Audio Engineering Society

Convention Express Paper 122

Presented at the 155th Convention

2023 October 25-27, New York, USA

Close and Distant Gunshot Recordings for Audio Forensic Analysis

Robert C. Maher

Electrical & Computer Engineering, Montana State University, Bozeman, MT 59717-3780

Correspondence should be addressed to rmaher@montana.edu

ABSTRACT
We describe contemporary forensic interpretation of multiple concurrent gunshot audio recordings made in acoustically complex surroundings. Criminal actions involving firearms are of ongoing concern to law enforcement and the public. The U.S. FBI annually lists 166,000 criminal incidents involving firearms each year. Meanwhile, over 80% of the large general-purpose law enforcement departments in the U.S. now use audio-equipped body-worn cameras (BWCs), more than 135 communities use ShotSpotter® gunshot audio detection systems, and tens of millions of private homes and businesses now have round-the-clock surveillance camera systems—many of which also record audio. Thus, it is increasingly common for audio forensic examination of gunshot incidents to include multiple concurrent audio recordings from a range of distances. A case study example is presented.

1 Introduction

More and more criminal justice proceedings involve audio forensic evidence in general, and gunshot audio recordings in particular, due to (i) the large number of law enforcement incidents each year in the United States involving firearm use [1], (ii) the widespread use of ShotSpotter® acoustic gunshot detection systems [2], (iii) the millions of newly installed residential security and doorbell camera systems with audio capability [3], and (iv) the fact that most general-purpose law enforcement agencies in the U.S. are now utilizing body-worn cameras (BWCs) for all routine work [4]. In fact, law enforcement investigators now routinely canvass neighborhoods and seek access to private surveillance recordings following shootings and other criminal and civil incidents involving gunshot sounds.

This work extends our previously reported research sponsored by the U.S. National Institute of Justice in nearfield gunshot acoustical analysis [5][6][7][8]. In our current research, we focus upon audio forensic gunshot audio cases in which concurrent recordings from close and distant locations are available. For example, a microphone within a few tens of meters of the firearm, such as from a body-worn camera or a nearby residential surveillance system, and a concurrent recording from a few hundred meters away obtained by ShotSpotter, a commercial gunshot detection and localization system. With the known "ground truth" from incident reports and the audio recorded near the incident scene, we compare and correlate information from ShotSpotter to develop general recommendations and best practices for improved audio forensic examination of gunshot sounds recorded in complex acoustical environments.

In this paper, we briefly review the acoustical properties of firearm sounds close to the gun and at greater distances. Next, we present a case study involving gunshots from an incident in which nearby home surveillance cameras capture the gunshot sounds, while simultaneously a ShotSpotter system also detects the gunshots. Finally, we offer several remarks and suggestions regarding this type of analysis, which we feel will become increasingly common in audio forensics for the reasons mentioned above.

2 Gunshot sounds

A conventional firearm uses a cartridge containing gunpowder affixed with a bullet. Triggering the primer in the cartridge rapidly combusts the gunpowder within the confining cartridge, and the hot combustion gases expand rapidly behind the bullet, abruptly forcing the bullet and a jet of gas out of the muzzle. The resulting *muzzle blast* causes an acoustic pressure wave forming a loud pop sound lasting only a few milliseconds. The muzzle blast is notably directional: the on-axis sound level is more intense than the level off to the side or toward the rear of the firearm.

In summary, gunshot muzzle blast sounds are (a) high amplitude (140+ dB SPL); (b) short in duration (2-3 milliseconds); (c) directional (sound level varies as a function of azimuth with respect to gun barrel); and (d) recorded with echoes and reverberation (acoustic “impulse response” of the physical surroundings).

The firearm itself may produce relatively subtle mechanical sounds before, during, and after the bullet is discharged. These sounds may include the mechanical action of the gun, such as the trigger and cocking mechanism, ejection of the spent cartridge, and positioning of new ammunition. These characteristic sounds may be of interest for forensic study if the microphone is located sufficiently close to the firearm to pick up the tell-tale sonic information.

If the bullet emerges from the barrel traveling faster than the speed of sound, the supersonic projectile creates a ballistic shock wave. The shock wave itself forms a radially propagating cone that trails the bullet as it travels down range. The expanding face of the shock wave cone moves through the air at the speed of sound, and the passage of the shock wave may be picked up by microphones located near the bullet’s path [9].

2.1 Reflections and reverberation

Forensic audio recordings of gunshots generally contain significant evidence of acoustical reflections and reverberation. Distant recordings may not be line-of-sight to the shooting location, so the effects of diffraction and multi-path interference is also expected.

As noted above, the muzzle blast sound produced by a firearm lasts only a few milliseconds, but the recorded gunshot sound often has energy lasting hundreds of milliseconds due to the reverberation of the recording scene. The acoustic clutter of the reverberation can sometimes provide good information about the acoustical surroundings, but, in general, the specific acoustical characteristics of the firearm itself are lost in the overlapping sound reflections.

2.2 Effects of audio coding

Many sources of user generated recordings (UGR) come from devices that use perceptual audio coding (e.g., MP3 or MP4) between the microphone and the digital storage system. Perceptual coders are designed to maintain the perceived audio quality of the original signal, but such coders are not intended to retain objective waveform information that would be desirable for forensic purposes, especially for impulsive signals such as gunshots. Nevertheless, perceptually-coded audio can provide useful timing information within the constraints of the block size and timing of the coding algorithm [10].

Currently, commercial gunshot detection systems such as ShotSpotter appear to use conventional uncompressed pulse-code modulation (PCM) to store recorded audio. These recordings should therefore have fewer concerns about waveform interpretation and timing.

3 ShotSpotter system

The ShotSpotter system from SoundThinking™ [2] is a proprietary commercial system intended to detect and localize the sound of a gunshot in the jurisdiction of a law enforcement agency that subscribes to the ShotSpotter service. The system consists of numerous microphones and computer processing systems (sensor nodes) installed and maintained by ShotSpotter on rooftops, poles, and other structures in the area of coverage. ShotSpotter’s dispatching service, which uses information derived from the acoustic sensors to make a judgement about the occurrence of a sound that might be a gunshot, provides an estimate of the geographic location of the sound source. The proprietary system uses signal processing algorithms and human listeners.

The geometric coordinates of the sensor nodes are determined by ShotSpotter via an integrated GPS receiver at each node. The local clock at each node is also time synchronized via the integrated GPS receiver. The sensor nodes include a processor system to record and analyze audio digitized from the acoustic sensor system, and a communication system links each sensor to ShotSpotter's central office.

The basic theoretical principle of the ShotSpotter system is that a loud, impulsive sound (such as a gunshot) will travel through the air in all directions at the speed of sound and will then be detected as the sound arrives successively at the acoustic sensor nodes in the vicinity. The time-of-arrival of the impulsive sound at the various sensors depends upon the distance between the sound source and the sensor, the speed of sound in the vicinity, the presence of structures, terrain, and other obstacles, and various atmospheric effects such as wind and temperature gradients.

The audio processing system at each sensor node includes computer firmware that uses various proprietary algorithms to estimate the likelihood that a received sound is a gunshot, and not some other common sound like a door slamming or a firecracker. If the algorithm makes the determination that a gunshot sound was observed, the system processes the recorded waveform and estimates the arrival time of the impulsive sound based upon the details of the microphone signal, and this information is sent to the ShotSpotter central office.

When the ShotSpotter central office has received several contemporaneous reports from the acoustic sensor nodes in a particular area, the central office system automatically uses the time reports and the known sensor locations to estimate the time difference between the sound's earliest arrival at a sensor—presumably the sensor closest to the sound source—and the arrival times at the other sensors. The known location of each sensor and the measured time-difference of arrival at each sensor can be used to compute an estimate of the sound source location. This mathematical computation is known as *multilateration* [11].

In order to notify the law enforcement agency about the shot detection so that the agency's officers can be dispatched, ShotSpotter reports an estimated location in two dimensions (e.g. latitude and longitude). The sound arrival times from at least three sensor nodes are needed for the mathematical multilateration to compute a set of possible sound source locations that are consistent with the observed time-differences of arrival, and at least four sensor nodes are needed to reduce the ambiguity of the possible source positions [12].

4 Case study: concurrent close and distant gunshot recordings

An example gunshot audio recording scenario comes from a case involving a ShotSpotter gunfire incident activation and several home surveillance recordings. The vicinity of the incident is shown in Figure 1.

The shooting incident appears to have started in the rear of a residential dwelling where a party crowd had gathered late in the evening. One of the surveillance recordings (Camera 3) included audio of the gunshot sounds, while two other surveillance recordings (Cameras 1 & 2, video only) captured images of firearms being discharged.

Based upon the available surveillance camera audio recording (Camera 3), a total of 17 gunshots are audible. The initial two gunshots, separated by about 5 seconds, occur clearly in the surveillance audio recording.

The location where the shooting incident took place was covered by a ShotSpotter gunshot detection system. In fact, the sequence of shots resulted in three separate ShotSpotter detection reports, as described below. Thus, the incident includes both close and distant recordings of the gunfire sounds.

In this incident, the entire sequence of gunshot sounds was captured by the audio recording from Camera 3, located approximately 50 meters north of the area of the party. The audio recording from Camera 3 is shown in Figure 2.

Figure 1: View of the incident scene showing the general location of the surveillance cameras and the path of the shooters moving through the area.

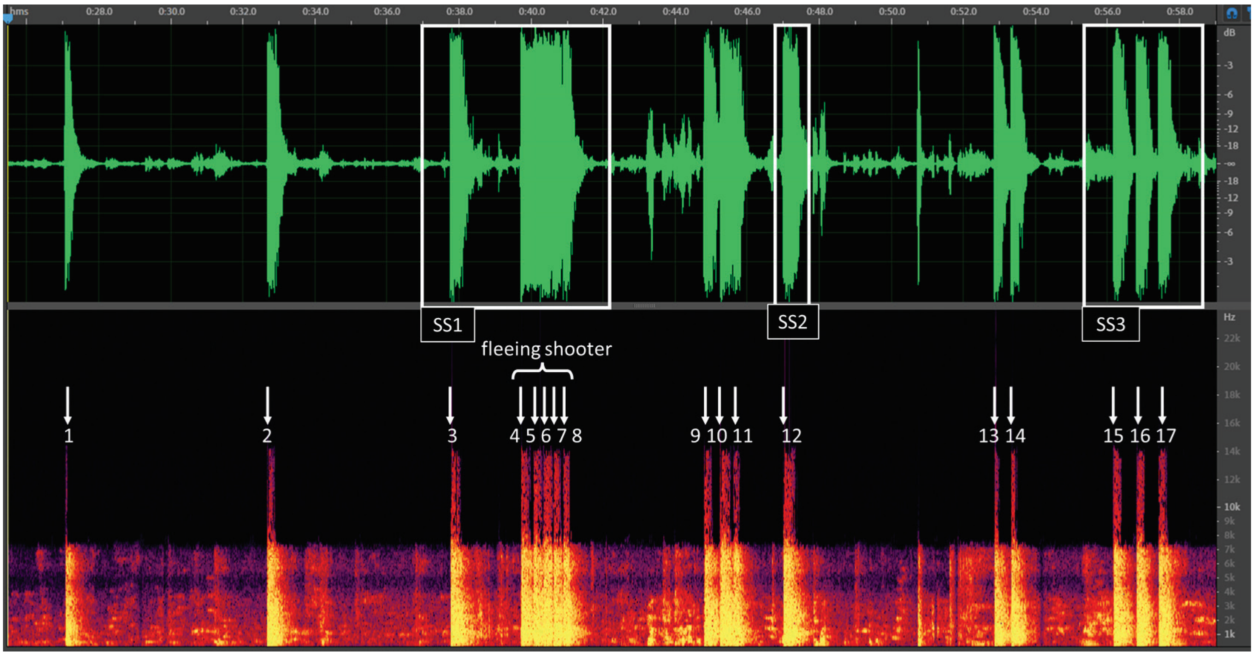


Figure 2: Audio recording and spectrogram from Camera 3 position. Total duration 33.5 seconds. Individual gunshot reports 1-17 indicated manually. The three ShotSpotter activation detections are shown (SS1, SS2, SS3).

By comparing the timing of the shots present in the Camera 3 recording to the incident reports from the ShotSpotter system, we can identify that the ShotSpotter system did not activate on the first two audible gunshots, and also missed shots 9, 10, 11, 13, and 14, as seen in Figure 2.

The Camera 3 position recorded the gunshot audio, but the accompanying video did not contain images of the shooting scene. The Camera 1 and Camera 2 locations did capture images with visible muzzle flashes and other evidence of the gunfire (e.g., Figure 3, Figure 4, and Figure 5), but Camera 1 and 2 did not have audio recording capability.

Figure 3: Frame from Camera 1, handgun muzzle flash (Shot 7 from Figure 2).

Figure 4: Frame from Camera 2, handgun muzzle flash (Shot 9 from Figure 2).

Based upon the Camera 3 audio recording, the shots visible in the Camera 1 and 2 videos, and other evidence reported from the incident scene, we can deduce the position of the firearms making all 17 audible shots. The shot locations are depicted in Figure 6.

Figure 5: Frame from Camera 1, handgun muzzle flash (Shot 10 from Figure 2).

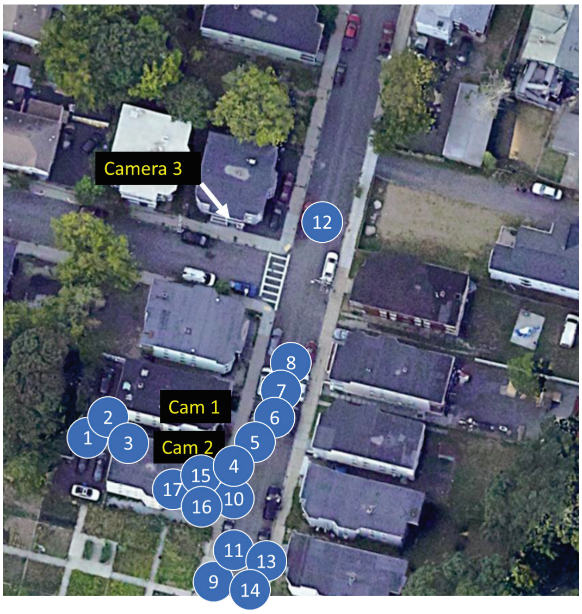


Figure 6: Approximate location of shots based upon video information (Camera 1 and 2) and related evidence.

The ShotSpotter system issued three real time incident reports related to this case. The incident report SS1 detected six shots, which correspond to shots 3-8 in the sequence. Incident report SS2 indicated one gunshot, corresponding to shot 12, and report SS3 indicated three shots, corresponding to shots 15-17. The locations identified by ShotSpotter's multilateration for the ten detected gunshots are shown in Figure 7.

In this case, it is clear that ShotSpotter did not detect all of the shots, and of the shots it did detect, the locations differ somewhat from the "ground truth" provided by the close cameras. While this may appear to call into question the value of ShotSpotter for comprehensive audio forensic assessment of a shooting incident, it is important to note that the *real time* detections and locations sent by ShotSpotter to the law enforcement agency happened within one minute after the incident, and the position information was adequate to dispatch officers to the correct residential block. So, while the accuracy of the ShotSpotter multilateration positions in this example show a noticeable discrepancy for crime scene reconstruction purposes, the system did fulfill the original intent of ShotSpotter, namely, to identify the incident location very quickly and with sufficient precision to dispatch the law enforcement responders to the scene.

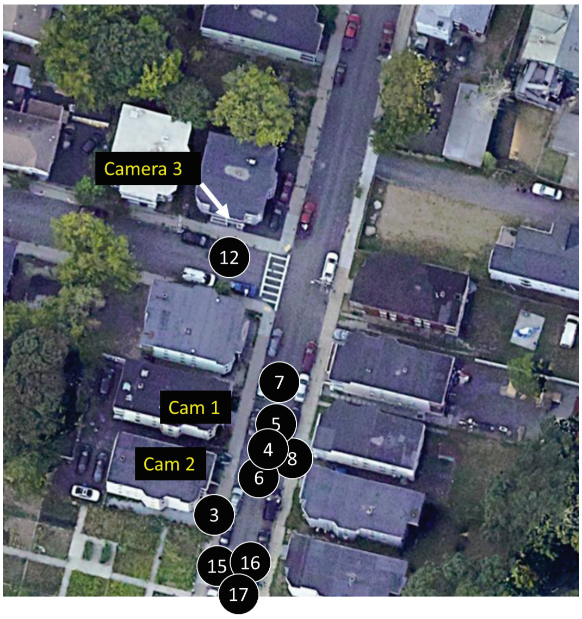
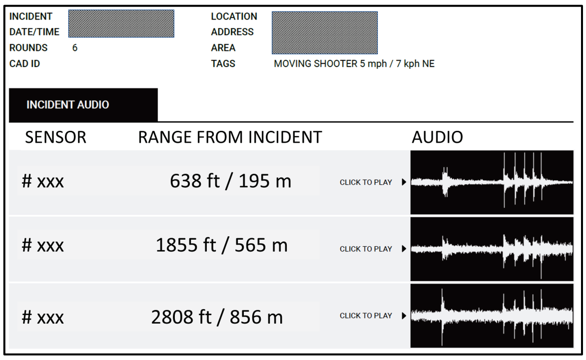


Figure 7: Location of shots from ShotSpotter detection reports. The ShotSpotter positions are within approximately 20 meters of the actual gunshot locations observed via Camera 1 and 2.

Figure 8 shows an example excerpt from ShotSpotter report SS1, identifying shots 3-8 of the incident. The report includes excerpts of the audio recorded at just three sensors, although it is assumed that ShotSpotter had additional sensors used in the multilateration calculation that are not listed in the summary report. For the three sensors included in the summary report, the table indicates the distance from each sensor to the shot location determined by the multilateration algorithm. ShotSpotter generally does not identify the precise location of its sensors due to concerns about vandalism. In any case, it is evident that the gunshot acoustical waveforms differ among the three sensors due to the different range, azimuth, elevation, obstructions, reflections, and other physical path differences.



INCIDENT
DATE/TIME [redacted]
ROUNDS 6
CAD ID

LOCATION
ADDRESS [redacted]
AREA
TAGS MOVING SHOOTER 5 mph / 7 kph NE

INCIDENT AUDIO

SENSOR	RANGE FROM INCIDENT		AUDIO
# xxx	638 ft / 195 m	CLICK TO PLAY	
# xxx	1855 ft / 565 m	CLICK TO PLAY	
# xxx	2808 ft / 856 m	CLICK TO PLAY	

Figure 8: Example excerpt from ShotSpotter incident report SS1 (shots 3-8).

In this example, the three sensors in the example were at distances 195, 565, and 856 meters, respectively (see Figure 9). The ShotSpotter recordings are significantly farther away from the firearms than the roughly 50 meters for the Camera 3 surveillance recording, and therefore the distant recordings will include the effects of acoustical diffraction/shadowing and reflections/reverberation due to the numerous buildings, trees, and other obstacles between the shooting location and the ShotSpotter sensor microphones.

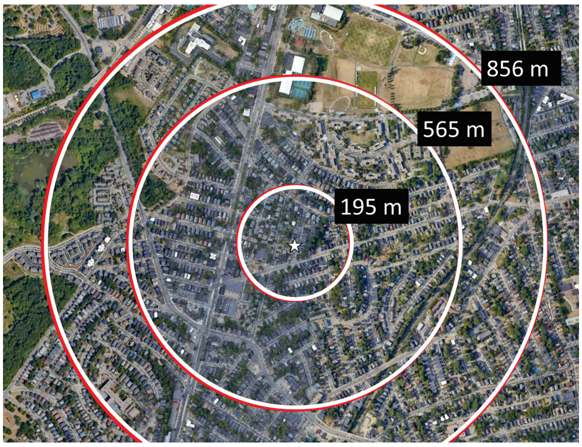


Figure 9: Aerial view of the neighborhood surrounding the shooting incident, with distances to the three ShotSpotter sensors indicated in summary report SS1 (Figure 8).

Comparing the audio recordings from the close microphone (Camera 3) and the more distant ShotSpotter sensors, the differences in amplitude and reverberation characteristics become apparent, as presented in Figure 10.

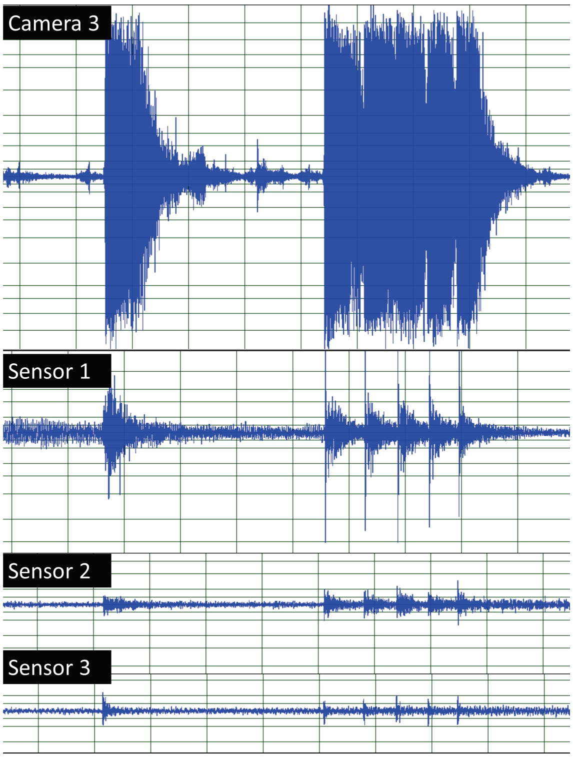


Figure 10: Comparison of the waveforms for shots 3-8 recorded by the close Camera 3 microphone and the three more distant ShotSpotter sensors.

5 Discussion and Conclusions

This case study presents an increasingly common situation in which a shooting incident is recorded by a microphone close to the scene and concurrently by microphones located some greater distance away. The "close" recording in this case was from a home surveillance system with audio recording capability, while the distant recordings were obtained by the commercial ShotSpotter gunshot detection system deployed in the particular jurisdiction.

The close recording is useful for distinguishing gunshots from various other sounds at the incident scene, and for establishing a reliable timeline.

The close recording is susceptible to amplitude clipping and other waveshape distortions due to the sound intensity of firearm muzzle blasts, and the likelihood that the recording was captured with perceptual audio coding (e.g., MP3, MP4, etc.).

The more distant recordings, such as the ShotSpotter recordings from 195, 565, and 856 meters available in this case, have the advantage that amplitude clipping is unlikely. The ShotSpotter system also provides multilateration calculations that can help estimate the location of the gunfire with respect to the precisely known positions of the sensors, along with precisely known timestamps. However, as was observed in this case, ShotSpotter's algorithms may not correctly identify every gunshot sound in a shooting incident, and the multilateration position uncertainty may be greater than the dimensions of the shooting scene, particularly if the shooters are in obstructed locations or are moving during the incident. Based on this case study, the audio forensic recommendation would be to compare and confirm the ShotSpotter information with the close recording, but clearly the timing and details of each shot in the incident is represented fully with the close recording, and incompletely with the ShotSpotter reports.

While this example includes ShotSpotter evidence, it is also increasingly common to have user-generated surveillance recordings or cellphone video from locations near and far from a shooting incident, or the close recording may come from a law enforcement body-worn camera system. The additional information provided by the multiple concurrent recordings may be important in answering questions about the circumstances of the incident, but the fact that the locations of user devices is usually not known very precisely, it is generally not feasible to get reliable multilateration estimates of the gunshot location using muzzle blast time-of-arrival in the various recordings. Nevertheless, there may be useful information available by comparing the timing of multiple shots in the incident recordings, such as addressing the sequence of shots if multiple shooters were present [13].

References

[1] Federal Bureau of Investigation, *Crime Data Explorer*, URL: https://cde.ucr.cjis.gov/LATEST/webapp/, September 1, 2023.

[2] SoundThinking™, ShotSpotter gunshot detection technology, URL: https://www.soundthinking.com/law-enforcement/gunshot-detection-technology/, September 1, 2023.

[3] Straits Research, "Doorbell Camera Market," URL: https://straitsresearch.com/report/doorbell-camera-market, September 1, 2023.

[4] U.S. National Institute of Justice, "Research on Body-Worn Cameras and Law Enforcement," URL: https://nij.ojp.gov/topics/articles/research-body-worn-cameras-and-law-enforcement, January 7, 2022.

[5] R.C. Maher and T.K. Routh, "Advancing forensic analysis of gunshot acoustics," Preprint 9471, Proc. 139th Audio Engineering Society Convention, New York, NY (2015).

[6] R.C. Maher and T.K. Routh, "Wideband audio recordings of gunshots: waveforms and repeatability," Preprint 9634, Proc. 141st Audio Engineering Society Convention, Los Angeles, CA (2016).

[7] R.C. Maher, "Shot-to-shot variation in gunshot acoustics experiments," elib 20461, Proc. 2019 Audio Engineering Society International Conference on Audio Forensics, Porto, Portugal (2019).

[8] B.F. Miller, F.A. Robertson, and R.C. Maher, "Forensic handling of user generated audio recordings," Preprint 10515, Proc. 151st Audio Engineering Society Convention, Online (2021).

[9] R.C. Maher and S.R. Shaw, "Deciphering gunshot recordings," Proc. Audio Engineering Society 33rd Conference, Audio Forensics—Theory and Practice, Denver, CO (2008).

[10] R.C. Maher and S.R. Shaw, "Gunshot recordings from digital voice recorders," Proc. Audio Engineering Society 54th Conference, Audio Forensics—Techniques, Technologies, and Practice, London, UK (2014).

[11] R.C. Maher and E.R. Hoerr, "Audio Forensic Gunshot Analysis and Multilateration," Preprint 10100, Proc. 145th Audio Engineering Society Convention, New York, NY (2018).

[12] R. Calhoun and S. Lamkin, "Determining the Source Location of Gunshots From Digital Recordings," Express Paper 18, Proc. 153rd Audio Engineering Society Convention, New York, NY (2022).

[13] R.C. Maher, "Forensic interpretation and processing of user generated audio recordings," Preprint 10419, Proc. 149th Audio Engineering Society Convention, New York, NY, Online (2020).