Audio Engineering Society

Convention Paper 10281

Presented at the 147th Convention
2019 October 16–19, New York, USA

Forensic Comparison of Simultaneous Recordings of Gunshots at a Crime Scene

Robert C. Maher and Ethan R. Hoerr

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

Correspondence should be addressed to Robert C. Maher (rob.maher@montana.edu)

ABSTRACT

Audio forensic evidence is of increasing importance in law enforcement investigations because of the growing use in the United States of personal audio/video recorders carried by officers on duty, by bystanders, and by surveillance systems of businesses and residences. These recording systems capture speech, background environmental sounds, and in some cases, gunshots and other firearm sounds. When there are multiple audio recording devices near the scene of a gunfire incident, the similarities and differences of the various recordings can either help or hamper the audio forensic examiner’s efforts to describe the sequence of events. This paper considers several examples and provides recommendations for audio forensic examiners in the interpretation of this gunshot acoustic evidence.

1 Introduction

The widespread use of handheld smartphones and other device capable of recording audio and video has resulted in the likelihood that user generated recordings (UGRs) may be presented as evidence in a criminal investigation. Combined with other recordings from law enforcement, the availability of user generated recordings may offer important audio forensic insights.

1.1 Simultaneous recordings

For this paper, we consider the increasingly common situation in which a gunshot is recorded simultaneously by several mobile recording devices in the area that the shooting took place. The mobile devices might be the dashboard camera systems of several police cruisers, a body camera system worn by a law enforcement officer at the scene, and perhaps a bystander making a video or engaged in a phone conversation with an emergency call center.

Multiple recordings can offer the forensic examiner options for multilateration, timing analysis, etc., but while gunshot incidents may have multiple simultaneous recordings obtained concurrently, the different recording positions, different device settings, and unsynchronized start/stop times may provide unexpectedly different information.

1.2 Characteristics of gunshot sounds

Prior work to create a database of quasi-anechoic gunshot recordings suitable for ongoing audio forensic research has emphasized the importance of taking into account the directional characteristics of muzzle blast sounds [1, 2]. While this prior work obtained gunshot recordings with specialized recording equipment and quasi-anechoic procedures, typical law enforcement investigations involve gunshot sounds recorded under less-than-ideal circumstances by mobile phones and personal audio recorders [3].

Analysis of gunshot sounds can involve a variety of audio forensic procedures [4]. Gunshot forensic audio examiners may be tempted to treat gunshot acoustical propagation as an impulse source emanating sound spherically. This is an appealing notion, because it would allow simulating the sound of a gunshot in a reverberant environment by convolving a single gunshot waveform with an estimate of the acoustical impulse response of the space in which the shot occurred. However, many firearms exhibit a broadband sound level difference of ~20 dB SPL or more between the level on-axis with the gun barrel and the same shot observed to the side or behind the barrel [5]. It is now clear that the directional acoustical characteristics of common firearms must be included if the forensic audio examiner expects to use sound propagation modeling [6].

In addition to the directional characteristics of the muzzle blast, the reflections and reverberation present at the scene of the shooting have a marked effect upon the recorded waveforms [7]. If the microphone is located off-axis from the gun barrel, the reflections and reverberation generally represent a significant portion of the total acoustic energy received.

Finally, the role of the recording device is very important. Unsurprisingly, gunshot recordings obtained with body camera systems and personal voice recorders typically show many limitations compared to carefully-controlled recordings with professional equipment [3, 8].

The current research reported in this paper compares multiple simultaneous audio recordings obtained by different devices located near the scene of a gunfire incident. While an audio forensic examiner might expect that these multiple concurrent recordings would show very similar results, we find that the differences among the various recordings are often significant, and this may complicate the audio forensic examiner's efforts to describe the sequence of events [9].

1.3 Simultaneous recording example

As an example, Figure 1 depicts evidence from an audio forensic case in which five different recording devices captured the sound of gunfire from the general vicinity of a shooting incident. The first panel shows the audio recorded by a body camera worn by a law enforcement officer close to the shooting position. The recorded waveform shows severe amplitude clipping and little detail regarding the gunshots. Panels two through five (Units H4, B1, B4, and C1) show waveforms recorded by microphones located inside the cabin of four police cruisers parked some distance away. Units H4 and B1 had all windows closed, while Units B4 and C1 had driver's side windows open.

Although the five recordings are shown in Figure 1 on a common time scale, the individual recording devices were not actually synchronized, so the time alignment was carried out manually. Moreover, the location of each microphone was not known with sufficient precision to allow determination of the expected time of arrival of the gunshot sound at each microphone. The unknown microphone positions and the lack of time synchronization limits the usefulness of interpreting the multiple concurrent recordings.

In this example, identifying the presence of approximately 18 individual overlapping gunshot muzzle blast reports is difficult to discern from the highly-clipped recordings such as the vest camera and open-window recordings of Units B4 and C1, but somewhat more clearly visible in the closed-cabin recordings of Units H4 and B1 due to the reduced signal distortion. However, the relatively quiet but audible pulses from a *TASER*® brand Electronic Control Weapon (ECW) are not discernable in any of the recordings except the vest camera that was very close to the incident scene. Clearly, if the audio forensic evidence in this example had only included one of the five recordings, the examiner would have had very limited options to interpret the subtle details of the incident.

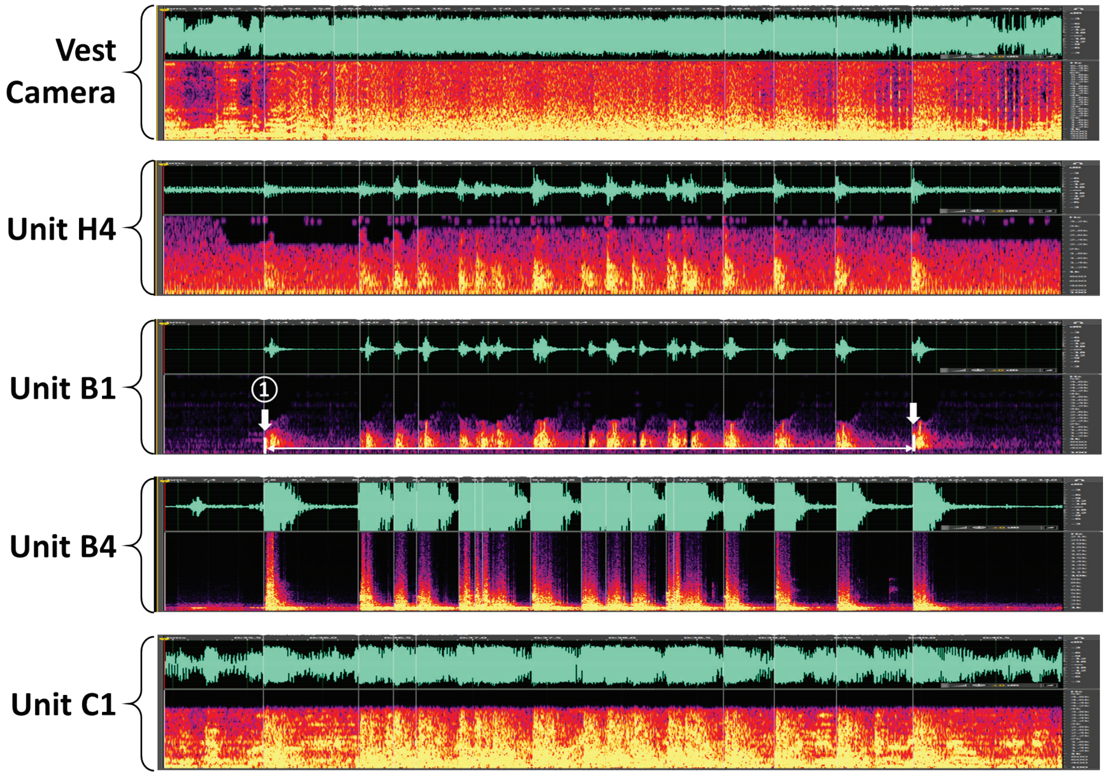

Figure 1: Example of five simultaneous recordings from an actual shooting incident. Each panel shows the time waveform above and the corresponding spectrogram below (6 seconds total duration).

1.4 Relevant research with unsynchronized recordings

Several researchers have studied the use of multiple unsynchronized user generated recordings (UGRs) of public events, such as live concert bootleg recordings, to create a composite mixture [10-15]. The prior non-forensic work has generally involved video sequences of an event that come from multiple vantage points, so the goal has been to create a post-produced video sequence with simulated "cuts" from one camera to another [16].

Among the key prior research questions is the need to deduce the temporal relationship between different UGRs. Sometimes the recordings may include metadata, such as the date and time the recorded file was created or last modified. However, the timing precision and reliability of metadata may not be of sufficient quality for forensic audio analysis, so one approach is to use an audio matching process to infer the best combination of the available recordings [9, 17, 18].

2 An experimental investigation

A controlled experiment at a closed shooting range helps demonstrate procedures suitable for audio forensic use. For this test, a firearms instructor from a local law enforcement agency made several successive shots using a variety of commercially-available guns. The sounds of the shots were captured simultaneously by fifteen microphones with various recording devices, as listed in Table 1.

The microphones were situated to the side or behind the shooting position. The marksman fired from a standing position.

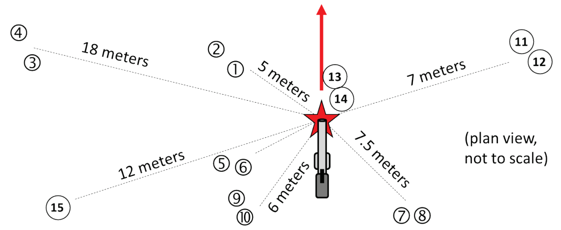

Figure 2: Example recording configuration

The microphones and recording devices were arranged at heights between 0.5 m and 1.5 m above the gravel surface of the shooting range, simulating a gunshot incident with user generated recordings. The devices had gain settings appropriate to record typical conversation at a distance of 1-2 m, so no attempt was made to accommodate the muzzle blast peak pressures, again simulating a scenario with UGRs.

For research purposes, a timing synchronization slate was made by initially having all the recording microphones close together, starting the recording devices, and making a sharp clap. With the recording devices still running, the microphones were moved to the specified locations. This synchronization information would not be present in a typical scenario with UGRs.

A summary photograph of the test configuration is shown in Figure 3. The plywood and metal roof canopy and the earthen side berms of the range caused a variety of natural reflections in the recording.

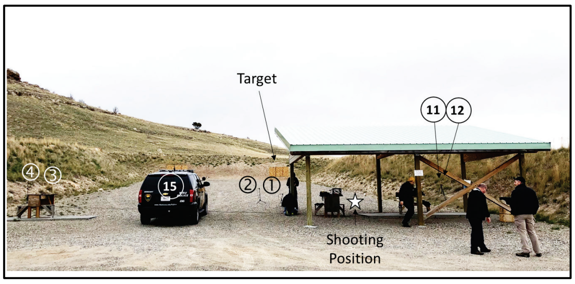

Figure 3: Photograph showing a summary of the shooting range test configuration.

Device	Mic Type	Channels
Tascam DR-680	Josephson C550F	1, 2
Tascam DR-680	DPA 4091	3, 4
Roland Edirol R-09HR	Built in	5, 6
Sony PCM-D50	Built in	7, 8
Olympus VN-702PC	Built in	9, 10
ZOOM H1	Built in	11, 12
Panasonic body cam	Built in	13
Olympus WS-852	Built in	14
Panasonic (in vehicle)	Built in	15

Table 1: Recording Configuration

A comparison of recorded channels *1-12* is shown in Figure 4 for a single shot from a Remington brand Model 870 shotgun with a 1-ounce lead slug shell. The prominent features of the shotgun slug recording in Figure 4 include the muzzle blast of the gun, followed approximately 200 ms later by the arrival at the microphone of the sound of the slug hitting the plywood target at the far end of the shooting range.

With the channels synchronized by the pre-recording slate, Figure 4 shows a relative time offset among the channels of the abrupt bang of the muzzle blast due to the acoustic path length differences. In a typical forensic investigation with unsynchronized user generated recordings, neither the time offsets nor the microphone positions would be available.

The important observation from Figure 4 is that the recorded waveforms differ significantly for the various recording devices and microphone positions. The gain settings and microphone sensitivities differ, causing more or less clipping of the waveform, while the very short duration of the actual muzzle blast (~3 ms) is extended by reflected sound energy, reverberating for 100 ms or more. This observation should lead to great caution for forensic determination of firearm type based on arbitrary UGR waveform analysis alone, as the available waveforms show considerable differences.

Another test example, shown in Figure 5, has multiple shots by a Glock brand G42 handgun, .380 caliber, recorded by channels *13*, *14*, and *15*, with channels *3* and *4* also shown for reference.

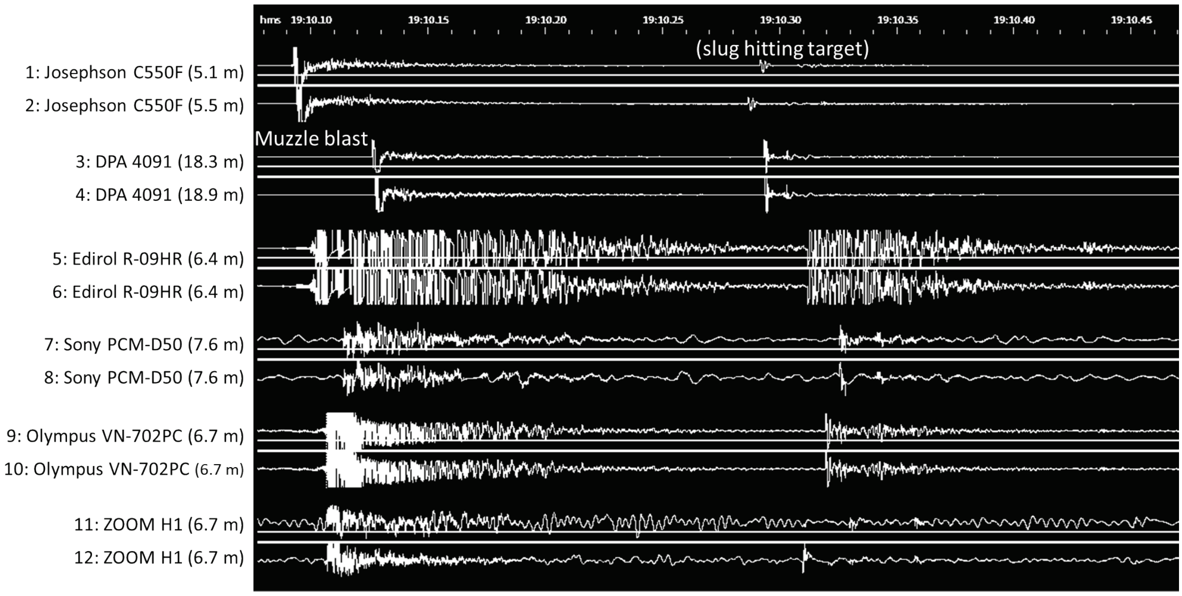

Figure 4: Simultaneous recordings of Remington brand model 870 shotgun, 1 oz lead slug cartridge.

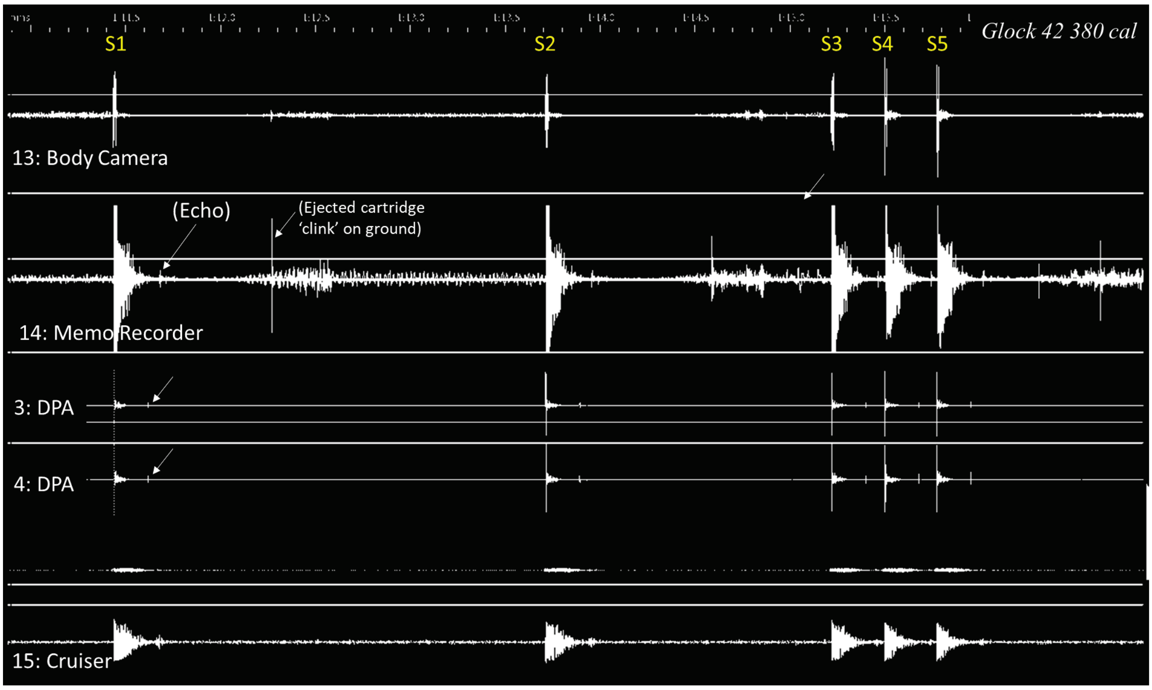

Figure 5: Simultaneous recordings of five shots (S1-S5) by Glock brand G42 handgun, .380 caliber.

As was observed for the shotgun example in Figure 4, the recorded handgun waveforms shown in Figure 5 differ significantly for the various recording devices and microphone positions. Again, caution is advised before a forensic examiner attempts determination of firearm type based on arbitrary UGR waveform analysis, because any expectation of a single "gunshot fingerprint" is not supported by actual observations, as shown here.

Two interesting interpretations can be seen in Figure 5. First, comparing signal *13* recorded by the body camera worn by the marksman, to signal *14* recorded by the memo recorder located about 20 cm away in the marksman's pocket, it is clear that the body camera system has a very fast-acting automatic gain control compared to the memo recorder. The body camera gain essentially shuts down at the high-amplitude report of the gun, while the memo recorder gain causes the waveform to clip, but still picks up the subtle sound of the spent cartridge ejected from the handgun falling and hitting the gravel ground surface of the shooting range approximately 900 ms after the shot. Even though both recorders are very close to the firearm and might be expected to show very similar data, the differences of the corresponding signals are evident. Second, while the amplitude decay envelope of each shot varies in detail from one recording to another, the onset timing of each distinct shot appears consistent across all of the recordings. This adds confidence that forensic examination involving shot-to-shot timing details, such as for multilateration, should be feasible from UGR waveforms.

As a final example, consider the telephone voicemail recording shown in Figure 6, which is from the same shotgun test shown previously in Figure 4. The voicemail recording was taken from a telephone call using an Apple® iPhone X placed to a corporate voicemail system (Cisco® Unity Connection). Channels *1* and *2* are also shown for reference.

The voicemail recording in Figure 6 shows again that waveform details are significantly altered by the communication and storage process: telephone microphone, speech coder, mobile phone channel, and voicemail storage coder. Care must be taken before drawing strong conclusions from gunshot sounds with many layers of processing.

3 Conclusions

The ever-increasing use of body cameras, handheld smartphones and other recording devices has resulted in the growing likelihood that multiple concurrent (but unsynchronized) user generated recordings (UGRs) may be presented as audio forensic evidence in a criminal investigation. As found in test gunshot recordings as well as in actual audio forensic evidence, recording the impulsive, high-amplitude report of a gunshot muzzle blast depends significantly upon the microphone type and position, the recorder settings, the acoustical surroundings, and the type of firearm discharged.

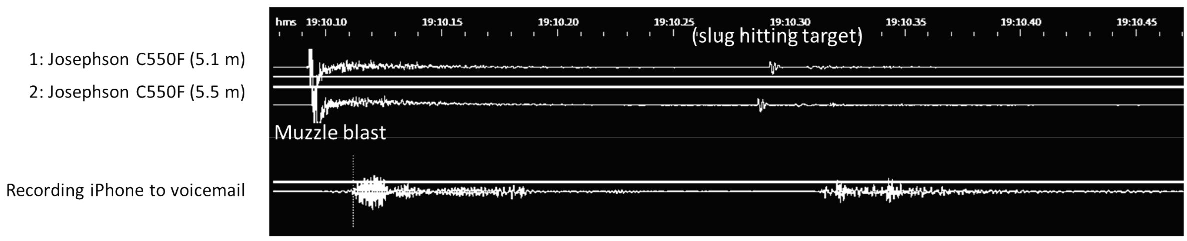

Figure 6: Remington 870 shotgun (slug) recording from a telephone call (Apple® iPhone X) to a corporate voicemail system (Cisco®Unity Connection). Channels *1* and *2* also shown for reference.

Our current recommendation for audio forensic examiners is to be aware of the potentially significant differences in the information content of the multiple recordings due to the spatial location, microphone sensitivity, bandwidth, automatic gain control, perceptual audio coding, and other peculiarities of the recording devices. This recommendation should be of particular importance to developers working to create automatic firearm classification systems: attempting to "train" a recognition system using a small number of gunshot recordings, or recordings obtained under a limited range of circumstances, is not advised.

4 Acknowledgements

This research work was supported in part by U.S. National Institute of Justice grants 2014-DN-BX-K034 and 2017-DN-BX-0179. The authors thank Montana State University Police officers Lt. Michael Stanley and Lt. Randy Crandall, and student research assistants Cameron Blegen and Matthew Blunt, for assistance with the gunshot recordings and interpretation. The authors also gratefully acknowledge the assistance of the anonymous reviewers in improving this paper.

References

[1] T.K. Routh and R.C. Maher, "Determining muzzle blast duration and acoustical energy of quasi-anechoic gunshot recordings," Preprint 9635, Proc. 141st Audio Engineering Society Convention, Los Angeles, CA, October, 2016.

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

[3] R.C. Maher, "Challenges of audio forensic evaluation from personal recording devices," Preprint 9897, Proc. 143rd Audio Engineering Society Convention, New York, NY, October, 2017.

[4] D.R. Begault, S.D. Beck, and R.C. Maher, "Overview of forensic audio gunshot analysis techniques," Proc. Audio Engineering Society International Conference on Audio Forensics, Porto, PT, June, 2019.

[5] S.D. Beck, H. Nakasone, and K.W. Marr, "Variations in recorded acoustic gunshot waveforms generated by small firearms," J. Acoust. Soc. Am., vol. 129, no. 4, pp. 1748-1759, April 2011.

[6] R.C. Maher and S.R. Shaw, "Directional aspects of forensic gunshot recordings," Proc. Audio Eng. Soc. 39th Conf., Audio Forensics—Practices and Challenges, Hillerød, Denmark, pp. 1-6, June 2010.

[7] R.C. Maher, "Acoustical modeling of gunshots using directional information and reflections," Preprint 8494, Proc. Audio Eng. Soc. 131st Conv., New York, NY, October 2011.

[8] 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, June 2014.

[9] S.D. Beck, "Who fired when: Associating multiple audio events from uncalibrated receivers," Proc. Audio Engineering Society International Conference on Audio Forensics, Porto, PT, June, 2019.

[10] P. Shrestha, M. Barbieri, and H. Weda, "Synchronization of multi-camera video recordings based on audio," presented at the Proceedings of the 15th ACM international conference on Multimedia, pp. 545–548, 2007.

[11] L. Kennedy and M. Naaman, "Less talk, more rock: Automated organization of community-contributed collections of concert videos," Proceedings of the 18th international conference on World Wide Web, pp. 311–320, 2009.

[12] C. Cotton and D. Ellis, "Audio fingerprinting to identify multiple videos of an event," Proc. ICASSP, pp. 2386–2389, 2010.

[13] J. Bryan, P. Smaragdis, and J. Mysore, "Clustering and synchronizing multi-camera video via landmark cross-correlation," Proc. ICASSP, pp. 2389 – 2392, 2012.

[14] J. Kammerl, N. Birkbeck, S. Inguva, D. Kelly, A. Crawford, H. Denman, A. Kokaram, and C. Pantofaru, "Temporal synchronization of multiple audio signals," Proc. ICASSP, pp. 4603–4607, 2014.

[15] S. Bano and A. Cavallaro, "Discovery and organization of multi-camera user-generated videos of the same event," Journal of Information Sciences, vol. 302, pp. 108–121, 2015.

[16] P. Shrestha, P. de With, H. Weda, M. Barbieri, and E. Aarts, "Automatic mashup generation from multiple-camera concert recordings," Proc. Int. Conf. on Multimedia, ACM, pp. 541–550, 2010.

[17] N. Stefanakis, Y. Mastorakis, A. Alexandridis, and A. Mouchtaris, "Automated mixing of user-generated audio recordings from the same event," J. Audio Eng. Soc., vol. 67, no 4, pp. 201-212, April 2019.

[18] R.C. Maher and E.R. Hoerr, "Audio forensic gunshot analysis and multilateration," Preprint 10100, Proc. 145th Audio Engineering Society Convention, New York, NY, October, 2018.