Graduate Program in Design Science: Audio and Acoustics
Faculty of Architecture, Design & Planning, The University of Sydney.
Unit of Study: DESC9115 - Digital Audio Systems

A set of pitch tracking, harmonically related resonance filters for monophonic signals.

Angus Gardiner
agar5694@uni.sydney.edu.au

1. Abstract:

This processor facilitates a range of timbral manipulations and resynthesis approaches for monophonic input signals. The input signal's fundamental frequency is tracked, and the signal is passed through a series of 11 resonance filters, each tuned to an integer multiple of the input signal's fundamental frequency over time. Alternatively, white noise can be harmonically filtered using the fundamental frequency data of an input signal. This creates a synthetic tone which mimics the input signal melodically, but whose timbral qualities are defined by the user input. This report includes a brief outline of the signal flow, and some discussion of the possible applications and usefulness of this processor.

2. Signal Flow Overview:

The processor has three distinct components: a pitch tracking component, a filtering component, and a summation component. A basic diagram is provided in Section 5 of this report.

2.1. Pitch Tracking Component:

The fundamental frequency of the input signal is estimated using a frequency domain analysis approach with phase-difference analysis and some post-processing to eliminate analysis of sounds without a singular pitch.

Firstly, the input signal is analysed in two distinct chains of 1024 point FFT windows, which are offset by R number of samples (henceforth referred to as a hop) as shown in Figure 1 below. The difference in phase between the two FFT analysis windows is used to improve the frequency resolution compared to what would be possible with a single 1024 point FFT.

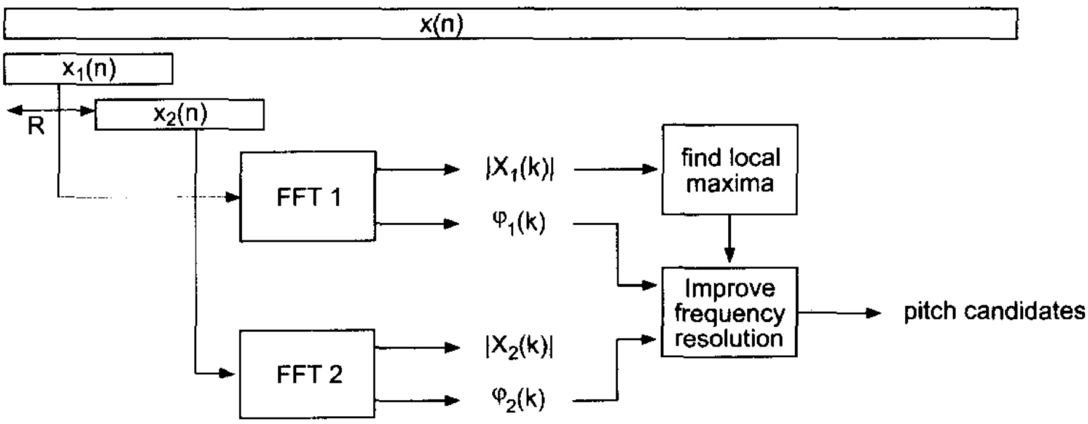

Figure 1 - Diagram of FFT pitch extraction method. Taken from DAFX, 2nd Edition.[1]

---

[1] Zolzer, U. *DAFX: Digital Audio Effects*. 2nd Ed. 2002. John Wiley & Sons. ISBN 0-470-84604-6

Secondly, post processing is applied in order to eliminate sonic elements without a single distinct pitch (eg: transient signals, noisy sounds, or unvoiced consonant sounds). A threshold of deviation of the fundamental frequency estimation from the mean of the values preceding and proceeding it is used to eliminate pitches estimations which are based on signals with no singular pitch.

Essentially, the output of this component of the processor is the fundamental frequency estimation for each FFT hop (100 samples) of the signal.

2.2. Filtering Component:

In this component, 11 resonance filters with time-varying centre frequencies; each an integer multiple of the fundamental frequency, are implemented to split the input signal into its harmonic components, and the resulting filtered signals are weighted with a user defined coefficient (which can be thought of as a gain). The filters can be thought of as multiple wah-wah effects with a basic point of difference: rather than the centre frequency for the filter being set by the movement of a foot pedal or by a triangular wave, the centre frequency is determined by the fundamental frequency of the wave itself.

The specific architecture is that which is recommended in Zolzer's *DAFX: Digital Audio Effects* for state variable filters[2]. This is a bandpass filter with time varying parameters of centre frequency and bandwidth. For this project, the bandwidth is set by the user for all the filterbands, and is not time varying, however the variable centre frequency variation possibility of this filter make it perfectly applicable for this project. The equation for the filterband's outputs is below, where $x(n) = the\ input\ signal$, $yl(n) = the\ lowpass\ filtered\ output$, $yb(n) = the\ bandpass\ filtered\ output$ and $yh(n) = the\ highpass\ filtered\ output$, and σ is a damping factor related the bandwidth of the filter.:

$$F_1 = 2sin(\frac{\pi f_c}{f_s})$$
$$Q_1 = 2\sigma$$
$$yl(n) = F_1 yb(n) + yl(n-1)$$
$$yb(n) = F_1 yh(n) + yb(n-1)$$
$$yh(n) = x(n) - yl(n-1) - Q_1 yb(n-1)$$

For our purposes, the bandpass filtered output of this filter is weighted with the user input gain coefficient for each filterband.

As previously mentioned, the processor also includes a feature where white noise can be generated and used as the input signal for the filtering component, rather than the monophonic input signal itself. This can have some interesting effects, including being able to fully sculpt the spectral shape of a synthetic tone, whilst assigning to it the precise melodic features of another tone. This precise separation of melody from timbre in audio signals has some intriguing possibilities for musical applications that are yet to be fully explored.

2.3. Summation Component:

In this component, the filtered signals resulting from each of the 11 resonance filters are combined. This combined signal is then mixed with the original, dry input signal at a mix ratio set by the user. The output signal is then normalised.

[2] Zolzer, U. *DAFX: Digital Audio Effects*. Chapter 2: Filters. Page 35-36. 2nd Ed. 2002. John Wiley & Sons. ISBN 0-470-84604-6.

3. Calling the Function in Matlab, Audio Examples and Discussion:

The function has 17 input arguments, being the file name (InputFileName), the Q factor for all the filters (Q), the output mix ratio between wet and dry signals (mix), the weighting coefficients for each of the filtered harmonic partials (g0 – g10), a switch for the white noise feature ('1' for on, '0' for off), and the frequency range for the fundamental frequency estimation (fmin,fmax) in Hz.

The function has three output arguments, being the original input wave (for comparison), the output wave, and the sampling frequency for ease of playback.

Four examples are given to show the range of tones which can be manipulated using this processor. All are based on a single input signal, being a short vocal melody[3]. Wave files for the audio input, and each of the examples can be found in the attached folder.

3.1. Subtle Timbral Manipulation

User Inputs:
InputFileName='Toms_diner.wav'; Q=20; mix = 0.2; g0=0; g1=0; g2=0; g3=0; g4=0; g5=2; g6=3; g7=5; g8=6; g9=6; g10=7; whitenoise=0; fmin=100; fmax=1000;

The timbre of the voice is subtly manipulated by boosting the upper harmonic partials (5^{th} – 10^{th} order) by varying amounts. This could be applied in a more nuanced fashion in order to help project a musical melody which was having trouble cutting through in a recording mix, for example.

3.2. Less Subtle Timbral Manipulation

User Inputs:
InputFileName='Toms_diner.wav'; Q=150; mix = 1; g0=0; g1=1; g2=0; g3=1; g4=0; g5=1; g6=0; g7=1; g8=1; g9=1; g10=0; whitenoise=0; fmin=100; fmax=1000;

The timbre of the cello is less subtly manipulated by increasing the mix to 1 (fully wet) and eliminating all but the 1^{st}, 3^{rd}, 5^{th}, 7^{th},and 9^{th} order harmonics. This gives the tone a more clarinet-like quality.

3.3. Resynthesis of Melody with Some Resemblance of Original Tone

User Inputs:
InputFileName='Toms_diner.wav'; Q=1000; mix = 1; g0=1; g1=1; g2=1; g3=1; g4=1; g5=1; g6=1; g7=1; g8=1; g9=1; g10=1; whitenoise=1; fmin=100; fmax=1000;

Here, white noise is filtered with harmonic resonance filters and has energy at every single harmonic partial up to the 10^{th} order harmonic. In this case, the tone retains some semblance of the original tone's timbral structure (i.e. has some significant energy at all harmonics).

3.4. Resynthesis Melody with No Resemblance of Original Tone

User Inputs:
InputFileName='Toms_diner.wav'; Q=1000; mix = 1; g0=1; g1=1; g2=0; g3=1; g4=0; g5=0; g6=0; g7=1; g8=0; g9=0; g10=0; whitenoise=1; fmin=100; fmax=1000;

Here, white noise is filtered with filter bands which are related by octaves (i.e. fundamental frequency, 1^{st}, 3^{rd}, 7^{th} order harmonic partials. This gives the output signal an organ-like tone colour.

---

[3] This melody has been sourced from the example sound files for Chapter 9 from the *DAFX* website. http://www2.hsu-hh.de/ant/dafx2002/DAFX_Book_Page_2nd_edition/chapter9.html

4. Conclusions:

This processor has clear musical applications for both the manipulation and enhancement of natural timbres, and for the synthesis of timbres whilst retaining a given melodic feature. At this stage, it's envisioned that it could prove to become and invaluable and flexible mix tool, given some more attention to detail and improvements, particularly in the pitch tracking area to eliminate some artefacts that are emerging at this early stage of development. The current filtering process can also be improved in order to minimise phase artefacts which currently exist.

5. Signal Flow Diagram

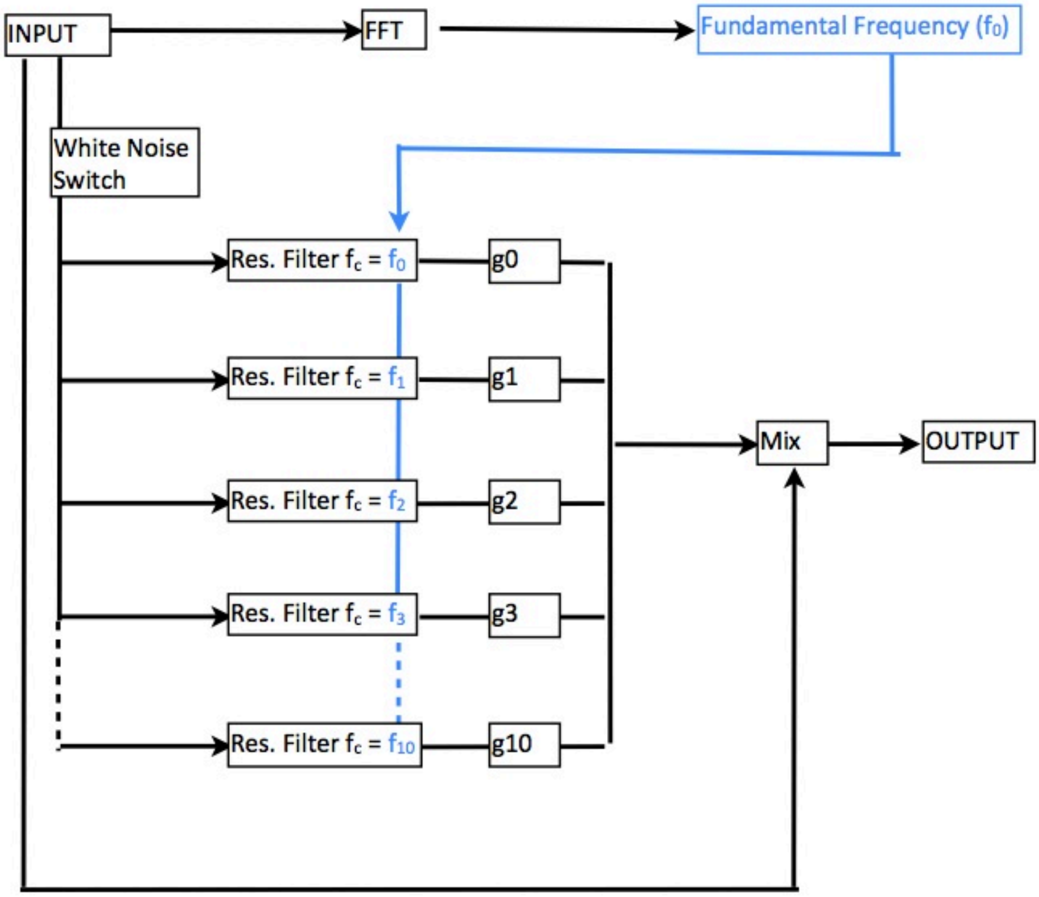

Figure 2 - Basic Signal Flow Diagram of the Processor

6 Bibliography:

Brown, J. *A High-Resolution Fundamental Frequency Determination Based on Phase Changes of the FFT*. Journal of the Acoustical Society of America, 91, 2375, (1992).

Chen, B C-L. *Final Report – The Wah Wah Effect* – Sydney University e-Scholarship Repository. http://ses.library.usyd.edu.au/bitstream/2123/7662/3/Lab%202%20Manual.pdf. Accessed 9th June

De Cheveigne, A. & Kawahara, H. *Yin, a Fundamental Frequency Estimator for Speech and Music.* Journal of the Acoustical Society of America, 111(4), 2002.

Gerhard, D. *Pitch Extraction and Fundamental Frequency: History and Current Techniques*. Technical Report. 2003. Department of Computer Science, University of Regina. ISSN 082-3494

Harris,F. & Brooking,E. *A Versatile Parametric Filter Using an Imbedded All-Pass Sub-Filter to Independently Adjust Bandwidth, Center Frequency and Boost or Cut*, Presented at the 95th Convention of the AES. 1993.

Kedem, B. *Spectral Analysis and Discrimination by Zero-Crossings*. Proceedings of the IEEE, 74(11: 1477-1493, November 1986

Maher, RC. *Tunable Bandpass Filters in Music Synthesis.* Presented at the 99th Convention of the AES. 1995.

Gareth Middleton, *Pitch Detection Algorithms.* OpenStax CNX. Dec 18, 2003 http://cnx.org/contents/8b900091-908f-42ad-b93d-806415434b46@2@2. Accessed on 24 March 2015.

Rossignol, S. Rodet, X. Soumagne, J. Collette, J-L. Depalle, P. *Features Extraction and Temporal Segmentaion of Acoustic Signals.* International Computer Music Conference, pages 199-202, 1998.

Schroeder, MR. *Period Histogram and Product Spectrum: New Methods for Fundamental-Frequency Measurement*. Journal of the Acoustical Society of America, 43, 829, 1968.