FAST ONSET DETECTION 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

We present a fast algorithm for onset extraction, based on spectral detection functions, peak picking with dynamic thresholding and silence detection. The algorithm is implemented in the C language and has been specifically tailored to output its result within short delays. A brief overview of the algorithm and its implementation are given, as well as speed measurements.

ALGORITHMS

This submission to the MIREX Contest is based on various work on onset extraction that have been reviewed in [1]. The onset decision has been implemented using spectral detection functions, along with a peak picking algorithm specially modified to obtain short latencies. Details specifics to this implementation have been described in [2].

The process of the onset extraction consists in the following steps:

- **silence detection**: a hard threshold is set to discard onsets detected within quiet regions
- **onset detection function**: a function is constructed based on successive spectral frames
- **dynamic thresholding**: a short window of the detection function is used to compute a dynamic threshold
- **peak selection**: onsets are selected if a peak is detected above the dynamic threshold and with sufficient loudness

The detection function are built using windows of 1024 samples with 256 samples overlap. The silence decision is performed using the same window length and a threshold of $-70dB$ – this ensures no onsets will be detected in background noise.

Various detections functions are available. All have been designed to maximise their amplitude variations when changes in the spectral content are detected – either in frequency, amplitude or phase. The Kullback Liebler distance, complex-domain detection function, High Frequency Content, phase based and other methods have been described in details in [1].

The dynamic threshold is computed using less than 10 frames of the detection window, about 50 ms at 44100 Hz. Details of the algorithm, which involves low pass filtering, moving median and moving mean, are presented in [2]. A parameter can be set manually to obtain over or under detection. Since only one frame ahead is used to take the final decision, the algorithm can be used in systems where minimal delays are critical.

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 fundamental frequency estimation 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 onset extraction tool, a command line utility, can be used to read audio either from files or from hardware inputs/outputs. Arguments to change the algorithm parameters are supported (peak picking and silence thresholds, window and overlap sizes) and different functions can be combined together. The two most successfull functions will be used for the contest submission.

The algorithms were tested on a 1.1GHz PowerPC system running Debian GNU/Linux. The C implementation, compiled using GCC, ran the analysis of 14 files of one minute of audio in 48.35 seconds - averaging 3.36 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, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. B. Sandler. A tutorial on onset detection in music signals. *IEEE Transactions on Speech and Audio Processing*, 2004.

[2] P. M. Brossier, J. P. Bello, and M. D. Plumbley. Real-time temporal segmentation of note objects in music signals. In *Proceedings of the International Computer Music Conference, ICMC 04*, Miami, Florida, 2004.