FAST MELODY EXTRACTION USING AUBIO (BROSSIER), MIREX 2005

Paul M. Brossier
Centre for digital Music
Queen Mary University of London
Mile End Road, London E1 4NS
`paul.brossier@elec.qmul.ac.uk`

ABSTRACT

In this submission to the Audio Melody Extraction Detection contest, we use a spectral matching comb filter to track the fundamental frequency of the signals. The final decision on the pitch value is taken within very few frames delay in order to be usable within short latencies. The different methods are implemented as a C library.

ALGORITHMS

The algorithm for this submission to the MIREX Audio Melody Extraction is thoroughly described in [1] and was used in [2] in the following implementation. The most important steps of this method are recalled below:

- **pre-filtering**: The signal is first filtered through and ARMA A-weighting filter to enhance medium frequencies and reduce high and low parts of the spectrum.
- **phase vocoder**: For a signal at 44.1kHz, windows of 4096 samples of the pre-filtered signal are sent to a phase vocoder with an overlap rate of typically 25%. The magnitude spectrum of each window is low pass filtered and normalised to smooth out the small peaks and favorise the wideband ones.
- **harmonic comb matching**: After pre-processing, peaks are detected in this spectral frame, and matched against a harmonic comb filter. The most likely fundamental frequency is choosen against the number of matched peaks in the comb filter and the energy measured in each of these harmonics.

After the pitch candidates have been selected, filtering of the candidates is done using a median filter algorithm and further rules to constrain the pitch track to a more continuous path. The operation requires a small number of consecutive candidates, so that the smoothing of the pitch can be made within short latencies – about 5 to 10 candidates. This process is further described in [2].

A loudness threshold is used to discard candidates in silence regions, so that silence regions will be labelled as unvoiced.

Three more algorithms are also provided along with the multi comb filtering. The YIN algorithm [3], a modified autocorrelation function, is proposed as an alternative method. Originally designed for voice, the algorithm has been modified to be suitable on more complex signals. Two less CPU expensive algorithms, an implementation of the Schmitt trigger and a fast version of the comb filter, are implemented in the program.

SOFTWARE IMPLEMENTATION

The algorithms are implemented as a command line utility, binded to a library of C functions. The library, named AUBIO, includes other functions such as onset and beat tracking. It was built to provide both a proof of concept of real time implemetations and a suite of high level extractors to embed in other systems.

The pitch extraction tool is a command line utility. It can be used to read audio from files or hardware inputs on outputs in real-time. Different command line argument are provided to modify the default parameters (pitch detection algorithm, window and overlap sizes, algorithm specific parameters).

The algorithms were tested on a 1.1GHz PowerPC running Debian GNU/Linux. The multi comb implementation (compiled using GCC) computed the analysis on 14 one minute files in 250.5 seconds – averaging 17.5 seconds per minute input file. The YIN algorithm needed 261.3 seconds to compute the same task - 18.2 seconds per minute processed.

ACKNOWLEDGEMENTS

Thanks to last year's MIREX organisation commitee for starting a captivating contest, and to this year commitee for organising such a sucessful event.

References

[1] J. P. Bello. *Towards the Automated Analysis of Simple Polyphonic Music*. PhD dissertation, Queen Mary, University of London, Centre for Digital Music, 2003.

[2] P. M. Brossier, J. P. Bello, and M. D. Plumbley. Fast labelling of notes in music signals. In *Proc. of the International Symposium on Music Information Retrieval (ISMIR)*, Barcelona, Spain, 2004.

[3] A. de Cheveigné and H. Kawahara. YIN, a fundamental frequency estimator for speech and music. *Journal of the Acoustic Society of America*, 111:1917–1930, 2002.