Audio Engineering Society

Conference Paper

Presented at the International Conference on
Audio Forensics
2019 June 18–20, Porto, Portugal

Overview of Forensic Audio Gunshot Analysis Techniques

Durand R. Begault[1], Steven D. Beck[2], and Robert C. Maher[3]

[1] *Audio Forensic Center, Charles M. Salter Associates, 130 Sutter St., Ste. 500, San Francisco, CA 94104*

[2] *Beck Audio Forensics, Austin, TX 78731 stevendbeck@alumni.rice.edu*

[3] *Department of Electrical & Computer Engineering, Montana State University, Bozeman, MT 59717 rmaher@montana.edu*

Correspondence should be addressed to Durand R. Begault (durand.begault@cmsalter.com)

ABSTRACT

Audio recordings of gunfire are often of evidentiary interest in legal proceedings, requiring the opinion of experts with combined knowledge of audio engineering, acoustics, and signal processing. The recordings originate from diverse instruments such as body-worn cameras, security camera systems or telephones, and can be complicated by multiple factors that affect signal analysis. This paper provides an overview of types of forensic audio gunshot analysis and an evaluation of the reliability and limitations of particular techniques that have been used by experts in the field of forensic audio.

1 Introduction

An audio forensic expert may be asked to both opine on the meaning and interpretation of gunfire audio recordings, as well as defend the bases of their opinions. They may also be asked to opine on the opinions of other experts, or laypersons. The present manuscript provides an overview of most types of forensic audio gunshot recording analysis introduced in U.S. Federal and State courts by experts during the last several decades, and gives an initial evaluation of the reliability and limitations of particular techniques that are the bases of these expert opinions.

Audio forensic expertise is considered here to be primarily based on knowledge of gunfire acoustics; transformation of the source signal by the environment; and transformation of the original acoustic waveform by the microphone and recording system. In other words, the area of expertise is fundamentally based on technical and scientific knowledge of acoustics and signal processing. A completely separate area of inquiry is analysis of non-acoustic factors in the context of a gunfire incident, such as motivation, adherence to procedure, or human factors issues such as perception-reaction time. Generally, these separate areas are outside the domain of audio forensic expertise; conversely, an expert in firearms or police practices would not necessarily qualify as an expert in audio analysis of gunshot recordings without specific training.

Gunfire acoustic analysis using audio recordings has its early beginnings with military technologies used to triangulate the location of an acoustic source [1, 2]. The police dispatch recordings from the Kennedy assassination and recordings of the Kent State shootings were evaluated in reports by researchers at Bolt Beranek and Newman in the 1970s, representing some of the first forensic applications of digital signal processing techniques [3, 4]. In 1989, Koenig et al. [5] reviewed prior work having an emphasis on firearm discrimination.

Recordings involving cellular telephones, tablet computers, or other digital computing devices are increasingly the focus of audio forensic examinations, while recordings originating with other devices such as hand-held digital recorders have decreased. Often, an audio recording is part of a video recording. Other sources of recordings include police vehicle "dashcam" transmitters (from the body and/or within a police vehicle); body-worn cameras (e.g., AXON); cameras attached to weapons (e.g., "Taser Cam"); dispatch center recordings (involving radio frequency, land line, and cellular telephone transmissions); voice mail recordings; audio from security cameras; and commercial gunshot detection systems.

All expert opinions in gunfire acoustic analysis are ultimately based on a collection of observations that should be made on the basis of scientific methodologies, including the techniques reviewed below. It is important to bear in mind that realistic scenarios matching the circumstances of a particular case may be difficult to evaluate using the scientific method *per se*, since the circumstances of recording can vary widely. This is a recognized challenge to most forensic analysis disciplines, as described in the National Research Council 2009 report "*Strengthening forensic science in the United States: a path forward*" [6]. While certain forensic techniques are now largely disparaged (e.g., bitemark analysis, dog scent) or determined to have limited value (e.g., hair fibre analysis, partial fingerprints), the scientific validity of forensic audio gunshot analysis depends on the particular question that the techniques attempt to answer. Accordingly, experts should strive to inform the trier of fact as to the limitations of specific techniques used and their intended application.

2 Corroborative Information

A typical forensic case might include extraneous information that may or may not be useful in a purely acoustic evaluation of the recording. We refer to this as "corroborative information." Examples include source/receiver locations, counts of persons involved, shell casings or other physical evidence; or video imagery, such as muzzle flashes, or bullet penetrations.

Corroborative information can in some cases increase the accuracy of an analysis and also the confidence of an expert in making certain conclusions. For example, a question regarding if and when gunfire occurred on an audio file might be answered in conjunction with inspection of a video file (taking into account the relative speed of propagation).

However, corroborative information can also be a source of undesirable bias in the decision process. This may include the role of persons involved in the incident (law enforcement vs. civilian) or reported facts of the case outside of the area of recorded gunfire. It should be recognized that opinions based on integration of extraneous evidence are likely to alter an examiner's criteria, while a completely "blind" assessment may cause an examiner to be extra cautious in making conclusions, i.e., to use a more conservative criteria in their decision process (see [7-8] for further discussion of bias and decision criteria). Forensic audio experts should clearly report when and how corroborative information was used to reach their conclusions.

3 Questions Addressed, Challenges

Legal questions regarding gunshot recordings and the accompanying analysis challenges for experts are characterized below. Experts should describe how these relevant challenges are met in a particular case.

<u>Timing and quantity of shots</u>. Due to the high amplitude of gunfire relative to most other sounds, this analytic challenge can be minimal. However, distortion of the signal by overmodulation, background noise, reverberation, other impulsive events and the signal chain can obscure the initial transient used to establish timing.

<u>Discrimination of gunfire from non-gunfire</u>. In some cases, the question arises whether a specific sound event is actually gunfire, or something else. Discrimination becomes increasingly challenging

depending on the quality of the recording, the distance of the gunfire from the microphone, and the potential similarity of firearm acoustic characteristics to non-gunfire sounds.

Firearm identification. It is sometimes presumed that a particular weapon has a "signature" that can be discerned on a recording (see [9]; ref. [10-11] for a refutation of "gun prints"). In situations where two weapons are in physical evidence, the answer to the question "who shot first" is sometimes requested of an expert. This can be particularly difficult to accomplish due to recording quality and the similarity or masking of available unique acoustic characteristics of different firearms in question after the recording has been convolved with non-linear components of the recording chain [12-14].

Gunfire acoustic analysis encompasses a range of weapon types and discharges. A single recording may include both lethal weapons (handguns, rifles) and "non-lethal" weapons (tasers, bean bags, flash-bang explosives). Discrimination between different weapon types is less challenging the greater the contrast between them (e.g., a shotgun with a non-lethal round compared to a pistol). Note in this latter example, additional cues such as the mechanical racking of a shotgun may assist in discrimination.

Location of shooter. Experts can sometimes estimate or identify the location of a shooter amongst a set of known choices, based on room reverberation or sound quality characteristics, and/or spatial information from multiple channels. Under ideal conditions (e.g., an uninterrupted acoustical path with known microphone locations), gunshot location information can be ascertained via multilateration methods [12, 15]. With adequate recording fidelity and timing analysis of reflection patterns from buildings or other surfaces, it is also possible to determine the relative location of multiple shooters from a single microphone [16].

Earwitness reliability. Laypersons or persons with shooting experience may testify regarding hearing gunfire: when or where it occurred, and/or the type of weapon heard. A first principles approach involves estimating the likelihood of detection in terms of an acoustic analysis of the signal-to-noise ratio. Beyond this, evaluation of earwitness testimony may require research psychology expertise in witness reliability in general, particularly regarding the accuracy of short-term vs. long-term memory [17-18].

4 Acoustic & Signal Chain Effects

Audio recordings of gunfire originate with the acoustics of the firearm itself. Multiple references are available that review gunfire acoustics as a function of directionality or weapon type (e.g., [19-21]). In most cases this work has been conducted under controlled, quasi-anechoic conditions, with high-quality calibrated data instrumentation. This can include digital recordings at 192 kHz sampling rate or greater, special 1/8" microphone cartridges to remove directional sensitivity and accommodate high sound pressure levels, and in some cases apparatus to simulate the head and ear canal resonance (Figure 1). Such recording systems are capable of recording not only the muzzle blast but also the non-linear shock wave ("N wave") that includes very high frequencies.

These studies demonstrate significant differences in peak levels and spectral balance as a function of firearm orientation. Measurements show peak levels in the range of 150 -168 dB peak at the ear of the shooter, implying that compared to other sounds, gunfire will be a significantly high-level event if the recording is made in relatively close proximity. The spherical spreading loss of gunfire means that as distance increases, its level will eventually be subsumed in the ambient background noise and therefore more likely to be masked.

Under realistic forensic conditions, the recorded gunshot signal is convolved with the acoustic and electronic *signal chain* that ultimately yields the recorded signal that is available for analysis. As might be expected, the signal chain can cause significant transformations compared to recordings under controlled conditions and can overwhelm differences between weapon types, peak levels or orientation (see, e.g., [5], comparing orientation, weapons, and recording devices).

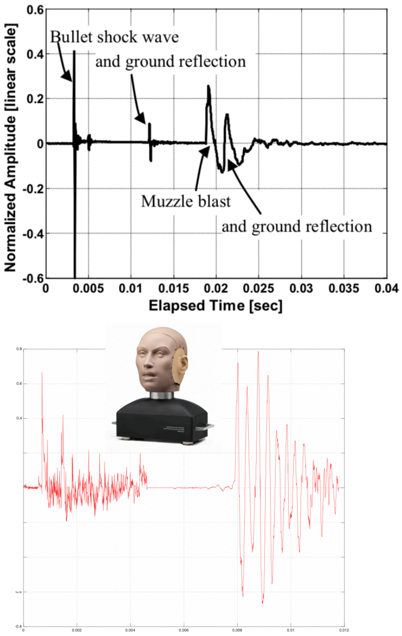

Figure 1. Top: from [12], showing ground reflections the supersonic shock wave, and muzzle blast. Bottom: gunshot recorded with measurement microphone (left) and transformation at the middle ear using mannequin microphone (right).

The acoustic portion of the signal chain involves the net effect of reflections from the ground or surrounding buildings, and the level of the signal relative to background noise [12, 14, 16, 21]. Other acoustic factors include local acoustic effects of the microphone (moving versus stationary; shielding) and, in some cases, the object receiving the gunfire.

The electronic portion of the signal chain involves the recording device itself, and in some cases the telecommunication of the signal. For example, the provenance of a recorded gunshot may ultimately be a file on a voicemail from a cellular telephone, a police dispatch recording system from a landline call or police radio microphone, or directly recorded to a body-worn camera. Each situation involves multiple successive transformations of the signal by systems that are typically optimized for speech communication intelligibility, not gunshots. Additional processing using lossy compression may further transform a recorded gunshot signal.

The combined effects of the acoustic and electronic signal chain cause significant signal transformation, distortion, and loss in the time and frequency domain. References [5, 14, 22] have examined the effect of personal recording devices on gunfire. Figure 2 compares the recordings of a balloon pop (134 dB peak at 1 m) made in an acoustically damped room, with early reflections extending out to ~30 ms. The upper time domain graph shows a calibration microphone recording; the lower graph shows the same event recorded by a cellphone voicemail system. The voice mail system delayed the amplitude peak by 18 ms, and a signal processing "echo" occurs from 70–160 ms. Similar effects were found in examination of a 911 dispatch recording, due to multiple signal processing stages.

The impact of signal chain transformation must be considered in expert evaluation of recorded gunshot signals. The root causes may be time-invariant or not, known or unknown. Separate testing of the signal chain may be necessary to buttress analytic conclusions.

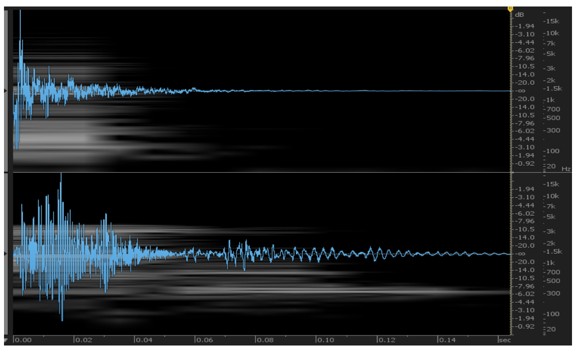

Figure 2. Balloon pop response. Top: measurement mic response. Bottom: iPhone 5 recorded to voice mail system. Abscissa 0-0.16 s; ordinate 20-20 kHz. From [14].

5 Evidence Analysis Techniques

In 2015, six different experts (incl. the current authors) produced independent gunfire reports for the same shooting incident [23]. The analysis techniques employed were remarkably consistent and mirrored the available literature from other experts [e.g., 3-5, 16]. The following attempts to categorize analysis techniques as they are applied to the legal questions reviewed in section 3, above, particularly for signals that have been impacted by the signal chain.

Not all analysis techniques are applicable or possible to every forensic situation. In some cases, no opinion can be rendered when evidence is "spoiled" to the degree that an expert cannot make a reliable opinion. For gunshot analysis, this is typically a function of signal distortion, or with regards to a subset of evidentiary questions that cannot be reliably answered (e.g., timing, yes, but not identification).

5.1 Pre-processing of signals

Definition. Application of signal processing methods to increase the signal-to-noise ratio of recorded gunshots.

How applied. The two most commonly used pre-processing methods are approximations to matched filtering and some form of denoising. From [20] we know the general shape of the muzzle blast spectrum, and that the peak spectral energy for small firearms is between 500 and 1000 Hz. The peak energy for a ballistic shockwave is often above 2000 Hz. Judicious use of filtering may mitigate the effects of broadband noise and narrowband interference. Many audio programs have a denoising tool that estimates a spectral model from a noise-only section of the signal and subtracts noise energy across the spectrum from the rest of the signal.

Considerations. Because gunshot waveforms are broadband, filtering will likely remove signal energy and thereby distort the signal. Filtering also creates group delay, phase changes, and ringing in the time domain waveform. Denoising also subtracts energy from the signal waveform, and this effect can be significant at low frequencies which propagate the farthest. Denoising also creates additional artifacts like "musical" noise during the reconstruction process, due to the transformations between the time and frequency domains.

5.2 Critical Listening

Definition. "Careful listening" to identify or compare likely gunshot events, usually with a time-domain waveform editor in conjunction with waveform analysis.

How applied. The expert develops criteria for identification, based on unique characteristics of the gunshot "sound object" such as impulsivity and timbre. The expert should be able to state the basis of the characteristics used for identification.

Considerations. In the absence of other types of analysis or corroborative information (section 2), critical listening is likely unreliable (e.g., firecrackers or other explosions can sound like gunfire at a distance). Forensic audio experts and triers of fact should keep in mind that there is no such thing as a "golden ear" when listening to a recording: it is not acceptable to claim that one hears something that others cannot. While it is possible to be "trained" to discriminate between some signal characteristics, the reliability of such discriminations must be demonstrated in some other way other than attestation or vigorous assertion.

5.3 Waveform Analysis

Definition. Inspection of the time domain representation of an acoustic waveform.

How applied. Usually, audio editing software is used for timing analysis, based on visual estimates of the initial pressure increase above the ambient. Audio editing software allows waveform analysis to be performed in conjunction with critical listening. It is also possible to use audio analyzer hardware, e.g., with a peak marker function within an oscilloscope-like display. Analysis of differences in peak amplitudes may enable discrimination between two shooter locations if their locations were sufficiently displaced.

Considerations. In the absence of other types of analysis or corroborative information, isolating or identifying gunfire waveforms can be difficult or impossible. Figure 3 illustrates the difference between ideal and field conditions. The effects of the signal chain can exacerbate the problem of determining which part of a waveform represents the initial gunshot transient, particularly in conjunction with background noise or other gunfire.

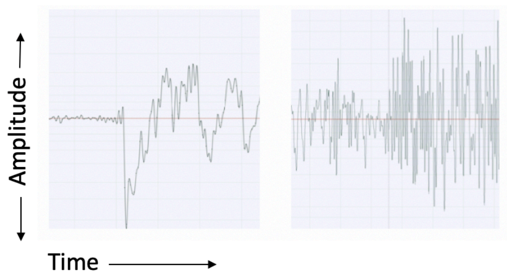

Figure 3. Gunshot waveforms under controlled (left) and forensic (body-worn camera) conditions.

Figure 4 shows timing analysis differences between four experts who independently examined the same recording. The outliers seen for one expert (D) are possibly a result of shifts in the decision criteria used during the examination.

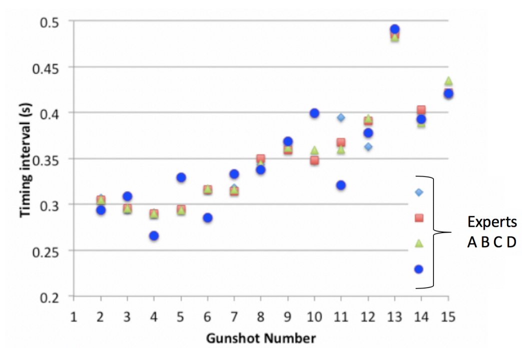

Figure 4. Differences in timing analysis from four experts (recording source: police car dash cam microphone). Note outlier data for expert D.

Timing analysis of successive gunshots (cadence) can also reveal whether or not more than a single weapon has been fired in a particular time period. For semi-automatic pistols, timing analysis of successive shots below a minimum of ~250 ms indicates the firing of multiple weapons; additional research specific to individual weapons is required [16, 24].

The three plots in Figure 5 show sequences of successive shots over a 5 second period, all with similar cadence timing. The top plot from the CBS news audio recording of the Reagan assassination attempt shows six shots (E1-E6) fired with a separation of 0.32 to 0.43 seconds. The second plot is one sequence from a controlled gunshot experiment, funded by a major network cable company, where multiple people were asked to fire a revolver as fast as they could pull the trigger while aiming at a target. All shooters fired at a rate between 0.28 and 0.42 seconds between shots (including misfires). The third plot shows the energy envelope of a 5 second sequence of shots from a forensic case the authors were asked to analyze. The timing between successive shots for these 12 events ranged from 0.29 to 0.40 seconds.

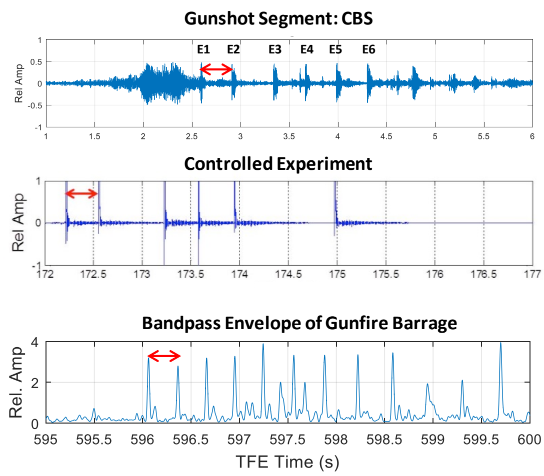

Figure 5. Gunshot cadence experiment.

5.4 Envelope Analysis

Definition. The envelope is defined as the root mean square of a time series signal, where every envelope sample has been integrated over a short time period. The bandpass envelope can be computed by integrating the energies from specific bands in every frame of a short time Fourier transform.

How applied. Envelopes are commonly used to characterize the start and decay characteristics of a transient signal, and for determining the average peak energy in the signal. Multiple events, including echoes and other interference, can often be identified by peaks in the envelope. Time differences between events are often computed from the peaks in the envelope. The levels of the energy peaks are also useful for showing consistency across multiple shots, as in a barrage.

Considerations. Figures 6-7 from [12] shows an example of a 100 ms envelope to multiple gunshots. Differences and similarities in the envelope patterns can be discerned in Figure 5. Figure 6 takes the same envelopes and overlaps them after correction to a common onset time, allowing more detailed comparison and exhibiting a difference between shots 1-2 and shots 3-6.

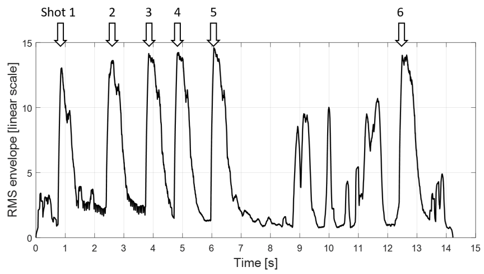

Figure 6. Amplitude envelope of six gunshots [12].

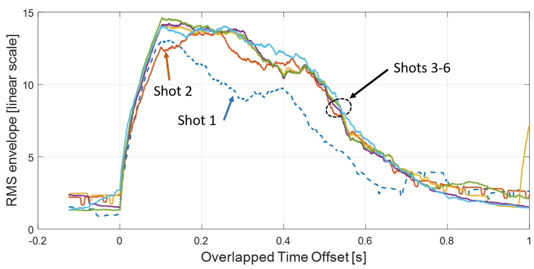

Figure 7. Comparison of amplitude envelopes [12].

Figure 8 shows the application of an envelope time curve (based on the Hilbert transform) to differentiate the echo pattern from surrounding surfaces between two gunshots (and thereby put the shooter at two different locations). In this case, the effect of echoes and reverberation are principally captured through comparative analysis. Interpretation of acoustic reflections from structures and rooms to determine shooter location or "who shot first" has been addressed in [3, 16, 25-27].

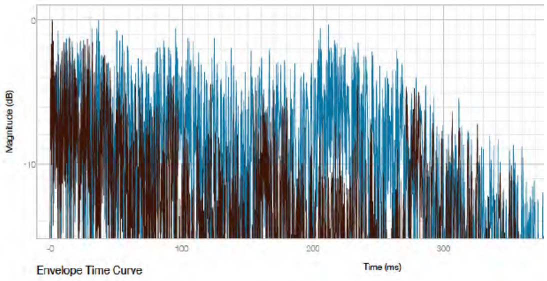

Figure 8. Two recorded gunshots with different echo patterns (blue, black), plotted as envelope time curves.

5.5 FFT and Spectrographic Analysis

Definition. The FFT (Fast Fourier Transform) is a well-established scientific technique used to display the frequency content, or *spectrum*, of a signal. A frequency range is depicted on the *x* axis and the magnitude of analyzed frequencies is shown on a *y* axis.

A *spectrogram* is effectively a series of FFT analyses over a period of time, depicting the short-time spectrum of successive and possibly overlapping short time intervals of an analyzed signal. The spectrogram presents a graph of audio signal energy, with the frequency scale as the ordinate (vertical axis) and the time scale as the abscissa (horizontal axis). It is customary to represent the third dimension—signal energy—with color or brightness. The high energy areas typically appear as bright colors and the low energy areas of the graph with darker colors.

How applied. A spectrum analysis can be applied to short time periods corresponding to a muzzle blast and signal chain effects (e.g., acoustic reflections). Figure 9 shows an analysis of three recorded gunshots and reflections. Variance in the analysis is minimized through the use of time averaging. The difference between the spectra in the upper plot (labeled 1) and the middle and lower plots (labeled 2 and 3) suggests different weapons or acoustical conditions. Specifically, plot 1 has less energy below 1 kHz compared to plots 2 and 3. Plot 1 also has a spectral peak at 2.4 kHz not seen in plots 2 and 3.

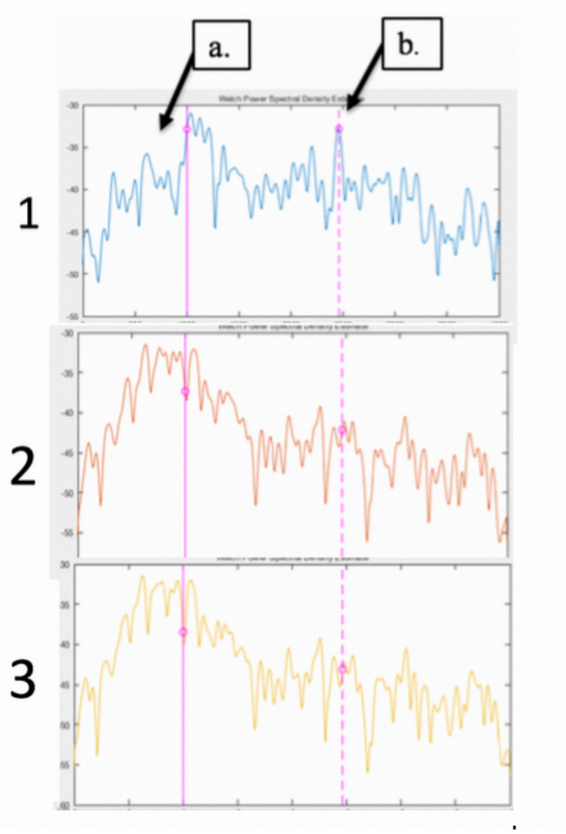

Figure 9. Welch periodogram analysis of three gunshots (150 ms). Plot 1 differs from plots 2-3.

The spectrogram of an impulsive sound such as a gunshot will appear as a bright but narrow vertical line, indicating that there is signal energy for a very short duration (narrow extent on the horizontal axis) but across all frequencies (broad extent along the vertical axis). Conversely, if the signal contains a continuous tone, the spectrogram will contain a narrow horizontal line, indicating that the signal energy is confined in frequency but continuous in duration. See example in Figure 10.

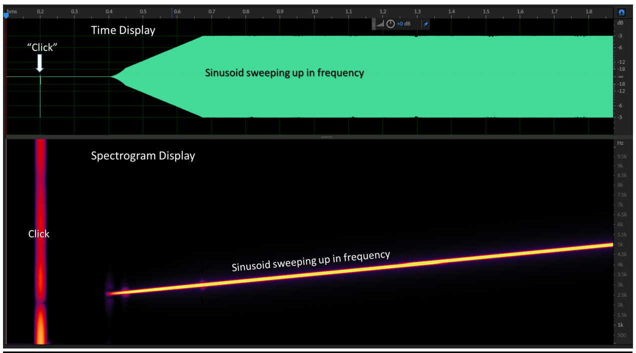

Figure 10. Example waveform (top) and spectrogram (bottom) showing an impulsive 'click' followed by a continuous sinusoid sweeping up in frequency

Time-frequency analysis, such as observations using a spectrogram, requires a choice of several parameters, including the *duration* of each short-time block, the *window* function (tapering) used for each block, and the *overlap* of adjacent blocks. The segmentation is sketched in Figure 11.

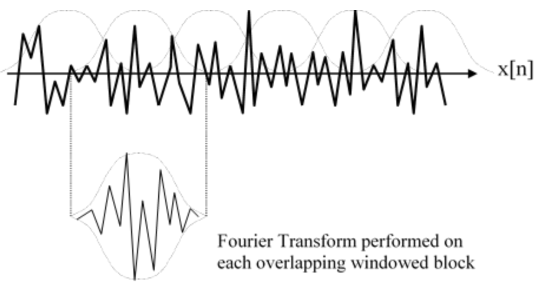

Figure 11. The spectrogram starts with the Fourier Transform of a short-time block of audio samples. The user must specify the block duration, window function, and overlap between adjacent blocks. This figure depicts 50% overlap between blocks.

Time-frequency analysis has an inescapable inverse tradeoff between the block length and the frequency resolution of the resulting spectrogram. Choosing a very short block length with minimal overlap between blocks will provide good time resolution: it will be clear when something occurred in the waveform. However, a short block length provides proportionately poor frequency resolution: it will be less clear how the frequency energy is distributed. Choosing a longer block length will blur out details of changes in the signal but will give better resolution of energy details along the frequency axis.

For very short and impulsive sounds such as gunshots, it is generally not very useful to examine fine details of the energy distribution in frequency, since the transient nature of the signal violates the basic premises of the Fourier transform. As an example, Figure 12 shows differences in spectrographic representations of the same weapon fired in three different locations in a room, caused by differences in proximity to reflective and absorptive wall surfaces. It can be concluded that the FFT cannot provide a unique spectral "thumbprint" for a particular weapon.

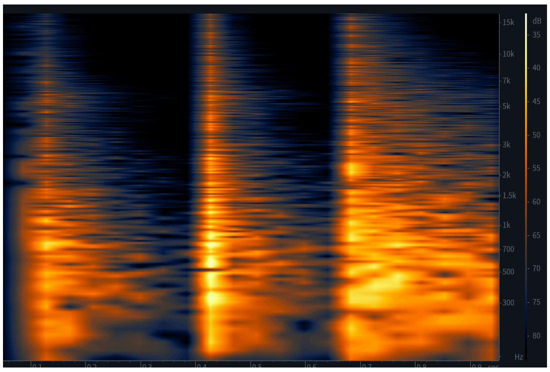

Figure 12. Spectrogram of three AR15 gunshots from the same weapon at three different locations in a large room. Quiet recording conditions using 1/8 in. measurement microphone; 8192 pt FFT. EDT_{20} = 20 decibel decay time of early reflections. Key: 1. located ~20 ft from absorptive wall surfaces, EDT_{20} =35 ms. 2. Same as 1, but ~5 ft from wall, EDT_{20} =70 ms. 3. Location ~20 ft from reflective wall surfaces, EDT_{20} =170 ms.

Nevertheless, the spectrographic interpretation can reveal shot-to-shot differences that may be attributable to different firearms being discharged, or a change in the position and/or orientation of the firearm with respect to the recording microphone. In other words, if a single firearm is fired more than once from a fixed location, we would expect the spectrographic representation to be unchanged, but if the microphone position changed between shots, the spectrographic record would also be expected to change.

An example involving several simultaneous sound sources is shown in Figure 13. Fiducial notations have been added manually. The upper portion of the figure shows the audio waveform, while the lower portion is the spectrogram. The overlapping sound sources, such as the gunshots, pulses from a Taser brand conducted electrical weapon (CEW), and human speech, can be interpreted effectively by observing the time waveform in conjunction with the spectrogram.

5.6 Cross-correlation

Definition. Cross-correlation measures the similarity between a time series signal and a time-shifted (lagged) version of another signal as a function of the lag. The sample cross correlation, or Pearson's correlation, is the covariance of the two signals normalized by the product of their standard deviations. The sample correlation coefficient r explains the percentage of variance between those signals as (r^2 x 100); for example, r = 0.8 explains 64% of the variance between the two compared waveforms.

How applied. Cross correlation is often used for measuring the similarity between gunshot recordings. In particular, cross correlation has been used in scientific studies to quantify the effects due to differences in sources (different firearm types and firing azimuth angle), differences in propagation channels, and differences in recording systems [4, 5, 21]. Cross correlation coefficients can be used to show consistency, or lack thereof, across a barrage of shots, and has been combined with peak energy levels in scatterplots to show clusters with similar features.

The other primary use of cross correlation is to find the time difference of arrival (TDOA) between two channels containing gunshots. Cross correlation has also been applied to envelope signals to find matching patterns between two sequences [27].

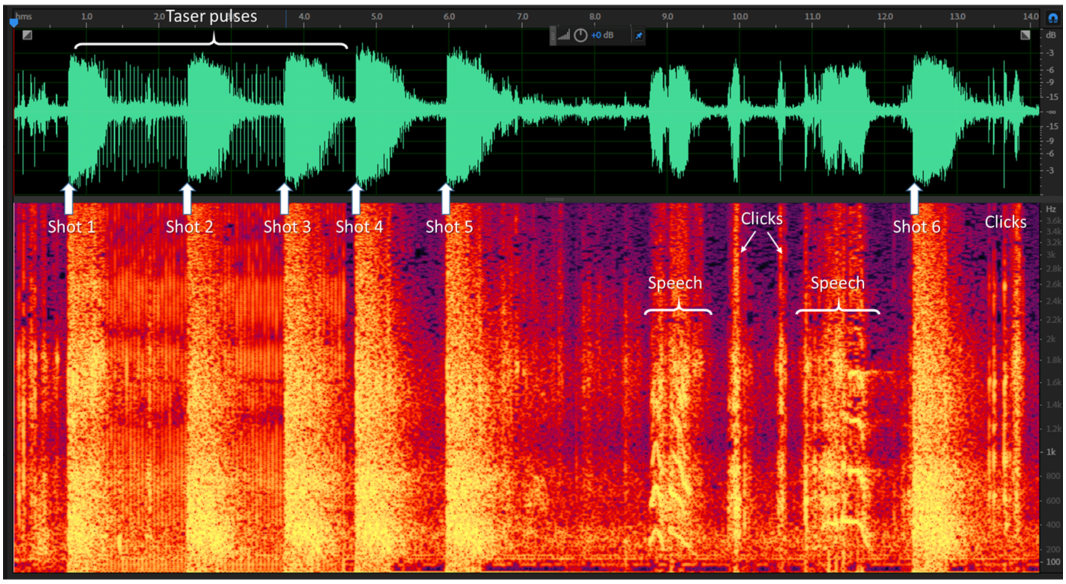

Figure 13. Example forensic audio recording waveform (top) and corresponding spectrogram (bottom), with manual notation identifying overlapping aspects of the audio scene.

Considerations. There are many variations in gunshot recordings that can adversely affect the value of the cross-correlation coefficient, and there is no one threshold that guarantees a match. For example, cross correlations of the same firearm but recorded at different azimuth angles can produce significantly different values. Recent studies have shown that shorter length signals or signals that have been smoothed will have (artificially) higher correlation coefficients [28]. Reference [29] discusses mitigation techniques for short-time cross correlation of noisy signals.

5.7 Location estimation

Definition. Sound source localization from multiple recordings typically employing time difference of arrival (TDOA) estimation and multilateration techniques.

How applied. Multiple recording devices may be activated during a gunshot, and therefore can in some cases be used to determine the location of a shooter. Since the speed of sound is relatively constant (~343 m/s), time of arrival differences can be interpreted as differences in linear distance. In such situations, if the location of microphones is known and there is an uninterrupted acoustical path, then the time of arrival differences between the microphones may be calculated, yielding a series of possible shooter locations. Figure 14 shows an example of resulting hyperbolic functions from three microphone sources; their intersections indicate possible source locations. Additional microphone sources can increase location resolution. References [15, 26] discuss techniques and examples in detail.

Considerations. In many cases, the signal chain can affect the precision of estimating gunshot time of arrival, as discussed previously in section 5.3, above. In particular, audio compression systems optimized for speech can affect the timing of a gunshot impulse. Propagation effects and reflections over long distances can also complicate timing estimation and discrimination of gunshots from other impulsive sound sources.

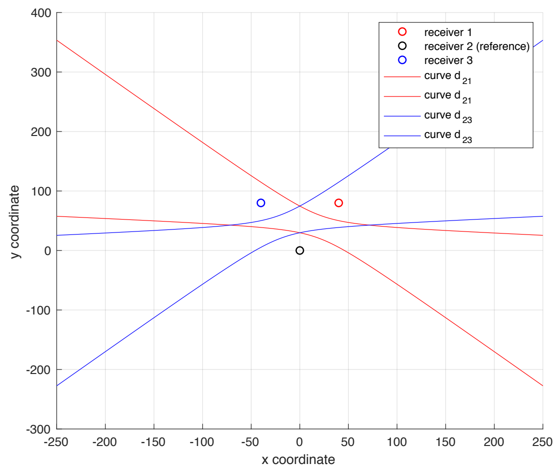

Figure 14. Multilateration plot.

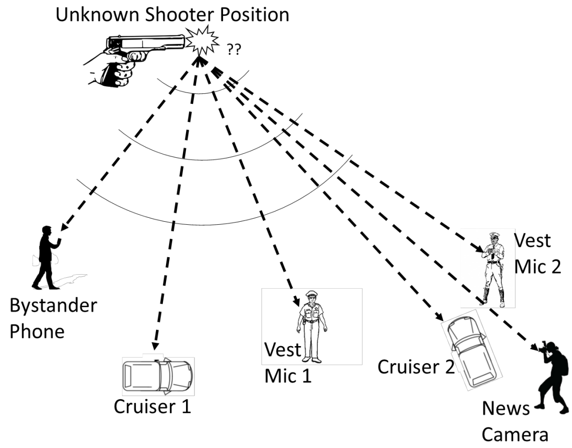

Figure 15. Audio recording sources involved in an example forensic situation (from [15]).

In many forensic situations, the position of microphones at the time of gunfire is unknown. TDOA calculations can also be complicated by movement of the microphone and/or shooter. Figure 15 illustrates an example scenario. Furthermore, the time synchronization between multiple sources must somehow be established. Radio frequency communications can be used if a microphone is at a relatively constant distance from a transmitter. In some situations, corroborative information such as video camera imagery can be used [15]. Finally, without a direct path of sound, TDOA estimates can only be made using multipath sound reflections, whose timing is usually difficult to estimate reliably.

Signal processing "machine" discrimination of gunfire from potential sound sources in an urban environment can have high false alarm rates, in part due to the relative cost of missing an actual gunshot. Some commercial gunshot detection systems utilize human listeners to confirm machine identification. However, human listeners are also susceptible to bias and false alarms in signal detection [30].

6 Conclusions

This paper has attempted to review fundamental techniques used in forensic gunshot analysis, along with an evaluation of the reliability and limitation of particular techniques. The methods have a scientific basis and can be considered reliable so long as the expert applies criteria that include options for indicating that an analysis cannot be conducted due to corrupted signals or the lack of required information. The unique conditions of each audio forensic recording make estimations of error rates impossible, but ongoing research in the field by academic and professional sources improves collective understanding of the techniques, their applicability, and their limitations.

References

[1] "Silence Visible: This Is What the Sound of WWI's Ceasefire 'Looked' Like", *Time* (November 18, 2014).

[2] "How Far Off Is That German Gun? How 63 German guns were located by sound waves alone in a single day", *Popular Science Monthly,* p. 39 (December, 1918).

[3] J. E. Barger, C. M. Gogos, C. I. Malme, S. Robinson, D. W. Steele, C. Warner, & J. Wolf, *Analysis of Recorded Sounds from the 1970 Shooting Episode at Kent State University*, Bolt Beranek and Newman, report 2781 (1974).

[4] *Report of the U.S.H.R. Select Committee on Assassinations* (March 29, 1979). USGPO.

[5] B. E. Koenig, S. M. Hoffman, H. Nakasone, & S. D. Beck, "Signal convolution of recorded free-field gunshot sounds", *Journal Audio Eng. Soc.*, v. 46(7/8) (1998).

[6] National Research Council, *Strengthening forensic science in the United States: a path forward.* National Academies Press (2009).

[7] J. S. Gruber & F. T. Poza, "Voicegram Identification Evidence", *American Jurisprudence: Trials,* v. 54, Lawyers Cooperative Publishing (1995).

[8] F. T. Poza & D. R. Begault, "Voice Identification and Elimination Using Aural-Spectrographic Protocols",, *Audio Eng. Soc. 26th International Conference on Audio Forensics, Denver, CO* (2005).

[9] H. Hollein, *The Acoustics of Crime. The New Science of Forensic Phonetics*, New York, NY: Plenum Press (1990).

[10] People v. Lincoln (Superior Court for California, Mendocino County). *Points and authorities in support of motion to exclude the acoustic signature testimony of Dr. Harry Hollien and for evidentiary hearing under evidence code section 402* (1997).

[11] People v. Lincoln (Superior Court for California, Mendocino County). *Order excluding opinion testimony of Dr. Harry Hollien* (1997).

[12] R. C. Maher. *Principles of Forensic Audio Analysis*. Cham, CH: Springer (2018).

[13] E. B. Brixen, "Acoustics of the crime scene as transmitted by mobile phones", *Audio Eng. Soc. 126th Convention, Munich, DE* (2009).

[14] D. R. Begault, A. L. Lu, and P. Perry "Forensic Sound Analyses of Cellular Telephone Recordings", *Audio Eng. Soc. 139th Convention, New York, NY* (2015).

[15] R. C. Maher and E. H. Hoerr, "Audio Forensic Gunshot Analysis and Multilateration", *Audio Eng. Soc. 145th Convention*, New York, NY (2018).

[16] B. E. Koenig, "Acoustic Gunshot Analysis, The Kennedy Assassination and Beyond", *FBI Law Enforcement Bulletin* vol. 52:11/12 (1983).

[17] E. F. Loftus & J. C. Palmer, "Eyewitness testimony", in P. Banyard and A. Grayson (ed.s), *Introducing Psychological Research.* London, UK: Palgrave (1996).

[18] A. D. Yarmey, "Earwitness speaker identification," *Psychology, Public Policy, and Law*, v. 1(4) (1995).

[19] R. C. Maher, & S. R. Shaw, "Directional aspects of forensic gunshot recordings", *Audio Eng. Soc. 39th International Conference on Audio Forensics, Hillerod, DK* (2010).

[20] P. Rasmussen, G. Flamme, M. Stewart, D. Meinke, & J. Lankford, "Measuring recreational firearm noise", *Sound and Vibration*, v. 43:8 (2009).

[21] S. D. Beck, H. Nakasone, & K. W. Marr, "Variations in Recorded Acoustic Gunshot Waveforms Generated by Small Firearms", *J. Acous. Soc. Am.* v. 129:4, pp. 1748-1759 (2011).

[22] R. C. Maher, & S. R. Shaw, "Gunshot recordings from digital voice recorders", *Audio Eng. Soc. 54th International Conference on Audio Forensics, London, UK* (2014).

[23] *State of Ohio v. Michael Brelo. Judgement entry of the verdict after a bench trial* (May 23, 2015).

[24] S. D. Beck, "Physics and Psychoacoustics Related to Forensic Ear Witness Accounts", workshop presentation, *AES International Conference on Audio Forensics, Washington, D. C.* (2017).

[25] A. H. Moore, M. Brookes, and P. A. Naylor, "Room identification using room prints", *Audio Eng. Soc. 54th International Conference on Audio Forensics, London, UK* (2014).

[26] R. C. Maher, "Gunshot recordings from a criminal incident: who shot first?", *J. Acous. Soc. Am.,* v. 139:4, pp. 2024.

[27] S. D. Beck, "Who Fired When: Associating Multiple Audio Events from Uncalibrated Receivers", *Audio Eng. Soc. 2019 International Conference on Audio Forensics, Porto, PT* (submitted 2019).

[28] D. S. Lacey, B. E. Koenig, and C. E. Reimond, "The Effect of Sample Length on Cross-Correlation. Comparisons of Recorded Gunshot Sounds", *Audio Eng. Soc. 54th International Conference on Audio Forensics, London, UK* (2014).

[29] S. D. Beck, "A short-time cross correlation with application to forensic gunshot analysis", *Audio Eng. Soc. 2019 International Conference on Audio Forensics, Porto, PT* (submitted 2019).

[30] D. M Green & J. A. Swets, *Signal Detection Theory and Psychophysics*, *volume 1*, New York, NY: Wiley (1966).