ADVANCED FOURIER DECOMPOSITION FOR REALISTIC DRUM SYNTHESIS

Tim Kirby * *and Mark Sandler*

Centre for Digital Music
Queen Mary University of London
London, UK
`t.c.kirby@qmul.ac.uk`

ABSTRACT

This paper presents a novel method of analysing drum sounds, demonstrating that this can form the basis of a highly realistic synthesis technique for real-time use. The synthesis method can be viewed as an extension of IFFT synthesis; here we exploit the fact that audio signals can be recovered from solely the real component of their discrete Fourier transform (RDFT). All characteristics of an entire drum sample can therefore be conveniently encoded in a single, real-valued, frequency domain signal. These signals are interpreted, incorporating the physics of the instrument, and modelled to investigate how the perceptual features are encoded. The model was able to synthesize drum sound components in such detail that they could not be distinguished in an ABX test. This method may therefore be capable of outperforming existing synthesis techniques, in terms of realism.

Sound examples available here.

1. INTRODUCTION

Closing the loop between analysis and synthesis helps us to better understand sound production in musical instruments, with clear application in the editing of audio and creation of virtual instruments.

1.1. Virtual Instruments

Research into synthesis methods is of significant value in the context of music production. Virtual instruments (VI's) offer many advantages over physical instruments and are employed in both professional mixes and live performances [1], as well as being used by the wider public in educational and recreational settings [2].

Sample-based VST's such as Toontrack's Superior Drummer 3.0 [3] are currently the most convincing way to replicate the sound of an acoustic drum kit, utilizing genuine recordings. Libraries are limited, however, by the scope of the sampling process. One is limited to the use of the specific drums that were recorded, each with a finite number of velocity layers and articulations. These restrictions limit both the creative potential of the instrument, the level of expression that can be conveyed, and the resulting realism of the performance.

Modelled instruments do not rely on samples, however, so are not subject to these limitations. It is possible to model drums of

---

arbitrary specification, including non-physical situations, such as a membrane with upwards pitch glide, or a gigantic drum.

Model parameters such as strike location and intensity can be varied in a continuous manner for enhanced playability. Subtle variations can be added to every sample, to avoid the "machine gun effect". The main barrier to these methods is the challenge of synthesising realistic sounds, which remains an open problem.

Plenty of synthesis methods have been applied to drums, including the classic additive, subtractive, and FM methods [4] found on famous synthesizers and drum machines. Other methods include digital waveguide meshes [5] [6], deterministic plus stochastic modelling [7], and the Karplus-Strong algorithm [8]].

While these techniques can create pleasing percussive sounds, which have found wide use in contemporary music, the sounds generated are clearly synthetic and can't compete with sample-based drum machines in terms of realism.

Finite difference methods [9], modal synthesis [10] and the functional transformation method [11], [12] have yielded improved results, such as those generated in the NESS project [13], but none that are truly convincing.

While there is potential for finite difference models to be further developed for improved realism, the huge computational cost rules out real-time performance in the foreseeable future [14]. This research, however, presents the basis for a highly realistic real-time method.

1.2. IFFT Synthesis using the Real Component of DFT

The Discrete Fourier Transform (DFT) is a valuable and widely used analysis tool in digital audio research, providing the frequency and phase spectra of an audio signal. Any audio signal can be transformed to the frequency domain via the FFT and recovered with the IFFT, without any loss of in-formation.

If you can model the spectra of a sound, you can therefore synthesise it. IFFT synthesis was introduced in 1980 [15], but has mainly been used as an efficient way of generating large ensembles of sinusoids for additive synthesis [16]. These sinusoids have fixed amplitude and frequency within a given window, so the FFT representations that are modelled are still relatively simple.

It is, however, possible to transform an audio signal of arbitrary length or complexity. The challenge is that it becomes harder to meaningfully interpret the FFT representation for more complicated signals, let alone edit or model them. If we are not dealing with a signal with a well-known Fourier transform, we can only compute the transform and investigate how the information is encoded. This paper demonstrates that entire drum samples transform in an informative manner and can be modelled in full, without the need for windowing.

One of the difficulties is that the FFT values are complex, so there are two signal components to interpret in tandem, whether
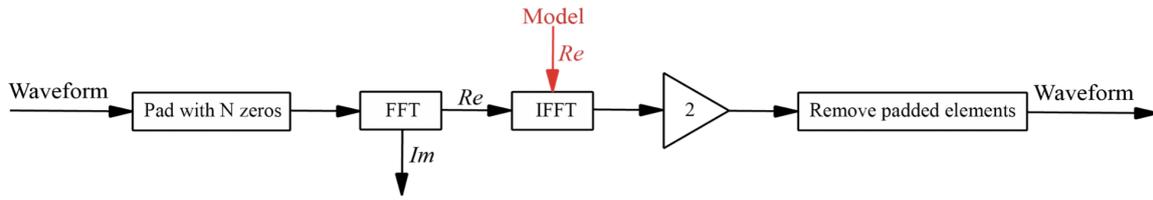
Figure 1: *Block scheme for the recovery of a drum sample from the real component of its Fourier transform (black), where N is the length of the original signal. The gain block of 2 compensates for the halving of amplitude which occurs when the imaginary component is discarded. The alternative scheme for synthesis begins with a model of the real component (red), bypassing all upstream stages.*

the real and imaginary components themselves, or the magnitude and the phase of the signal. In usual procedure, both the real and imaginary components are required to reproduce an audio signal, however, it is possible to reproduce a signal from solely the real (or imaginary) component (Figure 1), provided it has been sufficiently padded with zeros [17].

This makes it is possible to simplify the synthesis problem, to one of modelling a single, real-valued, frequency domain signal. This signal is the real component of the discrete Fourier transform (RDFT). The RDFT bears strong similarity with the Discrete Cosine Transform[18], so it might be conceptually helpful to view the proposed method in terms of modelling the DCT of a drum sample. This implementation uses the DFT, however, in case the imaginary components prove useful. These signals were investigated for tom-tom drum samples.

2. THEORETICAL BASIS

2.1. Physics of Sound Production in Tom-toms

When a drumstick collides with the batter head of a drum, the associated initial vibrations produce the attack component of the drum sound. A two-dimensional travelling wave moves through the membrane, reflecting at the bearing edge, to form a standing wave, responsible for the sustain component of the drum sound. This contains normal modes (Fig. 2), each classified by their number of nodal diameters, m, and their number of nodal circles, n. This is written as (m, n), where the fundamental frequency is (0, 1). Modal frequencies can be calculated for an ideal circular membrane as follows:

$$f_{m,n} = \frac{x_{m,n}}{2\pi r} \sqrt{\frac{T}{\sigma}} \qquad (1)$$

where $x_{m,n}$ in equation (1) is the nth non-trivial zero of the mth-order Bessel function, T is the surface tension in Newtons/metre, σ is the surface density in kilogram per square metre, and r is the radius of the membrane in metres [19]. It should be noted that measured values can vary from their ideal values [20].

The amplitude of each mode is dependent on the strike position. Central strikes excite circular modes (such as the fundamental) due to their central antinode. Strikes that are closer to the rim, also excite radial modes, causing the characteristic overtones to be heard. Modal frequencies can be calculated for an ideal circular membrane as follows:

Uneven tuning can cause audible mode-splitting due to the degeneracies of modal oscillation [21]. The presence of a shell or a second membrane also complicates the system, as the components resonate in a coupled fashion [22]. This creates additional modes

and can suppress or accentuate existing modes. When a drum is struck at high intensity, the local tension in the skin increases. This leads to the characteristic pitch glide found associated with drum sounds [23]. Finite Difference models have modeled this effect using a non-linear term from the von Kármán equations for thin plates [24].

Resonant modes are clearly identifiable in spectrograms as well defined curves (Fig. 3), with visible pitch glide. These make up the sustain component of the drum sound. The remaining energy is less well defined, and makes up the attack component that is produced by the initial collision. The sustain component dominates the bass and low mids, while the attack component dominates the higher frequencies. These components could be considered separately, for example, modeling them as deterministic and stochastic, respectively.
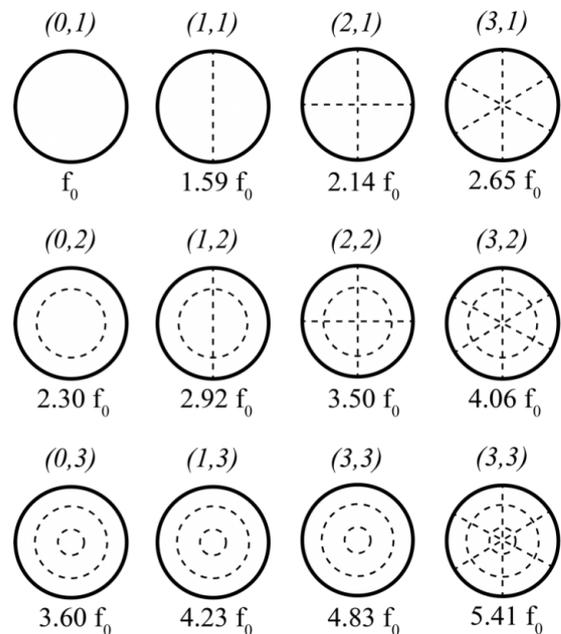
Figure 2: *Matrix of resonant modes of an ideal membrane, where (m,n) describes the number of nodal diameters, m, and the number of nodal circles, n.*
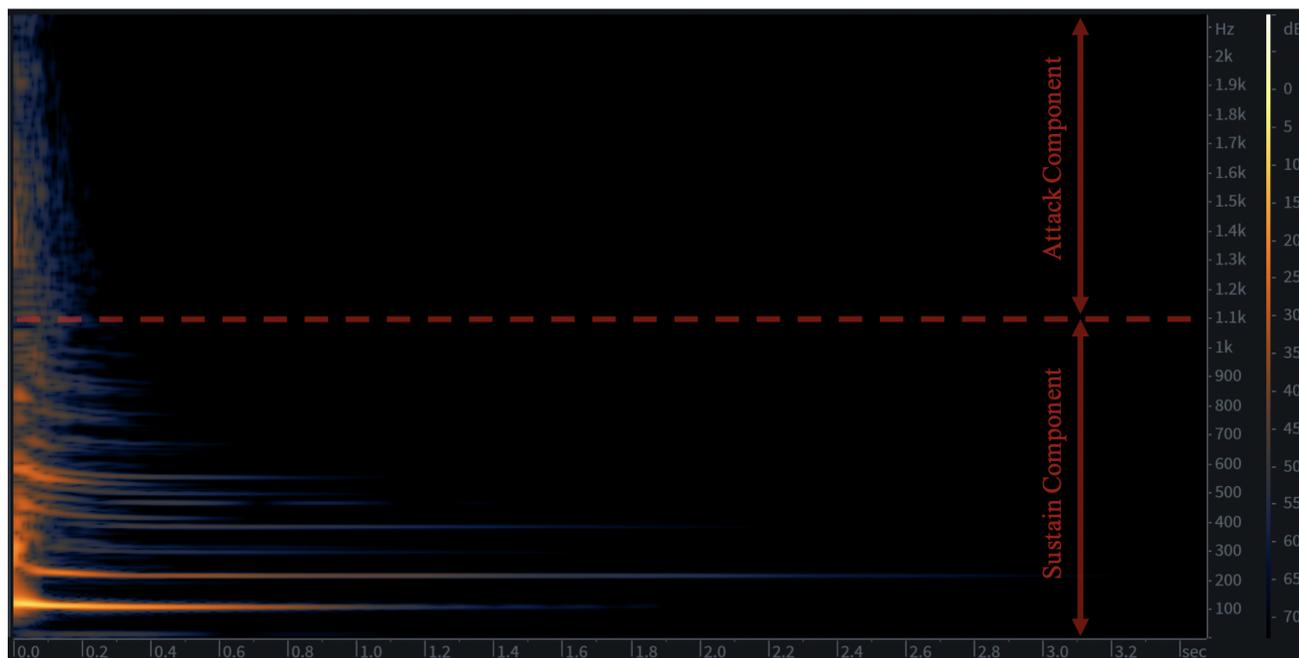
Figure 3: *Spectrogram of a close-miked 9x13" Ludwig Concert 3-Ply rack tom struck at high intensity, taken from Superior Drummer 3, and visualised in RX7. The dashed line indicates the frequency above which the attack component dominates the spectrogram. This was visually determined for illustrative purposes. It can be automated based on ridge detection and/or decay time.*

3. METHOD AND RESULTS

3.1. Data Set

416 close miked tom samples (44.1kHz) were taken from Superior Drummer 3, comprising 5 different models of kit (Yamaha Beech Custom, Pearl Masterworks, Premier Genista Birch Original, Ludwig 70's 3-ply Classic and Concert), 32 individual drums of various dimensions (including all common sizes), struck with drumsticks at 13 different velocities per drum. Each sample is 8s long, to ensure a full decay. This sample set provided a good selection of common tom sounds for analysis in MATLAB. While each drum has a unique sound, similar features were found across all drum types, in line with the physics of the system.

3.2. Modeling the Real Component of the DFT

The real component of the DFT (RDFT) of a drum sample is shown in Fig. 4. Although this is not a magnitude spectra, activity in the RDFT still indicates energy at a given frequency.

The linearity of the Fourier transform means that addition in the frequency domain is equivalent to addition in the time domain. This means that the RDFT can be viewed as the superposition of the (sparse) RDFT of each mode of the sustain component, and the RDFT of the attack component. These components were analysed independently across all drum samples, and consistent results were attained.

3.3. The Sustain Component

A distinctive feature is reliably observed in the RDFT at dominant modal frequencies. This is shown in Fig. 5, which is merely a zoomed in version of the RDFT of Fig. 4, along with the RDFT of three other samples of the same drum. These modal features did not overlap with any others, so no decomposition was required.

Each modal feature is a chirp-like frequency domain signal (Fig. 5), which encodes the entire time domain signal of a given mode, in up to 1000 times less samples (for an N-point FFT, where N is the number of samples per audio file).

The mapping between domains is best understood numerically:

1. Isolate the fundamental mode from a drum sample using a lowpass filter.

2. Calculate the Fourier transform contribution from each successive time domain sample, and inspect the effect that each has on the RDFT.

3. Notice that the initial samples correspond to a peak in the RDFT located at the tension modulated frequency. When successive samples contributions are introduced, the RDFT activity gradually shifts to lower frequencies, until the unmodulated frequency is reached, by which time a chirp signal is reliably obtained between these frequencies.

The width of the chirp signal, illustrated by the arrows on Fig. 5, corresponds to the magnitude of the pitch glide. The exact placement of the chirp on the frequency axis will determine the upper and lower frequency limits, which correspond to the initial tension-modulated frequency and the final unmodulated modal frequency respectively.

The exact trajectory of the pitch glide is encoded in the phase function, $\phi(x)$, that defines the chirp type (whether it is a linear chirp, an exponential chirp, or in this case, some other function).
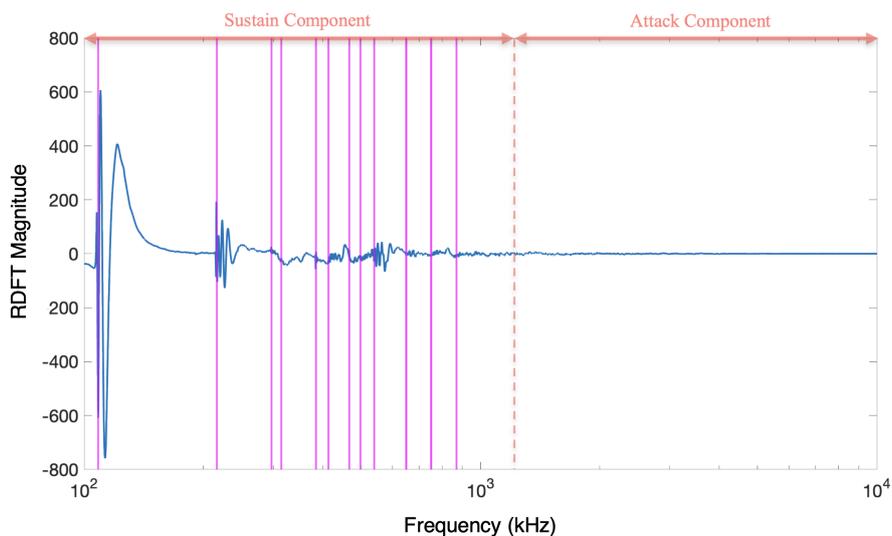
Figure 4: *RDFT of the active frequencies in a 9x10" Yamaha Beech Custom rack tom sample. The dashed line indicates the frequency above which the attack component dominates, determined as in Fig. 3. Solid vertical lines indicate unmodulated modal frequencies.*
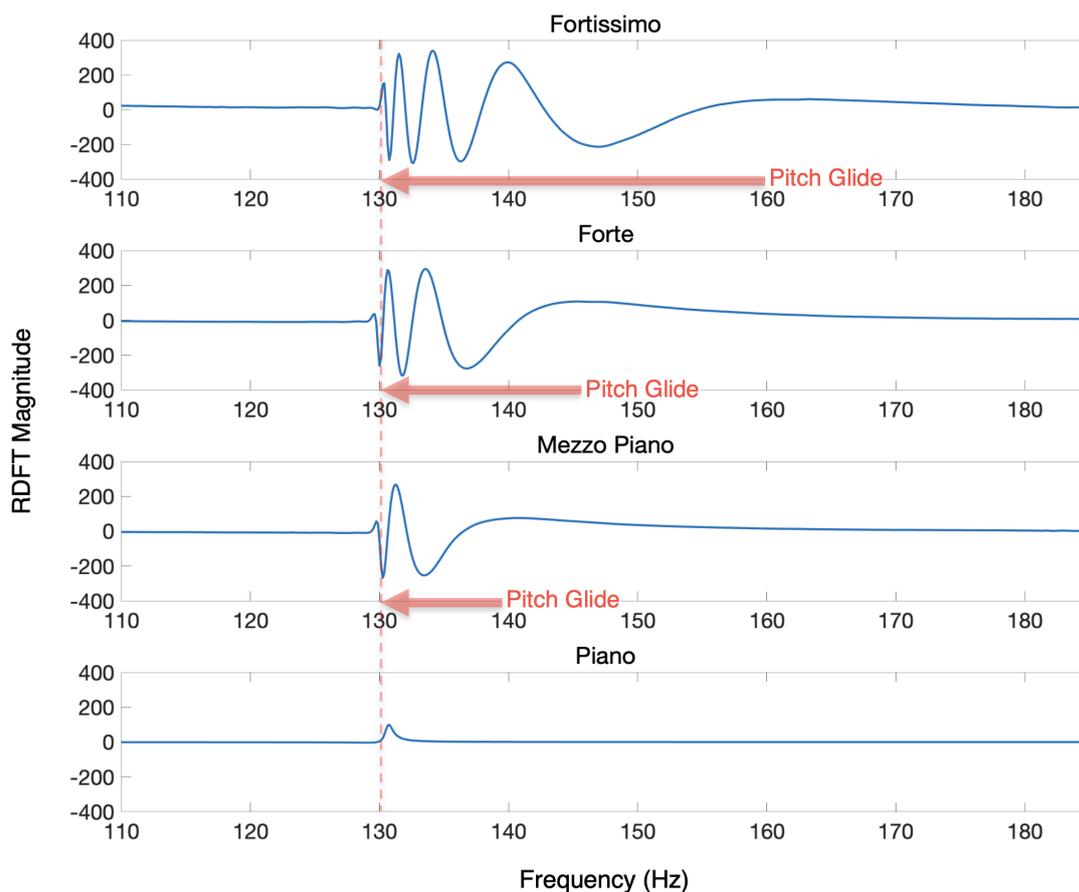

Figure 5: *RDFT representation of the fundamental mode of a 9x10" Yamaha Beech Custom rack tom, struck at four different dynamic levels. The dashed line indicates the unmodulated fundamental frequency of the drum, averaged across the four samples. The arrows indicate the direction and magnitude of the pitch glide, where present. This feature captures the full time-domain evolution of each mode; wider peaks correspond to earlier activity, and narrower peaks correspond to later activity.*
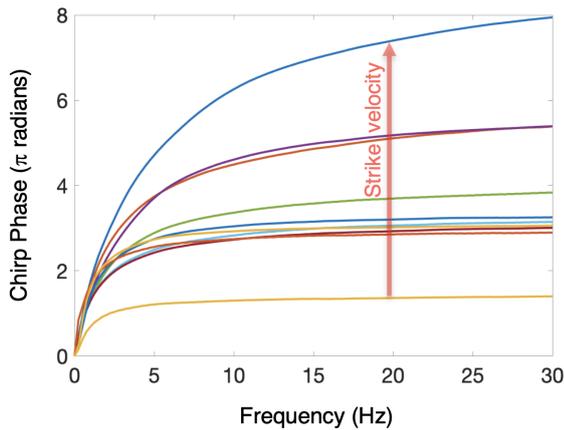
Figure 6: Plot of $\phi(x)$ (chirp phase as a function of frequency), for the RDFT chirp representation of the fundamental mode of a 9x10" Yamaha Beech Custom. Each curve corresponds to 1 of 13 samples with differing strike velocities. Frequency values have been shifted so that 0 Hz corresponds to the unmodulated fundamental frequency. The value of the horizontal asymptote of each curve corresponds to the number of peaks and troughs in the chirp. The point at which each curve levels off determines the width of the chirp. Higher strike velocities have more peaks and troughs, and a larger chirp width, as demonstrated in Fig. 5

The chirp is modeled in the frequency domain as:

$$y = A(x)\cos\left(\phi(x)\right) \qquad (2)$$

where y in equation (2) is the vertical axis that corresponds to RDFT magnitude, A is the amplitude envelope, and $\phi(x)$ is the phase function of the chirp, and x is the horizontal axis that corresponds to the RDFT frequency-domain.

The phase function can be extracted from a given chirp using the Hilbert transform:

$$\phi_{wrapped}(x) = \arg\left(H(y(x))\right) \qquad (3)$$

The phase values returned are wrapped to $-\pi > \phi > \pi$, but can be unwrapped to obtain a monotonic phase function, $\phi(x)$. MATLAB analysis determined $\phi(x)$ can be fitted, and well modeled with a linear rational function (as shown in Fig. 6):

$$\phi(x) = \frac{ax + b}{cx + d} \qquad (4)$$

Finally, $A(x)$ is a relatively simple envelope in the frequency domain, similar in shape to a positively skewed distribution function. Its peak value intuitively encodes the intensity of a given mode, and it can be shown experimentally that the exact shape of the envelope can subtly affect the time domain envelope of the mode.

While this may not initially be the most intuitive representation of a drum mode, it is complete, and as such, able to perfectly reproduce a drum mode, without any approximations. These modal features can be extracted and used for analysis, even if they overlap in the frequency domain. We can use our knowledge of the chirp shapes to separate overlapping modal features, through tactical filtering techniques, or by replacing the original chirps with
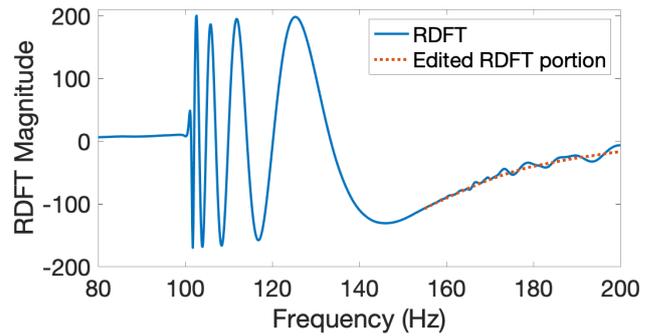

Figure 7: RDFT plot of overlapping modal features in a 7x12" Sugar Percussion Mahogany Stave sample (solid line). The large fundamental chirp spans $\sim$100-200 Hz, overlapping with the chirp representing next partial, which begins at $\sim$160 Hz. The dotted line shows the portion of the RDFT that has been edited through selective use of a moving average filter, to isolate the fundamental.
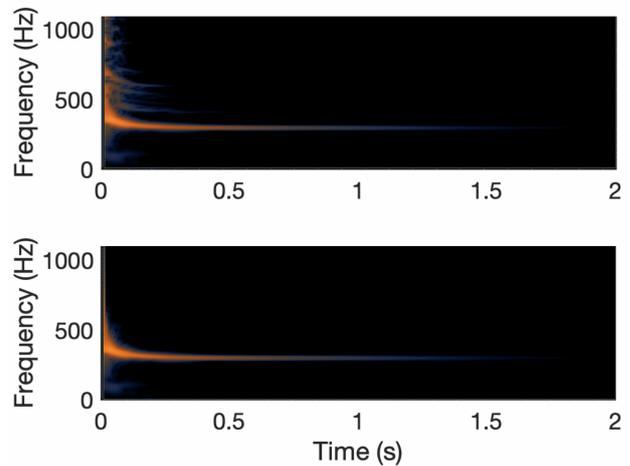

Figure 8: Spectrograms corresponding to the original RDFT from Fig. 7 (top), and the RDFT that has been edited in order to isolate the fundamental mode (bottom). The fundamental has been well isolated, without being altered, and no artefacts have been introduced.

modelled versions. In Fig. 7, a mode is isolated through selective use of a moving average filter. This exploits the specific way in which modal features overlap, as shown, and the fact that an isolated modal feature can be completely smoothed out. The modal features are superposed in the region of interest, and a smoothing filter can be used to remove the mode of higher frequency, isolating the lower mode. In the modeling approach, we would fit one modal feature (eg. the fundamental) with a chirp model, and then subtract the model RDFT from the reference, to isolate the remaining modes, and then repeat this process, for all desired modes. Once one sample has been modeled in high detail, the model can begin to be generalised. It is worth noting that certain modes are of much greater perceptual importance than others; initial tests have found that as few as 5-10 key modes could be sufficient to replicate convincing central strikes.
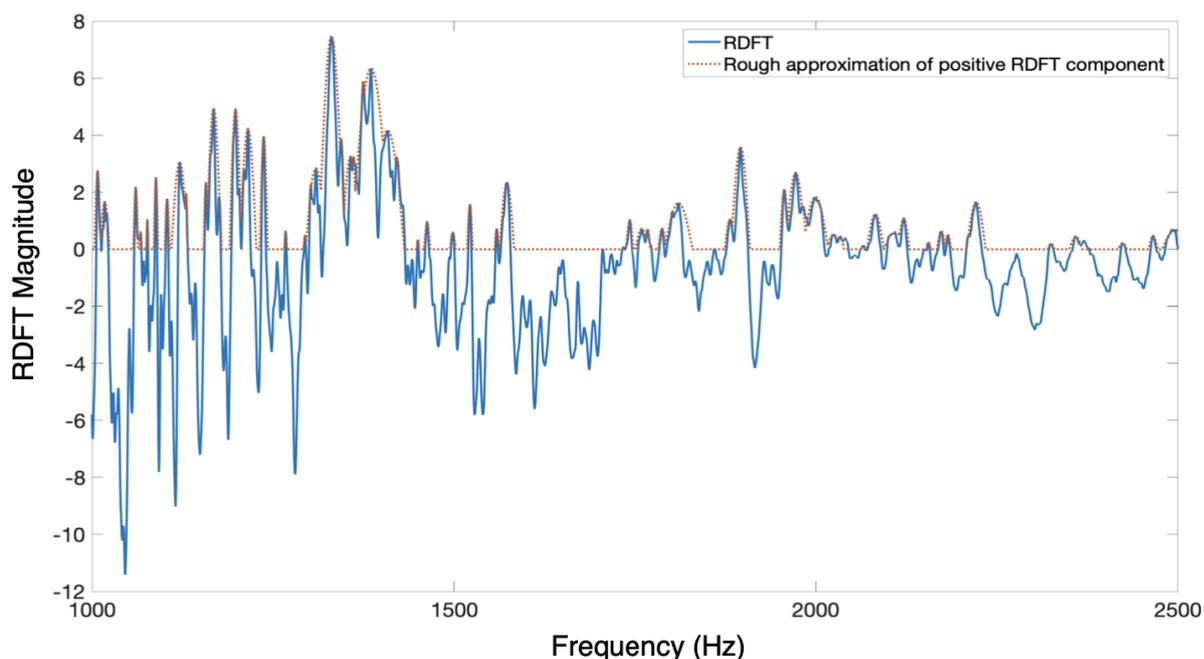
Figure 9: *RDFT of the 1-2kHz portion of the attack component for a 9x13" Ludwig Concert 3-Ply Tom struck at high intensity. The dotted line shows a rough approximation of the positive component of the RDFT, using a sinusoidal peak model. The width of the sinusoidal peaks have been widened slightly for illustrative purposes. The same method can be applied for the negative peaks and all peaks superposed for a full approximation of the attack component.*

These decomposition methods can isolate modal activity in a way that is not possible through mere bandpass filtering (Fig. 8), as you cannot use a band-pass filter to decompose sound components that overlap in frequency. This helps us to better decompose drum sounds, so that we can analyse the constituent parts more accurately, and better synthesise them. This synthesis can take place in the time domain, using extracted envelope and pitch trajectory information, or the chirp signals themselves can be modelled from scratch for ultra-realistic IFFT synthesis. The chirp signals can also be transformed to create non-physical sounds, such as mirroring the chirp to create an upwards pitch glide.

The fact that these phase functions have a shared form suggests that there is also commonality in the time-domain pitch glides themselves, which is in line with the physics of the system. The simple linear form allows phase functions to be well modelled, allowing for a continuous blending of strike velocity.

3.4. The Attack Component

The majority of drum synthesis models are focused on reproducing the sustain component of a drum sound, as this component of the sound has greater perceptual importance and is also better understood in terms of sound generation. The attack portion of a drum sound is rarely investigated or modelled in such detail, despite also being critical to a realistic drum sound.

Filtered noise bursts have been common for synthesis [25], in both research and music production contexts, but these are essentially a first-order approximation. There have been some attempts to model the stick-membrane interaction [26] using lumped collisions, but these models are not yet representative, nor practical for
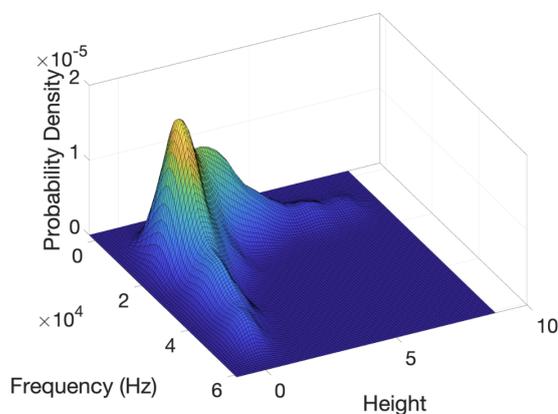

Figure 10: *Estimated probability density function for peak parameters from RDFT attack component model. This plot illustrates that we expect more peaks in regions of lower frequencies (within 1-8 kHz attack component limits for this 9x13" Ludwig Concert 3-Ply Tom sample) and that both average and maximum peak height decreases with frequency. This approach is more informative than a mere spectral envelope.*

real-time synthesis due to the extreme computational cost.

Here, we model the RDFT of the high frequency attack sound, using sinusoidal peaks. The attack portion of the RDFT is partially shown in Fig. 9. On visual inspection, there is little discernible

pattern to the peaky signal, other than a general decrease in amplitude with increasing frequency.

A better understanding of this signal is achieved by decomposing it into a series of positive and negative peaks. These peaks can be detected automatically, and parameterised by their height, width, and centre frequency. Once we return to the time domain, each peak corresponds to a quickly decaying partial, and their parameters encode the intensity, decay time, and centre frequency, respectively. It is found that this attack sound can be well synthesised by replacing the signal peaks with sinusoidal peaks of equivalent height, width, and centre frequency. This creates a reasonable approximation of the RDFT signal by eye (Fig. 9), which is sufficient to produce a synthesised attack sound that is perceptually very close to the original attack sound, and indistinguishable when combined with the membrane component in context.

Furthermore, the distributions of these peak parameters can be analysed for a specific drum sound (Fig. 10), so that novel attack sounds can be generated from these distributions, that have a similar character, but are intentionally distinguishable, to mimic the subtle variations in sound that one experiences when playing an acoustic drum set. The distributions themselves could also be edited, to model different striking tools, different drums, and even non-physical situations.

4. SYNTHESIS AND PERCEPTUAL EVALUTATION

4.1. Synthesis Proof of Concept

To provide an initial proof of concept for synthesis, a sound was part-synthesised using a chirp model, and another using a sinusoidal model. An ABX listening test was used to assess the perceptual realism of part-synthesised sounds. It is important to note that most (if not all) synthesised drum sounds are clearly distinguishable from their reference and as such would perform very badly in an ABX test (a scale based test such as a MUSHRA test would be more appropriate in those cases). The listening test was conducted using a Universal Audio Apollo Twin and Sennheiser HD800 headphones.

4.2. Scenarios and Results

5 drummers each performed 25 trials for each of 3 scenarios:

Sound examples available here

a) A sample from a 9x10" Yamaha Beech Custom rack tom, struck centrally, at maximum intensity was tested against an edited version, with a fundamental synthesised using the chirp model. The participants got only 62/125 correct (49.6%), so there is no evidence to reject the null hypothesis that the sounds are indistinguishable; the participants didn't even surpass the null expectation value of 50%.

b) A sample from a 15x16" Sugar Percussion Mahogany Stave floor tom, struck centrally, at maximum intensity, was tested against an edited version, where the high frequency content (>2kHz) was synthesised using a sinusoidal model. The participants got 67/125 correct (53.6%), this corresponds to a p-value of 23.7%, so we cannot reject the null hypothesis that the sounds were indistinguishable at even a 20% confidence interval. Even if the participants could occasionally distinguish the sounds, they clearly could not

do it at all reliably.

c) In order to illustrate the perceptual importance of the high frequency content, the reference drum sound from scenario b was tested against a low pass filtered version (2kHz). The filtered version can be viewed as a sort of anchor. The participants got 123/125 correct (98.4%), which is striking evidence to reject the null hypothesis, as expected in this scenario. It also shows that all participants engaged in the task and could listen critically enough to distinguish samples based on high frequency content.

These results are strong evidence that the part-synthesised sounds are highly realistic. The author is not aware of any drum synthesis methods that can produce truly convincing, indistinguishable results, and as such, this method may have the potential to outperform existing methods. Fully synthesised sounds will be generated and evaluated in future work to assess this claim.

4.3. Synthesis Engine Framework

The full RDFT model can be written as a single equation:

$$RDFT = \sum_{n=1}^{N} A_n(x - f_n) \cos\left(\phi_n(x - f_n)\right) \; + $$
$$\sum_{j=1}^{J} \Pi\left(\frac{1}{2}\left(\frac{x}{w_j} - f_j\right)\right) h_j \cos\left(\frac{\pi}{2}\left(\frac{x}{w_j} - f_j\right)\right) \tag{5}$$

where the top term in equation (**??**) is the sum of N modal chirps (as in equation (3)), but shifted along the frequency axis by f_n, the unmodulated frequency of a given mode). This sum corresponds to the sustain component. The bottom term is the sum of each sinusoidal peak in the attack model, where Π is the rectangular function, f_n is peak centre frequency, h_j is peak height $\pm$, w_j is half the full peak width.

These parameters can be obtained through analysis, physical modelling, or a hybrid of these methods, which is likely to be optimal.

A flexible general framework for synthesis would be as follows:

1. Calculate modal frequencies from equation 1, for a given membrane.

2. Calculate relative modal amplitudes based on strike position, by evaluating the membrane displacement of each modal surface, at that position. The relative amplitudes are scaled by strike intensity to determine the maximum values of $A_n(x)$.

3. Generate an RDFT chirp for each mode, using a modeled phase function, as in equation 5. (Multiple chirps could be generated from the same general phase function.)

4. Generate a series of sinusoidal peaks for the attack component of the RDFT, as in equation 5. This could be done deterministically through analysis of a reference sample, or a stochastic element could be introduced by using a probability density function.

5. Sum together all RDFT components as in equation 5, and follow the steps described in Figure 1.

5. CONCLUSION

It has been demonstrated that the RDFT of a drum sample reliably contains features that can be well modelled with chirps and peaks. This representation has been interpreted and found to be useful, with RDFT model parameters mapping well onto key performance parameters such as strike velocity, strike location and drum size. This method is therefore thought to provide an interesting parameter space in which to analyse, decompose, and ultimately synthesise drum sounds, as demonstrated in this research.

6. FUTURE WORK

This method can provide the basis of a realistic synthesis engine. Further analysis and development will be required in order to best construct and parameterise the engine, which can then be triggered by appropriate hardware and evaluated through listening tests and user feedback.

7. REFERENCES

[1] M. Collins, Ed., *A Professional Guide to Audio Plug-ins and Virtual Instruments*, Focal Press, New York, USA, 2003.

[2] A. Brown, Ed., *Music Technology and Education*, Routledge, New York, USA, 2014.

[3] Toontrack, "Superior Drummer 3," Available at https://www.toontrack.com/superior-line/, accessed April 18 2020.

[4] J. Risset and D. Wessel, *The Psychology of Music*, chapter Exploration of timbre by analysis and synthesis, pp. 37–39, Academic Press, New York, USA, 1982.

[5] J. Hsu T. Smyth and R. Done, "Toward a real-time waveguide mesh implementation," in *Proceedings of the 2017 International Symposium on Musical Acoustics*, Montreal, CA, June 2017, pp. 54–57.

[6] J. Laird, *The Physical Modelling of Drums Using Digital Waveguides*, Ph.D. thesis, University of Bristol, 2001.

[7] X. Serra, *A system for sound analysis/transformation/synthesis based on a deterministic plus stochastic decomposition*, Ph.D. thesis, Stanford University, 1989.

[8] K. Karplus and A. Strong, "Digital synthesis of plucked-string and drum timbres," *Computer Music Journal*, vol. 7, no. 2, pp. 43–45, 1983.

[9] S. Bilbao and C. Webb, "Physical modeling of timpani drums in 3d on gpgpus," *J. Audio Eng. Soc*, vol. 61, no. 10, pp. 737–748, 2013.

[10] F. Avanzini and R. Marogna, "A modular physically based approach to the sound synthesis of membrane percussion instruments," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 18, no. 4, pp. 891–902, 2010.

[11] R. Marogna and F. Avanzini, "Physically-based synthesis of nonlinear circular membranes," in *Proceedings of the 12th International Conference on Digital Audio Effects*, Como, Italy, Jan 2009, pp. 373–379.

[12] L. Trautmann, S. Petrausch, and R. Rabenstein, "Physical modeling of drums by transfer function methods," in *2001 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings*, Salt Lake City, USA, May 2001, vol. 5, pp. 3385–3388 vol.5.

[13] S. Bilbao, "Ness project, 3d percussion," Available at http://www.ness.music.ed.ac.uk/archives/systems/3d-embeddings, accessed March 08, 2006.

[14] V. Zappi et al., "Shader-based physical modelling for the design of massive digital musical instruments," in *NIME*, 2017.

[15] H. Chambelin, *Musical Applications of Microproccessors*, Haydens Books, New York, USA, first edition, 1980.

[16] X. Rodet and P. Depalle, "Spectral envelopes and inverse fft synthesis," in *Audio Engineering Society Convention 93*, Oct 1992.

[17] S. So and K. K. Paliwal, "Reconstruction of a signal from the real part of its discrete fourier transform," *IEEE Signal Processing Magazine*, vol. 35, no. 2, pp. 162–174, 2018.

[18] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Transactions on Computers*, vol. C-23, no. 1, pp. 90–93, 1974.

[19] S. Errede, "Vibrations of Ideal Circular Membranes (e.g. Drums) and Circular Plates," Available at https://courses.physics.illinois.edu/phys406/sp2017/Lecture_Notes/P406POM_Lecture_Notes/P406POM_Lect4_Part2.pdf, accessed April 18, 2020.

[20] E. Skrodzka et al., "Vibroacoustic investigation of a batter head of a snare drum," *Archives of Acoustics*, vol. 31, pp. 289–297, 2006.

[21] R. Worland, "Normal modes of a musical drumhead under non-uniform tension," *J. Acoust. Soc. Am.*, vol. 127, no. 1, pp. 525–533, 2010.

[22] S. Bilbao, "Time domain simulation and sound synthesis for the snare drum," *J. Acoust. Soc. Am.*, vol. 131, no. 1, pp. 914–925, 2012.

[23] F. Avanzini and R. Marogna, "Efficient synthesis of tension modulation in strings and membranes based on energy estimation," *J. Acoust. Soc. Am.*, vol. 131, pp. 897–906, Jan. 2012.

[24] A. Torin and M. Newton, "Nonlinear effects in drum membranes," in *ISMA*, Le Mans, France, 01 2014.

[25] J. Hsu and T. Smyths, "Percussion synthesis using loopback frequency modulation oscillators," in *SMC*, Malaga, Spain, 2019.

[26] A. Torin et al., "An energy conserving finite difference scheme for the simulation of collisions in snare drums," in *Proceedings of the 14th International Conference on Digital Audio Effects*, Erlangen, Germany, 01 2014.