Audio Engineering Society
Conference Paper 6
Presented at the AES 8th International Conference on Audio Forensics
2024 June 27–29, Denver, Colorado, USA

A Geometric Approach for Generating Synthetic Gunshot Acoustic Signals

Siddat B. Nesar, Bradley M. Whitaker, and Robert C. Maher

Electrical and Computer Engineering, Montana State University, Bozeman, MT

Correspondence should be addressed to Siddat B. Nesar (`siddatnesar@montana.edu`)

ABSTRACT

This paper presents a novel geometric approach using anechoic gunshot recordings to create large libraries of training data for classifying firearm sounds using machine learning. Realistic gunshot sounds require consideration of the type of firearm, its orientation and directionality, and at least the first-order effects of acoustic reflections from surrounding obstacles, but available gunshot sound libraries do not contain a sufficient variability in these factors to represent the wide range of conditions encountered in actual audio forensic investigations of gunshot sounds. To generate a more comprehensive set of training examples, we used a set of directional anechoic gunshot recordings and simulated geometrical transformations to achieve an arbitrary number of simulated gunshots representing different firearm-to-microphone configurations. This research advances the realism of gunshot simulation, generating sufficient synthetic data for training and evaluating gunshot classification methods.

1 Introduction

Gunshot acoustics are vital in forensic investigations, military training, and crime scene simulations. Accurate modeling of gunshot acoustics can give important information about the orientation and the location of the firearm when discharged. Standard analyses rely on recordings from law enforcement officers and/or nearby witnesses, which can later be verified using the gunshot acoustic model to corroborate the statement provided by the witness about the crime scene. It is challenging to simulate realistic gunshot sounds due to the unknown location and orientation of the firearm and the complex environmental echoes. This work proposes a novel geometric approach to generate realistic gunshot sounds from anechoic data while considering the layout and other properties of the surrounding environment.

It is impractical to collect controlled gunshot sound data covering all firearm types, orientations, and geometric configurations, so generating realistic synthetic gunshot sound data is helpful for simulating forensic evidence [1]. Scientifically realistic gunshot synthesis provides an abundance of data while preserving both security and privacy [2], which in some cases can be important concerns. Synthetic data is used for developing machine learning models to recognize gunshots produced by various firearms [3]. With the advancement of machine learning models, synthetic data generation for various applications is becoming popular in other fields, such as semantic segmentation in vision [4], text-to-speech for speech recognition [5], exploring the distribution of drug molecules for drug discoveries [6], and so on. However, little research has been done on generating synthetic data for gunshot sounds, especially in a complex environment. Other than using machine learning models, another widely used approach is to generate data based on the distribution of a subset of the dataset. The best distributions can be obtained using the Monte Carlo method or various machine learning models [7, 8].

Anechoic gunshot acoustic data is considered to be clean recordings without any environmental reflections. However, these high-quality data can be used to generate acoustic signatures of gunshot sounds incorporating spatial and environmental information. We leverage acoustic signal processing fused with geometric modeling of the surrounding environment to generate the complex sound waves that will occur when a gun is fired in similar circumstances. We consider the location's temperature as the speed of sound varies with temperature, and consider the reflections from the ground as the sound absorption capability is different for different kinds of materials. The temperature and the echoes from the ground reflections will affect the timing and the final sound waves that reflect in our approach.

In this work, we propose a novel geometric approach for generating gunshot sound waves considering the ground reflection and the temperature of the surrounding environment. First, we talk about the unique dataset containing anechoic sound waves from different firearms collected in a very controlled manner in Section 2. Then, section 3 explains the proposed model used to simulate the gunshot sound wave in a given scenario. Section 4 illustrates the generated sound waves in two scenarios for three different firearms. Finally, section 5 wraps up the work with an overall summary and future extensions.

2 Anechoic Data

We have collected anechoic data for this experiment to minimize the reflections and echoes, preserving the precise sound characteristics with a very high sampling frequency of 500 kHz. This work included the anechoic data of three commonly used firearms, Glock 19, AR15, and 308 Rifle. Twelve microphones (GRAS 40DP) were set up in a semi-circular arc 3 meters above the ground to create a quasi-anechoic environment [9]. Different azimuthal variations of the muzzle blast of all the firearms were collected using LabVIEW. The very high sampling frequency allows the precise capture of the muzzle blasts and the shock waves present at certain orientation angles with supersonic projectiles [10]. The raw data may then be bandlimited and downsampled to a sampling rate suitable for a particular purpose.

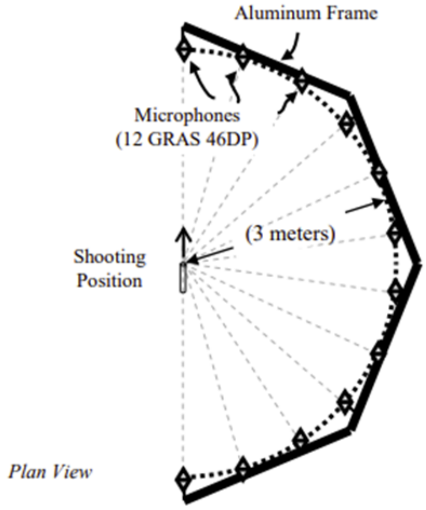

Fig. 1: Data collection setup showing the positions of recording microphones and the position of the firearm. The microphones are 3 meters above the ground and 3 meters away from the firing position.

The orientations of the twelve different microphones are illustrated in Fig. 1. They were arrayed directly in the line of the fire with an angular spacing of around 16 degrees. The exact angles of the microphones are 0.00, 16.36, 32.73, 49.09, 65.45, 81.82, 98.18, 114.55, 130.91, 147.27, 163.64, and 180.00 degrees respectively where the microphone at 0.00 degrees is on-axis, and the microphone at 180.00 degrees is directly behind the firearm.

We removed the ballistic shock wave, if any, from the anechoic data and used the resulting raw gunshot waveforms as the baseline for generating the other sound waves at different parameters. The cleaned audio wave is around 11 milliseconds long containing the muzzle blast and the reflections from the muzzle blast. Fig. 2 presents the collected anechoic data for three different firearms: (a) AR15, (b) Glock 19, and (c) 308 rifle. We have data from 12 different angles around a semicircular arc, and out of those 0 degrees (on-axis or in line of fire) and 180 degrees (directly behind the line of fire) are demonstrated in Fig. 2. The rifles AR15, and 308 exhibit high amplitude as expected on-axis, while

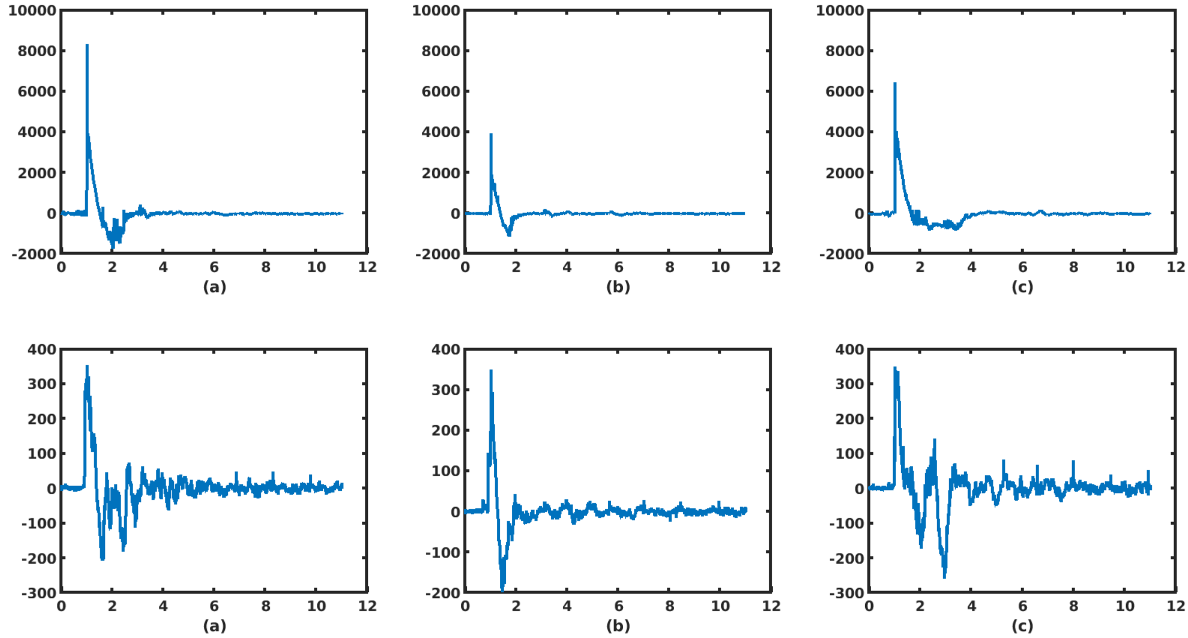

Fig. 2: The top three plots show the anechoic on-axis (0 degrees) gunshot sound waves, and the bottom three plots show the anechoic directly behind (180 degrees) gunshot sound waves. (a) presents the plots for AR15, (b) for Glock 19, and (c) for 308 rifles. The x-axis represents the time in ms, and the y-axis represents the acoustic sound pressure in pascals.

the amplitude of a Glock-19 is almost half. When the gunshot sound is collected directly behind (180 degrees), we see the amplitude for all three different firearms are quite similar but very noisy.

3 Methods

The generation model involves several acoustic signal processing techniques, image methods in optics and acoustics, and geometrical modeling for ground reflection. In our proof-of-concept model, we used the anechoic data from various firearms including a Glock 19, an AR15, and a 308 rifle. First, we set some of the parameter ranges such as the target range along X (T_x), Y (T_y), and Z (T_z) axes from 5 to 100 meters. The gun height, G_h was considered within a range of 0 to 10 meters from the ground. The temperature, T, and the reflection coefficient, Γ can vary, however, we will use $T = 10$, and $\Gamma = 0.98$ for consistency and explanation in this paper. Different variations of T_x, T_y, T_z, and G_h will give different sound waves as they will be heard in such circumstances.

Fig. 3 presents a simplified simulation scenario to generate the synthetic gunshot data as heard from the recorder's perspective. The diagram shows the geometric setup of a person recording the gunshot. The parameters shown in the figure are known and can be easily obtained from the scene of fire. The direct distance between the recording phone and the gun is d, while the gun and the recording phone are both h and m meters above the ground. α and β are the two angles that will be used in the estimation of the signal as received at the recorder's end leveraging the image method [11]. The ground-level reflection coefficient is Γ.

At higher temperatures, the air molecules vibrate faster and allow the sound waves to travel quickly. At 0 degrees Celsius, the speed of sound is 331.30 m/s. For any given temperature, T, in Celsius, the speed of sound, c, in meters per second, can be calculated using Eq. 1.

$$c = 331.3 \times \sqrt{1 + \frac{T}{273.15}} \tag{1}$$

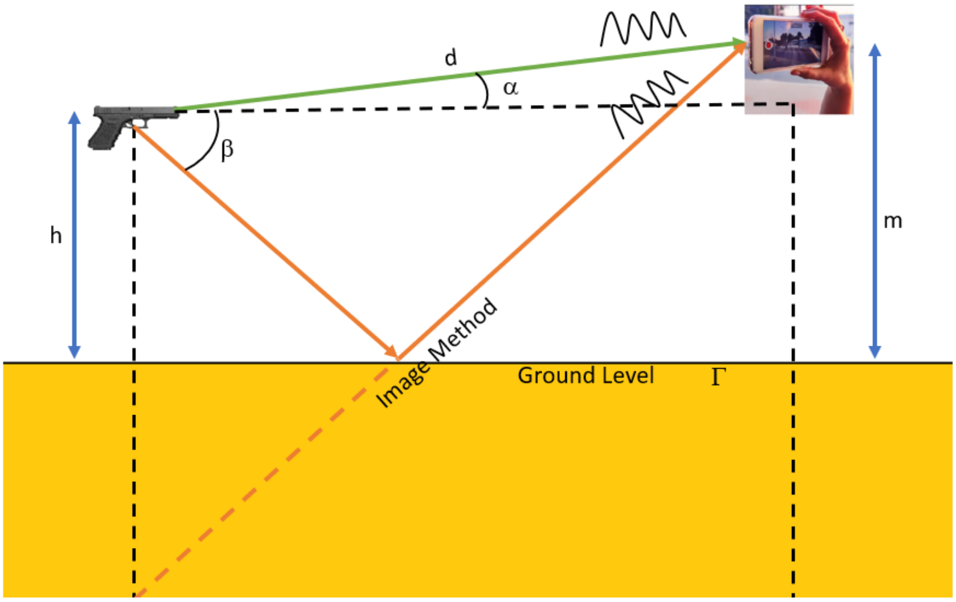

Fig. 3: Visualization of a two-dimensional geometric setup of the model. The shooter is *d* meters away from the recorder, where *d* is the direct distance between the gun and the phone. The recorder receives one direct signal traveling through the air and the echo reflecting from the ground. Our model estimates both the signals and combines them appropriately to generate the final signal recorded. Note that our algorithm considers three spatial dimensions, not just two.

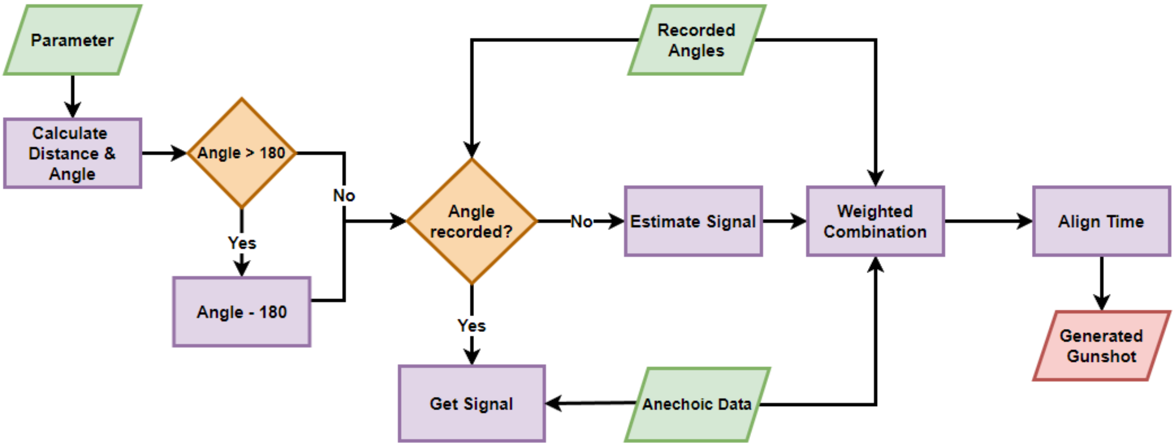

Fig. 4: Flow chart showing the steps of the proposed model to generate gunshots from anechoic data.

Fig. 4 represents the block diagram of the proposed model. We get the parameters from the scene of gunfire that include the witness's location, the fired gun's height, and the ground surface's reflection coefficient. We also note the temperature of the surroundings at the time of the fire. We then calculate the distance and angle of the direct and the echo signals. The next step is to check if the angle of the gunshot signal is over 180 degrees. If yes, we subtract 180 from the calculated angle as the gunshot sound is omnidirectional. Next, if

Algorithm 1 Gunshot Sound Generation Model

```
Require: params ← $G_h, T_x, T_y, T_z, \Gamma$
    $v \leftarrow [T_x, T_y, T_z - G_h]$
    $D_d \leftarrow [0,0,0], v$ use $Eq.2$
    $D_a \leftarrow [1,0,0], v$ use $Eq.3$
    $E_d \leftarrow [0,0,-2 \times G_h], v$ use $Eq.2$
    $v \leftarrow [T_x, T_y, T_z + G_h]$
    $E_a \leftarrow [1,0,0], v$ use $Eq.3$
    if $D_a$ or $E_a \geq 180$ then
        $D_a$ or $E_a \leftarrow 360 - D_a$ or $E_a$
    end if
    $[D_s, E_s] \leftarrow x, R, G_h, D_{da}, E_{da}, \Gamma$ use $Alg.2$
    $(D_t, E_t) \leftarrow (D_d, E_d)/c$
    $y, t \leftarrow$ Combine and align $D_s, E_s$ use $Alg.4$
```

Algorithm 2 Estimate the Signal and Echo

```
Require: params ← $x, R, G_h, D_{da}, E_{da}, \Gamma$
    if $D_a$ or $E_a$ is in $R$ then
        $j \leftarrow D_a$ or $E_a == R$
        $D_s \leftarrow x_j \times G_h / D_d$
        $E_s \leftarrow x_j \times G_h / E_d \times \Gamma$
    else
        $j \leftarrow$ find $((R > D_a$ or $E_a), 1)$
        $D_{st} \leftarrow x_{j-1}, x_j, R_{j-1}, R_j, D_a$ use $Alg.3$
        $E_{st} \leftarrow x_{j-1}, x_j, R_{j-1}, R_j, E_a$ use $Alg.3$
        $D_s \leftarrow D_{st} \times G_h / D_d$
        $E_s \leftarrow E_{st} \times G_h / E_d \times \Gamma$
    end if
```

Algorithm 3 Estimating Signals for Unknown Angles

```
Require: params ← $x_a, x_b, \theta_a, \theta_b, \theta_y$
    $\alpha \leftarrow \theta_y - \theta_a$
    $\beta \leftarrow \theta_b - \theta_y$
    $y \leftarrow (\beta \times x_a + \alpha \times x_b)/(\alpha + \beta)$
```

the calculated angle is present in our recorded dataset, then the estimated signal will be similar to that in the recorded dataset. If the angle of the signal is anything other than the recorded angles, then we use a weighted combination technique to estimate the gunshot signal. We then use the distances to align the direct signal and the echo signal to generate the combined final signal at the receiver's end.

The model proposed is summarized in the Alg. 1 which takes $G_h, T_x, T_y, T_z, \Gamma,$ and c as the input and gives the final signal along with the time axis. First, we find the relative position, v of the gun with respect to the recorder to calculate the direct distance, D_d, and the direct angle, D_a between them. We will use Eq. 2 to calculate the distance between two vectors in a 3-dimensional space. Eq. 3 calculates the angle between two vectors in a 3-dimensional space. We then find the relative position of the gun and the recorder to estimate the signals for echo. We leverage the image method to estimate the echo distance, E_d, and the echo angle, E_a using equations 2 and 3 respectively. As sound waves propagate omnidirectionally, we check for the direct and echo angles over 180 degrees.

$$d = \sqrt{\sum (\vec{v}_1 - \vec{v}_2)^2}. \tag{2}$$

$$\theta = \arccos\left(\frac{\vec{v}_1 \cdot \vec{v}_2}{||\vec{v}_1|| ||\vec{v}_2||}\right) \tag{3}$$

Now that we have all the required parameters calculated, we use Alg. 3 to estimate the direct signal and the echo signal as will be received at the recorder's end. Here, x is the anechoic data and R is the vector of all the recorded angles in the anechoic dataset. In our case, R is a vector of length 12 containing the angle values [0.00, 16.36, 32.73, 49.09, 65.45, 81.82, 98.18, 114.55, 130.91, 147.27, 163.64, and 180.00]. We check if the direct angle, D_a, and the echo angle, E_a are in R. If so, then we already have the signal we desire and can estimate the direct signal, D_s, and the echo signal, E_s by taking the product of the signal and the gun height and dividing the relative distance from the gun. For the echo signal, we need to take the product of the signal and the reflection rate, Γ to account for the signal lost due to the reflection from the ground.

If D_a and/or E_a are not present in R, then we estimate the signals using a weighted combination based on the differences of the angles as presented in Alg. 3. We take the nearby known signals for an unknown angle and find the difference between the unknown angle from the lower known angle and the higher known angle. The angle differences α and β are used as weights to estimate the new signal. We still have to take the product of the signal with the gun height and divide the respective direct distance, and for the echo signal, we divide by Γ to account for the signal lost due to the absorption or diffraction on the ground. Thus we estimated the direct signal, D_s, and the echo signal, E_s.

Algorithm 4 Align Signals and Echo

```
Require: params ← $D_s, E_s, D_t, E_t, f_s$
    $d \leftarrow (E_t - D_t) \times f_s$
    $D_p$ or $E_p$ ← zero pad $D_s$ or $E_s$ with $d$
    $C_s$ ← stack $D_p, E_p$
    $C_t \leftarrow (1 : \mathrm{len}(E_p))/f_s + D_t$
    $y$ or $t$ ← zero pad $C_s$ or $C_t$ with $D_t \times f_s$
```

We divide the direct distance, D_d, and the echo distance, E_d, by the speed of sound, c to obtain the start time of the direct signal and the echo signal as received on the recorder's end. We now use the Alg. 4 to merge the direct signal, D_s, and the echo signal, E_s, and align them properly as they will be received. The echo signal will be delayed by some samples, d, that can be obtained by the product of the sampling frequency, f_s, and the difference between the echo signal start time, E_t, and the direct signal start time, D_t. d needs to be rounded to the nearest integer and d zeroes are padded to get the exact time for both the signals. We then add them to get the final signal. We use the sampling frequency to get the timing aligned as well.

4 Results

Our proposed model can generate the signal at any given coordinates, when the gun height, temperature, and ground reflection coefficient are known. We need anechoic data at different angles to be used as a baseline for the desired firearm to be able to generate a realistic signal at the asked location. Say, the outdoor temperature is 10 degrees Celsius and the ground reflection coefficient is 0.98. We present the signals that will be received at two different positions for the three different firearms we experimented with in Fig. 5 and Fig. 6.

We present two scenarios with the gun height as 1 meter and 1.5 meters for the three firearms to simulate someone standing on the ground holding a handgun or rifle. We also see the differences in the generated plots with the same parameters but a lower reflection coefficient to correspond to a higher absorption or refracting surface. For the position (22, 46, 1) meters and the gun being shot from a height of 1 meter, the top row in Fig. 5 illustrates the generated gunshot waves for AR15 (a), Glock-19 (b), and 308 rifles (c) respectively. The AR15 and 308 rifles exhibit similar acoustic signatures, and we see two very close peaks in all three plots in the top row. The first peak is the muzzle blast heard at the desired location, while the second peak which is a bit low in amplitude is due to the echo reflecting from the ground because of the strong reflection coefficient simulated here for clarity (0.98). Similar characteristics are seen for the Glock-19 handgun, however the amplitudes are much lower compared to the rifles. The delay in the echoed signal is due to the long distance traveled by the sound wave and since the distance traveled by the direct signal and the echo signal is almost the same, the two peaks are less than 1 milliseconds apart. For the mentioned scenario, the distance traveled by the direct signal and the echo signal is 50.99 m and 51.03 m respectively. The angle of the mentioned position to the gun for the direct signal and the echo signal is 64.44 degrees and 64.46 degrees respectively. The sound is heard around 152 milliseconds after the shot was fired and lasts till 163 milliseconds.

The bottom plots in Fig. 5 represent the generated gunshot waves for the position (13, 7, 10) meters, and the gun was shot from 1.5 meters above the ground. We see two high peaks in the generated gunshot waves, where the first peak is from the muzzle blast received directly, and the second peak is due to the echo from the ground. We see similar characteristics in the peak amplitudes compared to the scenario aforementioned. However, we see the echo signal is quite delayed, and the first high sound is heard sooner. In this given scenario the distance traveled by the direct signal and the echo signal is 17.04 meters and 18.72 meters respectively. The angle of the direct signal and the echo signal is 40.27 degrees and 46 degrees respectively. Since the receiver is quite close to the source from where the gun was shot and the gun is at a greater height, the echo signal travels a longer distance and gets delayed from being received at the recorder's end. We observe the signals are received at around 50 milliseconds and last till 67 milliseconds. We notice a delay of around 5 milliseconds in the echo signal.

Fig. 6 represents similar results as reported in Fig. 5, however, with a lower reflection coefficient of 0.8 to simulate the received signal in a rough surface. A lower reflection coefficient results in a lower amplitude in the sound signals by around 100 pascals. Other than that, the trace of the gunshot signal is almost the same as the scenario reported for Fig. 5.

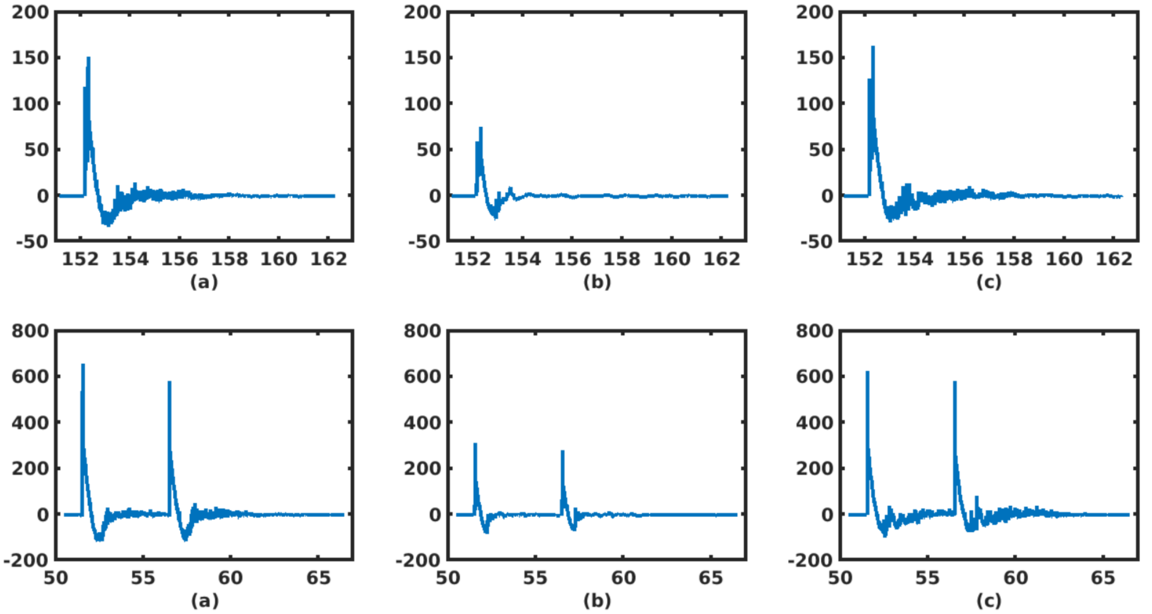

Fig. 5: Some examples of generated gunshot waves at different parameters using the model proposed. The top three plots show the estimated signal received at position (22, 46, 1) meters while the gun was fired from a height of 1 meter. The bottom three plots show the estimated signal received at position (13, 7, 10) meters while the gun was fired from a height of 1.5 meters. (a) presents the plots for AR15, (b) for Glock 19, and (c) for 308 rifles. The x-axis represents the time in ms, and the y-axis represents the sound amplitude in pascals.

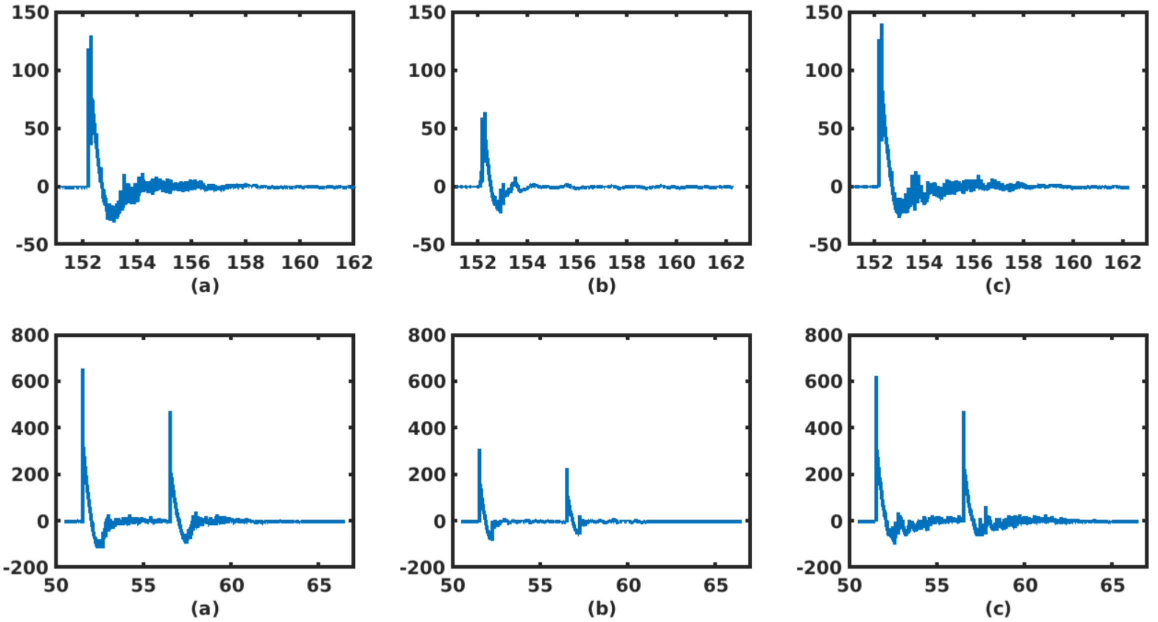

Fig. 6: Same parameters are used as in Fig. 5, however with reflection coefficient of 0.8. Note that the amplitude of the echo signal is lower in all the cases.

5 Conclusion

In conclusion, the main contribution of this research is presenting a unique geometric approach to generate gunshot sounds from anechoic acoustics data along with accurate rendering of ground reflections. We incorporated the geometric properties of acoustics and the image method for simulating room acoustics to generate real-world-like gunshot sounds capturing the complex interactions of the surrounding environment. Our method produces realistic gunshot sounds with high fidelity by accurately modeling the environmental factors, such as the temperature of the surroundings and the reflections and diffractions from the ground surface. This research is a significant step toward diverse applications ranging from training simulations for law enforcement and forensic reconstructions and investigations. Looking ahead at future research, further refinement of the geometric approach considering a more complex environment could lead to more realistic gunshot simulations. Comparing the simulated gunshot sounds using the same circumstances as of real-world collected gunshot sounds could shed light on the performance quality of the simulations. In summary, we present an important step forward in the field of gunshot acoustics simulation, by offering an efficient geometric approach to generate gunshot sounds with ground reflections.

References

[1] Busse, C., Krause, T., Ostermann, J., and Bitzer, J., "Improved gunshot classification by using artificial data," in *Audio Engineering Society Conference: 2019 AES International Conference on Audio Forensics*, Audio Engineering Society, 2019.

[2] Lucini, F., "The real deal about synthetic data," *MIT Sloan Management Review*, 63(1), pp. 1–4, 2021.

[3] Nesar, S. B., Whitaker, B. M., and Maher, R. C., "Machine Learning Analysis on Gunshot Recognition," in *2024 Intermountain Engineering, Technology and Computing (IETC)*, pp. 1–5, IEEE, 2024.

[4] Sankaranarayanan, S., Balaji, Y., Jain, A., Lim, S. N., and Chellappa, R., "Learning from synthetic data: Addressing domain shift for semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 3752–3761, 2018.

[5] Fazel, A., Yang, W., Liu, Y., Barra-Chicote, R., Meng, Y., Maas, R., and Droppo, J., "Synthasr: Unlocking synthetic data for speech recognition," *arXiv preprint arXiv:2106.07803*, 2021.

[6] Gómez-Bombarelli, R., Wei, J. N., Duvenaud, D., Hernández-Lobato, J. M., Sánchez-Lengeling, B., Sheberla, D., Aguilera-Iparraguirre, J., Hirzel, T. D., Adams, R. P., and Aspuru-Guzik, A., "Automatic chemical design using a data-driven continuous representation of molecules," *ACS central science*, 4(2), pp. 268–276, 2018.

[7] Figueira, A. and Vaz, B., "Survey on synthetic data generation, evaluation methods and GANs," *Mathematics*, 10(15), p. 2733, 2022.

[8] Miok, K., Nguyen-Doan, D., Zaharie, D., and Robnik-Šikonja, M., "Generating data using Monte Carlo dropout," in *2019 IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP)*, pp. 509–515, IEEE, 2019.

[9] Routh, T. K. and Maher, R. C., "Recording anechoic gunshot waveforms of several firearms at 500 kilohertz sampling rate," in *Proceedings of Meetings on Acoustics*, volume 26, AIP Publishing, 2016.

[10] Maher, R. C., "Acoustical characterization of gunshots," in *2007 IEEE Workshop on Signal Processing Applications for Public Security and Forensics*, pp. 1–5, IEEE, 2007.

[11] Allen, J. B. and Berkley, D. A., "Image method for efficiently simulating small-room acoustics," *The Journal of the Acoustical Society of America*, 65(4), pp. 943–950, 1979.