View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Queen Mary Research Online

Sound Effect Synthesis

David Moffat, Rod Selfridge and Joshua D. Reiss

May 13, 2018

Abstract

Sound effects add vital cues for listeners and viewers across a range of media content. The development of more realistic, interactive, immersive sound effects is an exciting and growing area of research. In this chapter we provide a comprehensive overview of sound effect synthesis, including definitions, classifications, techniques and examples. The contextual reason and importance of sound effects are presented, including how these are sub-categorised as well as the importance of Foley artists. An in-depth review of the wide range of sound effect synthesis techniques is given, highlighting the strengths and weaknesses of different synthesis methods. Evaluation techniques are described along with reasons why evaluation is essential when deciding on which sound effect synthesis method to use and how research will develop in the future. We also looked at the definition of procedural audio, drawing attention to why this is an active development area for games and virtual reality environments. An example design process is given for a physically inspired sound synthesis model which can be integrated as a procedural audio effects.

1 Introduction

Sound effects are commonly defined as non-musical, non-speech sounds used in some artificial context, such as theatre, TV, film, video game or virtual reality. The purpose of a sound effect is typically to provide a diegetic context of some event or action, that is a sound that exists within the narrative of the story line. A 1931 BBC White Paper proposed that there were 6 types of sound effects (BBC, 1931)

Realistic, confirmatory effect The convincing sound of an object that can be seen, to directly tie into the story, eg. the sound of a gunshot when we see a gun being fired

Realistic, evocative effect A convincing sound within the landscape, that cannot be directly seen eg. in a forest, a bird tweeting off screen

Symbolic, evocative effect Sounds that don't actually exist within the narrative, designed to create an emotion within the listener, eg. a swelling sound to build suspense

Conventionalised effect A sound that though not entirely realistic, is perceived as realistic, due to overuse and hyper-realism eg. the ricochet after a gun shot in a western film

Impressionistic effect creating a general feeling or indication of an occurrence without an exact realistic example eg. a cartoon punch sound

Music as an effect producing a sound effect through some musical means, eg. chimes to represent a transformation

From this, sound effects can often be the linchpin of a sound scene, and different sounds and styles will vary drastically dependent on the style and design of the medium, among other factors.

Sound synthesis is the technique of generating sound through artificial means, either in analogue or digital or a combination of the two. Synthesis is typically performed for one of three reasons;

- facilitate some interaction or control of a sound, whether for a performance or direct parameter driven control of a sound, e.g. Heinrichs *et al.* (2014); Wilkinson *et al.* (2016)
- facilitate a sound designer searching for a suitable sound within a synthesis space rather than through a sound effect library, e.g. Hendry and Reiss (2010)
- to create something that does not exist, such as creating artificial sci-fi sounds or repairing damaged sound files, e.g. Puronas (2014).

Public demand is increasing for instantaneous and realistic interactions with machines, particularly in a gaming context. Farnell (2007) defines Procedural Audio (PA) as “non-linear, often synthetic sound, created in real time according to a set of programmatic rules and live input”. As such PA can be viewed as a subset of sound synthesis, where all sounds are produced in real time, with a particular focus on synthesis control and interaction. PA is fundamental to improving human perception of human computer interactions from an audible perspective, but there are still many unanswered questions in this field (Fournel, 2010). Böttcher and Serafin (2009) demonstrated subjectively that in an interactive gameplay environment, 71% of users found synthesis methods more entertaining than audio sampling. Users rated synthesised sound as higher quality, more realistic and preferable. From this, it is clear that user interaction is a vital aspect of sound synthesis.

Foley sound was created in the 1920’s by Jack Foley. The premise is that a Foley artist or ‘performer’ can perform a particular sound using any objects that may create the idea of the sound, rather than just use a recording of a real sound. The story goes that someone was looking for a bunch of chains to rattle to create a prison scene, and Jack just pulled out a bunch of keys, rattled them in front of a microphone recording the sound. When they listened back, they were happy with the results and the concept of Foley sound was born. The emphasis on creating a ‘larger than life’ sound was one of the key founding aspects of Foley work. A sound does not need to be real, it just needs to be convincing. This has resulted in the idea of ‘hyper-realism’, which is commonplace in much of Hollywood sound design (Puronas, 2014). Hyper-realism is the idea that a sound must be bigger, more impressive and ‘more real’ than the real world sound, so as to create a level of excitement or tension (Mengual *et al.*, 2016). This is particularly common in many TV and film explosion and gunshot sounds, where a real world recording of a gunshot is considered too boring and mundane compared to the artificial gunshot, which is often some combination of much bigger sounds, such as a gunshot, an explosion, a car crash, a lion roar and a building collapse. Foley attempts to provide a similar idea with some performance sounds, where each action or idea is significantly over-performed and every action is made larger than the real world case. Foley grew into an entire field of work, and professional ‘Foley artists’ can still be found worldwide. Foley sound became prominent since it allowed a sound designer to perform, act or create the desired sound, and easily synchronize it with the action. The level of control that a Foley artist had over the sound was greater than ever before.

Much in the same way that Foley sound allowed for control, interaction and performance of a sound, sound synthesis can allow for control over digital sounds. Previously, the only way to digitise a sound was to record it. Now we can model a sound and control its parameters in real time. This creates a much more naturally occurring sound, as controls can be derived directly from the physical parameters, and thus the expectation of the listener is satisfied, when every small detail and interaction they have produces a predictably different sound. As such, in many ways Sound Synthesis can be considered digital Foley

The key advantages of a synthesised sound effect over a recording is the ability to control and interact with the sound. This interaction creates a feeling of a realistic world Heinrichs and McPherson (2014). Immersion is a key goal in game design. A player feels more immersed in a game, if they feel like they are actually situated in the game environment. Immersive sound can be created either through the use of 3D sound, or by creating realistic interactions with sonic objects. Creating an immersive sound, is an important aspect, as it will draw a user into the virtual environment, and make them feel more included as part of the game, rather than simply watching the game through a window. Realistic sonic feedback is a vital part of producing a believable and consistent immersive world.

2 Sound effect synthesis

There are many methods and techniques for synthesising different sound effects, and each one has varying advantages and disadvantages. There are almost as many sound synthesis classification methods, but the most prominent was produced by Smith (1991). Sound synthesis can generally be categorised into these categories:

2.1 Sample Based Synthesis

In Sample based synthesis, audio recordings are cut and spliced together to produce new or similar sounds. This is effective for pulse-train or granular sound textures, based on a given sound timbre.

The most common example of this is granular synthesis. Granular synthesis is the method if analysing a sound file or set of sound files and extracting sonic 'grains'. A sound grain is generally a small element or component of a sound, typically between 10-200ms in length. Once a set of sound grains have been extracted, they can then be reconstructed and played back with components of the sound modified, such as selecting a subset of grains for a different timbre, to changing the grain density or rate to change the pitched qualities of the sound.

2.2 Signal Modelling Synthesis

Signal Modelling Synthesis is the method where sounds are created based on some analysis of real world sounds, and then attempting to resynthesise the waveform sound, not the underlying physical system. The premise of signal modelling, is that through comparing and reproducing the actual sound components, we can extrapolate the control parameters and accurately model the synthesis system. The most common method of signal modelling synthesis is Spectral Modelling Synthesis (SMS) Serra and Smith (1990). SMS assumes that sounds can be synthesized as a summation of sine waves and a filtered noise. Spectral modelling is often performed by analysing the original audio file, selecting a series of sine waves to be used for resynthesis, and then creating some 'residual' noise shape, which can be summed together to produce the original sound (Amatriain *et al.*, 2002). SMS performs best on simple harmonic sounds. For less harmonic sounds, other methods such as nonnegative matrix factorisation Turner (2010) or latent force modeling Wilkinson *et al.* (2017)

2.3 Abstract Synthesis

Sounds are created from abstract methods and algorithms, typically to create entirely new sounds. A classic example of abstract synthesis is Frequency Modulation (FM) Synthesis (Chowning, 1973). FM Synthesis is a method derived from telecommunications. Two sine waves are be multiplied together to create a much richer sound. These sounds can be controlled in real time, as computation is low, to create a set of sounds that do not exist in the natural world. A lot of traditional video game sounds and 1980's keyboard sounds were based on FM synthesis.

2.4 Physical Modelling Synthesis

Sounds are generated based on modelling of the physics of the system that created the sound. The more physics is incorporated into the system, the better the model is considered to be, however the models often end up very computational and can take a long time to run. Despite the computational nature of these approaches, with GPU and accelerated computing, physical models are beginning to be capable of running in real time. As such, physical models are based on fundamental physical properties of a system and solving partial differential equations at step sample (Bilbao, 2009).

2.5 Synthesis Methods Conclusion

There are a range of different synthesis methods, that can produce a range of different sounds. From abstract synthesis techniques that are lightweight and can be implemented on old 80's hardware, to physical modelling techniques that require optimisation and GPU and even still, are only just able to operate in real time. There are a range of different synthesis methods and each one has its advantages and disadvantages. Misra and Cook (2009) performs a rigorous survey of

Sound Type	**Synthesis Method**
Sci-Fi / Technology Sounds	Abstract Synthesis
Environmental Sounds	Sample Based Model / Signal Models
Impact Sounds	Physical models / Signal Models
Voiced Sounds	Signal Models
Sound Textures / Soundscapes	Sample Based Models

Table 1: Recommendation of Synthesis Method for Each Sound Type

synthesis methods, and recommends different synthesis techniques for each type of sound to be produced. Abstract synthesis is great for producing artificial sounds, sounds of the 80's and some musical sounds. Signal modelling can produce excellent voiced sounds and environmental sounds. Physical models are great for impact or force driven sounds, such as the pluck of a string. Where as sound textures and environmental sounds are often best produced by sample based models. A summary of recommendations as to a method of synthesis that would work for each type of sound class can be found in Table 1.

3 Evaluation

The aims of sound synthesis are to produce realistic and controllable systems for artificially replicating real world sounds. Evaluation is vital, as it helps us understand both how well our synthesis method performs, and how we can improve our system. Without a rigorous evaluation method, we cannot understand if our synthesis method performs as required or where it fails. Evaluation of a sound synthesis system can take many different forms. Jaffe (1995) presented ten different methods for evaluation of synthesis techniques. There are many examples of these evaluation methods being employed in literature, including evaluation of controls and control parameters (Rocchesso *et al.*, 2003; Merer *et al.*, 2013; Selfridge *et al.*, 2017b), human perception of different timbre (Merer *et al.*, 2011; Aramaki *et al.*, 2012), sound identification (Ballas, 1993; McDermott and Simoncelli, 2011), sonic classification (Gabrielli *et al.*, 2011; Hoffman and Cook, 2006; Moffat *et al.*, 2017) and sonic realism (Moffat and Reiss, 2018; Selfridge *et al.*, 2018a, 2017c).

Evaluation methods can be broken down into one of two methods

3.1 Evaluation of Sonic Qualities

One of the most important aspects of evaluating a synthesis method is evaluating the sonic quality of the sound produced. Does the produced sound actually sound as intended? If you cannot create the sound you prefer, then no quantity of sound interaction will make a synthesis model effective. Generally, this evaluation needs to be performed with human participants, where recorded samples of a given sound can be compared to samples rendered from a synthesis method, and the two compared by users in a multi-stimulus perceptual evaluation experiment (Moffat and Reiss, 2018; Bech and Zacharov, 2007). This evaluation comparison method will evaluate synthesised sounds and compare them against recordings, in the same contextual environment. This method of evaluation can be applied to a range of different sounds Mengual *et al.* (2016); Selfridge *et al.* (2018a, 2017a,b,c,d).

It is important that similar sounds are compared, and that participants are asked suitable questions. Generally participants are asked to evaluate how real or how believable a given sound is. This is important as although participants may have a strong idea of what a sound is, this does not mean that their impression of a real sound is correct. It has often been the case that a participant will rate a synthetic sound as 'more realistic' than a real recording of a sound, especially in less common sounds. This is due to the hyper-realism effect discussed earlier. As people are generally expecting explosions and gunshots to be 'larger than life', when they hear a real recording vs a synthesised sound, the recording just seems flat and boring compared to a synthesised sound (Mengual *et al.*, 2016).

However, despite this, there is rarely effective perceptual evaluation of synthesis methods. Schwarz (2011) noted in a review of 94 published papers on sound texture synthesis that only 7 contained any perceptual evaluation of the synthesis method.

3.2 Evaluation of Control and Interaction

Evaluating the control and interaction of a synthesis engine is a vital aspect of understanding in which environment the sound can be used. Much in the same way a foley is the performance of 'analog' sounds, synthesis is the performance of digital sounds, and the control interaction is key. However, in most cases, the physical interaction that creates the sound will not be suitable for directly driving the individual synthesis parameters, and as such, some mapping layer for parameters and physical properties of a game will be required (Heinrichs *et al.*, 2014; Heinrichs and McPherson, 2014). There are numerous methods for evaluating these sonic interactions, and in many cases, the control evaluation has to be designed bespoke to the synthesis methods and parametric controls (Heinrichs *et al.*, 2014; Heinrichs and McPherson, 2014; Turchet *et al.*, 2016; Selfridge *et al.*, 2017b). User listening tests, where participants are able to interact with the synthesis engine, through some mapping layer, can be performed to evaluate a series of criteria. Key aspects of synthesis control systems to evaluate are

- **Intuitive** - How intuitive and interpretable are the controls, can a user easily find the exact sound they want.
- **Perceptible** - How much can someone perceive the impact each control makes, at all times, so as to understand what each control does.
- **Consistency** - Do the controls allow for consistent reproduction of sound, or is there some control hysteresis.
- **Reactiveness/Latency** - Do the controls immediately change the sound output, or is there a delay on control parameters, that impact the ease of usability. Typically 20ms of latency is acceptable, in most cases, so long as the latency is consistent Jack *et al.* (2016).

4 Example Design Process

A number of synthesis techniques have been identified and here we illustrate how to apply these principles to design our own sound effect. We looked at designing a sword sound effect, initially answering a number of questions:

- What synthesis technique shall we use to implement the effect?
- Are we going to design from scratch or based on samples?
- Do we want real-time operation?
- Are we going to use specialist hardware?
- What software will we implement the effect on?
- How do we want to control the effect?

For this example, we wanted our sound effect to be able to be used as part of a procedural audio implementation and to be able to capture elements of natural behaviour. This meant some sort of physical model was preferred. Such physical models generally involve synthesis from scratch, since they are based on the physics that produces the sound rather than analysis or manipulation from a sound sample. From the definition of procedural audio, real-time operation is key to enabling the effect to adapt to changing conditions.

Table 2: Table highlighting different synthesis methods for swing sounds.

Reference	Synthesis Method	Parameters	Comments
Marelli *et al.* (2010)	Frequency domain signal-based model	Amplitude control over analysis and synthesis filters	Operates in real time
Böttcher and Serafin (2009)	Granular	Accelerometer speed	Mapped to playback speed
	Sample-based	Accelerometer speed	Triggered by threshold speeds
	Noise shaping	Accelerometer speed	Mapped to bandpass centre frequency
	Physically inspired	Accelerometer speed	Mapped to the amplitude of frequency modes
Dobashi *et al.* (2003)	Computational fluid dynamics	Length, diameter and swing speed	Real-time operation, but requires initial off-line computations

The use of specialist hardware, Graphics Processing Units (GPUs) or Field Programmable Gate Arrays (FPGAs), are mostly used for musical instruments rather than sound effects Bilbao (2009). Due to the complex nature of the computations, these are necessary for real-time operation. It was not our intention to require specialist hardware in order for the model to operate in real-time which indicated that the model should avoid highly complex computations. However, simplifications which result in far weaker audio quality or realism, as deployed for dynamic level of detail Durr *et al.* (2015), should also be avoided.

The choice of software to implement the effect was based on a number of factors including, programming experience, licence required or open source, complexity of the model and efficiency of the language. As recommended in Farnell (2010), the open source programming language Puredata has proven to be excellent at developing the sound effects via a graphical syntax.

When developing a sound effect it is of value to look at other state of the art synthesis techniques used to create similar sound effects. A number of sword models have been developed and are listed in Table 2. A signal-based approach to a variety of environmental sound effects, including sword whoosh, waves and wind sounds, was undertaken in Marelli *et al.* (2010). Analysis and synthesis occur in the frequency domain using a sub-band method to produce narrow band coloured noise.

Four different sword models were evaluated in Böttcher and Serafin (2009). Here, the application was for interactive gaming, and the evaluation was focused on perception and preference rather than accuracy of sound. The user was able to interact with the sound effect through the use of a Wii Controller. One model was a band-filtered noise signal with the centre frequency proportional to the acceleration of the controller. A physically-inspired model replicated the dominant frequency modes extracted from a recording of a bamboo stick swung through the air. The amplitude of the modes was mapped to the real-time acceleration data.

The other synthesis methods in Böttcher and Serafin (2009) both mapped acceleration data from the Wii Controller to different parameters; one using the data to threshold between two audio samples, the other a granular synthesis method mapping acceleration to the playback speed of grains. Tests revealed that the granular synthesis was the preferred method for expression and perception. One possible reason that the physical model was less popular could be the lack of correlation between speed and frequency pitch, which the band-filtered noise had. This may also be present in the granular model.

A physical model of sword sounds was explored in Dobashi *et al.* (2003). Here, off-line sound textures were generated based on the physical dimensions of the sword. The sound textures were then played back with speed proportional to the movement. The sound textures were generated using off-line computational fluid dynamics software (CFD), solving the fundamental fluid dynamics equations. In this model Dobashi *et al.* (2003), the sword was split into a number of compact sound sources, spaced along the length of the sword. As the sword was swept thought the air, each source moved at a different speed; therefore, the sound texture for each source was adjusted accordingly. The sounds from each source were summed and output to the listener.

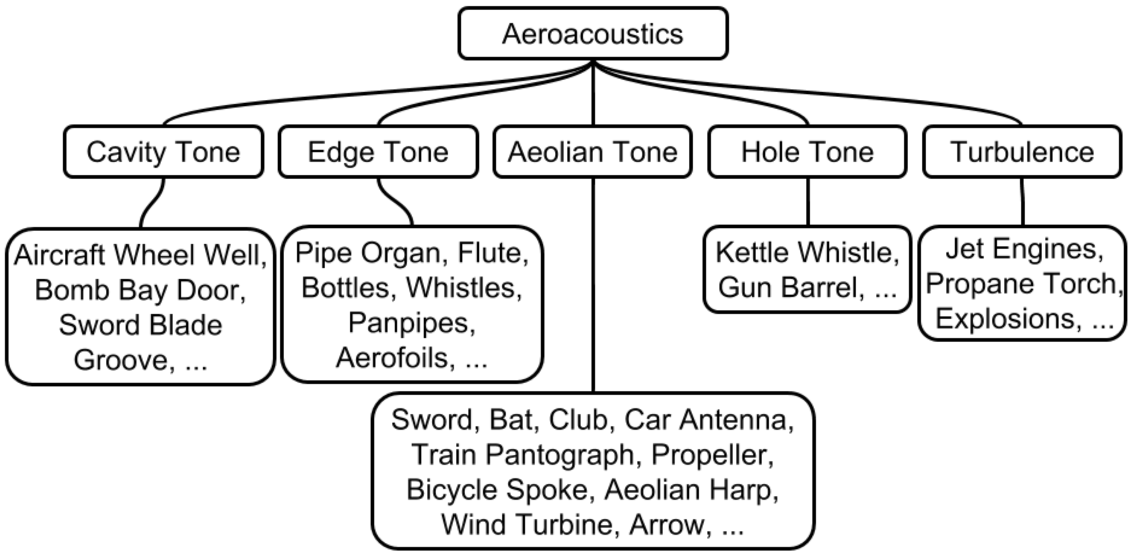


Figure 1: Simplified taxonomy of fundamental aeroacoustic sounds including examples of where each can be found

The desire for our sword example was to create a sound effect in which parameter changes modified the output sound in real-time, requiring no off-line processing, while simultaneously producing plausible sounds within a procedural audio context. These requirements meant that a physically inspired synthesis model was the most suitable.

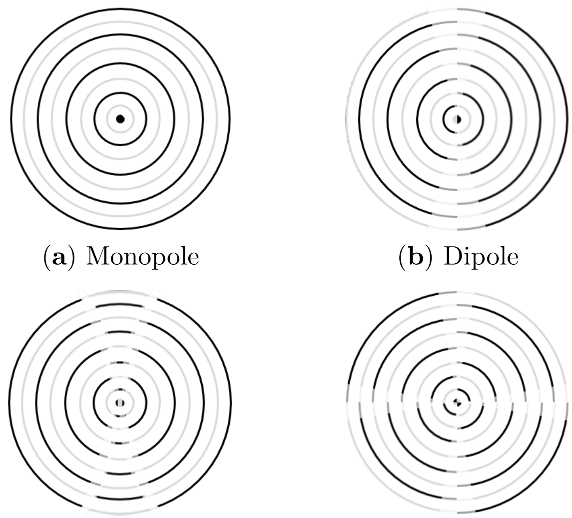
4.1 Aeroacoustics

The sound of a sword swinging through the air comes from a class of sounds called aeroacoustics. Aeroacoustics is the name given to the field of study that determines the sounds produced by air, for example a boiling kettle or helicopter rotor. Aeroacoustics has a number of fundamental tones which individually and collectively can replicate a number of common sound effects. A basic taxonomy of these is shown in Fig. 1. We can see from this that the main fundamental tone that a swinging sword produces is the Aeolian tone.

It was shown in Curle (1955) and confirmed in Gerrard (1955) that aeroacoustic sounds, in low flow speed situations, could be modelled by the summation of compact sources, namely monopoles, dipoles and quadrupoles. An acoustic monopole, under ideal conditions, can be described as a pulsating sphere, much smaller than the acoustic wavelength. This is shown in Fig. 2a. A dipole, under ideal conditions, is equivalent to two monopoles separated by a small distance but of opposite phase, shown in Fig. 2b. Quadrupoles are two dipoles separated by a small distance with opposite phases. A longitudinal quadrupole has the dipoles axes in the same line while a lateral quadrupole can be considered as four monopoles at the corners of a rectangle, Crighton *et al.* (2012). These are shown in Figs. 2c and 2d.

The sword model created by Dobashi *et al.* (2003) made use of compact sound sources to generate their aeroacoustic sounds. This required complex off-line CFD calculations to generate sound textures for each compact source and then concatenated together with playback speed mapped to the speed of the sword swing.

In our model, we also used compact sound sources but instead of off-line calculations we carried out research into equations, known as *semi-empirical equations*, where assumptions and generalisations have been made to simplify calculations or to yield results in accordance with observations. Although many of these equations may at first appear complicated, once all the relevant parameters are known, they produce exact results with errors only due to the approximations made during the equation derivation.



(**a**) Monopole (**b**) Dipole

(**c**) Longitudinal Quadrupole (**d**) Lateral Quadrupole

Figure 2: Ideal radiation pattern for a monopole, dipole, lateral quadrupole and longitudinal quadrupole

4.2 Aeolian Tone

To implement a physically inspired model, basic knowledge of the physics behind the fundamental tone that was being generated was required. Determining how much knowledge of physics the model possesses is often a balance between computationally complexity, choice of software and hardware the model used for implemented, and perceptual relevance. The first step was to gain basic understanding of the Aeolian tone and the parameters upon which it depends Selfridge *et al.* (2016).

The Aeolian tone is generated when air flows around an object. The vast majority of research on this tone has been carried out on cylindrical objects and we modelled our sword based on these results. When air passes around a cylinder vortices are generated and then released, or shed, from the back of the cylinder. This is depicted in Fig 3 where we can see what is known as a vortex street behind the cylinder. It can also be seen that vortices are shed from alternate sides of the cylinder. This causes an oscillating lift force perpendicular to the flow. Normal to the flow there is a drag force at twice the frequency of the lift force. We are able to model each of the oscillating forces and their harmonics by dipole compact sound sources.

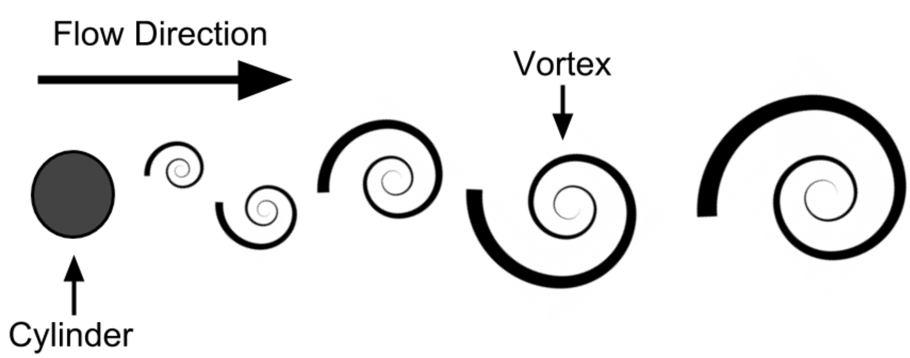


Figure 3: Example of vortices being shed from opposite sides of a cylinder known as a vortex street

4.2.1 Frequency Calculations

In 1878 Czech physicist Vincenc Strouhal carried out one of the first important studies into the frequency of a tone produced as air flows around a cylinder. He defined the formula given in Eqn. 1:

$$S_t = \frac{fd}{u} \tag{1}$$

where S_t is the Strouhal number, f is the tone frequency, d the cylinder diameter and u the airspeed. The fluctuating lift force caused by the vortex shedding is dominated by a fundamental frequency f_l with $S_t \approx 0.2$, the drag force is dominated by a fundamental frequency f_d which is $2f_l$. From Eqn. 1 we can therefore calculate the tone frequencies due to the lift, drag and any harmonics.

4.2.2 Acoustic Intensity Calculations

The time-averaged Acoustic Intensity $\overline{I_{1l}}$ (W/m^2) of the Aeolian tone lift dipole source was derived in Goldstein (1976). The full derivation is beyond the scope of this publication and the reader is referred to Goldstein (1976) if they wish to find out more. The Acoustic Intensity $\overline{I_{1l}}$ is proportional to:

$$\overline{I_{l1}} \propto u^6 \sin^2\theta \cos^2\varphi \tag{2}$$

where θ is the elevation angle and φ the azimuth angle.

4.2.3 Tone Bandwidth

The bandwidth around the fundamental frequency is affected by the airspeed and diameter, the higher the airspeed or diameter, the wider the tone bandwidth.

4.2.4 Wake Noise

When the airspeed or diameter increases the vortices produced by shedding diffuse rapidly, merging into a turbulent wake. The wake produces wide band noise modelled by quadrupole sources whose intensity varies with the flow speed to the power of 8, $\overline{I_w} \propto u^8$ Etkin *et al.* (1957). There is very little noise content below the lift dipole fundamental frequency f_l, Etkin *et al.* (1957). Above this frequency the roll off of the turbulent noise amplitude is $\frac{1}{f^2}$, Powell (1959). The wake is modelled by a range of combinations between the various longitudinal and lateral quadrupoles.

4.3 Implementation of a Compact Sound Source

In Section 4.2 values for the frequency of the lift dipole and drag dipole have been identified along with the acoustic intensity values for the lift dipole and the wake noise. We have also highlighted the fact that as the airspeed or diameter increases that the bandwidth around the tone frequencies increases. Considering this in relation to deciding which method to implement our physically inspired model, it is judged that subtractive synthesis would provide a suitable method.

4.3.1 Control Parameter Sample Rate

As control parameters can be modified during operation of our synthesis method, we need to look at sampling the variables over discrete time $[n]$. We can record the control parameter at audio rate (44100Hz), but this requires performing every single calculation for every audio sample. This can result in a very computationally heavy process, which requires a lot of CPU to process. Alternatively we can update control parameters at a lower rate, such as every 1000 samples (441Hz), however we then introduce the possibility of some parameter jumps causing audio glitches.

This can potentially be resolved with smoothing control parameter values. In our Puredata model, a number of variables were sampled at audio rate of 44100 Hz. This allow real-time performance but overuse can put a strain on the audio buffers causing drop-outs. A balance between accuracy and perception has to be achieved if this occurs. A list of sampled variables were:

- Airspeed $u[n]$
- Elevation $\theta[n]$
- Azimuth $\varphi[n]$
- Distance $r[n]$

With these variables measured at discrete time $[n]$ we were able to calculate the discrete parameters based on the semi-empirical equations described above. The list is certainly not exclusive and variables like length and diameter can be sampled to allow real-time morphing of the sword. The properties of air could also be varied to create realistic sounds depending on weather changes or even due to alien atmospheres.

4.3.2 Lift and Drag Dipoles

Subtractive synthesis is based around shaping the frequency response of a noise source by filtering. For our purposes, bandpass filtering was employed to obtain the required sound of the dipoles. Bandpass filters require a centre frequency and the width of the peak. Using Eqn. 1 the centre frequencies of bandpass filters representing the lift and drag dipoles were be calculated.

In signal processing, the relationship between the peak frequency and bandwidth is called the Q value, ($Q = f_l/\Delta f$). As stated in Section 4.2.3, there is a bandwidth around the tone, and this is related to the Reynolds number. Data available in Norberg (1987) enabled a definition for the required Q values to be defined.

The output for the lift dipole is given as the output from a bandpass filter whose input was a white noise source, with centre frequency f_l and predicted Q. Likewise, the drag dipole output and third harmonic were generated in the same manner.

4.3.3 Wake Quadrupole

The final aspect added into the compact sound source was the noise associated with the wake. The required noise profile required a roll off of $1/f^2$ which is known as a brown noise source. There is little wake contribution below the fundamental frequency Etkin *et al.* (1957). Therefore, a high pass filter was applied with the filter cut-off set at the lift dipole fundamental frequency, f_l. This produces the turbulent noise profile required.

4.3.4 Final output

The output was obtained by adding lift dipole, drag dipole, harmonic and wake gives the final output, as shown in Fig. 4.

4.4 Modelling a Sword

4.4.1 Physical Model of a Sword

The knowledge gained from understanding some of the aeroacoustic sound generating process assists in design decisions on how to model the sword. If we wished to capture the physics of the entire sword we could place a number of compact sound sources all the way from the tip to the hilt. Dobashi *et al.* (2003) determines the output in exactly this manner with the output form each compact source obtained through off-line CFD calculations. To copy this in our model would have increased the complexity and the chance of audio drop outs. Instead we looked at Eqn. 2 and the other intensity equations while appreciating the characteristics of a swinging sword.

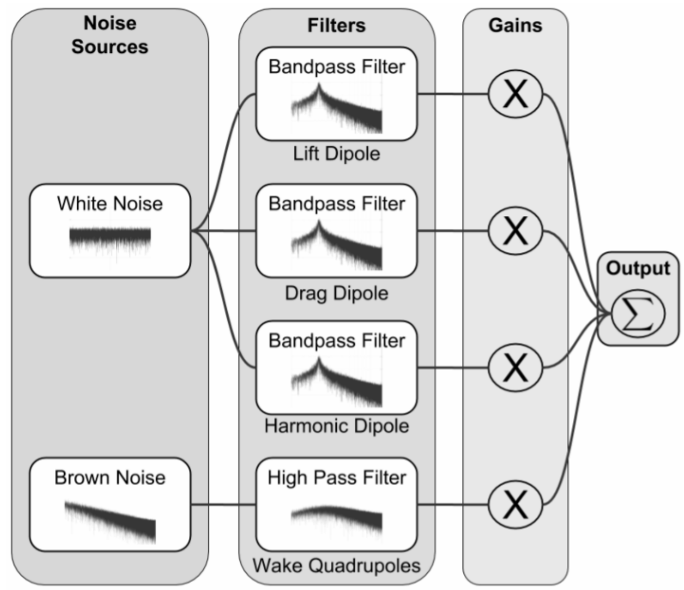


Figure 4: Flow diagram outlining the signal processing chain implementing the Aeolian tone compact sound source

We can see that for the Aeolian tone the acoustic intensity of the dipoles is proportional to u^6 and the wake quadrupoles to u^8. This means that the majority of sound generated will be in the area that u is the greatest. In the case of a sword swinging this will be the tip. In our model we placed six compact sound sources at the tip of the sword as well as one and the hilt and one halfway between the hilt and the last compact source at the tip. This is illustrated in Fig. 5.

4.4.2 Modelling the Behaviour of a Sword

Modelling the behaviour of a sword swing is an important aspect of achieving a believable physical model sound effect. To model the sword behaviour a number of design decisions were taken and implemented to give a limited range of motion to the user for the swing. It is feasible to attach the compact sound source models to a game object and calculate the sword dimensions from the graphics and airspeed from the animation. This can cause problems under some circumstances which we shall discuss later.

The speed of the sword at the start and end of the swing were set to 0m/s, with the top speed (set by the user) at the halfway point. The track of the swing was set to be circular for ease of programming - in reality a swing may probably more like a variety of arcs. The length of the forearm was added to the length of the sword with the elbow joint at the centre of a coordinates system, shown in Fig. 5.

With these conditions imposed on the sword swing action the user was able to set a number of parameters prior to the swing. The parameters available to the user were:

- Position of the observer in 3-dimensions
- Start position of the sword - azimuth and elevation
- Thickness of the tip and hilt of the sword - diameter of each source extrapolated from these
- Length of the sword
- Top speed the tip will reach during the swing

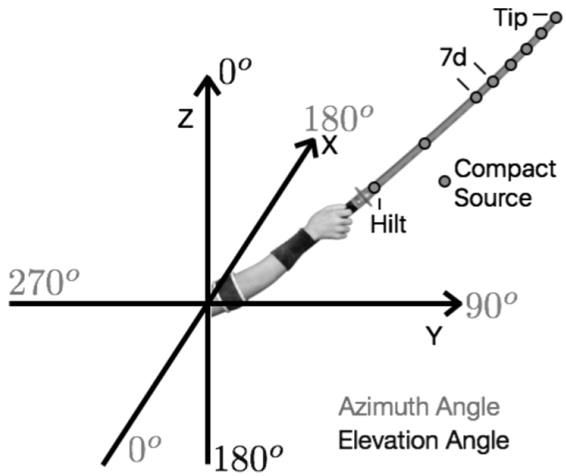


Figure 5: Position of 8 compact sources and coordinates used in the sword model

It is important when modelling the swing of an object that the Doppler effect is taken into consideration as well as panning if a stereo sound effect is being produced. By adjusting the diameters of the compact sound sources we are able to replicate the geometry of a number of other objects. In order to replicate objects as accurately as possible measurements of the dimensions of a number of objects were taken. We took the measurements of a baseball bat, 3-wood and 7-iron golf clubs, a broom handle, a wooden sword and a metal sword and programmed the physical model to give pre-set values which replicate these.

4.5 Evaluation of the Physical Model

4.5.1 Objective Evaluation

The objective evaluation focussed on the compact sound source. There has been a number of studies giving results of the frequency of air passing a frequency at known airspeeds and diameters. These are either experimental from wind tunnel measurements or simulated from CFD calculations. The average absolute error for our model was 4.66% while CFD has an average absolute error of 18.11%. The absolute error is calculated as the absolute difference between the recording or simulation and the model we are measuring.

4.5.2 Subjective Evaluation

To subjective evaluate the sword model listening tests were carried out asking participants rated the sword sounds on how plausible they believed them to be. The Web Audio Evaluation Tool, Jillings *et al.* (2015, 2016) was used to build and run listening tests in the browser. Each participant was presented with a page for each of the pre-set sound effects. The sounds generated by the physical model were compared to real recordings and alternative synthesis methods, including spectral modelling synthesis (SMS) Amatriain *et al.* (2002) for all objects and samples from Böttcher and Serafin (2009), and Dobashi *et al.* (2003) for the metal sword.

To obtain recorded samples as close as possible to those we were attempting to replicate, recordings were captured by the authors of the actual objects we had measured and used to programme our pre-sets. These were recorded within the Listening Room, Electronic Engineering and Computer Science Department, Queen Mary University of London. They were recorded on a Neumann U87 microphone placed approximately 20 cm from the midpoint of the swing and at 90 degrees to the plane of the swing. The impulse response of the room was captured and applied to all other sounds in the listening test so that the natural reverb of the room would not influence the results (except samples from Böttcher and Serafin (2009); Dobashi *et al.* (2003)).

The anchors were created from a real-time browser-based synthesis effect to allow a thorough comparison of how plausible the synthesis method is compared to the recorded sample. It was expected that a low pass filtered sample, as used in the MUltiple Stimuli with Hidden Reference and Anchor (MUSHRA) standard, would still be considered plausible, whereas a low-quality anchor would encourage the full use of the scale and allow for better understanding as to the effectiveness of the synthesis method.

Rating the plausibility of sound from a physical model was the preferred judgement in Castagné and Cadoz (2003), stating a plausible sound as one that listeners thought "was produced in some physical manner". Box plots for all five objects are shown in Figure 6. The box plots are a visualisation of the distribution of the result. It shows the median of the data, in the centre black line, the upper and lower bounds of the box show the first and third quartiles. This means the boxed area shows 50% of the data within the box, and the 'whiskers', or the lines from the box show the last 25% of the data, on each edge. Identified outliers are marked with an o. It can be seen from Figure 6, that our physical model outperforms the alternative synthesis methods on all of the objects except the metal sword. The metal sword performed poorly for plausibility in this test, with the model with added cavity tones performing slightly better.

4.6 Discussion

We can see from Fig. 6 that results from the listening test indicate that overall our model performs well compared to other synthesis models. It has exceptional performance for the broom handle, baseball bat, golf club, and wooden sword objects, where participants found sounds generated by our model were as plausible as real recordings. The exception to this was the metal sword physical model sound effect. It is important that once we have designed the sound effect, evaluated it through objective and subjective test, to try and understand the difference between results and recorded sounds. Having an understanding of this or hypothesis of the reasons behind the differences, gives areas where the model can be improved in the future.

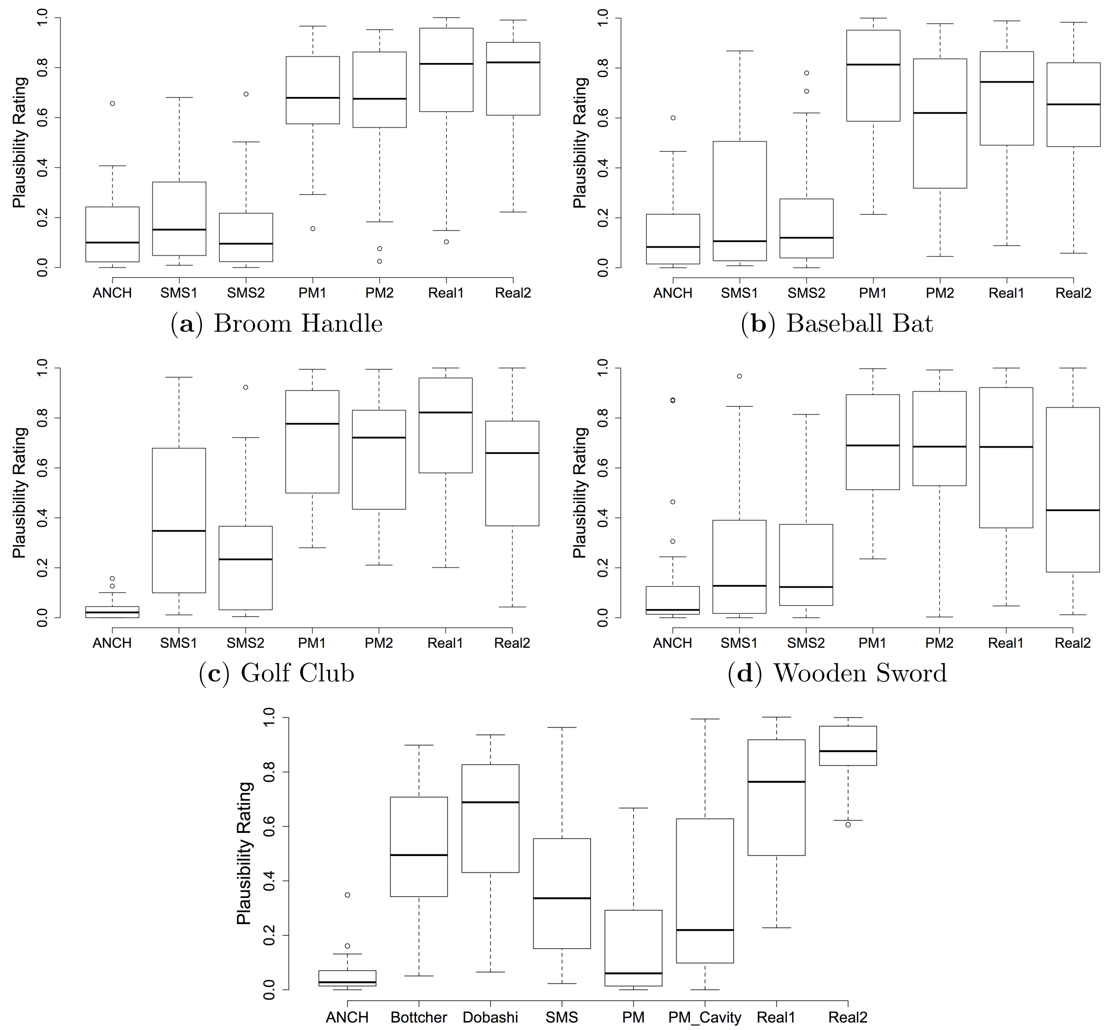
The broom handle, baseball bat and golf club objects were all cylindrical with thickness to width ratios of 1:1. For the wooden sword this ratio decreases to approximately 0.37:1 and for a metal sword to approximately 0.14:1. The Aeolian tone model is designed around vortex shedding from cylindrical objects and it is reasonable to assume additional discrepancies may exist when there is a deviation from the thickness to width ratio of a cylinder.

The metal sword clearly has the poorest performance than the other modelled objects. Objects thicker than the sword thicker have greater wake noise which may have influence over the plausibility of the sounds. Spectral Modelling Synthesis analyses a recording and extracts sinusoidal components. Thinner objects produce sounds closer to pure tones and hence are better synthesised using SMS than thicker objects.

A possible reason for the poor rating of the metal sword object compared to the other objects is that the number of participants who have swung a real sword and heard the sound may well be less than those who have perhaps swung a golf club and the other objects. Memory plays an important role in perception Gaver and Norman (1988). If participants have heard a Foley sound effect for a sword more often than an actual sword sound, this may influence their perception of the physical model.

In contrast, it can be argued that participants will have more likely heard the actual sounds of a golf club at a live sporting event or within sporting broadcasts and hence their memory of these sounds would be closer to the physical model. Since all participants were from the UK the baseball bat would most likely not be as familiar to them as other objects and hence they might not have as strong a memory of the sound made by this object. This would make the difference between a memory of a Foley sound and an actual sound diminish.

We developed a method for modelling the behaviour of the sword which includes mapping the compact sound sources directly to graphical objects in a game engine. Care has to be taken in relation to this style of mapping as often the movement of the graphics extend what is physical possible. An example of this is when a person swinging a sword generates speeds which are much



(**a**) Broom Handle (**b**) Baseball Bat

(**c**) Golf Club (**d**) Wooden Sword

(**e**) Metal Sword

Figure 6: Box plots showing plausibility results for the preset objects. (ANCH, Anchor; SMS, Spectral Modelling Synthesis; PM, Physical Model; PMCavity, physical model including cavity tone; Real, recorded sample.)

higher than physically possible. Remembering that acoustic intensity of the dipoles is proportional to u^6, the signal level goes extremely high and will clip at the output.

The model presented here offers a unique and novel approach to synthesising aeroacoustic sound effects. Instead of starting by modelling the sound of the entire sword, we investigate and model compact sound sources which can then be manipulated to model the sword but also a range of objects which would not otherwise have been possible. As well as the modelling of clubs and bats the compact sound sources can be extended to model other items. In Selfridge *et al.* (2017a) the compact sound sources are used to generate the vortex sounds which are an integral component of the sound of a propeller spinning. The coupling of vortex shedding and mechanical vibrations was illustrated in a physical model of an Aeolian harp, Selfridge *et al.* (2017d), where the output sound is mainly caused by the vibration of strings. In this model, the vortex shedding causes the strings to vibrate with the compact sound source providing control data for this process.

The development of further aeroacoustic compact sound sources offers an increase in the sounds and objects that can be modelled. The cavity tone model, Selfridge *et al.* (2017e), was used to add a grooved profile to the sword tested in our listening tests. A physically inspired synthesis model of an edge tone was presented in Selfridge *et al.* (2018b), using an approach which illustrates the use of machine learning techniques to provide information in circumstances where techniques the physics are not fully understood. The use of compact sound sources can be extended to a number of other sound effects. A number of these are listed in Chanaud (2010) and given in Table 3.

Table 3:

Source Type	**Sound Effect**
Monopole	pistons, exhaust flows, propane torch (combustion), weapon discharges (explosions), drums (membranes), automobile tyre sounds, bubbles, splashes, waterfalls, electrical sparks, kettle whistles, corrugated pipe tone
Dipole	airfoil sounds, propellers, bicycle spokes, exhaust flows, Aeolian tone, cavity tone, ring tone, edge tone, vortex whistle, bottles, police whistle, Levavasseur whistle, screech tone (supersonic jets)
Quadrupole	subsonic jets, wakes, supersonic jets

5 Conclusions

In this chapter we have given a comprehensive overview of sound effect synthesis and control methods. The contextual reason and importance of sound effects have been given including how these are sub-categorised as well as the importance of Foley artists. An in-depth review of the wide range of sound effect synthesis techniques has been given, highlighting the strengths and weaknesses of different synthesis methods.

It has been seen that control of a synthesised sound effect is paramount in generating the desired sound and exploring the nuances of the range of sounds a model can produce. Evaluation techniques of synthesis models have been described along with reasons why evaluation is essential when deciding on which sound effect we should use and how research will develop in the future.

We have looked at the definition of procedural audio, drawing attention to why this is an area being developed for games and virtual reality environments. An example design process has been given for a physically inspired sound synthesis model which can be integrated as a procedural audio effect, including how much understanding of physics can be incorporated into the model and how this can influence design techniques. This is one such approach that can effectively model a given sound. This approach can extend to a range of physically derived synthesis models, though this approach is not guaranteed to work in every single case, with modifications, a similar approach can be used for developing synthesis methods.

It is never expected be the case that every single sound, within an game, will be produced by a synthesis approach. The subtle aspects of sound design will result in a fusion of different types of sounds and tools, and synthesis is one such approach that can be integrated into the sound designed.

It is clear that sound effects add vital cues for listeners and viewers across a range of media content. The development of more realistic, interactive, immersive sound effects is an exciting and growing area of research, with many research questions still to be answered.

References

Xavier Amatriain, Jordi Bonada, Alex Loscos, and Xavier Serra. Spectral processing. In Udo Zölzer, editor, *DAFx: Digital Audio Effects*, chapter 10, pages 373–438. John Wiley and Sons, Ltd., Chichester, UK, 2002.

Mitsuko Aramaki, Richard Kronland-Martinet, and Sølvi Ystad. Perceptual control of environmental sound synthesis. In *Speech, Sound and Music Processing: Embracing Research in India*, pages 172–186. Springer, 2012.

James A Ballas. Common factors in the identification of an assortment of brief everyday sounds. *Journal of experimental psychology: human perception and performance*, 19(2):250, 1993.

BBC. *The BBC Year Book 1931*, chapter "The Use of Sound Effects", pages 194–197. British Broadcasting Corporation, 1931.

Søren Bech and Nick Zacharov. *Perceptual audio evaluation-Theory, method and application*. John Wiley & Sons, 2007.

Stefan Bilbao. *Numerical Sound Synthesis: Finite Difference Schemes and Simulations in Musical Acoustics*. Wiley Online Library, 2009.

Niels Böttcher and Stefania Serafin. Design and evaluation of physically inspired models of sound effects in computer games. In *Audio Engineering Society Conference: 35th International Conference: Audio for Games*, London, UK, 2009.

Nicolas Castagné and Claude Cadoz. 10 criteria for evaluating physical modelling schemes for music creation. In *Proceedings of the 6th International Conference on Digital Audio Effects (DAFx03)*, 2003.

Robert C Chanaud. Tools for analyzing sound sources. *Essex, CT: CCR Associates, LLC*, 2010.

John M Chowning. The synthesis of complex audio spectra by means of frequency modulation. *Journal of the audio engineering society*, 21(7):526–534, 1973.

DG Crighton, Ann P Dowling, JE Ffowcs Williams, MA Heckl, and FA Leppington. *Modern methods in analytical acoustics: lecture notes*. Springer Science &; Business Media, 2012.

N Curle. The influence of solid boundaries upon aerodynamic sound. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, pages 505–514, 1955.

Yoshinori Dobashi, Tsuyoshi Yamamoto, and Tomoyuki Nishita. Real-time rendering of aerodynamic sound using sound textures based on computational fluid dynamics. In *ACM Transactions on Graphics (TOG)*, volume 22, pages 732–740, 2003.

Gabriel Durr, Lys Peixoto, Marcelo Souza, Raisa Tanoue, and Joshua D Reiss. Implementation and evaluation of dynamic level of audio detail. In *56th Audio Engineering Society Conference: Audio for Games*, London, UK, 2015.

Bernard Etkin, GK Korbacher, and Roger T. Keefe. Acoustic radiation from a stationary cylinder in a fluid stream (aeolian tones). *The Journal of the Acoustical Society of America*, 29, 1957.

Andy Farnell. An introduction to procedural audio and its application in computer games. In *Audio Mostly Conference*, pages 1–31, Ilmenau, Germany, 2007. ACM.

Andy Farnell. *Designing sound*. MIT Press Cambridge, UK, 2010.

Nicolas Fournel. Procedural audio for video games: are we there yet? In *Game Developers Conference 2010*, San Francisco, California, 2010. Sony Computer Entertainment Europe.

Leonardo Gabrielli, Stefano Squartini, and Vesa Välimäki. A subjective validation method for musical instrument emulation. In *131st Audio Engineering Society Convention*, New York, USA, 2011.

William W Gaver and Donald A Norman. *Everyday listening and auditory icons*. PhD thesis, University of California, San Diego, Department of Cognitive Science and Psychology, 1988.

JH Gerrard. Measurements of the sound from circular cylinders in an air stream. In *Proceedings of the Physical Society*, page 453, 1955.

Marvin E Goldstein. *Aeroacoustics*. New York, McGraw-Hill International Book Co., 1976.

Christian Heinrichs and Andrew McPherson. Mapping and interaction strategies for performing environmental sound. In *1st Workshop on Sonic Interactions for Virtual Environments at IEEE VR 2014*, 2014.

Christian Heinrichs, Andrew McPherson, and Andy Farnell. Human performance of computational sound models for immersive environments. *The New Soundtrack*, 4(2):139–155, 2014.

Simon Hendry and Joshua D. Reiss. Physical modeling and synthesis of motor noise for replication of a sound effects library. In *Audio Engineering Society Convention 129*, Los Angeles, CA, USA, 2010.

Matthew D Hoffman and Perry R Cook. Feature-based synthesis: A tool for evaluating, designing, and interacting with music IR systems. In *International Symposium on Music Information Retrieval (ISMIR)*, pages 361–362, 2006.

Robert H Jack, Tony Stockman, and Andrew McPherson. Effect of latency on performer interaction and subjective quality assessment of a digital musical instrument. In *Proceedings of the Audio Mostly 2016*, pages 116–123, Norrköping, Sweden, 2016. ACM.

David A Jaffe. Ten criteria for evaluating synthesis techniques. *Computer Music Journal*, 19(1):76–87, 1995.

Nicholas Jillings, Brecht De Man, David Moffat, and Joshua D Reiss. Web audio evaluation tool: A browser-based listening test environment. In *Proceedings of the 12th Sound and Music Computing Conference*, Maynooth, Ireland, 2015.

Nicholas Jillings, Brecht De Man, David Moffat, and Joshua D. Reiss. Web audio evaluation tool: A framework for subjective assessment of audio. In *Proc. 2nd Web Audio Conference*, Atlanta, Georgia, USA, April 2016.

Damián Marelli, Mitsuko Aramaki, Richard Kronland-Martinet, and Charles Verron. Time-frequency synthesis of noisy sounds with narrow spectral components. *IEEE Transactions on Audio, Speech and Language Processing*, 18, 2010.

Josh H McDermott and Eero P Simoncelli. Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron*, 71(5):926–940, 2011.

Lucas Mengual, David Moffat, and Joshua D. Reiss. Modal synthesis of weapon sounds. In *Proc. Audio Engineering Society Conference: 61st Audio Engineering Society International Conference: Audio for Games*, London, February 2016.

Adrien Merer, Sølvi Ystad, Richard Kronland-Martinet, and Mitsuko Aramaki. Abstract sounds and their applications in audio and perception research. *Exploring music contents*, pages 176–187, 2011.

Adrien Merer, Mitsuko Aramaki, Sølvi Ystad, and Richard Kronland-Martinet. Perceptual characterization of motion evoked by sounds for synthesis control purposes. *ACM Transactions on Applied Perception (TAP)*, 10(1):1–24, March 2013.

Ananya Misra and Perry R Cook. Toward synthesized environments: A survey of analysis and synthesis methods for sound designers and composers. In *Proceeding of the International Computer Music Conference(ICMC)*, Montreal, Canada, 2009.

David Moffat and Joshua D. Reiss. Perceptual evaluation of synthesized sound effects. *ACM Transactions on Applied Perception (TAP)*, 2018.

David Moffat, David Ronan, and Joshua D. Reiss. Unsupervised taxonomy of sound effects. In *Proc. 20th International Conference on Digital Audio Effects (DAFx-17)*, Edinburgh, UK., September 2017. DAFx-17.

Christoffer Norberg. Effects of reynolds number and a low-intensity freestream turbulence on the flow around a circular cylinder. Chalmers University, Goteborg, Sweden, Technological Publications, 1987.

Alan Powell. Similarity and turbulent jet noise. *The Journal of the Acoustical Society of America*, 31, 1959.

Vytis Puronas. Sonic hyperrealism: illusions of a non-existent aural reality. *The New Soundtrack*, 4(2):181–194, 2014.

Davide Rocchesso, Roberto Bresin, and Mikael Fernstrom. Sounding objects. *IEEE MultiMedia*, 10(2):42–52, 2003.

Diemo Schwarz. State of the art in sound texture synthesis. In *14th International Conference Digital Audio Effects (DAFx)*, pages 221–231, Paris, France, 2011. Digital Audio Effects (DAFx).

Rod Selfridge, Joshua D. Reiss, Eldad J Avital, and Tang Xiaolong. Physically derived synthesis model of an aeolian tone. In *141th Audio Engineering Society Convention*, Los Angeles, CA, USA, 2016.

Rod Selfridge, David Moffat, and Joshua D. Reiss. Physically derived sound synthesis model of a propeller. In *ACM Audio Mostly Conference*, London, UK, August 2017.

Rod Selfridge, David Moffat, and Joshua D. Reiss. Real-time physical model for synthesis of sword swing sounds. In *International Conference on Sound and Music Computing (SMC)*, Espoo, Finland, July 2017.

Rod Selfridge, David Moffat, and Joshua D Reiss. Sound synthesis of objects swinging through air using physical models. *Applied Sciences*, 7(11):1177, 2017.

Rod Selfridge, David Moffat, Joshua D. Reiss, and Eldad J. Avital. Real-time physical model for an aeolian harp. In *International Congress on Sound and Vibration*, London, UK, July 2017.

Rod Selfridge, Joshua D. Reiss, and E Avital. Physically derived synthesis model of a cavity tone. In *Proceedings of the 20th Digital Audio Effects Conference*, pages 5–9, Edinburgh, UK, 2017.

Rod Selfridge, David Moffat, Eldad J. Avital, and Joshua D. Reiss. Creating real-time aeroacoustic sound effects using physically derived models. *Journal of the Audio Engineering Society (to appear)*, 2018.

Rod Selfridge, Joshua D. Reiss, and E Avital. Physically derived synthesis model of a edge tone. In *Proceedings of the 144th Audio Engineering Society Convention (to appear)*, Milan, Italy, 2018.

Xavier Serra and Julius Smith. Spectral modeling synthesis: A sound analysis/synthesis system based on a deterministic plus stochastic decomposition. *Computer Music Journal*, 14(4):12–24, 1990.

Julius O Smith. Viewpoints on the history of digital synthesis. In *Proceedings of the International Computer Music Conference*, 1991.

Luca Turchet, David Moffat, Ana Tajadura-Jiménez, Joshua D. Reiss, and Tony Stockman. What do your footsteps sound like? an investigation on interactive footstep sounds adjustment. *Applied Acoustics*, 111:77–85, October 2016.

Richard E. Turner. *Statistical Models for Natural Sounds*. PhD thesis, Gatsby Computational Neuroscience Unit, UCL, 2010.

William Wilkinson, Dan Stowell, and Joshua D Reiss. Performable spectral synthesis via low-dimensional modelling and control mapping. In *Proceedings of the Digital Music Research Network (DMRN)*, 2016.

William J Wilkinson, Joshua D Reiss, and Dan Stowell. Latent force models for sound: Learning modal synthesis parameters and excitation functions from audio recordings. In *Proceedings of the 20th International Conference on Digital Audio Effects*, Edinburgh, UK, 2017.

Author Biographies

David Moffat

David Moffat is a post-doctoral researcher within the Audio Engineering Group of the Centre for Digital Music at Queen Mary University London. His research focuses is on intelligent and assistive mixing and audio production tools through the implementation of semantic tools and machine learning. His PhD, from Queen Mary University London, focused on sound effect synthesis and evaluation. He has since developed new methods and approaches for sound synthesis, from this understanding of the field. He received an MSc in Digital Music Processing from Queen Mary University of London in 2014, and graduated from Edinburgh University with a BSc in Artificial Intelligence and Computer Science in 2011.

Rod Selfridge

Rod Selfridge is a PhD researcher within the Media and Arts Technology Doctoral College at Queen Mary University of London. His research is focused on creating real-time synthesis models of aeroacoustic sounds which are derived from fundamental fluid dynamics principles. His PhD research has produced two Journal articles and six peer-reviewed conference papers, including two best paper awards. Currently in the write-up stage of the PhD, he also works full time as a Research Software Developer at The University of Edinburgh, restoring historical buildings, including the acoustics, within a virtual reality environment, enabling historic musical performances to be recreated. In 1995 Rod obtained a First Class Honours Degree in Communication Engineer from Edinburgh Napier University and in 2011 completed a Masters with Distinction in Digital Music Processing at Queen Mary University of London.

Josh Reiss

Josh Reiss is a Professor of Audio Engineering with the Centre for Digital Music at Queen Mary University of London. He has published more than 200 scientific papers (including over 50 in premier journals and 5 best paper awards), and co-authored the textbook Audio Effects: Theory, Implementation and Application. His research has been featured in dozens of original articles and interviews on TV, radio and in the press. He is a former Governor of the Audio Engineering Society (AES), chair of their Publications Policy Committee, and co-chair of the Technical Committee on High-resolution Audio. His Royal Academy of Engineering Enterprise Fellowship resulted in founding the high-tech spin-out company, LandR, and he recently co-founded a second start-up, FXive. He has investigated psychoacoustics, sound synthesis, multichannel signal processing, intelligent music production, and digital audio effects. His primary focus of research, which ties together many of the above topics, is on the use of state-of-the-art signal processing techniques for professional sound engineering. He maintains a popular blog, YouTube channel and twitter feed for scientific education and dissemination of research activities.