WAVESET TRANSFORMATIONS: SOURCE CHARACTERISTICS AND TRANSFORMATION PECULIARITIES PRODUCING HARDLY PREDICTABLE SOUND RESULTS

Fellipe MARTINS (fmartins@ufmg.br) (0000-0002-2686-6203) [1] and
José Henrique PADOVANI (jhp@ufmg.br) (0000-0002-8919-7393) [1]

[1] *School of Music*, **Federal University of Minas Gerais (UFMG)**, MG Brazil

ABSTRACT

The wavesets are defined as a signal portion between three consecutive zero-crossings and were proposed by Trevor Wishart as a form to manipulate and transform sound files. For simple sounds, wavesets tend to coincide with the wave cycle, but for complex and polyphonic sounds the wavesets can represent a small portion of longer oscillations or retain a slow oscillation superposed by faster ones. As wavesets' shape, duration and amplitude are strongly unique for each sound, modifying a waveform based on this granular criterion can lead to unique sonorities. Firstly, we discuss the main issues that cause highly different sound results when applying the same waveset transformation to different waveforms (e.g., DC, low-frequency content, beats, phase, and others), secondly, we selected several waveset transformations to evaluate the range of sonorities produced as well as contextualizing them in the field of well-known digital audio effects. In the end, we evaluate the major pitfalls of working with this technique and report some enhancements that are currently being studied.

1. INTRODUCTION

Wavesets were proposed by Trevor Wishart [1] as a criterion to execute granular transformations on recorded audio samples. A waveset is defined as the signal portion between three consecutive zero-crossings, which corresponds to a full cycle on a purely sinusoidal oscillation. This definition can be interpreted as an attempt to select grains whose edges tend to zero thus reducing the introduction of high-frequency content due to discontinuities and fast transitions, in a similar fashion to the windowing process used on more traditional granular transformations [2].

The first computational implementation of waveset transformations was released on the Composers Desktop Project (CDP) [3] – a suite of functions for transforming audio files in a non-realtime fashion. As the method requires an analysis stage to find three zero-crossings and several of the transformations deal with the combination of a group of consecutive wavesets, real-time implementations entail a reasonable amount of delay. Real-time implementations available include a VST plugin implementation [4], a Csound operator [5], and several using the SuperCollider environment. The first SuperCollider implementation was made by de Campo [6, 7] as a Quark that analyzes a buffer in non-realtime and then applies transformations in a real-time fashion. An updated version of this Quark was made by de Campo, Bovermann and Rohrhuber [8]. There are also four other implementations Hochherz [9], Nishino [10], Mayer [11] and Seidl [12], the latter working exclusively in real-time.

Waveset transformations are categorized on the CDP as "distortion", a term with broad meaning regarding sound results, ranging from the loss of information (bitcrushing) to the increase of harmonic content through clipping. There is a brief discussion about the wavesets usage, source considerations, sound results, and predictability on [2, 3], but we could not find elsewhere an extensive discussion that aiming the musical usage. The closest related aspects to this topic, which an abundant amount of works can be found, are the evaluation of zero-crossing rate for time series analysis and stochastic processes [13], pitch, noisiness, voice/unvoiced signal detection [14, 15].

2. WAVESETS CONTENT

For stationary harmonic signals, the content of a waveset depends on the number of harmonics, their intensity, and their phase relationship. The first two are well known for defining the perception of timbre, pitch, and beats, however, for signals that have the same number and intensity of harmonics, the phase relationship between them change the wave shape drastically, but it does not interfere on the auditory perception of pitch and timbre, see Fig. 1. As a consequence, when executing waveset transformations on different waveforms that sound the same, it is possible to obtain utterly different graphical and sound results.

As stated by the Fourier transform, a general shape signal can be decomposed as a sum of sinusoidal waves. Therefore we might investigate how the properties of adding two elementary oscillations affect the content of a waveset. As a consequence of a trigonometric identity, the sum of two oscillations is identical to the product of two other oscillations: one proportional to the sum of the frequencies and the other proportional to the difference. Perceptually, the addition of two sine waves can result in four different auditory phenomena: beats, roughness, timbre, and polyphony. In terms of the waveset content, each case will present its

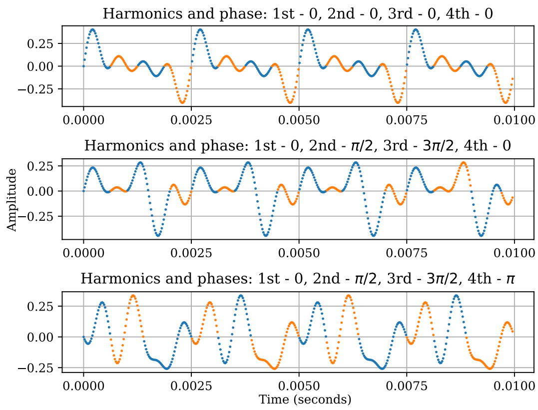

Figure 1. Three tones with four harmonics in the same frequency relationship, same amplitude but different phase relationships.

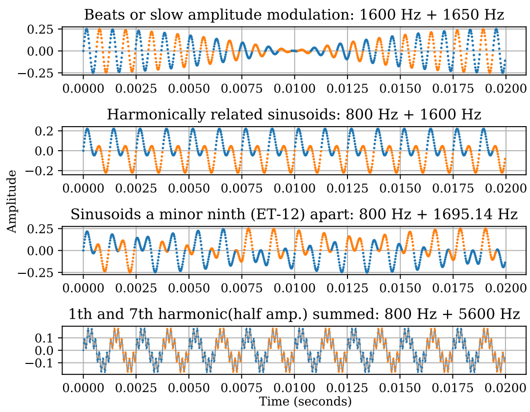

Figure 2. Four auditory effects of sinusoidal addition: beats, harmonically related sinusoids, polyphony and amplitude modulation.

own particularities.

When beats or slow amplitude modulations occur Fig. 2, the content of each waveset tends to retain the carrier shape (i.e. the high-frequency content). However, the slow amplitude variations will change the amplitude of each waveset over time and overmodulation can promote more complex patterns. In the case of harmonically related sinusoids, we will have a constant waveset content within the fundamental period Fig. 2, but there will be faster fluctuations inside this period. As a consequence, the wavesets will lose the characteristic shape of sinusoidal cycles and a full fundamental period will be a compound of more than two wavesets. The third case on this figure shows that when the oscillations of two sinusoids are not harmonically related (e.g. polyphonic signals) the wavesets' content and period will vary over time in a fashion that is strongly dependant on the frequency content of its derivated waves. The last example shows the summation of two highly spaced frequencies, which can be perceived as a single pitch with a complex timbre or as two

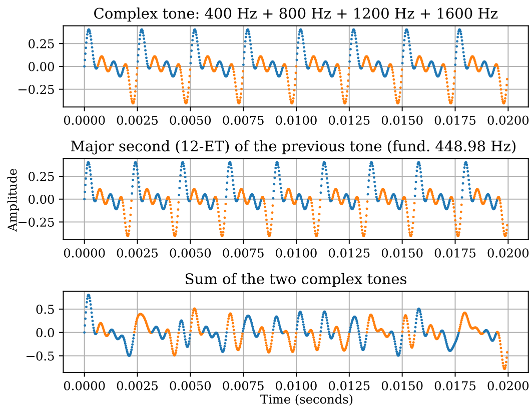

Figure 3. Two complex tones, a 200 cents apart, and the summation of both.

different tones played together. Graphically, we will have a situation of small period wavesets pushed upwards by a slower oscillation.

The last situation described can occur quite often, due to DC offset and non-audible low-frequency content presence in the analyzed signal. In this case, a single waveset may hold a considerable amount of high-frequency content although retaining a long waveset period (e.g. an exaggeratedly long waveset full of high-frequency content displaced upwards due to a very slow frequency component).

One step further regarding complexity, when evaluating the resulting wavesets formed by two spectrally complex tones, separated by a major second and played together, as displayed in Fig. 3, it is possible indeed to recognize its deterministic structure, although one might image several problems related with the transformation of such variable segments. Here, the wavesets' graphical content gets strongly complex and this example, however, is far from the complexity of recorded sounds, illustrating how the transformation of wavesets faces intricate issues. All these temporal variations on the waveset content will increase complexity when basic sound phenomena and properties like energy envelope, reverberation, background noise, etc, are present on the contemplated signal.

3. DISTORT TRANSFORMATIONS

3.1 Overall

As shown in Table 1, the CDP classifies waveset transformations under the umbrella term "distortion" [3]. In general terms, the distort functions available on the CDP either modify the wavesets order (shuffle, repeat, omit) or modify their content (average, clip, etc).

These functions were mainly named "distort" because, in general terms, they introduce high-frequency content and noise to the signal, which is mainly perceived as distortion. Although this nomenclature can be meaningful for sound sources like human-voice and musical instruments, in which radical modifications of the content are perceived by the listener as a degradation of the signal message (or

Function	Description
Average	Average the waveshape over N 'wavecycles'
Cyclecnt	Count 'wavecycles' in soundfile
Delete	Time-contract soundfile by deleting 'wavecycles'
Divide	Distortion by dividing 'wavecycle' frequency
Envel	Impose envelope over each group of cyclecnt 'wavecycles'
Filter	Time-contract a sound by filtering out 'wavecycles'
Fractal	Superimpose miniature copies of source 'wavecycles' onto themselves
Harmonic	Harmonic distortion by superimposing 'harmonics' onto 'wavecycles'
Interact	Time-domain interaction of two sounds
Interpolate	Time-stretch file by repeating 'wavecycles' and interpolating between them
Multiply	Distortion by multiplying 'wavecycle' frequency
Omit	Omit A out of every B 'wavecycles', replacing them with silence
Overload	Clip the signal with noise or a (possibly timevarying) waveform
Pitch	Pitchwarp 'wavecycles' of sound
Pulsed	Impose regular pulsations on a sound
Reform	Modify the shape of 'wavecycles'
Repeat	Timestretch soundfile by repeating 'wavecycles'
Repeat 2	Repeat 'wavecycles' without time-stretching
Replace	The strongest 'wavecycle' in a cyclecnt group replaces the others
Replim	Timestretch by repeating 'wavecycles' (below a specified frequency)
Reverse	Cycle-reversal distortion in which the 'wavecycles' are reversed in groups
Shuffle	Distortion by shuffling 'wavecycles'
Telescope	Time-contract sound by telescoping N wavecycles into 1

Table 1. Description of distort functions in the CDP Documentation [3].

distortion), it can be inaccurate or misleading when dealing with electronic sounds and some instruments like guitars, percussion, electric pianos, etc. For this reason, the documentation sometimes uses the term "constructive distortion" for referring to distinct sound results obtained when using these functions. In the following sections, we are going to delineate some of the broad results covered by the category "distortion".

The identification of wavesets is done through zero-crossings counting. However, the zero-crossing rate is also an indication of *pitchness* and *noisiness* of the signal. When the zero-crossing rate is stable at a low value, we have a rough indication that the signal holds a steady pitch characteristic, on the contrary, when the zero-crossing rate is high (being stable or not in this case) it strongly indicates that we have the presence of noise, harshness and signal instability. In this way, a criterion for selecting the minimal waveset length usually helps to avoid operating on the noise content of the sound file and reduces the introduction of harsh and noisy characteristics on the transformed signal.

3.2 Distortion and non-linearity

The term distortion is more commonly used by the audio community as a synonym of harmonic distortion, when a system applies a non-linear [1] curve to its inputs, especially clipping and soft clipping caused by amplifiers, filters, valves, pedals, and other devices, introducing harmonics which were not present in the original signal [16].

Another use for the term *distortion* in the audio community refers to procedures in which information is lost or degrade. Two commonly used types of this effect are sample-rate reduction (samples are discarded thus reducing the bandwidth) and resolution reduction (amplitude quantization bits are discarded).

Besides these two main types, several other digital audio effects operate with strongly non-linear characteristics: compressors, limiters, de-essers, analog simulators (tape, valve, amps, etc), exciters, enhancers, etc. Purely linear conditions are often a too hard constraint, and some distortion is always expected in any kind of audio system.

3.3 Wavesets transformation introducing non-linearities

Operations with wavesets are essentially non-linear procedures because each waveset is going to be processed individually, so each block is subjected to the same wave manipulation, independently of its content. Instead of uniformly transform the signal, sample-by-sample, we are introducing discontinuities due to the block processing. As a frequent result, this may either generate subsequent wavesets with amplitude gaps on their transitions or may produce radically different waveset content varying at a higher rate – which can be perceived as noise.

As a simple test for this second effect, we produced a sinewave with two different amplitude values for each subsequent waveset, therefore each full wave-cycle would be alternating between these two amplitude values. Under low background noise conditions, for a frequency of 440 Hz when the amplitude difference between the two wavesets was greater than around 0.17dB, it is possible to hear the introduction of new harmonic content. This provides a rough parameter about how sensible our auditory system is to modification in such short fragments.

3.4 Other artifacts introduced

One strategy to avoid the previous problem is to operate on a group of wavesets, therefore introducing discontinuities

[1] A system is called linear if it satisfies the property of superposition, expressed in $Ax_1(n) + Bx_2(n) \rightarrow Ay_1(n) + By_2(n)$ in which two given inputs x_1 and x_2 produce the output y_1 and y_2 with its correspondent scalar multiplication.

at a lower rate and reducing the high-frequency content introduction. Furthermore, another option is to execute modifications that are based on the rearrangement of the wavesets, instead of modifying their shape. These strategies partially solve the "distortion" problem but usually introduce other artifacts, especially a mechanical repetition which is perceived as grains of steady pitch or unnatural periodicity. In general, these strategies do not consider the long-term variations of the sound wave (envelope, vibrato, tremolo, *allure*) and tend to break them cyclically, producing repetitions.

Another procedure suggested on the CDP documentation [3] is to "filter" the application of a waveset transformation, which consists of applying transformations only on wavesets of a certain length. As the random and the harmonic part of the recorded signal are generally fully blended, this procedure leads to utterly unpredictable sound results (or from another point of view, leads to results that are strongly source-dependent). In some cases, this process can modify only a portion of a given signal while, in other cases, the resultant sound may consist of sections in which transformed and non-transformed blocks alternate.

Summarizing, the transformation of wavesets – particularly when made without considering other temporal parameters or without applying any smoothing correction – tends to be highly non-linear, introducing several audible artifacts. Moreover, numerous factors modify the wavesets' shape, duration, and rate of recorded sounds (reverberation, DC component, microphone position, intrinsic source characteristics, etc). As a result, the wavesets transformations presented here tend to produce highly unpredictable and source-dependent sound results.

3.5 Traditional approaches

Typical approaches for managing this issue involve some form of overlap and add (OLA) using tapering windows. The most related method is the pitch-synchronous overlap and add (PSOLA), which analyses the pitch of a signal to place the window so that it fits the pitch wavecycle [16]. Therefore, PSOLA is strongly dependent on the quality of its pitch detection algorithm as well as on the degree of pitch stability, noisiness, and inharmonicity. This method is used mainly for pitch shifting and time scale modification, although for more radical transformation, like those proposed by the CDP, it is not commonly used.

4. TRANSFORMATION CASES

4.1 Time Stretch

Waveset-based time-stretching transformations are made by repeating either a single or a group of wavesets. Sonically, the result is far from the traditional harmonic distortion and can lead to peculiar effects. When single wavesets are repeated in a small amount it can range from an introduction of a sound similar to rubber friction. As the repetition rate increases, the static pitch of each waveset emerges. When a group of wavesets is considered, details of the inner structure of the sound file (e.g. like grains, *allures*, small pitch variations) tend to be revealed due to their repetition. It can also produce steady pitches, but as the reproduction rate tends to not be transformed, it is more common to result in repetitions akin to short loops. Further comparisons and audio examples can be found in [17].

4.2 Wavesets alternation (substitution)

When alternating wavesets between two audio files (CDP's *distort reform* function), especially when working with bigger groups of wavesets, it is possible to obtain some unique mixtures between two sounds. This can be used to emulate the phenomena like multiphonics and crosstalk [17]. Moreover, it is a form of achieving multiphonics-alike sounds for instruments that do not easily allow this technique or to produce cross-over sounds made of wavesets extracted from distinct sources. Additionally, this technique produces several amplitude modulation effects which are strongly source-dependent.

4.3 Some effects on vocal sources

The resultant sonorities from the application of wavesets transformation on vocal sources are remarkably varied and cannot be exhausted. Here we inevitably enter the realm of speech perception, which brings a big amount of new topics to the investigation. To point out a few aspects, three CDP functions promote utterly different results when compared with harmonic distortion: *distort pitchwarp*, *distort replace*, and *distort average*. The first introduces a quaver quality to the voice, approximating it to the vocal characteristics of elder speakers. The second increases this approximation by adding a rough/harsh quality. Moreover, it also introduces some features which allude to a speaker with breathing difficulties. The last function is the one that is more similar to harmonic distortion, although it can be more related to the production of a hoarse voice than to vocal overdrive [18].

4.4 *Distort harmonic* and instrumental usage

Some waveset transformations can operate mainly emphasizing qualities of the sound sources, rather than changing or destroying them. The *distort harmonic* function acts like an additive synthesizer, summing the content of a waveset with its correspondent *waveset harmonic* – that is, integer multiples of that waveform. The main sonic result of this procedure is similar to a resonant filter, enhancing and focusing some spectral information as well as promoting sometimes the sensation of pitch shifting. When tested with orchestral instruments, the results could be aurally related to the change of instrument materials (e.g. wood marimba to glass marimba) or to the effect of highlighting amplitude modulations (e.g. more intensity on the vibrato present at the attack time).

5. CONCLUSIONS

Waveset transformations allow artists to explore constructive distortions and numerous other artifacts as creative resources. Furthermore, it also enables new types of combinations and mixtures between different sounds.

In general, few computer music resources provide waveset transformations, therefore there is still plenty of space for new improvements and propositions. The transformation of wavesets, although seems a straightforward task, demands detailed and careful operations to avoid the excessive introduction of high-frequency content and noise. We could not find in our bibliographical review a criterion or a curve-fitting method to reduce noise, harshness, and distortion. As the proposed auditory experiment showed, it is possible to outline some limits, but more complex transformation strategies still need to be tested and discovered.

6. FURTHER DEVELOPMENTS

Currently, we are developing a Python-based library that implements waveset functions and techniques. The main purpose of the library is to provide regular CDP transformations along with new strategies of waveset manipulations. Moreover, we are investigating the usage of different audio descriptors to inform the waveset transformations by means of audio features.

A second strategy, that has shown interesting initial results, is the application of waveset processes in conjunction with other DSP pre-processing and transforms (FFT, DCT, etc) as an intermediate step of analysis/re-synthesis procedures.

7. REFERENCES

[1] T. Wishart, *Audible Design: A Plain and Easy Introduction to Sound Composition*. Orpheus The Pantomime Ltd., 1994. [Online]. Available: http://www.trevorwishart.co.uk/AuD.html

[2] C. Roads, *Microsound*, 1st ed. The MIT Press, 2001.

[3] C. D. P. L. (CDP). (2015) CDP DISTORT Functions. CDP Documentation Home Page. [Online]. Available: http://www.ensemble-software.net/CDPDocs/html/cdistort.htm

[4] M. Norris. (2014) SoundMagic FX - Effect Documentation. [Online]. Available: https://web.archive.org/web/20140113044635/http://www.michaelnorris.info/soundmagic/effects.html

[5] J. ffitch. (2001) Waveset. Orchestra Opcodes and Operators. [Online]. Available: http://www.csounds.com/manual/html/waveset.html

[6] A. de Campo, "Wavesets," SuperCollider Quarks, 2020. [Online]. Available: https://github.com/supercollider-quarks/Wavesets

[7] ——, "Microsound," in *The SuperCollider Book*, 1st ed., ser. The MIT Press, S. Wilson, D. Cottle, and N. Collins, Eds., 2011. [Online]. Available: https://mitpress.mit.edu/books/supercollider-book

[8] A. de Campo, T. Bovermann, and J. Rohrhuber, "WavesetsEvent," musikinformatik, 2020. [Online]. Available: https://github.com/musikinformatik/WavesetsEvent

[9] O. Hochherz, "SPList, a Waveset synthesis library and its usage in the composition "draussen"," 2008, p. 6. [Online]. Available: http://lac.linuxaudio.org/2008/download/papers/19.pdf

[10] H. Nishino, "LC: A Mostly-strongly-timed Prototype-based Computer Music Programming Language that Integrates Objects and Manipulations for Microsound Synthesis," 2014. [Online]. Available: https://scholarbank.nus.edu.sg/handle/10635/78945

[11] D. Mayer, "miSCellaneous_lib," 2020. [Online]. Available: https://github.com/dkmayer/miSCellaneous_lib

[12] F. Seidl, "Granularsynthese mit wavesets für live-anwendungen," 2016. [Online]. Available: https://www2.ak.tu-berlin.de/%7Eakgroup/ak_pub/abschlussarbeiten/2016/Seidl_MasA.pdf

[13] B. Kedem, "Spectral analysis and discrimination by zero-crossings," vol. 74, no. 11, pp. 1477–1493, 1986.

[14] R. Bachu, S. Kopparthi, B. Adapa, and B. Barkana, "Voiced/Unvoiced Decision for Speech Signals Based on Zero-Crossing Rate and Energy," in *Advanced Techniques in Computing Sciences and Software Engineering*, K. Elleithy, Ed. Springer Netherlands, 2010, pp. 279–282.

[15] F. Gouyon, F. Pachet, and O. Delerue, "On the use of zero-crossing rate for an application of classification of percussive sounds," 2000. [Online]. Available: files/publications/dafx00-gouyon.pdf

[16] U. Zolzer, Ed., *DAFX: Digital Audio Effects*, edição: 2nd ed. John Wiley & Sons, 2011. [Online]. Available: http://dafx.de/DAFX_Book_Page_2nd_edition/index.html

[17] F. Martins, "Estudo exploratório de processos de transformação sonora a partir de trevor wishart: reinvenção e tradução para o ambiente supercollider," 2020. [Online]. Available: https://musica.ufmg.br/lapis/?p=1141

[18] A. Endrich. (2016) CDP Tutorial Workshop 1. CDP Tutorial Workshop 1. [Online]. Available: https://www.composersdesktop.com/workshops.html