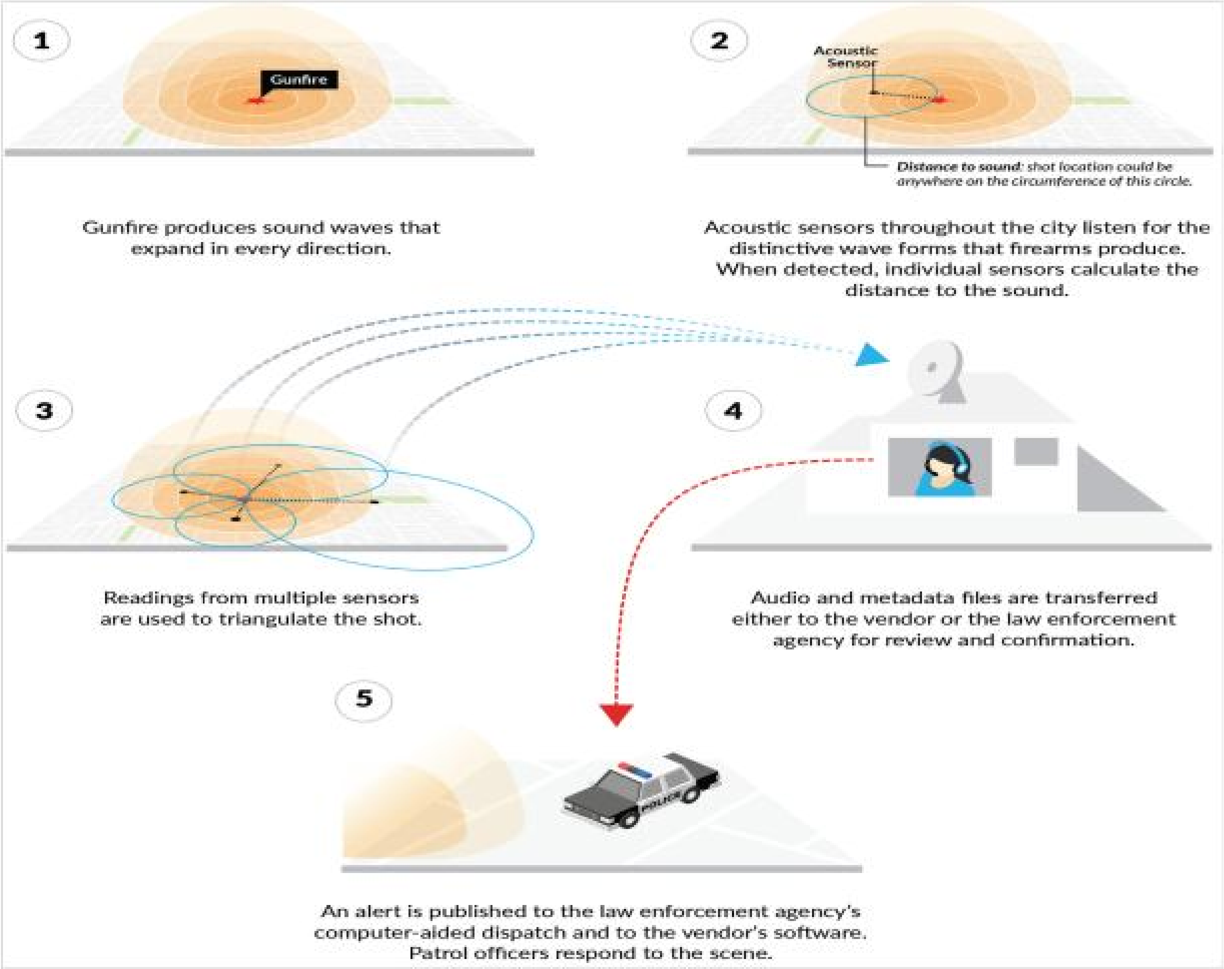
Back in the United States, agencies nationwide have deployed gunshot detection and localization systems in cities and other urban areas that are prone to gunfire-related crimes and random gunfire incidents. These systems are receiving more consideration as a significant contributor to community safety and law enforcement success and offer increased response capabilities and potential video evidence, incorporating video capture components in the system designs. While some critics raise concerns, including costs, privacy issues, and accuracy, gunshot detection and localization systems used in American cities have significantly impacted how authorities identify and respond to criminal activity.

3.0 System Design

How does gunshot detection work?

A sensor is used to listen for the acoustic properties unique to the sound created when a gun is fired and then triggers an alarm, usually auto dispatching the police. In most single-purpose devices, this is all the gunshot detector does. "Once gunfire is detected, an alert is sent to our 911 centre, and a text message is sent to our city police administrators and patrol supervisors. Most of this is done before anyone can even think about picking up a phone and dialling security number due to a chaotic situation, thus improving response time."

A dedicated gunshot detection system added benefits like pinpointing a shooter's location through triangulation. Adding both cameras and gunshot detectors can be costly for any customer. While unable to triangulate, a dual-purpose camera can track a shooter's direction if installed throughout a school.



Gunfire produces sound waves that expand in every direction.

Acoustic sensors throughout the city listen for the distinctive wave forms that firearms produce. When detected, individual sensors calculate the distance to the sound.

Readings from multiple sensors are used to triangulate the shot.

Audio and metadata files are transferred either to the vendor or the law enforcement agency for review and confirmation.

An alert is published to the law enforcement agency's computer-aided dispatch and to the vendor's software. Patrol officers respond to the scene.

Fig 1.3 Gunshot Detection Technology Data Flow

Methodological Processes

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.

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. 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.

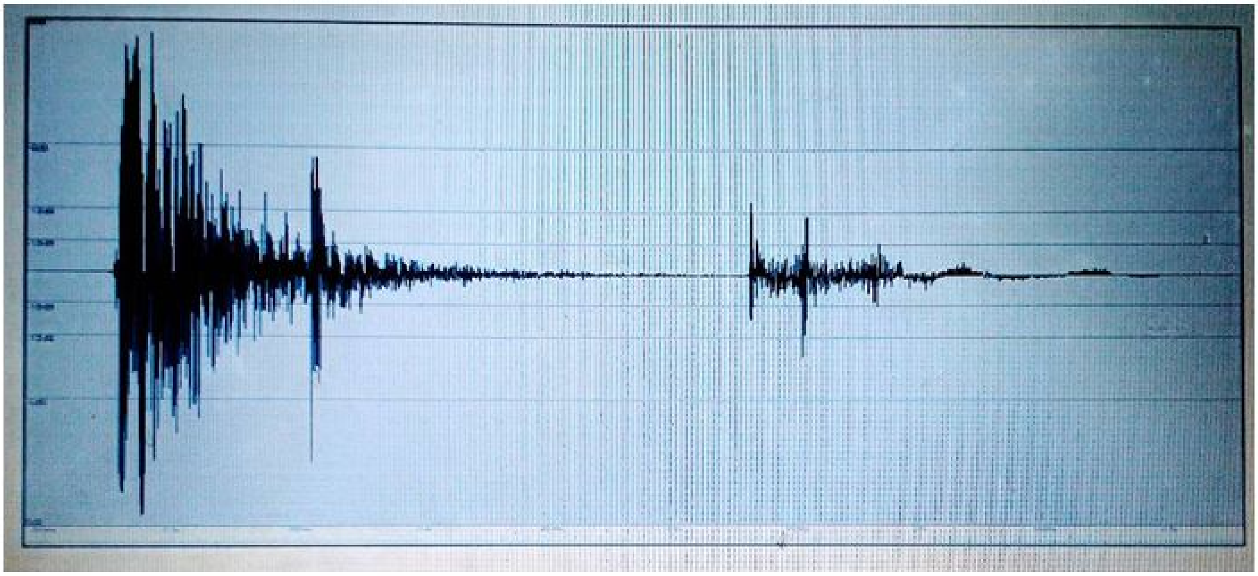


Figure 2.1 Waveform of .22 Caliber Rifle Gunshot with Reflection

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, most likely the “response” of the original sound event bouncing off the back 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 travelling wave maintains a constant speed over a known distance, the elapsed time will also be constant. 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 closest 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

Figures 2.2 through 2.5.

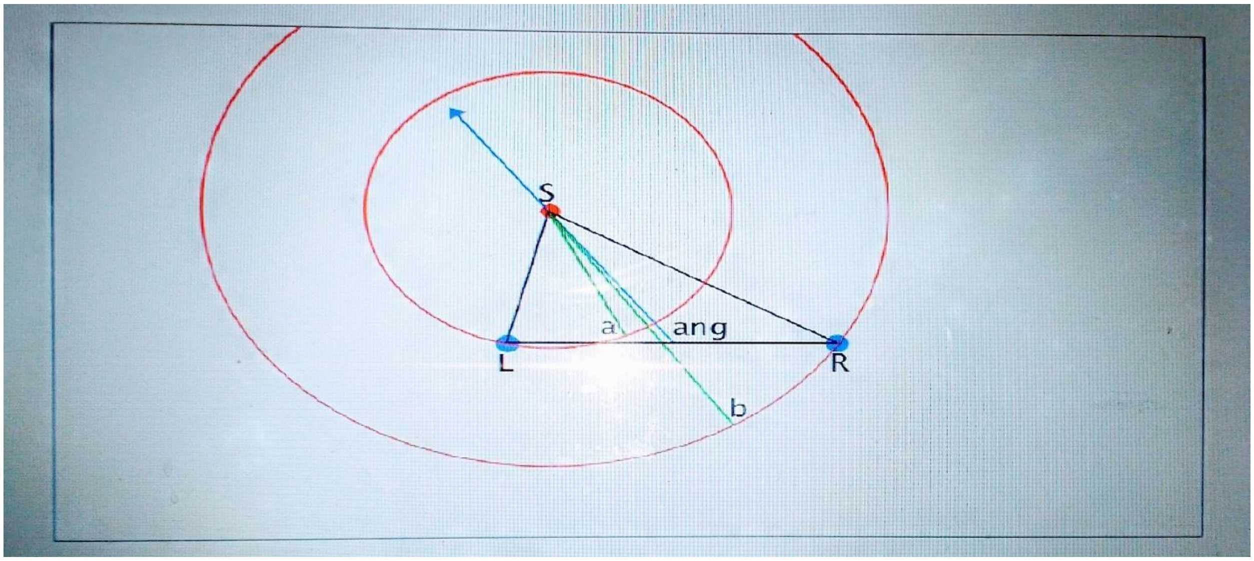


Figure 2.2 Angle Determination from Sound Delay between Two Microphones

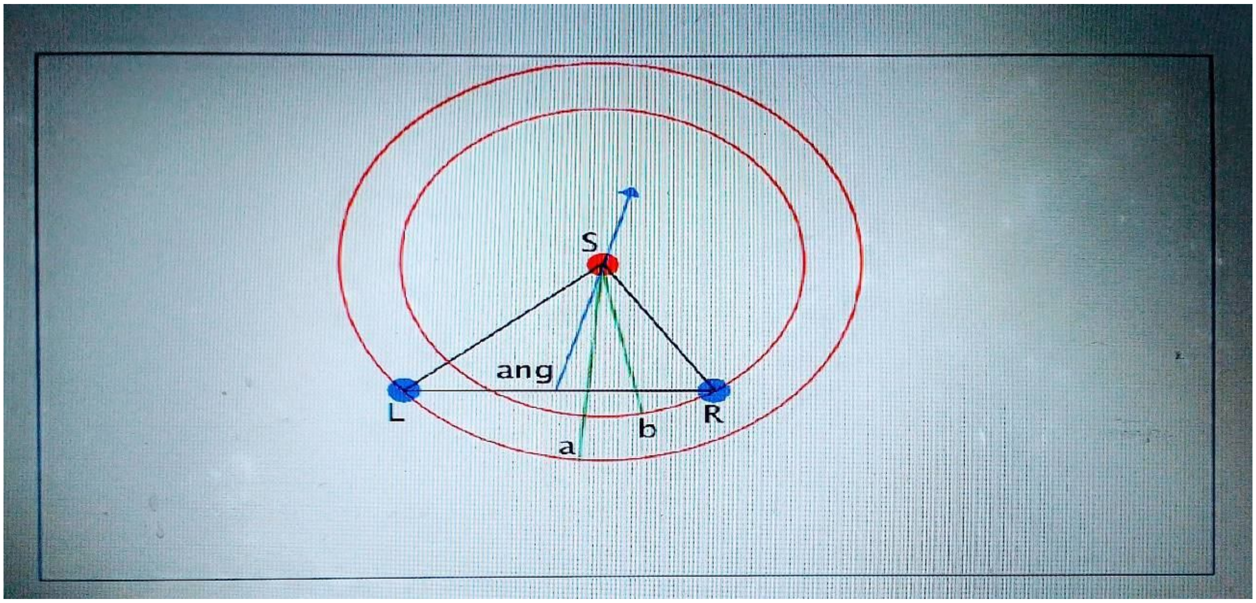


Figure 2.3 Angle Determination from Sound Delay between Two Microphones (II)

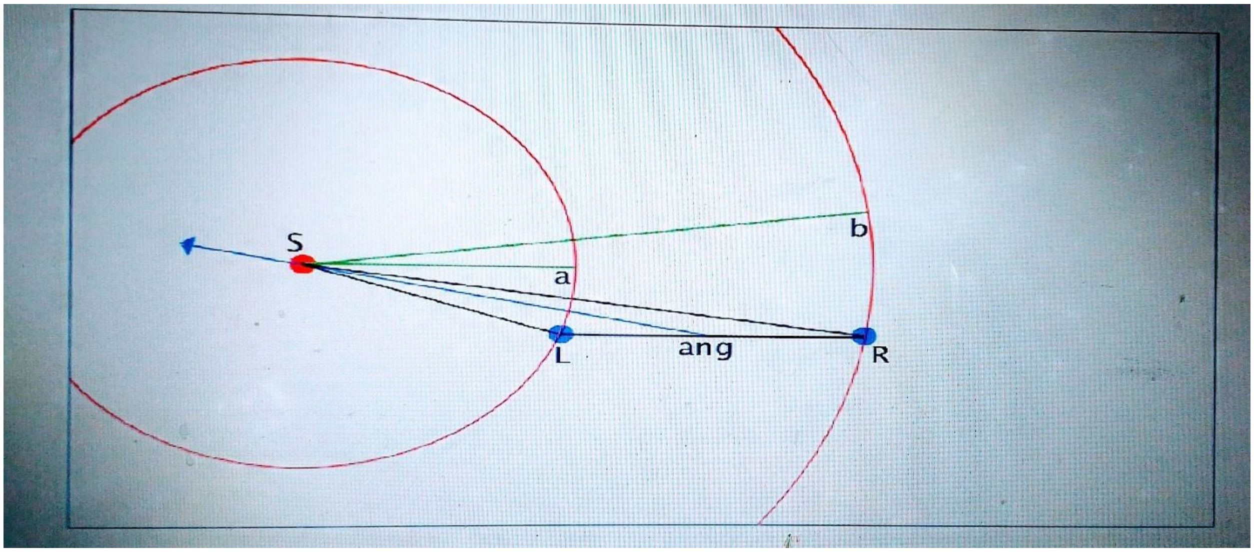


Figure 2.4 Wide Angle Determination from Sound Delay between Two Microphones

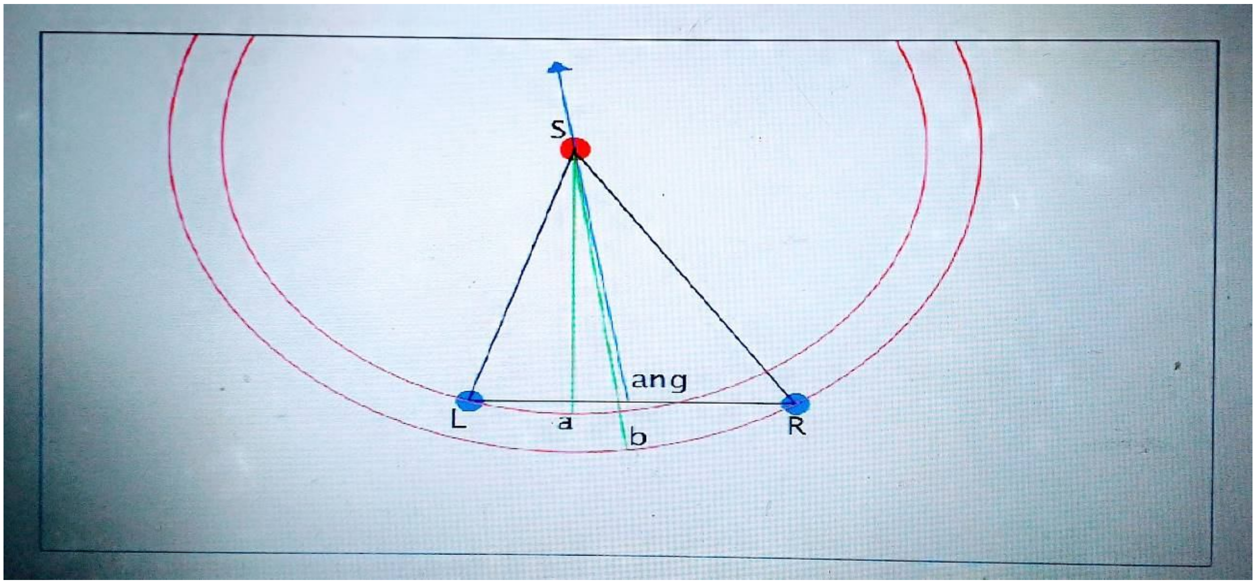


Figure 2.5 Distanced Angle Determination from Two-Microphone Sound Delay

With sound source *S*, the 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 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 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 the 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 an actual operation. Still, an inexpensive pair of smaller microphones are 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 the 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 connected back to the laptop computer for display purposes.

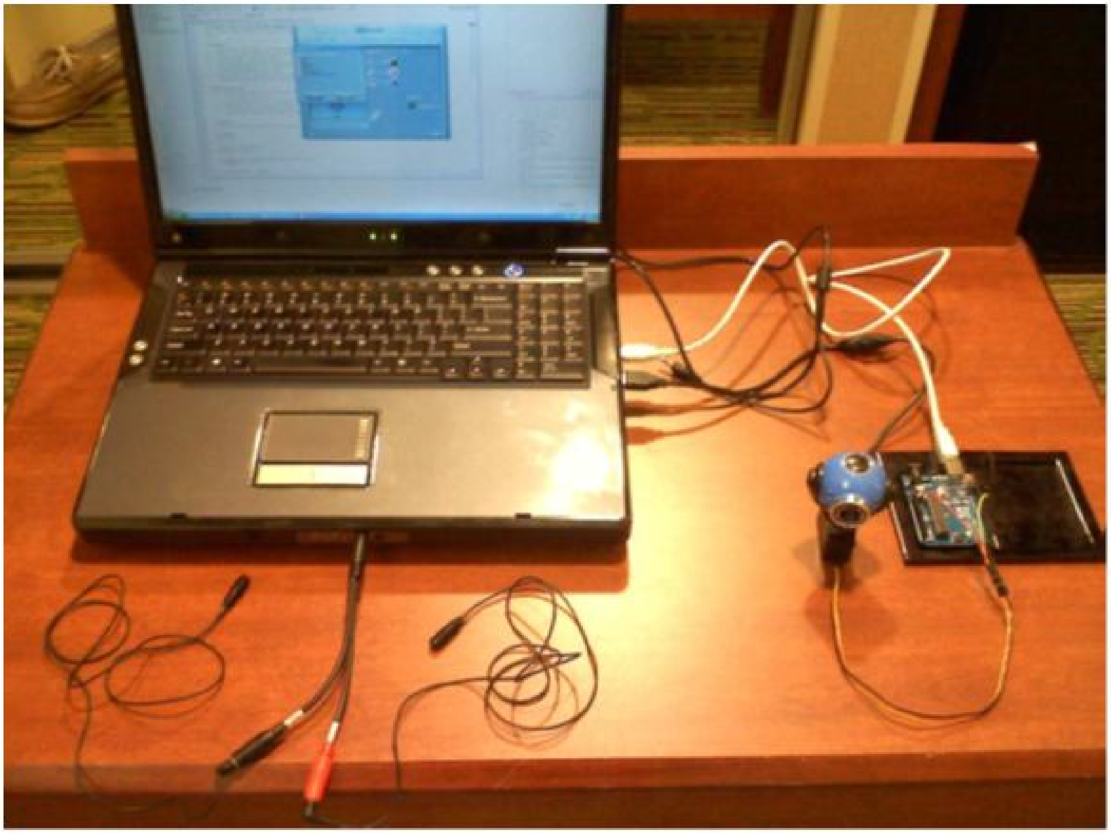


Figure 3.1 Gunshot Detection and Localization System, Basic Design – Overall

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 it 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 a 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 computer through another USB connection for display and potential recording purposes. This workflow is outlined in Figure 3.2 below.

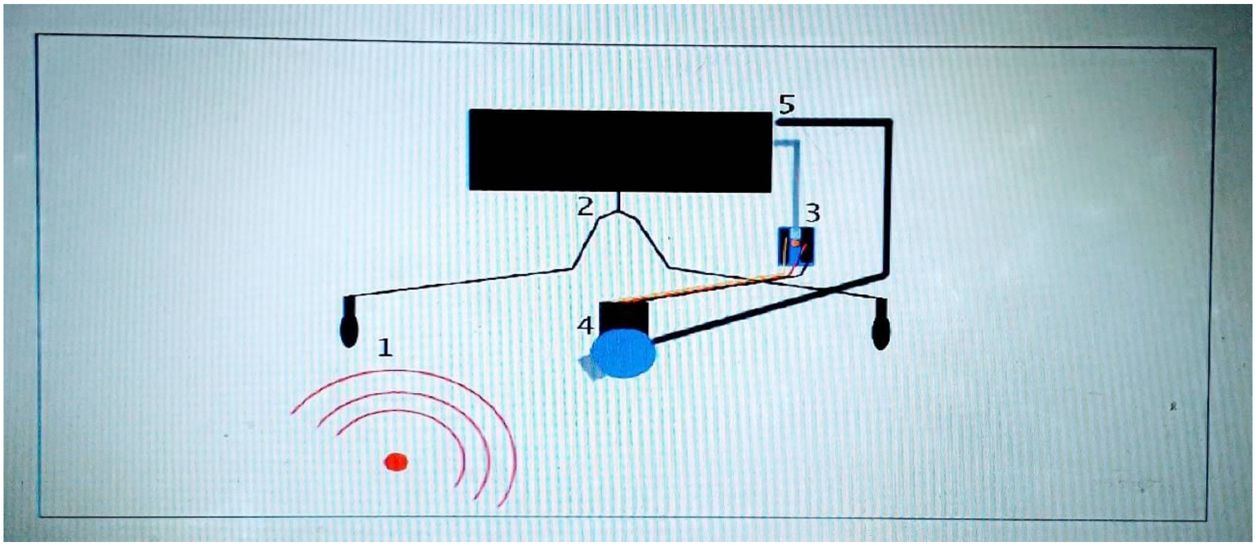


Figure 3.2 Basic System Workflow

1) 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; 4) Arduino sends voltage to the 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

The system is configured correctly, and the Arduino and MatLab must be appropriately programmed. 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 motors. Pdf file must be uploaded to the Arduino IDE, and the appropriate AFMotor.h and AFMotor.cpp files must be appropriately allocated. For instructions on performing these steps, refer to the forum on MatLab's home page.

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

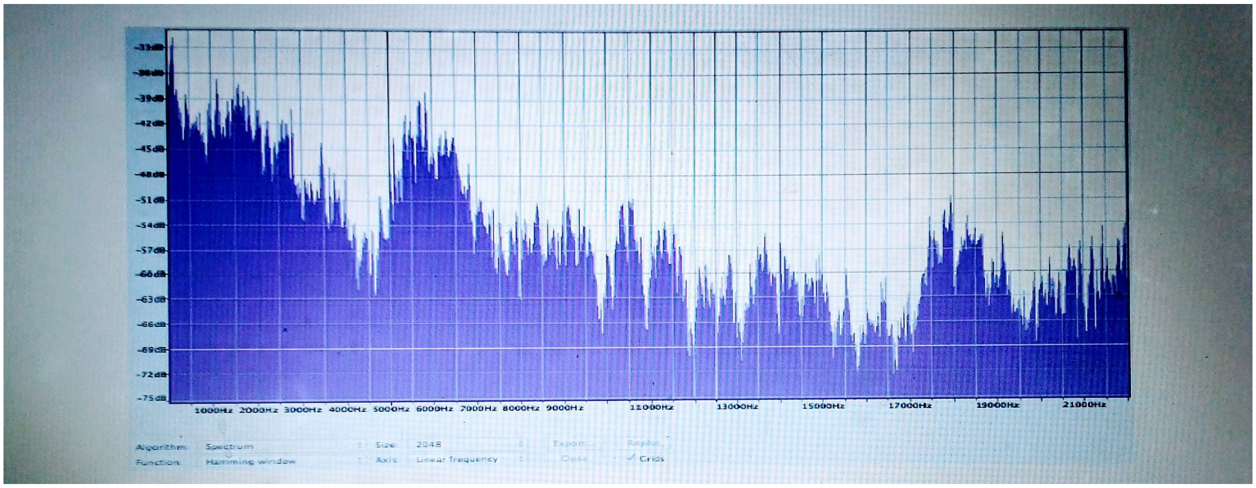
Audio Detection and Image Response Script

```
% 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 camera 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 Yadapter
%% 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)
```



Algorithm: Spectrum
Function: Hamming window
Size: 2048
Axis: Linear frequency

Fig 3.3:

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.

Thresholds in Audio Detection and Image Response Script

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

According to the script above, which refers to the threshold element of the script, the “th,” threshold is 1200 quantization levels of relative signal power. This setting depends on multiple factors, including microphone gain settings, expected distance from a sound source to microphones, and expected background noise. Due to these many factors, the “th” setting requires careful calibration for each deployment. Corresponds to a threshold of 200 samples audio is 48 kHz

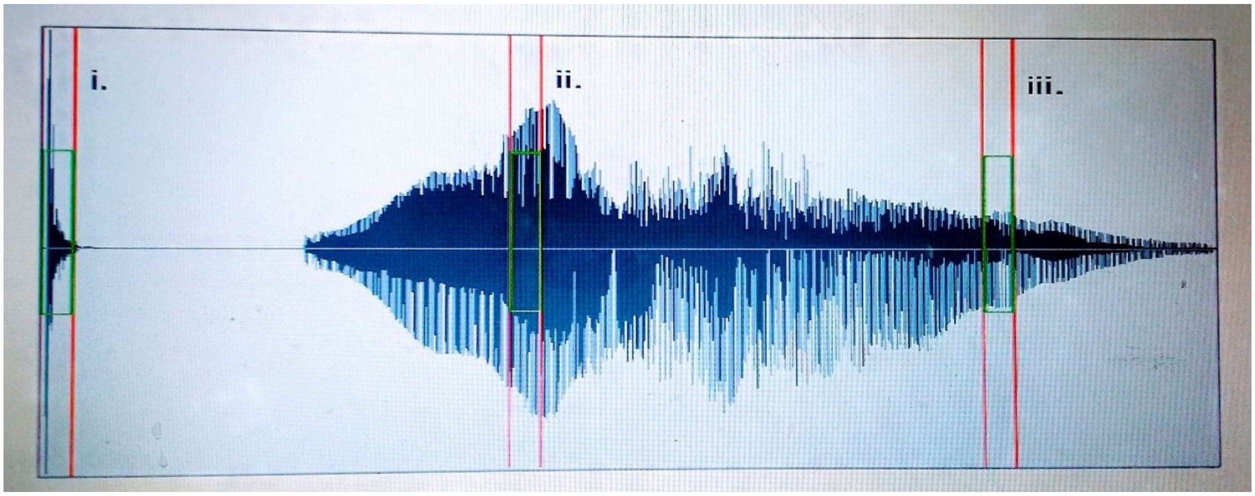


Figure 3.4 Discriminatory Thresholds for Audio Event Exclusion

Figure 3.4 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, to provoke a system response. The other significant portion of the project script is the delay calculation and angle conversion portions (Below Script).

Channel Delay and Angle Calculation in Detection and Response Script

```
[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)
```

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 includes 0 and spans from 0 to 179.
Further explanation is necessary for the script's delay and angle calculation portions. 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.

Assuming the speed of sound is approximately 350 meters per second, a sound wave would travel 6 feet (or about 1.829 meters) in around .0053 seconds. Six feet is the prescribed distance between microphones used in the system, and .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)

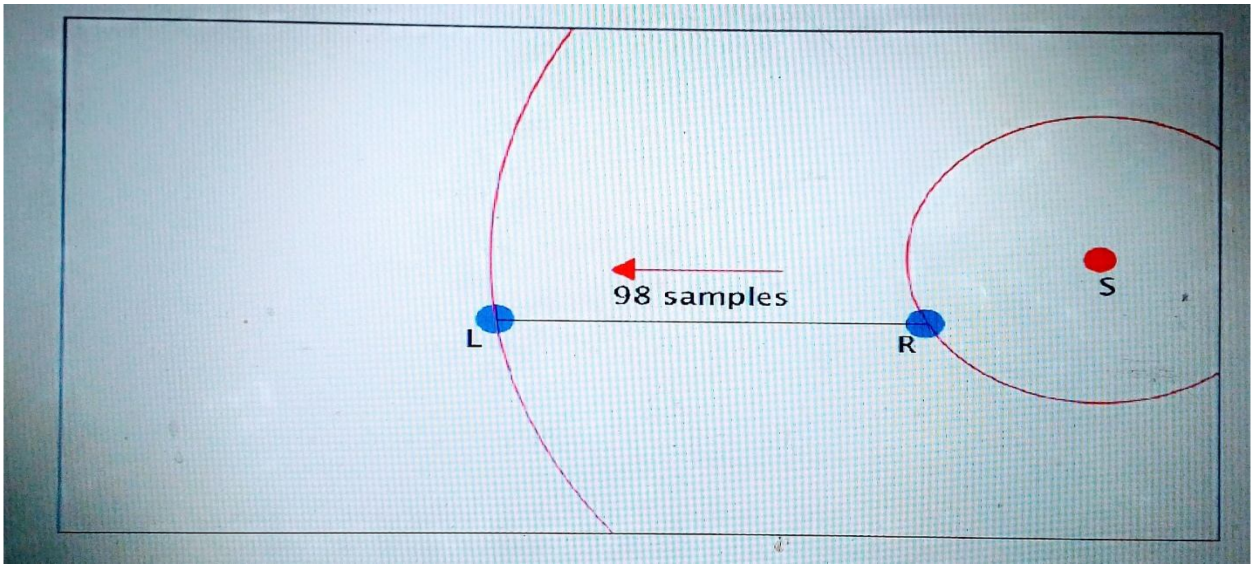


Figure 3.5 Delay and Angle Calculation Example 1

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):

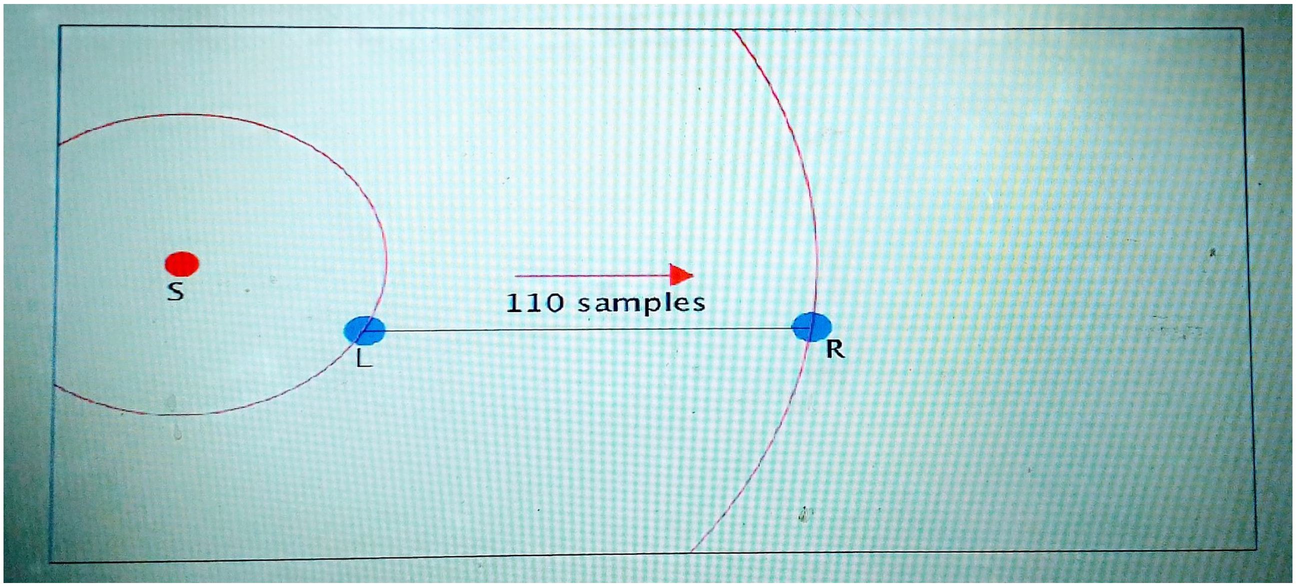


Figure 3.6 Delay and Angle Calculation Example 2

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 *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.

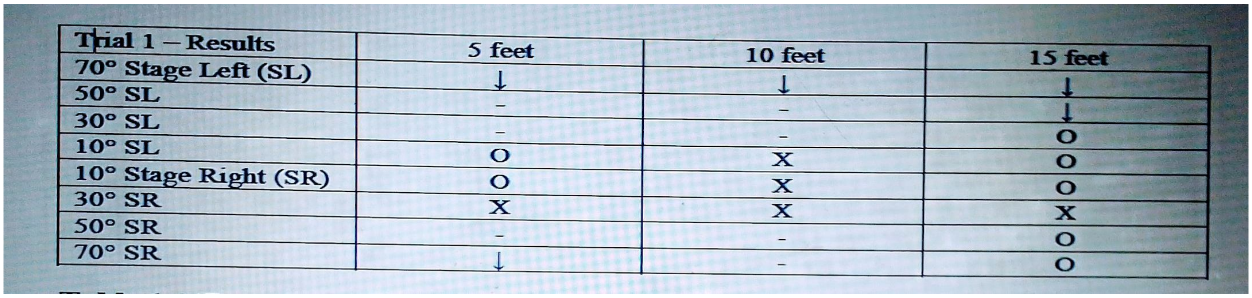
At the time of the test, an actual firearm was not an available sound source. Instead, a loud, sharp clap of the hands was utilized at each marker. The overall waveform shape of a hand clap could properly simulate a gunshot because both events can be characterized with high intensity and short duration. 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 ↓ 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)

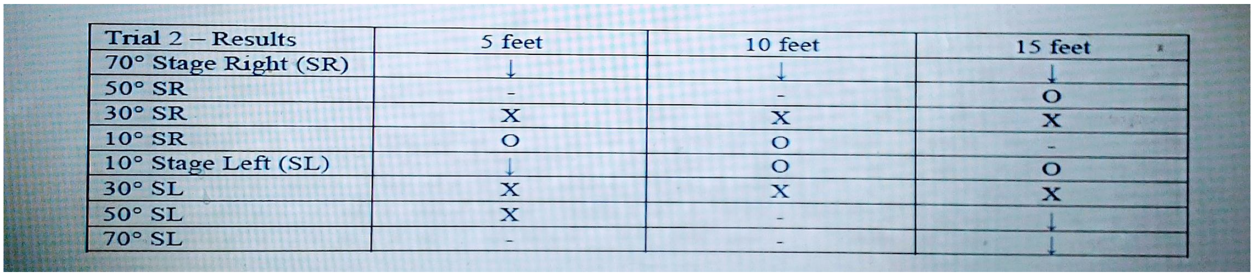
The responses were determined after some camera movement in response to the claps, or after a maximum 5 clap attempts at the marker.

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.



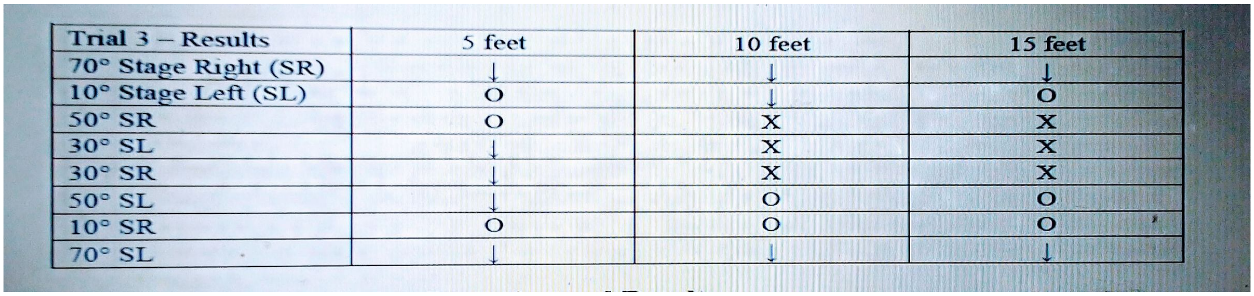
Trial 1 – Results	5 feet	10 feet	15 feet
70° Stage Left (SL)	↓	↓	↓
50° SL	-	-	↓
30° SL	-	-	O
10° SL	O	X	O
10° Stage Right (SR)	O	X	O
30° SR	X	X	X
50° SR	-	-	O
70° SR	↓	-	O

Table 1.1 Test Trial 1 Configuration and Results



Trial 2 – Results	5 feet	10 feet	15 feet
70° Stage Right (SR)	↓	↓	↓
50° SR	-	-	O
30° SR	X	X	X
10° SR	O	O	-
10° Stage Left (SL)	↓	O	O
30° SL	X	X	X
50° SL	X	-	↓
70° SL	-	-	↓

Table 1.2 Test Trial 2 Configuration and Results



Trial 3 – Results	5 feet	10 feet	15 feet
70° Stage Right (SR)	↓	↓	↓
10° Stage Left (SL)	O	↓	O
50° SR	O	X	X
30° SL	↓	X	X
30° SR	↓	X	X
50° SL	↓	O	O
10° SR	O	O	O
70° SL	↓	↓	↓

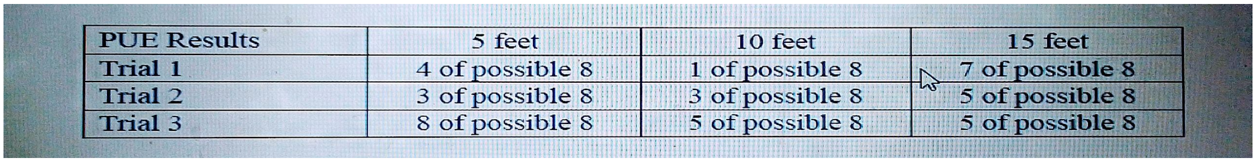
Table 1.3 Test Trial 3 Configuration and Results

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).

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 camera's viewpoint outward towards the markers).

Discussion

At first glance, the results from the test 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 Evidence" (or "PUEs") were created. These are defined within the test realm as camera responses that end with the location of the sound source somewhere in the frame, either centred or not centred. 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 the frame, either centred or not. Within those parameters, the test resulted in 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 eight possible instances for PUE response refer to each marker at the given distance.

This table of PUE results helps clarify precisely what 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 helpful evidence increases in Trial 3 from the others, Trials 1 and 2 exhibit increases in PUEs as the space grows. The opposite results occur in Trial 3, where the PUE decreases and stays the same at each distance. This is a counterintuitive result since one would assume that events at further distances are more likely to occur within a camera's frame of view.

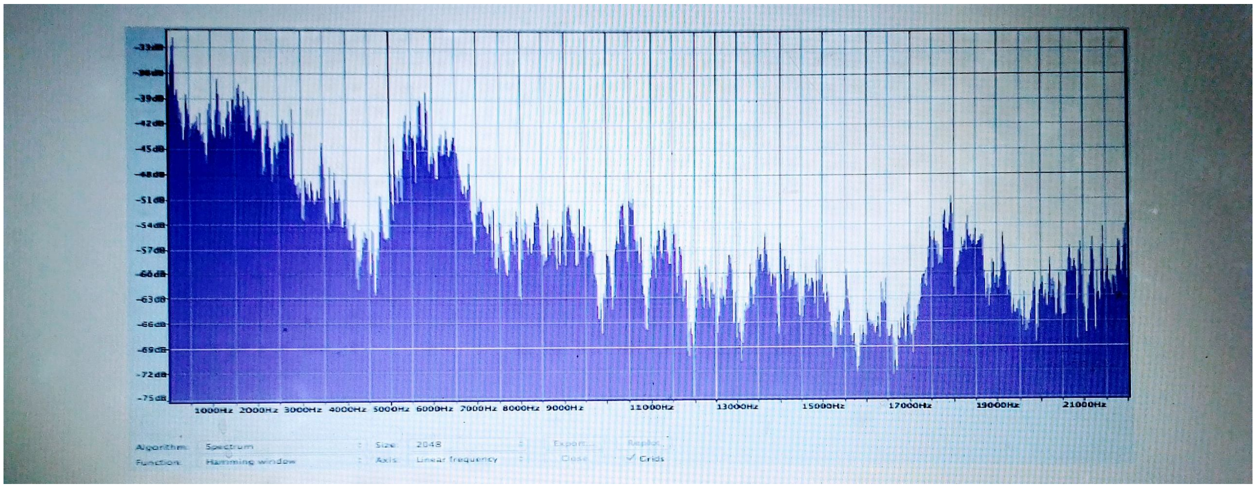
The overall inconsistency in the system can be attributed to multiple factors. First, the microphones used may not be most suitable for the discrimination task. Due to the concise duration of most gunshot sounds, microphones should be sensitive enough to accurately define the incoming audio information to the point that 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 significant 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 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. 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 adjust for changes in the noise floor, etc., automatically. Real-time adaptive filters might also work in 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.

These instructions, tests, and results are useful in describing the basic detection and response process, but the design carries an underlying flaw. Any short and loud impulse 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. These ranges are standard for various other sounds, making discrimination much more complicated.



Algorithm: Spectrum
Function: Hamming window
Size: 2048
Axis: Linear frequency

Figure 3.7 FFT of .22 Caliber Rifle Gunshot with Reflection

Using Fast Fourier Transform to bring a signal into the frequency domain, Figure 3.7 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 incorporate some “environmental acoustics” due to the reflection in the recording, but that would be expected for most audio events in realistic scenarios. To correctly discriminate from other sounds, the best approach at this point is an algorithmic learning strategy, such as those proposed by Morton and Collins 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 regarding 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 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.

5.0 CONCLUSION

While school shootings are unfortunately becoming a common stance in our country, implementing a gunshot detection solution could help prevent undue casualties by being the first link in the chain of events to dispatch police and warn staff of the imminent danger to the students.

If a growing number of gunshot detection systems are being installed in schools and other buildings throughout the country, instant detection of gunshots in mass shooting situations could save lives.

Though gunshot analytics are certainly not perfect — a car backfiring or a book being slammed on the ground could trigger a false alert — the idea is to use these alerts and have a human immediately review the video to verify a threat or lack thereof.

As with using a central station for intrusion video verification, instead of automatically dispatching police, integrators can work with their significant stations as a go-between to only dispatch when video supports escalating the situation for immediate response.

In testing records, the cost is always a factor, and the dedicated systems can be as much as 70% less to install and maintain and require live monitoring from a central station. In addition, these systems provide real-time updates directly to law enforcement, including the firearm type and the shooter's location indicated on a floor plan of the building. Each facility is different, and requirements vary.

REFERENCE

[1] Endsley, M. R. (1987). SAGAT: A methodology for the measurement of situation awareness (NOR DOC 87-83). Hawthorne, CA: Northrop Corporation.

[2] Endsley, M. R. (1988). Situation awareness global assessment technique (SAGAT). In Proceedings of the IEEE 1988 National Aerospace and Electronics Conference – NAECON 1988 (Vol. 3, pp. 789-795). New York: Institute of Electronical and Electronics Engineers.

[3] Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. Human Factors, 37, 65-84.

[4] Aljazeera "140 students missing after gunmen raid Nigerian school" Name of Newspaper, 5 Jul 2021, https://www.aljazeera.com/news/2021/7/5/140-students-missing-after-gunmen-raid-nigerian-school

[5] Peter W. Greenwood, Ph.D., Director, Criminal Justice Research Program, The RAND Corporation: Three Strikes, You're Out: Benefits and Costs of California's New Mandatory-Sentencing Law.

[6] 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.

[7] 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.

[8] Christian Pfeiffer, Ph.D., Director, Kriminologisches For schungs institute Niedersachsen: Sentencing Policy and Crime Rates in Reunified Germany.

[9] Arthur L. Kellermann, M.D., M.P.H., Director, Center for Injury Control, and Associate Professor, Emory University: Understanding and Preventing Violence: A Public Health Perspective.

[10] Locating Surveillance and Target Acquisition Association. Locating Artillery, Overview. http://www.locatingartillery.org/overview.htm

[11] Varshney L. Ground Surveillance Radars and Military Intelligence. North

[12] 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

[13] ShotSpotter, History. http://shotspotter.com/company/history

[14] Hsu J. Tracing a Bullet Back to Its Gun. ScienceLine.org. 2007 Jan.

[15] Defense Industry Daily. Sniping at US Forces Beginning to Boomerang. http://www.defenseindustrydaily.com/sniping-at-us-forces-beginning-toboomerang-01128/

[16] 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] Syracuse (NY): Syracuse Research Corporation; 2002 Dec. http://scienceline.org/2007/01/tech-hsu-guns/DefenseIndustryDaily.com. 2011 Feb.

[19] Endsley, M. R., & Kris, E. O. (1995). The out-of-the-loop performance problem: Impact of level of control and situation awareness. Human Factors and Ergonomic Society, 37(2), 381-394.

[20] Endsley, M. R. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), Automation and human performance: Theory and applications (pp. 163-181). Mahwah, NJ: Lawrence Erlbaum.

[21] Endsley, M. R., & Strauch, B. (1997). Automation and situation awareness: The accident at Cali, Colombia. Proceedings of the Ninth International Symposium on Aviation Psychology (pp. 877-881). Columbus, OH: Ohio State University.

[22] Endsley, M. R., Garland, D. J., Shook, R. W. C., Coello, J., & Bandiero, M. (2000b). Situation awareness problems in general aviation (SAT Tech 00-01). Marietta, GA: SA Technologies, Inc.

[23] SecuritySales, Education week. http://securityscales.com/

Alfred Blumstein, Ph.D., Professor, Carnegie Mellon University: Youth Violence, Guns, and Illicit Drug Markets.