

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/315862064>

Procedural Synthesis of Gunshot Sounds Based on Physically Motivated Models

Chapter · March 2017

DOI: 10.1007/978-3-319-53088-8_4

CITATIONS

3

READS

4,499

1 author:



[Huseyin Hacihabiboglu](#)

Middle East Technical University

76 PUBLICATIONS 794 CITATIONS

SEE PROFILE

Procedural Synthesis of Gunshot Sounds based on Physically-motivated Models

Hüseyin Hacıhabiboğlu

Abstract Generation of content for games is one of the major bottlenecks in terms of the effort required and the resources to be committed. A typical AAA game contains tens of thousands of sound files as audio assets which include spoken dialogue as well as sound effects. Procedural content generation (PCG) provides a cost effective alternative to recording these sounds in the studio or in the field. While some sound effects can be recorded fairly easily given the necessary time, effort, and resources, some others such as gunshot sounds are not easy to record. Since many games and simulations incorporate firearms, parametric sound synthesis which is essentially a PCG technique can be used to alleviate the need to record gunshot sounds. This chapter describes a physically-motivated parametric gunshot sound synthesis model. The model is based on a deconstruction of the gunshot sound event into its constituent parts and uses parameters such as the barrel length, bullet type, and muzzle velocity to synthesise the sounds of different firearms. A subjective evaluation which investigates the perceptual relevance of the proposed model is also presented.

1 Introduction

Sound effects are essential components of computer games. While it is difficult to estimate the average number of audio assets included in a game, it is reasonable to expect that potentially tens of thousands of sounds could be needed in a AAA title. For example, Battlefield 4 with its expansion packs has in excess of 78000 sound files [1]. Considering how hard it is to record so many sounds, procedural audio generation emerges as a possible solution. Many different sounds can be synthesised using procedural means instead of being recorded, and this would both reduce the size of the game and the substantial effort required for the task. Besides, some

Hüseyin Hacıhabiboğlu
Graduate School of Informatics, Middle East Technical University (METU), Ankara, TR-06800,
Turkey, e-mail: hhuseyin@metu.edu.tr

sounds such as wind or rain are very hard if not impossible to record, due to physical constraints. Procedural sound generation provides a viable solution for such cases, too.

Shooters remain one of the most popular video game genres constituting 24.5% of the video games sold in the USA in 2015 [24]. Shooters such as games in the Battlefield series [6] or the serious game, America's Army [27] require a variety of different weapon sounds. Game mechanics that involves shooting requires the playback of the sounds of a variety of firearms. For this reason, recording or generation of high quality weapon sounds is essential. As an example, Battlefield 4, praised for its sound design, has over 3900 recorded sound effects for weapons.

The problem with weapon sounds is even more profound with games which incorporate procedural generation. The games, Borderlands [10], and Stack Gun Heroes [26] are good examples which allow the parametric generation of a huge number of weapons, all having different properties. In such games, sounds of weapons, which typically will not exist in real life, also have to be generated.

The aim of this chapter is to present a physically-motivated, generalised procedural gunshot sound synthesis model which can be used to generate the sounds of a wide variety of classes and types of firearms with subsonic and supersonic projectiles. The presented synthesis model is simple, flexible, and modular so that it can be used as part of a larger chain of procedural sound generation algorithms. The perception of the synthetic sounds obtained from the model are compared with real recordings in a subjective experiment and the results indicate that procedurally generated gunshot sounds can possibly replace real gunshot recordings without substantially reducing the perceived realism.

2 Recording Gunshot Sounds

Generating a large and generic database of recordings of gunshot sounds for use in games, simulations, and other virtual reality applications may seem as a reasonable action to take. However, there are several drawbacks. First and foremost, legal difficulties may prohibit anyone interested in developing such a database from possessing an arsenal of different firearms especially in countries where strict gun control laws are in place.

Leaving this difficulty aside, several other technical issues make it difficult to record firearm sounds. The first difficulty is the sound pressure levels (SPL) observed with gunshot sounds. Depending on the gun and ammunition properties, peak signal level can easily exceed 150 dB SPL. Such a level is generally not handled well by general-purpose recording microphones used by sound effects artists and is likely to result in saturation and clipping rendering the original signal unrecoverable. The diaphragm size of the employed microphone, the maximum sound pressure level that it can handle, and its sensitivity are thus very important in obtaining a recording that is free of artefacts [21].

Another important difficulty is the inclusion of the acoustics of the environment in the recording. If the sound of the weapon is recorded anywhere other than an anechoic chamber, the reflections and the reverberation pertaining to the environment cannot be excluded from the recorded signal. If the same gun can be used in different environments in the game world, such as in a warehouse, a forest, a city canyon, or an open field, the mismatch between the acoustics of the recording venue and that of the virtual environment in the game would negatively impact its realism. Consider, for example, the case where the recordings are made in an outdoor shooting range and the game world is primarily designed as indoors.

Care has to be taken at every level of the audio recording chain while recording gunshots. An important difficulty which is often overlooked is the sampling rate. As will be discussed below, components of gunshot sounds have very short durations and thus gunshots will in general have very wide frequency ranges. A sufficiently high sampling rate should be used to capture a good representation of the gunshot.

Another aspect of gunshots is that they are highly directional [2]. Largest sound pressure levels are observed at the boreline direction and the sound level decreases progressively as the recording direction moves away from the boreline direction. Behind the firing position, the body of the person firing the gun shadows the sound wave. For supersonic projectiles the composition of the gunshot sound also changes with direction. The shock wave may or may not be observed based on the recording position. This makes it necessary to obtain multiple simultaneous recordings around the shooting position, which is impractical most of the time.

For the reasons stated above, gunshot recordings are in general not suitable for use in serious games where good levels of authenticity and realism are required.

3 Synthesis of Gunshot Sounds

Gunshot sounds experienced in real life and as used in games and movies are profoundly different. A gunshot sound in a game acts as an expressive device to elevate the excitement, and possesses diegetic and expectational relevance but not necessarily realism. Simulations, on the other hand, require authentic or near-authentic gunshot sounds due to the training requirements, and the need to satisfy the expectations of the users who are already accustomed to how a specific firearm sounds like.

3.1 *Properties of Gunshots*

Sound due to a gunshot is a transient, high-energy signal which has well-defined properties. These properties depend not only on the physical features of the firearm and the ammunition which generates the sound, but also on the environment in which the gun is fired and the direction of the microphone with respect to the

firearm [16]. More specifically, a gunshot sound consists of the muzzle blast and the shock wave. Muzzle blast is localised at the muzzle and occurs due to the explosion of the propellant in the bullet. Shock waves (i.e. also known as sonic boom or N-wave) occur only for bullets or projectiles which have supersonic speeds. These waves will have distinct propagation patterns following the projectile and spectra with a wide bandwidth due to the discontinuity they possess.

The sound generated by a firearm is also affected by i) the observation distance, ii) the direction of observation, and iii) the environment. The primary effect of distance is by attenuating the level, and absorbing the high-frequency content due to thermal relaxation process [15]. The direction of observation is an important factor in whether or not the shock wave will be audible and also in the relative level of the muzzle blast sound with respect to its front direction [18]. If the gunshot occurs in open air with no nearby reflectors apart from the ground, the effect of the environment will be limited by a single ground reflection. Any other environmental setting such as a room or a forest will also add reverberation. A realistic presentation of a gunshot should thus incorporate the sonic properties of the gun as well as the properties of the environment.

3.2 *Parametric Synthesis of Gunshots*

Parametric sound synthesis is a very active field of research and many different models have previously been proposed for different purposes [8]. For example, there exist models for footsteps [20], bird calls [14], and vehicle sounds [4]. Research into affective synthesis of speech is also an active field [22] which will, in the near future, make it possible to tie conversational agents to NPCs and to create games with dynamically evolving scripts.

A crucial component in the procedural generation of a sound is the availability of computationally tractable models of the processes which are involved in the generation of that sound in real life. Such models can be physical, perceptual, or statistical.

The sound generated by a gunshot is above all a physical phenomenon. There exist well-established but approximate theoretical models for different portions of a gunshot sound and the properties of the acoustical environment discussed above.

Regardless of the properties of the firearm being used, muzzle blast has a distinctive shape with a sharp onset followed by positive and negative phase portions, respectively [2]. While this waveform is common across many different types of firearms, the durations of the positive and the negative phases as well as the peak pressure at the onset will depend on the caliber, the form factor, and the length of the projectile, as well as the length of the barrel. The peak pressure also depends on the direction of observation around the firearm.

The pressure signature of the shock wave due to a supersonic bullet depends mainly on the dimensions and the speed of the bullet. While an exact analytical

Fig. 1 Ultra-high speed photo of bullet fired out of a SW revolver photographed with an air-gap flash by Niels Noordhoek, <https://www.scopus.com/authid/detail.uri?authorId=6603387174> (CC BY-SA 3.0)



solution of the pressure signature is not feasible, simplified models can be used to predict the shape of the N-wave [2].

The effects due to distance can be simulated by using two distinct attenuation models for the muzzle blast and the shock wave. The high-frequency attenuation can be simulated using specially designed air absorption filters [11]. In addition, the relative level of the shock wave will depend on the observation angle and should be simulated accordingly. There exist several different artificial reverberation algorithms which can be used to simulate different environments (e.g. an enclosure [5] or a forest [23]). If an open field is assumed, the only environmental effect is the ground reflection which can be simulated by adding to the original signal a delayed and attenuated version of itself.

Physical models of firearm acoustics as well as a simple parametric gunshot sound generator are described in the following section.

4 Physically-motivated Synthesis of Gunshot Sounds

A gun is a single stroke steam engine designed to throw a bullet in a desired direction with a very high speed. The bullet forms the tip of the cartridge which also consists of the casing, the propellant, and the primer. The propellant is typically a combination of some oxygen and chemicals which when ignited will produce rapidly expanding gas to propel the bullet through the barrel and out from the muzzle. Fig.1 shows an ultra-high speed photograph of a bullet exiting the muzzle of Smith & Wesson 686 .38 Special revolver pistol. Muzzle blast is clearly visible in the photo. Parts of an automatic handgun and ammunition are shown together with the parameters that affect the pressure signal generated due to the gunshot in Fig. 2.

Sounds generated during a gunshot can be separated into three: i) ballistic sounds such as muzzle blast and shock wave, ii) mechanical action sounds, and iii) environment effects such as reflections and reverberation. The first group is what is known as the gunshot and consist primarily of the muzzle blast and shock wave. The second group consist of sounds generated by the internal mechanics of the firearm before,

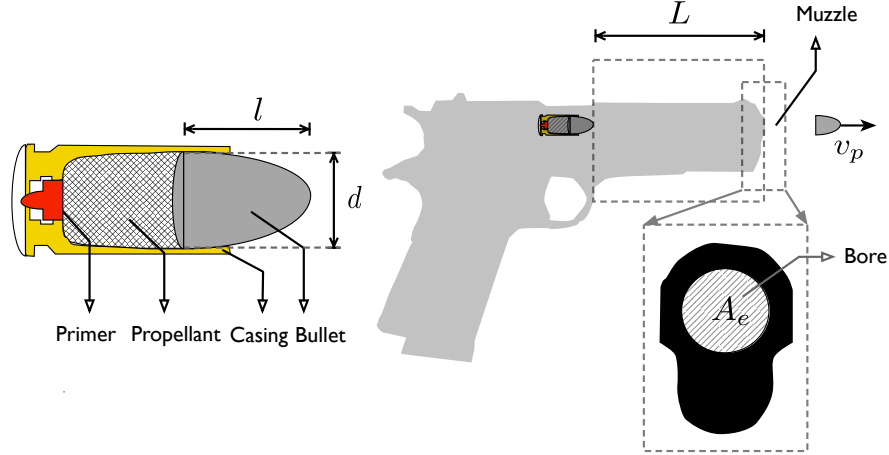


Fig. 2 Diagrams of a projectile and a M1911 handgun. Several different parameters used in the synthesis model are also shown.

during, and after the gun is fired. The third group includes the effects of the environment such as reverberation and attenuation of high frequencies.

This chapter is concerned mainly with the first and third group of sounds which will be discussed below. The second group consists of sounds such as the trigger action and bullet loading and in general has a much lower signal level rendering them practically inaudible sufficiently far away from the firearm location.

4.1 Muzzle Blast

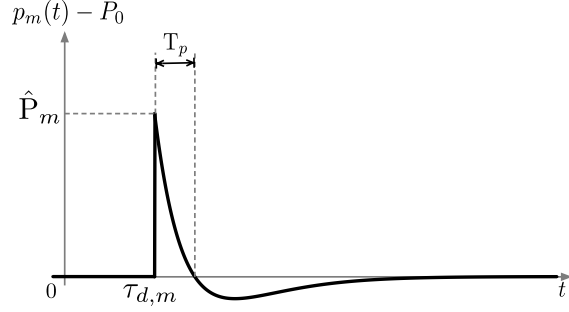
Muzzle blast is the explosive sound which occurs as the expanding gases rapidly escape from the muzzle. The associated pressure signal has a distinct shape which depends on many factors. A generic muzzle blast waveform is given in Fig. 3. It may be observed that the waveform has a very sharp onset immediately reaching a peak pressure level, \hat{P}_m followed by an exponentially decaying positive pressure phase and a negative pressure phase at which the pressure level falls below the ambient pressure, eventually reaching the ambient pressure level, $P_0 \approx 101 \text{ kPa}$.

The muzzle blast waveform is a typical example of a Friedlander wave [9] and can be represented using the following parametric form:

$$p_s(t) - P_0 = \hat{P}_m(r, \theta) \left(1 - \frac{t - \tau_{d,m}}{T_p} \right) e^{-(t - \tau_{d,m})/T_p}, \quad (1)$$

where P_0 is the ambient pressure, $\hat{P}_m(r, \theta)$ is the peak overpressure at a given distance r and a direction θ with respect to the boreline, $\tau_{d,m}$ is the delay of arrival of the blast wave at the observation position, and T_p is the duration of the posi-

Fig. 3 An ideal muzzle blast waveform showing the pressure due to the gunshot as a function of time. \hat{P} is the peak overpressure, $\tau_{d,m}$ is the delay, and T_p is the positive phase duration.



tive phase. There are several different studies investigating the properties of gunshot sounds. An adaptation of the model given by Fansler et al. in [7] is used in this chapter for calculating T_p and $\tau_{d,m}$.

The employed model is based on the concept of *scaling length* defined for the boreline direction of the gun as:

$$l_s \propto \sqrt{(dE/dt)/(cP_0)} \quad (2)$$

where dE/dt is the *energy deposition rate* due to the blast and is a function of the propellant properties, peak muzzle pressure, speed of the propellant at exit, the Mach number at projectile ejection, M_e ¹, and the bore area of the barrel. These parameters are typically provided by weapon and ammunition manufacturers.

The directional properties of the muzzle blast are modelled by weighting the scaling length with respect to the angle of the observation point from the boreline direction, θ :

$$l_{sn}(\theta) = \left(\mu \cos \theta + \sqrt{1 - \mu^2 \sin^2 \theta} \right) l_s \quad (3)$$

where μ is the *momentum index* defined as the ratio of the sound pressure at the front and at the rear of the firing position.

The peak overpressure was defined using data obtained from actual measurements as a function of observation distance, r , and the weighted scaling length, l_{sn} , such that:

$$\hat{P}_m(r, \theta) \approx 0.89 \frac{l_{sn}}{r} + 1.61 \left(\frac{l_{sn}}{r} \right)^2 \quad (4)$$

The time of arrival, $\tau_{s,m}$, and the duration of the positive phase portion of the muzzle blast, T_p define the general behaviour of the wave and are modelled as functions of scaled distance, $r_s = r/l_s$, and the weighted scaling length, l_{sn} :

$$\tau_{d,m} = \frac{l_{sn}}{c} [X(r_s) - 0.52 \ln(2X(r_s) + 2r_s) - 0.56] \quad (5)$$

¹ Mach number is the ratio of the projectile speed to the speed of sound in air under the same atmospheric conditions, i.e. $M = v_p/c$

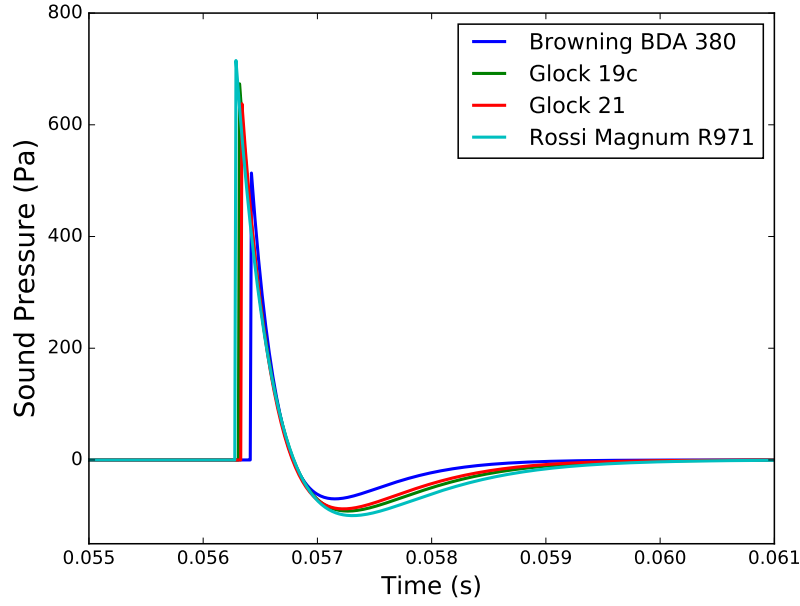


Fig. 4 Muzzle blast signal for different pistols and ammunition obtained using the described model for a distance of 20 m away from the firing position.

where

$$X(r_s) = \sqrt{r_s^2 + 1.04r_s + 1.88},$$

The model in [7] defines the positive phase duration differently for $r_s < 50$ and $r_s \geq 50$ such that:

$$T_p = \begin{cases} \frac{r}{c} - \tau_{d,m} + G, & r_s < 50 \\ \frac{l_{sm}}{c} \left[2.99 \sqrt{\ln(33119r_s)} - 8.534 + G \right], & r_s \geq 50 \end{cases} \quad (6)$$

where G is a correction parameter obtained from measurements.

Fig. 4 shows the muzzle blast signals calculated for different pistols used in this chapter. It may be observed that while the shapes of the signals are similar, the peak overpressure and positive phase durations are different.

4.2 Shock Wave

Projectiles that travel faster than sound generate pressure waves that have very sharp onsets and offsets. These waveforms also have strong positive and negative phase

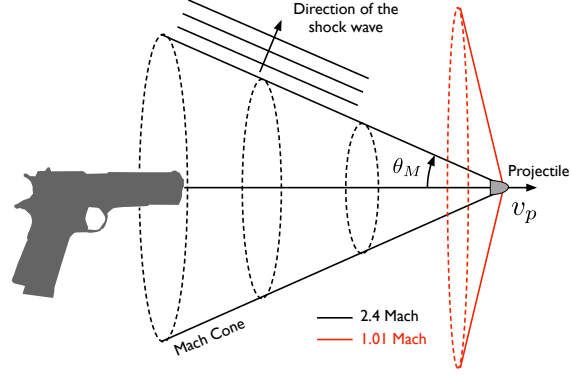


Fig. 5 Mach cone due to a projectile travelling at supersonic speeds. Two Mach cones for projectile speeds of 2.4 and 1.01 Mach are shown (adapted from [17]).

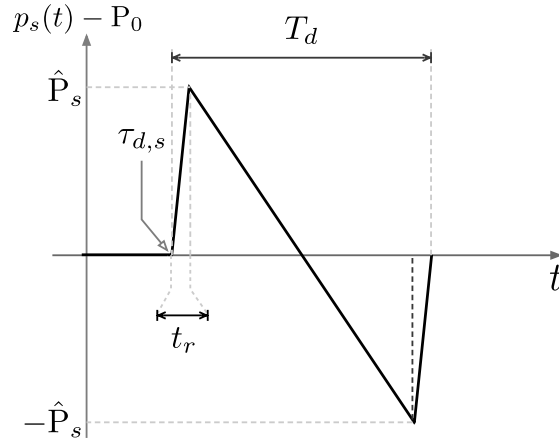


Fig. 6 Waveform of an ideal sonic boom showing the pressure due to a supersonic gunshot as a function of time. \hat{P}_s is the peak overpressure, $\tau_{d,s}$ is the time delay, t_r is the rise time, and T_d is the total duration of the N-wave.

parts and a distinctive shape. They are called an *N-waves* and the accompanying sound is called a *sonic boom*.

The production of the shock wave and its propagation characteristics is above all a function of the projectile speed. For projectiles travelling at speeds higher than the speed of sound (i.e. supersonic speeds), the shock wave will travel outwards on the conical surface trailing the projectile. This is known as the *Mach cone*. Half of the opening angle of the cone is called the Mach angle, θ_M and depends on the Mach number M_e such that:

$$\theta_M = \sin^{-1} M_e^{-1}. \quad (7)$$

Mach cone and mach angle are shown in Fig. 5. It should be noted in order for the shock wave to be observed, the observation point needs to be outside the Mach cone. For example, the person firing the gun will not hear it.

Fig. 6 shows the ideal waveform due to an N-wave. The wave is symmetric about $\tau_{d,s} + T_d/2$ so it can be defined by four distinct parameters: i) rise time, t_r , ii) total duration of the wave, T_d , iii) peak overpressure, \hat{P}_s , and iv) the time delay of ar-

rival, $\tau_{d,s}$. The time delay of arrival will be calculated in the next section. The other parameters are described below [31, 25].

The peak overpressure is the maximum pressure level that the shock wave will achieve:

$$\hat{P}_s = 0.53P_0 \frac{(M_e^2 - 1)^{1/8} d}{L_m^{3/8} l^{1/4}} \quad (8)$$

where d is the diameter of the projectile (in m), l is the length of the projectile (in m), and L_m is the miss distance (i.e. the geometric distance of the observation point to the trajectory of the projectile).

The total duration of the N-wave is given as:

$$T_d = \frac{1.82M_e L_m^{1/4} d}{c(M_e^2 - 1)^{3/8} l^{1/4}} \quad (9)$$

and the rise time is given as:

$$t_r = \frac{\lambda P_0}{c \hat{P}_s}, \quad (10)$$

where λ is the molecular mean free path. For air under normal conditions $\lambda \approx 6.8 \times 10^{-8}$ m. Typically, rise time is very short in the microseconds range, making it necessary to record or synthesise the gunshot sounds using a high sampling rate.

It is then possible to express the N-wave as a piecewise linear function given as:

$$p_s(t) = \begin{cases} 0, & t \leq \tau_{d,s} \quad \text{and} \quad t > \tau_{d,s} + T_d \\ \hat{P}_s \frac{t - \tau_{d,s}}{t_r}, & \tau_{d,s} < t \leq \tau_{d,s} + t_r \\ \hat{P}_s \left(1 - 2 \frac{t - \tau_{d,s} - t_r}{T_d - 2t_r} \right), & \tau_{d,s} + t_r < t \leq \tau_{d,s} + T_d - t_r \\ \hat{P}_s \left(\frac{t - \tau_{d,s} - T_d + t_r}{t_r} - 1 \right), & \tau_{d,s} + T_d - t_r < t \leq \tau_{d,s} + T_d. \end{cases} \quad (11)$$

This parametric form also allows a straightforward way to obtain the pressure value at any point in time, simplifying the sound synthesis process.

Fig. 7 shows the shock wave signals obtained using the model described above. The effect of projectile diameter on the shape of the shock wave is shown for three different projectiles with the same velocity of 388 m/s for demonstration purposes. It may be observed that the major effects are on the peak overpressure and duration of the shock wave and the strength of the shock wave increases with the projectile diameter.

4.3 Geometry of Propagation

The gunshot sound signal under anechoic conditions can be expressed as a combination of the sound signal due to the muzzle blast and the shock wave, when the latter is present. The geometry of propagation for these two components are different.

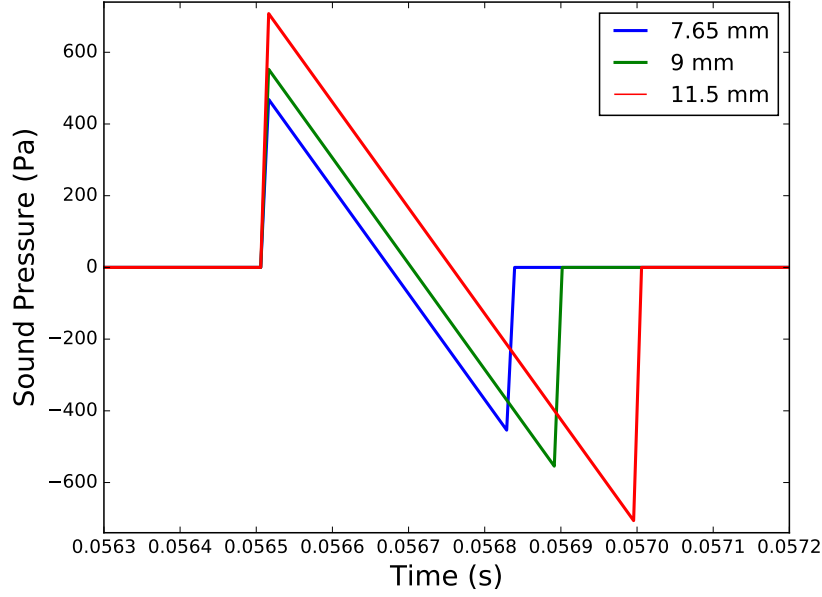


Fig. 7 Shock wave signals calculated using the described model for different projectile diameters calculated at a distance of 20 m for a theoretical gun with a barrel length of 10 cm, projectile length of 1 cm, exit pressure of 2000 kg/cm² and an exit velocity of 380 m/s for different projectile diameters.

The geometry of the gunshot sound propagation scenario is shown in Fig.8. The two components will travel for different durations and arrive at the microphone at different incidence angles. The time delay of arrival of the muzzle blast was discussed previously. The time delay of arrival for the shock wave can be calculated as:

$$\tau_{d,s} = (\langle \mathbf{x}_m - \mathbf{x}_g, \mathbf{n}_t \rangle + L_m \cos \theta_m) / c \quad (12)$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of two vectors, x_m is the microphone position, x_g is the muzzle position, \mathbf{n}_t is the unit vector in the direction of the trajectory, θ_m is the Mach angle, and L_m is the miss distance.

4.4 Effects of the Environment

In the absence of atmospheric effects such as wind or temperature gradients, three main environmental effects on gunshot sounds can be considered. The components of the gunshot sound will be reflected, their high frequency components will be

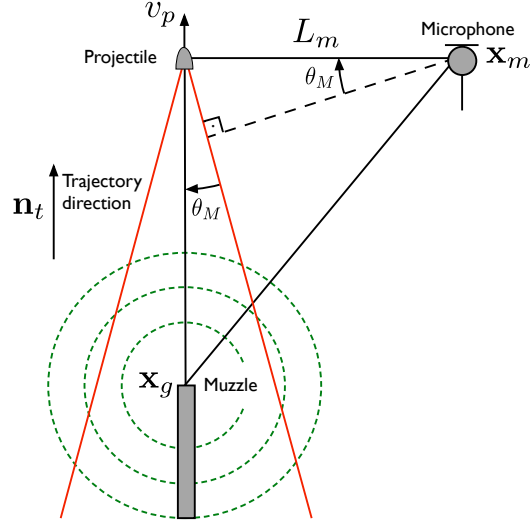


Fig. 8 The geometry (top view) of the gunshot sound propagation scenario showing the muzzle blast and the shock wave components.

attenuated at greater distances, and in enclosures such as rooms, gunshot sound will be reverberated.

In an open field, the sound recorded at the observer position will include a reflection incident from the ground in addition to the original gunshot sound. It should be noted that as with the direct path from the gun to the observation point, the ground reflection paths will also be different for the muzzle blast and the shock wave. The delays of the two reflections are given as:

$$\tau_{mb,g} = \|\mathbf{x}_g - 2h_g\mathbf{u}_z - \mathbf{x}_m\|/c, \quad (13)$$

$$\tau_{sb,g} = \|\mathbf{x}_{sb} - 2h_{sb}\mathbf{u}_z - \mathbf{x}_m\|/c, \quad (14)$$

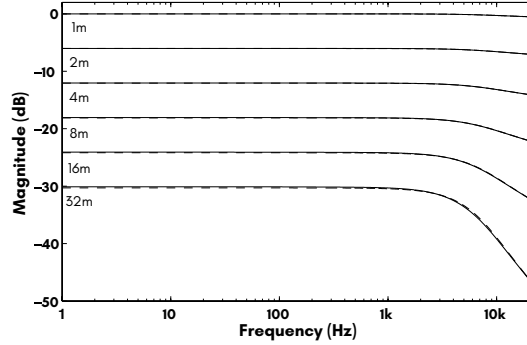
where $\mathbf{x}_{sb} = \mathbf{x}_g + L_m(\cot\theta_{m,g} - \tan\theta_M)\mathbf{n}_t$ is the shock wave source position, $\mathbf{u}_z = [0 \ 0 \ 1]$ and h_g and h_{sb} are the height of the muzzle and the height of the shock wave source position from the ground.

The effect imposed by the ground on the sound wave depends on the surface properties as different ground surfaces (such as grass, concrete, or gravel) will have different absorption characteristics. Most surfaces, however, will absorb more of the high frequency content, smoothing the original gunshot sound. Absorption characteristics tabulated elsewhere [30] can be used to design digital filters [11] to simulate acoustical absorption.

Two other important environmental effects are attenuation and air absorption which cause the overall sound level to decrease with distance and the higher frequency components to be attenuated, respectively.

Pressure due to a spherical wave decreases with the radial distance. This is known as the *spherical spreading law* or the $\frac{1}{r}$ -law. The employed firearm acoustics model

Fig. 9 Air attenuation and frequency-dependent air absorption. The solid curves show the physical model and the broken curves show the filters designed to emulate the given distances. It may be observed that larger distances correspond to higher absorption and lower levels for high frequencies.



already takes this attenuation into account. It is thus unnecessary to model these losses separately.

Other than attenuation, air absorbs a significant amount of high frequency energy from a sound wave. The main reason for this is the physical process called *thermal relaxation* [15]. A closed form expression for frequency dependent air absorption as a function of humidity is given in [12] and it was shown [11] that this functional form can be used for designing simple filters for simulating air absorption. Fig. 9 shows the air absorption and attenuation for different source distances. The magnitude responses of 4th-order infinite impulse response (IIR) filters are also shown with broken lines. It may be observed that air absorption is mostly effective at higher frequencies and at larger source distances. In the context of gunshot sounds, this corresponds to a significant reduction of higher frequencies at large distances.

When a weapon is fired indoors, the acoustical effect of the enclosure will also have a profound effect on the generated sound. Environments with many reflecting and scattering objects such as rooms, streets, or forests are multipath environments. In such environments, first-order reflections of sound sources are observed together with higher-order reflections and reverberation. In order to simulate the effects of different environments, existing artificial reverberators can be used. While one of the most commonly used reverberation algorithms is the Jot's reverberator [13], more recent algorithms such as scattering delay networks (SDN)[5] and *treeverb* [23] are more suitable as they provide means to simulate position-related aspects of the gunshots more easily and provide environment-specific reverberation. Readers interested in artificial reverberator algorithms are referred to two recent review articles [29, 28].

4.5 Synthesis Model

The synthesis model used in this chapter uses two sound synthesis modules, one for the muzzle blast and one for the shock wave. These modules use parameters from ballistic databases that can be bundled with a game, simulation, or a VR application.

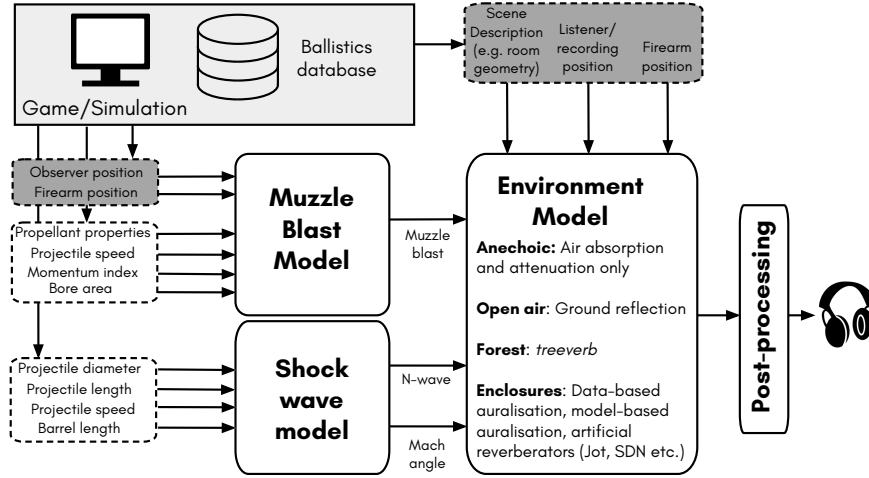


Fig. 10 The parametric gunshot sound synthesis model described in this chapter.

The positions of the listener (player) and the firearm can be fed to the model from the game at runtime for the muzzle blast and shock wave calculations. The simulation of the acoustics of the environment is handled in a second stage. Scene description including the listener and firearm positions can also be obtained at runtime from the game. Fig. 10 shows the synthesis model.

It should be noted that just a small number of parameters is sufficient to synthesise the sounds of a wide variety of firearms. Since in an interactive environment positions of the shooter and the observation position are always changing some parameters like the delays of each component of the gunshot sound need to be refreshed at runtime. However, since gunshot events are likely to be sparse, and also that the calculation of the mentioned parameters is trivial, the computational cost would be negligible. The described algorithm, as with most other PCG methods, will also reduce the disk access overhead significantly. The described algorithm (except the environment model) is provided as an open-source Python module at <https://github.com/metu-sparg/pygunshot>.

5 Evaluation

A subjective evaluation was carried out to assess the realism of synthetic gunshot sounds in comparison with real recordings.

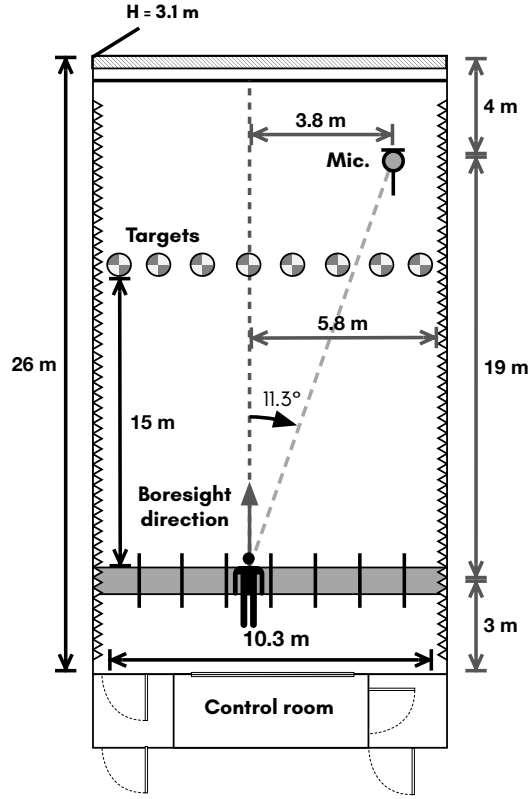


Fig. 11 The recording setup used for the real sound samples described in this chapter. The shooting position is indicated by the human figure. Position of the microphone is also indicated.

5.1 Measurements

A set of measurements were carried out in the indoor pistol shooting range located at the head offices of Mechanical and Chemical Industry Corporation (MKEK) which is a major firearm and ammunition producer based in Ankara, Turkey. Three repeated recordings from four different pistols were made². Table 1 gives details of the pistols and the ammunition used. It may be observed that two of the recorded pistols (Glock 19c and Rossi Magnum R971) nominally generate shock waves along with the muzzle blast.

The shooting range has a shoebox geometry with the dimensions $W = 10.3$ m, $L = 26$ m and $H = 3.1$ m. The room is acoustically treated with porous rubber acoustic tiles on the side walls, and with tilted steel backstops covered by a thick rubber curtain at the end of the range. The reverberation time averaged across frequencies is $T_{30} \approx 0.68$ s. Fig 11 shows the recording geometry.

² The recorded and synthetic gunshot sounds are made available at <https://figshare.com/s/89d90977887b0bb8f54c>

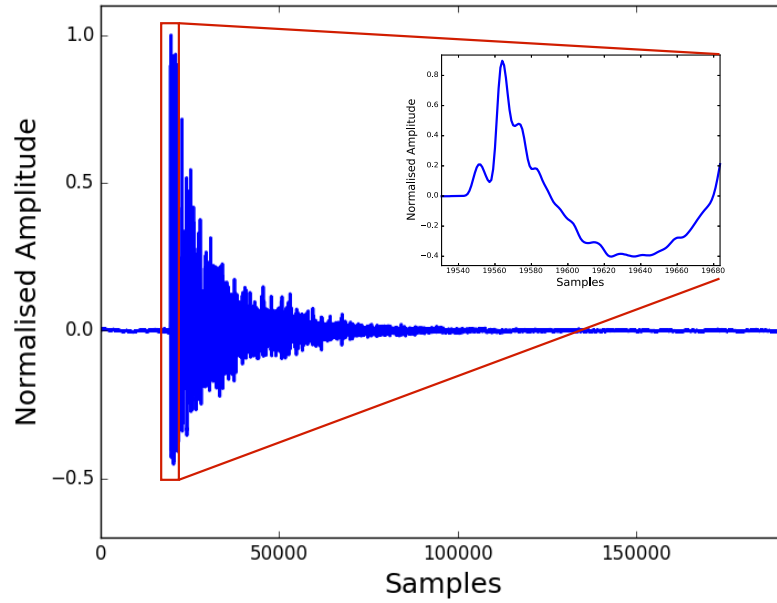


Fig. 12 Pressure signature of the Browning BDA-380 recorded in the indoor firing range. While the larger figure mainly shows the reverberation response of the room, the inset clearly shows that the direct path component has the Friedlander wave shape.

Table 1 Types and properties of guns and ammunition used in the recordings.

Type	Barrel Length	Cartridge	Muzzle Velocity	Muzzle Pressure
Browning BDA-380	97 mm	.380 ACP	300 m/s	1511 kg/cm ²
Glock 19C (with compensator)	102 mm	9 × 19 mm	361 m/s	2460 kg/cm ²
Glock 21	117 mm	.45 ACP	291 m/s	1476 kg/cm ²
Rossi Magnum R971	101.6 mm	.357 Mag	388 m/s	2817 kg/cm ²

An omnidirectional microphone (Alctron M6) with a nominal maximum SPL value of 149 dB SPL connected to the microphone preamps of a MOTU 896 Mk3 sound card was used to record the gunshots. In order to increase its dynamic range to prevent clipping, the microphone was muffled with a tight latex cover. This resulted in a 15 dB improvement in the dynamic range allowing higher levels to be recorded without clipping. The microphone was positioned 19 m away from the firing line and 3.8 m away from the target at a height of 1.56 m. The sampling rate used in the recordings was 192 kHz.

An example sound signal from the recording of the Browning BDA-380 pistol is shown in Fig 12. The Friedlander wave is clearly visible in the inset figure.

5.2 Stimuli

The sounds used in the subjective experiments were the recordings made in the indoor firing range as described above and samples synthesised using the algorithm described in this chapter.

The synthesis algorithm by itself does not incorporate the reverberation due to the recording environment. In order to add reverberation, an SDN-type reverberator was used. SDN allows the selection of the room size and absorption characteristics as well as the source and microphone positions. It was made sure that the simulated reverberation corresponded well to the acoustics of the firing range.

Three additional but trivial steps of processing were also applied on the synthetically generated sounds: Despite precautions the response of the microphone was observed to saturate for negative sound pressure values. The normalised synthetic signals were clipped between -0.8 and 1 to emulate this effect. The lack of low-frequency background noise in the synthetic gunshots was compensated for by adding Brownian noise. Equalisation was carried out on the synthetic signals to simulate the resonant properties of the physical gun which also affects how it sounds.

5.3 Methodology and Subjects

A pilot AB test with 10 subjects indicated that the differences between real and synthetic gunshot sounds are clearly discriminable. It was therefore deemed appropriate to assess the perceived realism of the synthetic gunshot sounds in comparison with real recordings.

The listening task involved presenting a hidden reference (i.e. the real recording), a hidden anchor (i.e. a processed version of the reference low-pass filtered at 1.5 kHz) and the stimulus under test (i.e. the synthetic gunshot sound generated as described above).

The subjects' task in each trial was to listen to all stimuli and rate them for their realism using three sliders on a continuous scale from 0 (Very unrealistic) to 100 (Very realistic). Intermediate markers were also provided: 25 (Somewhat unrealistic), 50 (Neither realistic nor unrealistic), and 75 (Somewhat realistic). The order of different tested gun types as well as the presentation order in each trial were randomised. Subjects were instructed to use the headphones that they typically use while playing games. The collected data is normalised and reported in the range 0 to 1 . Fig. 13 shows the test interface employed in the listening test.

The evaluation was carried out online using the Web Audio Evaluation Toolbox [19] over a paid experiment crowdsourcing service. A total of 60 subjects (42 male and 18 female) between the ages of 18 and 47 ($M = 27.7$, $SD = 6.4$) with self-reported normal hearing participated in the test. The online screening process allowed the selection of subjects who regularly play video games. The selected subjects self-reported playing video games 17 hours per week on average.

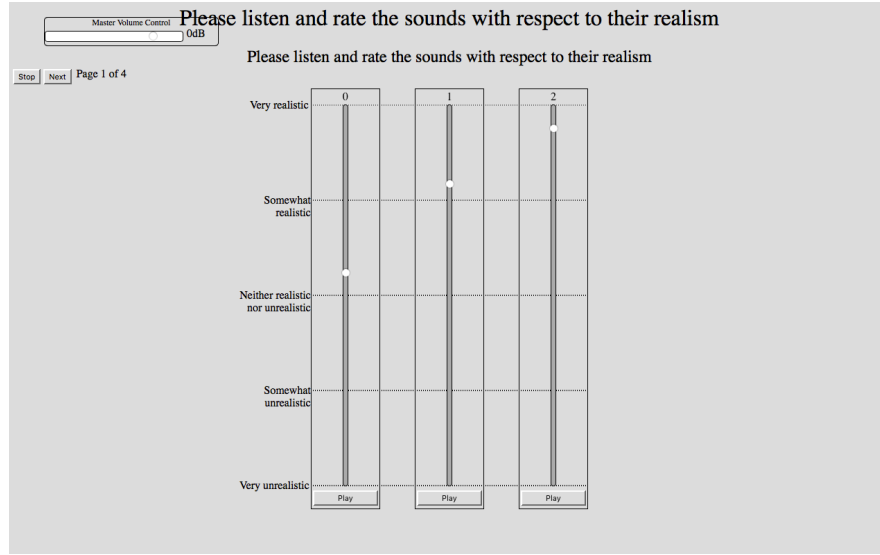


Fig. 13 The interface used in the listening experiment

5.4 Results

It was noticed during the analysis that one of the subjects reported only two out of the presented four gun types. This subject was removed from the analysis. This meant that there were $N = 59$ dependent samples for each comparison.

The marginal means for the real recordings (REAL), synthetic sounds (SYNTH), and the anchor (ANC) are 0.6372 (s.d. 0.2452), 0.4793 (s.d. 0.2495), and 0.3428 (s.d. 0.2646), respectively. A paired t-test was carried out to compare the means of REAL, SYNTH, and ANC. The test indicated that the differences between the means of the three different cases are statistically significant at $\alpha = 0.05$ level with $t(235) = 7.466$, $p < 0.001$ (REAL-SYNTH), $t(235) = -11.929$, $p < 0.001$ (ANC-REAL), and $t(235) = -5.681$, $p < 0.001$ (ANC-SYNTH). This shows that real recordings were rated higher on average than synthetic sounds, which in turn were rated higher than the anchor.

A second set of analyses were carried out to find out whether this finding holds also for different gun types. The analyses indicated that except for the case of Browning BDA 380 (REAL-SYNTH) and Rossi Magnum R971 (SYNTH-ANC), differences between REAL, SYNTH, and ANC were statistically significant. In the case of Browning BDA 380, the difference between means for REAL ($M = 0.6076$, $SD = 0.2648$) and SYNTH ($M = 0.5281$, $SD = 0.2485$) was not-significant $t(58) = 1.964$, $p = 0.054$. In the case of Rossi Magnum R971 the difference between means

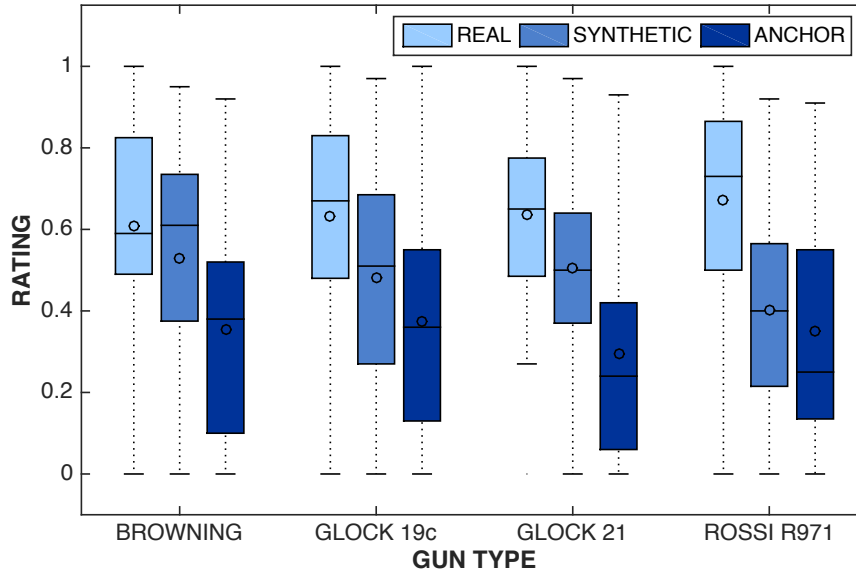


Fig. 14 Boxplots showing the results of the listening test.

for SYNTH ($M = 0.4019$, $SD = 0.2552$) and ANC ($M = 0.3490$, $SD = 0.2603$) was not-significant $t(58) = 1.089$, $p = 0.281$.

A third set of analyses was carried out to find out whether synthetic gunshot sounds with a shock-wave component (Glock 19c and Rossi Magnum R971) were rated differently than those without (Browning BDA 380 and Glock 21). The results of an independent samples t-test indicate that the mean responses between these groups for were not statistically significantly different at $\alpha = 0.05$ level for both REAL and ANC. However, SYNTH was rated differently and sounds with a shock wave component ($M = 0.4419$, $SD = 0.2551$) were rated lower than those without ($M = 0.5168$, $SD = 0.2390$) with $t(234) = 2.328$, $p = 0.021$. The results obtained in the listening test are summarised in Fig. 14.

6 Discussion and Conclusions

The results of the listening test are promising in the sense that while the reference stimuli (REAL) were rated higher than the test stimuli (SYNTH), the difference was not very high. This is an indicator that the described procedural synthesis algorithm can generate usable weapon sounds and can eliminate the need for costly and time consuming recordings. Another advantage is the ability to separate the source from the environment making it possible to use and re-use content in different virtual environments.

There were some surprising results from listening test: The reference signals were rated on average as being even less than “Somewhat realistic”. Since these sounds are actual recordings, the expectation was that the average should have been closer to “Very realistic”. This may indicate that the subjects have different expectations of real gunshot sounds.

The synthesis model has certain limitations, however. Ballistic information is neither readily, nor comprehensively available from ammunitions manufacturers. Similarly, different modifications to the guns (such as the Glock 19c in this study which has a compensator for reducing recoil) may make it harder to tune the model to obtain which can match the real recordings exactly. It is also impossible to measure the resonance characteristics of the gun with any accuracy.

Evaluation of procedurally generated gunshot sounds also have some limitations. A real gunshot sound can easily exceed 150 dB SPL at reasonable recording distances. Reproduction of such a sound level over typical consumer-grade headphones or loudspeakers is both technically impossible and ethically unacceptable as this could potentially damage hearing. Another important limitation is the fact that the two components of the gunshot have very short durations and they cannot reliably be tested without adding reverberation which may have a confounding effect on the results.

Despite these limitations, the proposed method is a good starting point for more elaborate algorithms that can also take into account additional components of gunshot sounds such as the mechanical sounds.

Acknowledgements The author wishes to thank MKEK shooting range manager, Mr. Tarık Anıl, as well as indoor shooting range personnel for their help and MKEK for allowing the usage of their facilities for the gunshot recordings used in the subjective experiments. The author also wishes to thank Mr. James D. Johnston for the advice on the recording setup.

References

1. BF4 audio files. Symtic Game Science Forums (2014). URL <http://forum.symthic.com/battlefield-4-general-discussion/6591-bf4-audio-files/>
2. Beck, S.D., Nakasone, H., Marr, K.W.: Variations in recorded acoustic gunshot waveforms generated by small firearms. *J. Acoust. Soc. Am.* **129**(4), 1748–1759 (2011)
3. Carlucci, D.E., Jacobson, S.S.: *Ballistics: theory and design of guns and ammunition*. CRC Press (2013)
4. Cascone, K., Petkevich, D.T., Scandalis, G.P., Stilson, T.S., Taylor, K.F., Van Duyne, S.A.: Apparatus and methods for synthesis of internal combustion engine vehicle sounds (2005). US Patent 6,959,094
5. De Sena, E., Hacıhabiboğlu, H., Cvetković, Z., Smith, J.O.: Efficient synthesis of room acoustics via scattering delay networks. *IEEE/ACM Trans. on Audio, Speech, and Language Process.* **23**(9), 1478–1492 (2015)
6. Electronic Arts: Battlefield (2016). URL <https://www.battlefield.com>
7. Fansler, K.S., Thompson, W.P., Carnahan, J.S., Patton, B.J.: A parametric investigation of muzzle blast. Tech. Rep. ARL-TR-227, US Army Research Laboratory, Arlington, VA, USA (1993)

8. Farnell, A.: *Designing sound*. MIT Press (2010)
9. Friedlander, F.G.: The diffraction of sound pulses. I. Diffraction by a semi-infinite plane. *Proc. Royal Soc. London A: Mathematical, Physical and Engineering Sciences* **186**(1006), 322–344 (1946)
10. Gearbox Software: *Borderlands* (2009). URL <http://borderlandsthegame.com/>
11. Huopaniemi, J., Savioja, L., Karjalainen, M.: Modeling of reflections and air absorption in acoustical spaces: a digital filter design approach. In: *Proc. IEEE Workshop on Appl. of Signal Process. to Audio and Acoust. (WASPAA'97)* (1997)
12. International Organization for Standardization: *Acoustics-Attenuation of sound during propagation outdoors—Part 1: Calculation of the absorption of sound by the atmosphere (ISO 9613-1:1993)* (1993)
13. Jot, J.M., Chaigne, A.: Digital delay networks for designing artificial reverberators. In: *Proc. 90th Convention of the Audio Eng. Soc.*, p. Preprint # 3030. Paris, France (1991)
14. Kahrs, M., Avanzini, F.: Computer synthesis of bird songs and calls. In: *Proc. Conf. on Digital Audio Effects (DAFx-01)*. Limerick, Ireland (2001)
15. Kuttruff, H.: *Room acoustics*. Crc Press (2009)
16. Maher, R.C.: Audio forensic examination. *IEEE Signal Process. Mag.* **26**(2), 84–94 (2009)
17. Maher, R.C., Shaw, S.R.: Deciphering gunshot recordings. In: *Proc. 33rd Int. Conf. of the Audio Eng. Soc.: Audio Forensics-Theory and Practice*, pp. P–2. Audio Engineering Society, Denver, Co, USA (2008)
18. Maher, R.C., Shaw, S.R.: Directional aspects of forensic gunshot recordings. In: *Proc. 39th Int. Conf. of the Audio Eng. Soc.: Audio Forensics: Practices and Challenges*, 4-2. Audio Engineering Society, Hillerod, Denmark (2010)
19. Nicholas Jillings Brecht De Man, D.M., Reiss, J.D.: Web audio evaluation tool: A browser-based listening test environment. In: *Proc. 12th Sound and Music Computing Conference*. Maynooth, Ireland (2015)
20. Nordahl, R., Turchet, L., Serafin, S.: Sound synthesis and evaluation of interactive footsteps and environmental sounds rendering for virtual reality applications. *IEEE Trans. Visualization and Computer Graphics* **17**(9), 1234–1244 (2011)
21. Rasmussen, P., Flamme, G., Stewart, M., Meinke, D., Lankford, J.: Measuring recreational firearm noise. *Sound and Vibration Mag.* **43**(8), 14–18 (2009)
22. Schröder, M.: Expressive speech synthesis: Past, present, and possible futures. In: J. Tao, T. Tan (eds.) *Affective Information Processing*, pp. 111–126. Springer London (2009)
23. Spratt, K., Abel, J.S.: A digital reverberator modeled after the scattering of acoustic waves by trees in a forest. In: *Proc. 125th Convention of the Audio Eng. Soc.*, p. Preprint # 7650. Audio Engineering Society, San Francisco, CA, USA (2008)
24. Statista Inc.: *Genre breakdown of video game sales in the united states in 2015* (2016). URL <https://www.statista.com/statistics/189592/breakdown-of-us-video-game-sales-2009-by-genre/>
25. Stoughton, R.: Measurements of small-caliber ballistic shock waves in air. *J. Acoust. Soc. Am.* **102**(2), 781–787 (1997)
26. Team Stack Studios: *Stack gun heroes* (2015). URL <http://www.stackgunheroes.com>
27. US Department of Defense: *America's army*. URL <https://www.americasarmy.com>
28. Välimäki, V., Parker, J., Savioja, L., Smith, J.O., Abel, J.: More than 50 years of artificial reverberation. In: *Proc. 60th Int. Conf. of Audio Eng. Soc.: DREAMS (Dereverberation and Reverberation of Audio, Music, and Speech)*, pp. K–1. Leuven, Belgium (2016)
29. Välimäki, V., Parker, J.D., Savioja, L., Smith, J.O., Abel, J.S.: Fifty years of artificial reverberation. *IEEE Trans. on Audio, Speech and Language Process.* **20**(5), 1421–1448 (2012)
30. Vorländer, M.: *Auralization: Fundamentals of acoustics, modelling, simulation, algorithms and acoustic virtual reality*. Springer Science & Business Media (2007)
31. Whitham, G.B.: The flow pattern of a supersonic projectile. *Comm. on Pure and Appl. Math.* **5**(3), 301–348 (1952)