Audio Engineering Society

Convention Paper

Presented at the 129th Convention
2010 November 4–7 San Francisco, CA, USA

Physical Modeling and Synthesis of Motor Noise for Replication of a Sound Effects Library

Simon Hendry[1] and Josh Reiss[1]

[1] *Centre for Digital Music, Queen Mary University of London*

Correspondence should be addressed to Simon Hendry (`mail@simonhendry.co.uk`)

ABSTRACT

This paper presents the results of objective tests exploring the concept of using a small number of physical models to create and replicate a large number of samples from a traditional sound effects library. The design of a DC motor model is presented and this model is used to create both a household drill and a small boat engine. The harmonic characteristics, as well as the spectral centroid were compared with the original samples and all the features agree to within 6.1%. The results of the tests are discussed with a heavy emphasis on realism and perceived accuracy and the parameters which have to be improved in order to humanise a model are explored.

1. INTRODUCTION

Traditionally, sound effect libraries are comprised of many sampled sounds stored as uncompressed wave files. Due to the versatile nature of sound, an object may have up to 100 samples in the library associated with it. Each of these samples will differ very slightly in order to give the sound designer a wide variety of sounds to choose from. Despite the vast selection of samples available to sound designers today, it is often the case that the perfect sound is not found and the designer has to record his own effect. This leads both to valuable production time being wasted and sample libraries quickly becoming both very large in size and unorganised.

Procedural audio is defined in [1] as a 'non-linear, often synthetic sound, created in real time according to a set of programmatic rules and live input'. Sound effects can be synthesised from a small number of starting blocks, for example a random number generator and sine wave generator, in much the same way that physical models of musical instruments have been developed in [2], [3] and [4] by impulses and digital waveguide synthesis. By using this methodology, a vast number of contrasting sounds can be created by modifying a small number of parameters made available to the user in the physical model.

1.1. Related Work

In recent years, procedural audio has been developed for use primarily in the computer game industry where memory for storing samples is limited and sample repetition such as a gun shot sound soon become tedious and unrealistic. Subjective tests carried out in [5] showed that 64% of the users preferred synthesised effects to sampled sounds in a simple game designed specifically for the test. With advances in game controllers such as the Wii remote, procedural audio can not just improve the quality of the gaming experience but can actively change the way in which the gamer plays. Although these techniques are present, but rarely seen in modern computer games, a talk given at this years Gaming Developers Convention entitled 'Procedural Audio for Video Games: Are we there yet?' by Nicolas Fournel (Principal Audio Programmer, Sony Computer Entertainment Europe), proves that this technology is starting to be accepted more into commercial applications. By far the most comprehensive collection of physical models can be found in [6] and it is on one of these models that the work in this paper is built.

1.2. Aims

The aim of this research is to find out if synthesis is a realistic alternative to sample based sound effects libraries. The research has two distinct questions to be answered. Firstly, is it possible to synthesis a sound effect to a high enough degree of accuracy that one is not able to distinguish between the sounds produced by a model and a recorded sample? Secondly, how many models would be needed to successfully replicate a large percentage of a traditional sample library? It is proposed that procedural audio methods can be used by sound designers in the theatre industry. The methods employed in this paper differ from those used in the gaming industry as here, the parameters are controlled by a human, as opposed to a machine in the gaming world. This work is primarily performed for use in the theatrical sound design industry where there is often a tight budget without the luxury of time. It is also an industry where, unlike music production, the work of the sound designer often has to be controlled and accepted by a director who has limited or no acoustic knowledge. For this reason, it is important that the model can be controlled using everyday high-level descriptors such as 'smashy-ness' or 'rumble' which the sound designer can modify, in real time, for the acceptance of the director.

2. DC MOTOR MODEL

To test the concept of a largely dynamic model, a model of a DC electric motor was built in Max/MSP. The model was based on [6] and is focused around three main sound sources; the rotor (the soft iron cylindrical core connected to the axis), stator (the hollowed out permanent magnet) and comutator (the connection between the brushes and the slip rings). All three attributes were controlled by a master maximum speed parameter. The model is fully designed and implemented in MAX/MSP primarily for use in real time.

2.1. Implementation

The brush and rotor components are intrinsically linked, as they would be in a real electric motor. The brush noise is simply made up of white noise with a fully parametric three-band equaliser. This is then summed with the rotor level in the signal domain and multiplied with a phasor to represent each half turn of the axis. The frequency of the phasor is controlled by the maximum speed parameter that can be given in either cycles per second or revolutions per minute. The stator sound is created from the magnetic force between the housing and the core. Here, it acts as a harmonic resonance with a frequency of four times the fundamental frequency with a quarter of the width of a cosine wave. This is achieved through the function

$$y = \frac{2}{cos^2(x) + 1} \tag{1}$$

This is then summed with the brush and rotor noise before passing through both high and low pass filters. A flow diagram of the model is presented in Figure 1.

2.2. Time Envelope

The time envelope of a DC motor is very complex

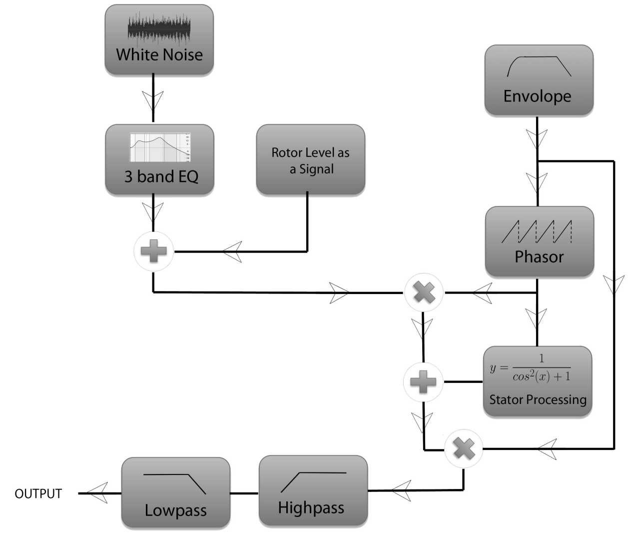

Fig. 1: Block diagram showing the methodology and simplified signal flow of the DC motor model

and is derived in [7]. The torque of a rotating axis can be expressed as

$$T = T_s(1 - \frac{w}{wf}) \tag{2}$$

where T is the torque of the system, T_s is the stall torque, defined as the torque produced by the system when $w = 0$ and w_f is the final angular velocity. If we then define the torque to be $T = J\dot{w}$, equation 2 can be rearranged as the differential equation of motion for the angular acceleration

$$\dot{w} = \frac{T_s}{J}(1 - \frac{w}{w_f}) \tag{3}$$

The solution to this differential equation is

$$w = w_f(1 - e^{-\frac{T_s}{Jw_f}t}) \tag{4}$$

thus the motor accelerates exponentially proportional to the load supplied to the axis. Theoretical convergence to the final angular velocity for three values of the moment of inertia is shown in Figure 2.

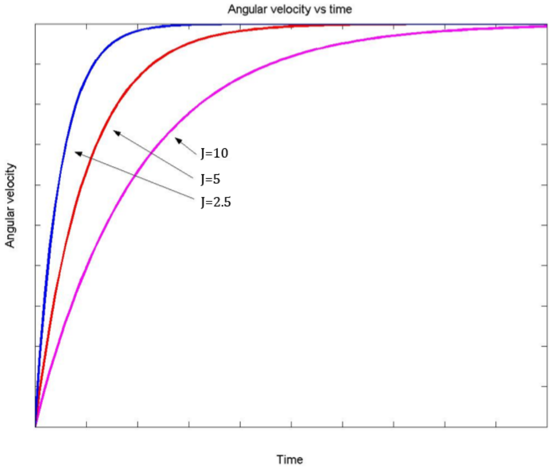

Fig. 2: Graph of angular velocity against time showing the theoretical convergence to the final angular velocity of a DC motor for three different values of J, the moment of inertia [7]

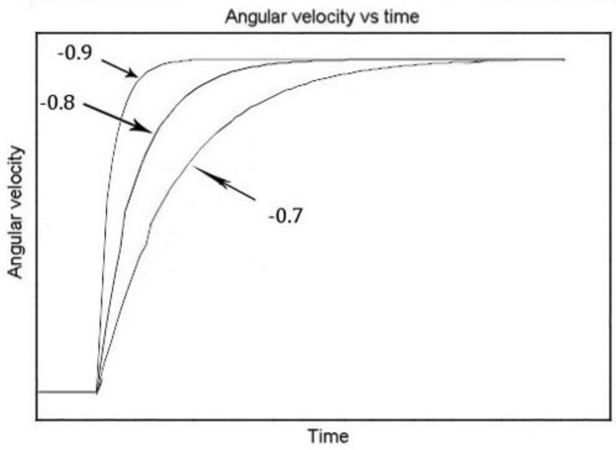

Fig. 3: Graph of angular velocity against time showing three curves created in Max/MSP with the *curve* object using curve parameters of -0.7, -0.8 and -0.9

In the interest of including only high-level features in the model, the attack envelope was replicated by using the *curve* object in Max/MSP. This ramp function will generate any iteration from a straight line to a very steep exponential curve depending on its input parameters. Various curves that can be created with this object are show in Figure 3 and are clearly comparable with the theoretical curves in Figure 2.

2.3. Humanisation

As we encounter everyday objects on a regular basis, it is more of a challenge to create convincing sound effects models than, say, a model of a musical instrument which we are less familiar with.*Humanising* the physical model with both long- and short-term variations in the sound may be the key to creating an accurate model. Here, a random walk is applied to the maximum speed, volume of the brush sound and gain parameters for the three bands in the parametric equaliser. This gives the sound a subtle human quality that is not immediately evident in the objective testing but creates a much more realistic sound to the listener which is not immediately identifiable through its machine-like static qualities.

3. THE EXPERIMENT

To test the concept of a diverse model to simulate multiple samples from an effects library, two vastly contrasting samples, a household drill and a small boat engine, were selected from the BBC sound effects library [8]. A replication of each sample was then attempted using the motor model discussed above. Note that the model was completed prior to choosing the two samples and was not modified in any way to achieve either of the sampled sounds, thus the number of parameters available to replicate both objects was limited to 16 (comprised of three mix level parameters, attack time parameter, maximum angular velocity parameter, high and low pass filters and nine parameters for the 3-band equaliser). The angular velocity parameter was taken from the fundamental frequency on a Fourier transform plot of the original sample and the relative levels and equalisation settings were set both subjectively from listening and visual comparisons with the spectrograms and by trying to match the spectral centroid with the original sample. Figures 4 and 5 show spectrograms of both the drill and boat samples on the left with the attempted reconstruction using Max/MSP on the right. A summery of the findings for both objects can be found in Table 1.

These results show that both contrasting samples can be estimated well by a single model of a DC motor. The spectral centroid for both objects agree well, with a 2.45% error on the drill sound and a 6.09% error for the boat. As expected, the harmonic frequencies correspond very well with all synthesised harmonics agreeing within 0.4% for the drill and 3.8% for the boat. This drop in harmonic accuracy for the boat is caused by the sample being slightly inharmonic, a feature that the model does not allow for. Being able to control the relative amplitudes of harmonics is another feature that the model cannot do and this is particularly evident in the drill sample where the first harmonic is stronger than the fundamental. Note also that the fundamental frequency of the boat sample is one order higher than the fundamental set as the maximum angular velocity of the model. The choice for this was made purely from a listening perspective, and although this introduces harmonics that are not present in the sample, the sound that is created results in a far better perceived likeness. The zero-crossing rate is a feature that the model does not accurately represent with a 28% and 33% error on the drill and boat samples respectively. The spectral centroid however, which here is recreated to a much higher degree of accuracy, gives a much better indication of the perceived similarity in tone between the sample and model.

4. FURTHER WORK

The results of this experiment are very promising and suggest that it is indeed possible to recreate a large number of samples from a single dynamic model. There is still a great deal of improvement however to increase the accuracy of the model. The current model accommodates a single rotor loop with two slip rings whereas in reality, there are often many rotor loops with comutators for each set [9]. For this particular improvement, the rotor sound should remain unchanged but both the stator and the brush sections of the model would be revised. Improvements are also needed to increase the realism of the model. The random walk did indeed help the perceived realness of the model but it is still easy to identify sample from synthesis. [5] suggest the use of granular synthesis as a method of randomising the sound. Another way to improve the perceived accuracy of the model is to introduce some environmental effects, as these are subconsciously expected to be present by human auditors as discussed in Section 2.3. The model is currently running with one channel creating a monophonic sound thus spatial field could be widened to create a stereo effect. This

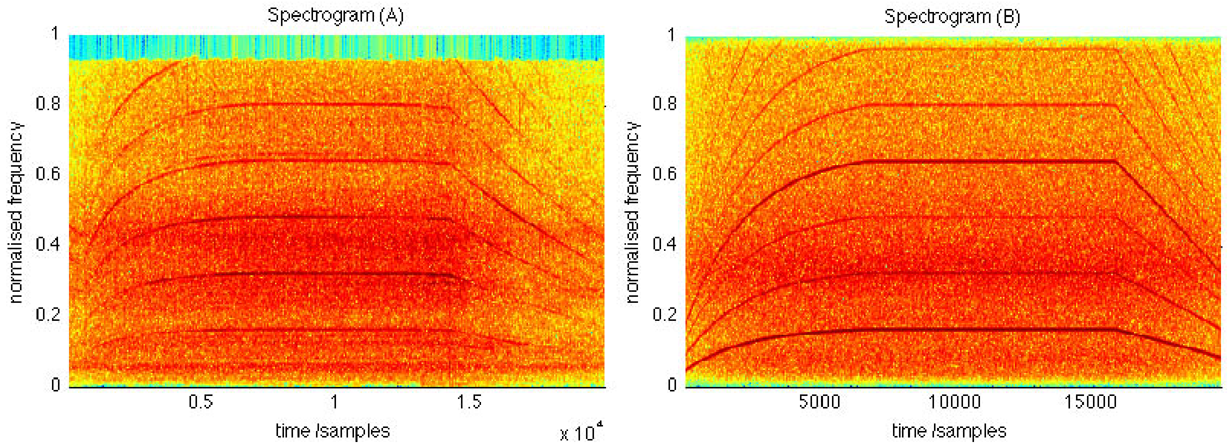

Fig. 4: Stereograms of the drill sample (left) and an estimation of the sound using the DC motor model in Max/MSP (right). Strong similarities are present including attack and decay curve features, spectral colour of noise and harmonic content. Note that the relative amplitudes of harmonics is not accuratly modeled

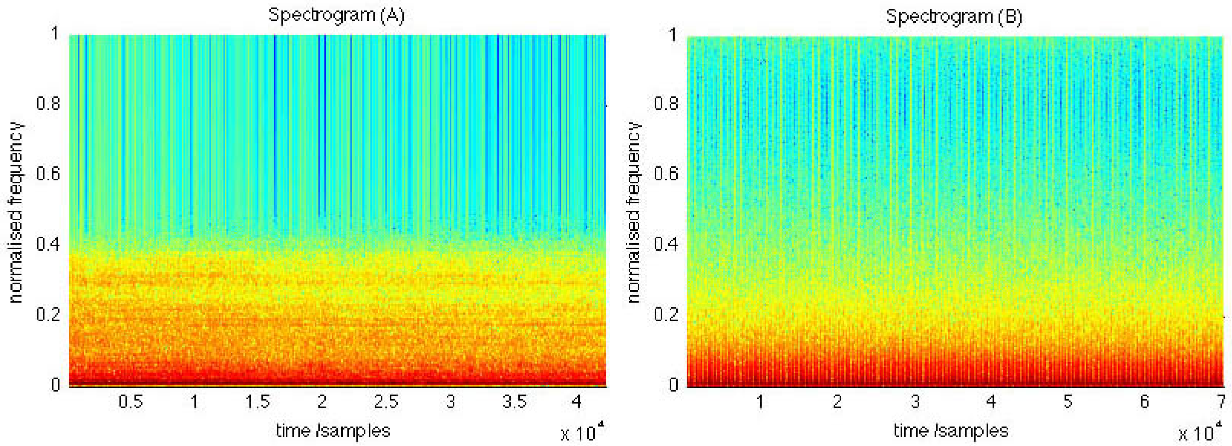

Fig. 5: Stereograms of the boat sample (left) and an estimation of the sound using the DC motor model in Max/MSP (right). The two plots share similar features, in particular the vertical streaks that are present due to the low angular velocity. The model fails to replicate the broadband noisy horizontal streaks probably caused by resonances in the boat motor.

Table 1: Summery of features extracted from the drill and boat objects for both the sampled sound and the model including percentage errors for each feature.

	Fundamental (Hz)	1st harm (Hz)	2nd harm (Hz)	3rd harm (Hz)	4th harm (Hz)	Centroid (Hz)	Zero-crossing (/sec)
DRILL							
Sample	3531	7063	10590	14080	17610	8672	16404
Model	3531	7062	10593	14124	17655	8465	12799
% error	**0.00%**	**0.01%**	**0.03%**	**0.31%**	**0.25%**	**2.45%**	**28.17%**
BOAT							
Sample		172.3		323		1665	565
Model	83	166	249	332		1773	838
% error		**3.80%**		**2.71%**		**6.09%**	**32.58%**

could be done by applying different equalisation filters to two versions of the same model, introducing a sample-level delay to one of the models, or integrating head related transfer functions (HRTFs) into the model. There is also an opportunity to imply movement with model for replicating objects such as remote controlled cars. This could be achieved by specifying a second speed parameter for horizontal motion and then modeling the Doppler effect.

4.1. Subjective Testing

We hope to perform subjective testing to further this research. The test will be performed on theatrical sound designers and will consist of two parts. Firstly, the subject will be asked to define, in his or her own terms, features of sound effects generated by models which they consider to be either particularly lifelike or sound particularly synthesised. This part of the test will take the form of a repertory grid [10]. Secondly, the subjects will be asked to perform a variable of ABX test in which sampled and synthesised sounds effects will be compared and indentified. The subjects will then asked to scale the synthesised sound on quality and usability. For example, does the fact that the model is easy to use make up for the lack of realism?

4.2. Sample Matching

One possible direction that this research could take is to combine the model with methods of feature extraction. Using extraction methods, a computer could determine parameter values from a sample such as fundamental frequency, relative harmonic content, harmonic to noise ratio and attack and decay times and parameters. These values could then be used in the motor model as a way of transposing a particular sample to a synthesis based sound as a starting point for a sound designer. Alternatively, machine-learning methods could be used to vary parameters sequentially and then compare the outcome to the target sample. Over many repeats of this process, the various parameter values should converge to create a synthesised sound that is as close to the target sample as is possible within the restrictions of the model.

5. CONCLUSION

In conclusion, the paper presents the idea of procedural audio and synthesis as a tool for theatrical sound designers. The design for a model of a DC motor is looked at in detail and the results of the experiment are presented. These results show that a single model can successfully estimate the sound of many contrasting samples from a traditional sound effects library with a small number of high-level parameters available to the user. The results also suggest that this methodology could be carried forward to other models for other effects such as weather, combustion motors and bangs, crashes and smashes. The results of the forthcoming listening test should supply us with an insight into which direction to take further research with the overall aim of humanising the sounds created by our physical models.

6. REFERENCES

[1] Andy Farnell. An introduction to procedural audio and its application in computer games. `http://obiwannabe.co.uk/html/papers/proc-audio/`, 2007.

[2] Kevin Karplus and Alex Strong. Digital synthesis of plucked-strong and drum timbres. *Computer Music Journal*, 7(2):43–55, 1983.

[3] Balaza Bank and Laszlo Sujbert. A piano model including longitudinal string vibrations. In *Proc. of the 7th Int. Conference on Digital Audio Effects*, Italy, 2004.

[4] Christope Vergez and Xavier Rodet. Trumpet and trumpet player: Model and simulation in a musical context. In *Proc. of International Computer Music Conference*, La Habana, Cuba, 2001.

[5] Niels Bottcher and Stefania Serafin. Design and evaluation of physically inspired models of sound effects in computer games. In *Proc. of the Audio Engineering Society 35th International Conference*, London, 2009.

[6] Andy Farnell. *Designing Sound*. Applied Scientific Press, London, 2008.

[7] Raul Rojas. Models for DC Motors. Unpublished document.

[8] British Broadcast Corporation Sound Effects Library discs 1-50. On loan from Ed Borgnis.

[9] John Chiasson. *Modeling and High Performance Control of Electric Machines*. Wiley-IEEE Press, 2005.

[10] Robert Neimeyer. Repertory grid methods. `http://www.pcp-net.org/encyclopaedia/repgrid-methods.html`, February 2004.