MWD: MULTIBAND WAVESHAPING DISTORTION

Pablo Fernández-Cid (*) — Univ. Europea de Madrid, Dpto. Electrónica y Teoría Circuitos, 28670, Villaviciosa de Odón, España, pablo@gaps.ssr.upm.es

Javier Casajús Quirós — Univ. Politécnica de Madrid, ETSIT, Ciudad Universitaria s/n, 28040-Madrid (SPAIN), javier@gaps.ssr.upm.es

Pedro Aguilar — pedro.aguilar@iece.es

(*) Communications about this paper should be addressed to Pablo Fernández-Cid

ABSTRACT

A new architecture for musical distortion is proposed. Based on WaveShaping as the distortion generation element, a multiband front-end is used in order to extract simple (ideally monotonal) non-full band signals. Distortion pattern can be adjusted per band, with the benefit that intermodulation distortion is kept low, balancing the end result towards 'harmonic' distortion rather than 'metallic' or 'ringing'. Many parameters can be adjusted by the end user in a musically meaningful domain, allowing the creation of rich, detailed and highly personal sounding distortions to be imposed over real world signals.

1. INTRODUCTION

Many audio effects have already been successfully translated from their original analog implementations into digital incarnations, with well known benefits in terms of quality, programmability and repetitiveness. This is particularly true for linear systems (like delay, reverb, chorus, flanger, etc.) for which a large corpus of design theory exists and has been applied for years in many different fields.

On the other hand, heavily non-linear units are far less well known. There is still a kind of magic surrounding some analog distortion units that fails to be captured on their digital counterparts.

Some commercial multieffect units include distortion algorithms that sound dull and rough, particularly when applied to non-monophonic signals (inharmonic distortion is easier to perceive and usually less pleasant). These units can not devote too much processing power for the distortion generator and use very simple techniques.

Other devices are fully devoted to distortion recreation. This is particularly true for guitar amp simulators, where some degree of so-called modelling of original devices has been applied, demanding much more processing power but achieving a more pleasant sound.

In either case, the user of available systems has little control about the distortion quality, basically the 'amount' and 'colour' parameters, that are too general and do not allow careful, detailed adjustment of distortion according to the musician's particular needs and taste. But distortion needs not to be a black-box untweakable system.

Our proposal uses processing power not to model 'classical' milestones of distortion, but instead to build up a flexible distortion engine. Musicians can tweak parameters in order to 'synthesize' the distortion pattern they are looking for (from a resemblance of a classical unit to a completely new one).

2. DISTORTION BASICS

'Distorted' is not necessary a pejorative adjective when applied by a musician. Playing a central role for some musical styles, distortion is also used sometimes for subtle transformations that enhance expressiveness or help to stand up a sound out of a mix.

A lot of debate continues about what makes a good distortion unit or what makes distortion 'A' sound better than distortion 'B' (lets say two different valves, for example).

While not trying to compete in the field of valve emulators or detailed old equipment response modelling, we feel that a highly programmable distortion system would be welcome, particularly if the programming model uses musically meaningful parameters, easy to adjust by musicians.

Our objective then can be defined as a 'distortion synthesizer'.

Distortion applied to musical instruments generates upper partials or harmonics from lower ordered ones. In the same way that spectral shape is linked to timbre, the relative heights of the newly generated partials define very different final distorted results.

It is also known (though usually not modelled due to the increase in system complexity) that distortion pattern is not the same for low level and high level signals, usually 'hot' signals include greater distortion than lower ones.

In a similar way, the distortion pattern seems to vary with the frequency of the original signal, in such a way that a real world signal (with many partials) sees each of its harmonics subject to a different distortion.

Finally, distortion applied on multi-instrument mixes tends to blur the whole signal, creating a wall-of-sound where each instrument looses its identity and can not be heard in isolation. The fact that distortion generates intermodulation partials is largely the responsible for this.

We can resume our goals for a good programmable distortion:

a) Allow the user to define the exact shape of distortion in terms of height of newly generated harmonics
b) Allow different shapes according to signal level
c) Allow different shapes for different spectral regions
d) Limit the intermodulation distortion to favour mostly 'harmonic' distortion

Appreciate that we are not saying anything about the phase of the distortion components. In fact, harmonic distortion (like the one we are looking for) is related to 'steady state' signals (as opposed to transients). It is a well known fact that phase is only of perceptual concern on transients but not on large time scale components (like those that build up the harmonics of musical notes). This kind of 'phase unsensitiveness' of distortion is very welcome, because waveshaping based distortion does not provide easy control of the phase of newly generated partials (see section 3).

3. WAVESHAPING BASICS

WS (WaveShaping) first appeared on the late 70s ([1],[2]), providing a simple way to generate new waveforms from conventional sinusoidal oscillators. One of the benefits of WS if that generation of the many overtones can be made by simple polynomial evaluation instead of (computational intensive) sum-of-sins (additive synthesis).

The aim of waveshaping [1] [2] is to change the waveform of a sinusoid, instead of the amplitude and/or phase. It is well known than a sinus through a clipping memory-less distortion produces a clipped sine, richer in harmonics. Likewise, the shape chosen for the transfer function defines the ultimate combination of overtones that is generated for sinusoidal input.

The transfer function F(x), describing a memory-less distortion, is called the shaping function. It associates with each input value the corresponding output value independent of time. If the input is *x(t)*, the output is

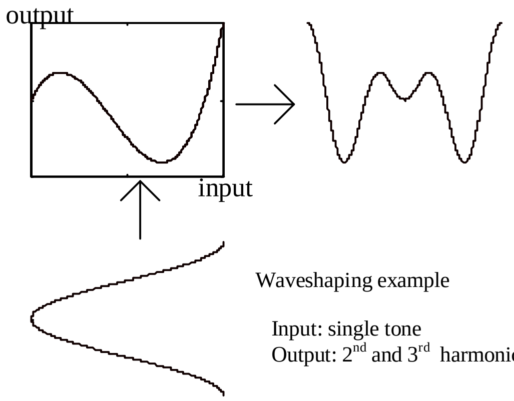
$$S(t) = F(x(t))$$

We are interested in shaping functions that can generate overtones from a cosine. The following equation has to be solved to find them:

$$F(\cos(2\pi ft)) = \cos(2\pi kft)$$

The analytical solution to this equation is known: the k-th order Chebyshev polynomial. It is possible to obtain a wide range of harmonics using a linear combination of a set of those polynomials (the result is just a new polynomial). In this way is easy to define the set of polynomial coefficients that generates a particular combination of overtones. It is possible to control even and odd, upper and lower, and closely spaced and widely dispersed harmonics without affecting other individual partials.

It can be proved that an n-th order polynomial will only generate overtones up to the n-th, but no further. Using a limited-degree polynomial for the shaping function allows band-limited output, making it possible to predict whether aliasing will occur and avoid it.



Waveshaping example

Input: single tone
Output: 2nd and 3rd harmonics

This formulation precludes detailed phase control on the newly generated overtones (if a phase shift is desired the transfer function is no longer single-valued, but instead looped, Lissajous-like figures are produced). According to our previous discussion this limitation is not of particular concern for our application.

Computational advantages of the WS are clear. There are also noise advantages in the WS polynomial approach opposed to the stored-table WS or the sum-of-sins: new distortion overtones are expected to be of mid-high frequencies. A mathematical closed-form to generate them (like the WS polynomial) warranties noise-free computation (up to the resolution used for the operations), while other schemes (like ROM storage of a sinus or waveshape, or interpolation of an stored sine table) would produce artifacts and noise (interpolation of close-to-Nyquist tones, like upper partials, needs a very large filter [3]).

4. WAVESHAPING OVER REAL WORLD SIGNALS

WS can be seen as a distortion of a sinusoidal waveform in order to obtain new shapes, harmonically richer than the sinus. But application of WS to real world signals in order to achieve distortion (generation of upper partials) is not obvious.

First of all, WS wouldn't work as expected (even for true sinusoidal input) if the envelope is not perfectly flat and adjusted to swing the waveshape from extreme to extreme (basically normalized to track the interval –1,+1).

Then (and most notably) the input for waveshaping must be a known-in-advance waveform, usually a sine. If the input is not the correct waveform, the output will not correspond to the expected combination of partials. Even for a perfectly monophonic and periodic (but not sinusoidal) signal, full band WS would give unproper results because the signal is not a sine. It would be necessary to separate each partial, apply WS individually to them, and the sum up the resulting signals. Only with independent WS we can keep control of the distribution of the newly generated partials.

Of course, in the case of a multi-instrument mix, all of these problems increase.

5. DEFINITION OF WS COEFFICIENTS

Typical users are expected to be musicians who don't want to deal with the polynomial coefficients, but instead would be happy to define the expected distortion shape in term of overtone amplitudes. The transformation between the amplitude domain description and the coefficients domain characterisation is well known and can be hidden from the user [1] [2].
In our system the user is allowed to specify the height desired for the overtone series, and automatic calculation of coefficients is carried out previously to signal processing.

6. MULTIBAND WAVESHAPING DISTORTION (MWD)

6.1. General Structure

To achieve the goals 'c' and 'd' from section 2, and also to alleviate some of the problems stated in section 4, we use multiband decomposition to obtain various simple signals from the original complex full-band signal. Each of the bands is distorted independently and then only the new, distortion components are added to the original signal to form the output.

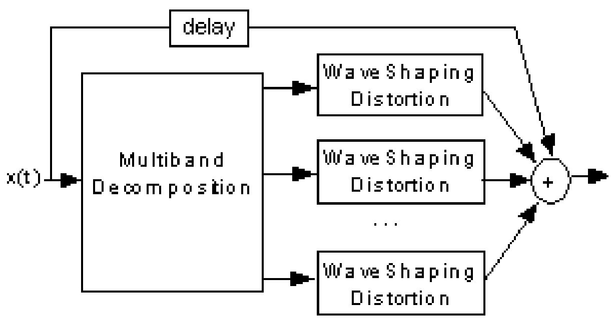


Figure 1. *Block Structure of MWD.*

We use two different structures for the decomposition. One is a fixed filterbank (and leads to a simple solution for the whole system). The other is a sinusoidal front end that can identify the most relevant partials at each moment (discards noisy spectral components and allows real monotonal input to WS units at the cost of increased computation and delay to perform the analysis).
Note that original input is passed directly to the output, while the distortion system only adds new harmonics. In this way the original is kept unaffected, and the multiband decomposition does not need to be a perfect reconstruction system (this fact will be later exploited).

The original signal can be optionally delayed to be time aligned with the distortion system output (delay in the multiscale decomposition can be compensated).

6.2. Details of a band

Inside each band waveshaping is used for distortion generation. The fact that each band contains a short bandwidth of the whole audio spectrum of the original signal makes the single band signal simpler and closer to the ideal monotonal input for waveshaping, keeping intermodulation distortion under control.
The requirement of a flat envelope is achieved by an envelope follower. The envelope is applied inversely over input (effectively nulling envelope variations) and reapplied over output (to restore original dynamics). This envelope flatting is also responsible of the necessary signal normalization previous to the WS process.
The flat envelope signal is subject to WS, which generates upper components to be mixed with the original in-band signal.

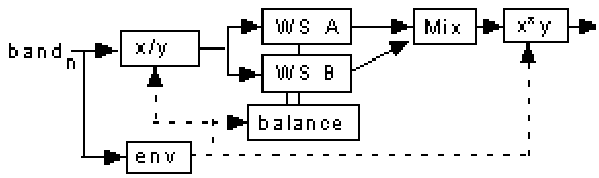


Figure 2. *Details for a single channel.*

Initially we used a fixed (though programmable) WS pattern for each band. On the results we missed the control of the distortion from signal level (goal 'b' of section 2). In order to change from a WS 'A' at low levels to a WS 'B' at high levels, the right way to go is to interpolate (or to balance) between the outputs of both WSs. The interpolation of the coefficients of 'A' and 'B' is not only unnecessary, but troublesome (the relation between the coefficients and the waveshape is nonlinear; also during the interpolation the waveshape can take values well outside the normalized range). Given the non-memory characteristics of the WSs, there are no concerns about phase problems in the mix of the outputs of WS 'A and WS 'B', hence crossfading between them achieves a gradual, free from artifacts transition between both WS patterns.
This is also true for the combination of distortion from different bands: phase misalignment can only be produced at the multiband decomposition stage (WSs themselves are memoryless).

6.3. Filterbank

There is no need to extend the filterbank over the whole audio spectrum. Very high frequency components will generate overtones well outside the audio spectrum. An upper limit of 10 KHz for the filterbank is to be used, but also a crosspoint in the 6-8 KHz region would be fine (upper partials in original signal will usually be very of a low level and subject to instability and noise). On the other side, very low tones (say under 500Hz) are better not to be included for distortion (their overtones would overcrowd the mid-high frequencies).

For tests, we have used a filterbank to split the 500 to 7000 region in several individual bands. The filters are of the FIR kind in order to be able to adjust constant group delay at all frequencies and in all bands (outputs of the filterbank are time aligned when necessary). In this way the whole system can still act as memory-less distortion, though one with delayed output (that's why the original signal path includes a delay).
The higher frequency limit is also useful to avoid unwanted aliasing: new distortion-produced overtones would otherwise easily extend beyond Nyquist frequency. In fact, the Matlab implementation checks each band upper limit and user defined overtones pattern and warns about the possibility of generation of overtones that would cause aliasing. The user is given the option to limit the overtone generation on higher bands, to those components that fit to the Nyquist criterium.

7. SINUSOIDAL FRONT-END FOR MWD

The second approach for the multiband decomposition is to use a sinusoidal model. In this approach the signal must be broken in frames. An 8-times oversampled FFT (with enhanced frequency resolution) for each frame is calculated, and peaks (possible tonal components) are validated according to a similarity measure with the shape expected for sustaining tones [4].

Every non-too-low-levelled peak in the FFT is a partial candidate. A 'quality of fit' for the peak is measured as:

$$\frac{\sum_{\xi}\left|S(k+\xi)-A\cdot W(\xi)\right|^2}{\sum_{\xi}\left|S(k+\xi)\right|^2} \quad \text{where } A=\frac{\sum_{\xi}\left|S(k+\xi)\cdot W(\xi)\right|}{\sum_{\xi}\left|W(\xi)\right|^2}$$

S is the signal complex oversampled spectrum, W is the window complex oversampled spectrum, k is the bin index of the peak, and the sum spans a range ξ of bins that corresponds to the width of the main lobe of W. This formula constitutes a kind of quality of fit between the measured peak and the expected ideal peak, and is related to the (energy normalized) least square difference between real and ideal peaks.
If the quality is not good enough, maybe the partial is subject to disturbing influence of nearby partials. If another peak lies closer than the window main lobe width, a new opportunity for the peak to be confirmed is given. Instead of trying to emulate the candidate peak nearbies with a single window main lobe, two are summed located at the candidate peak and at the interfering peak (with the corresponding amplitude and phase). Then, the previous formula is reevaluated with the sum instead of W.

Only peaks that conform to a quasistationary pattern are accepted, and are tracked through time (identifying the birth, evolution, and death of each tonal partial in the signal). The parameters (amplitude and frequency through time) of each track feed a sin generator. The output of each sinusoid generator is then subject to WS distortion.
This kind of modelling is much more complex and computationally intensive than the filterbank approach, but makes possible true single-tone input for WSs, and automatically rejects non-tonal signal components (avoiding distortion of non tonal components and transients).

For the MWD application the model is applied in a non-exhaustive way, so that only the most relevant partials of the signal at each moment are considered and subject to distortion.

8. BENEFITS OF MWD

Multiband Waveshaping Distortion is a new way to achieve a highly programmable distortion system. It can cope with goals 'a' to 'd' stated in section 2, and through the use of polynomial expressions for waveshaping the whole system is of low complexity and easy to perform with conventional DSP architectures (basically multiply-and-accumulate instructions), while being an analytical solution free from noise.

The filterbank approach is attractive, though suboptimal, for real-time operation and can take benefit of filterbanks already applied at other points in the whole system (like perceptual coding, equalization, etc.). Much more computationally intensive, an approach exists (sinusoidal analysis front-end) that can overcome the pitfalls of the filterbank approach making it possible true sinusoidal flat-envelope input to the WaveShaping elements.

Additional features can be included at little cost with the multiband approach, like group delay based enhancement (where slight time misalignment between bands is used to change the perception of the sound during transients), or multiband compression (as a part of the whole system the amplitude envelope of each band is obtained).
Thanks to the multiband approach with different distortion patterns per band, the user can exclude prone-to-noise bands (like those occupied by sibilant or cymbal) from distortion.
The balancing between two WS patterns according to the in-band signal level makes it possible a kind of noise control: when the in-band signal is of low level (hence with worse S/N ratio) it can be less distorted.
It would also be possible to include 'side-chains' to allow the signal-level based changes to react not only to the 'local' band signal, but also to other channel signals. This is not unlikely the valve behaviour, where an strong signal component will put the valve 'hot' and affect the sound of other components as well.

9. CONCLUSIONS

We continue to apply multiband techniques to conventional effect units, in order to advance to a higher degree of detail and programmability. Techniques described in this paper allow the design of 'distortion synthesizers' to be used by musicians to exactly adjust the distortion to their needs and taste, in a way opposed to the traditional black-box magic, non-adjustable design of conventional distortion units.

A Matlab implementation of the proposed architecture will be publicly available at the COST G6 web site, and example results will be played at the conference.

10. REFERENCES

[1] Arfib, D. "*Digital Synthesis of Complex Spectra by means of Multiplication on Non-linear Distorted Sine Waves*" ICMC'78 Proceedings .

[2] Le Brun, M. "*Digital Waveshaping Synthesis*" JAES, Vol. 27, nº 4. 1978

[3] Fernández-Cid, P., Casajús-Quirós, F.J. "*Enhanced Quality and Variety for Chorus and Flange Units*" DAFX'98.

[4] Fernández-Cid, P., Casajús-Quirós, F.J. "*Multi-Pitch Estimation for Polyphonic Musical Signals*" ICASSP'98.

[5] Roads,C. "A Tutorial on Non-Linear Distortion or *Waveshaping Synthesis*" Computer Music Journal, Vol. 3, nº 2. 1979