MEL-FILTER INTEGRATION FOR ONSET DETECTION

Kris West
School of Computing Sciences
University of East Anglia
Norwich, UK
kw@cmp.uea.ac.uk

1 INTRODUCTION

This extended abstract describes a submission to the Onset Detection task for MIREX2005. This onset detection function is based on the combined phase and amplitude onset detection function described in [1]. In this work 23ms analysis frames, overlapped by 50%, of 22 kHz signals are used to calculate the onset detection functions

2 ENERGY-BASED ONSET DETECTION

Positive differences in a signal's envelope or the envelope of sub-bands of that signal are taken to indicate the onset of events in the audio stream. In this work Energy-based onset detection functions are calculated by summing the positive differences in a 256 bin Fast Fourier Transform of the signal.

3 PHASE-BASED ONSET DETECTION

When performing spectral analysis of a signal, it is segmented into a series of analysis frames and a Fast Fourier transformation (FFT) is applied to each segment. The transform returns a magnitude $|S(n,k)|$ and a phase (n,k) for each bin. The unwrapped phase, $\tilde{}(n,k)$, is the absolute phase mapped to the range $[-\ ,\ ]$. Energy based techniques consider only the magnitude of the FFT and not the phase, which contains the timing information of the signal.

A musical event can be broken down in to three stages; the onset, the sustained period and the offset. During the sustained period of a pitched note, we would expect both the amplitude and phase of the FFT to remain relatively stable. However during a transient (onsets and offsets) both are likely to change significantly.

During attack transients, we would expect to see a much higher level of deviation than during the sustained part of the signal. By measuring the spread of the distribution of these phase values for all of the FFT bins and applying a threshold we can construct an onset detection function. Peaks in this detection function correspond to both onset and offset transients so it may need to be combined with the magnitude changes to differentiate onsets and offsets.

4 SUB-BAND ONSET DETECTION

In this work, both the Energy and Phase-based onset detection functions are based on 256 bin FFTs of the signal. A combined onset detection function can be produced by multiplying the absolute values of each bin of the phase deviations by the raw values of each bin of the energy differences. The final onset detection function is formed by half-wave rectifying the results and summing them.

5 MEL-FILTER INTEGRATION FOR ONSET DETECTION

The function described above contains significantly less noise than either of the two functions it was produced from. The remaining noise in this function can be further reduced by passing the both the Energy and Phase FFTs through a Mel-scale filter-bank prior to integration. This filtering virtually eliminates noise introduced by timbres and intonation (such as vibrato) that fluctuate, moving the energy in the function from one FFT bin to another. An example of the effect of this filtering is shown in Figure 1. In the (limited) optimization experiments performed, we found no benefit in using a threshold of the onset detection function in peak-picking. Therefore, a simple peak-picking window is used to select local maxima as onsets and an isolation window, half the size of the peak-picking window, is used to prevent multiple detections.

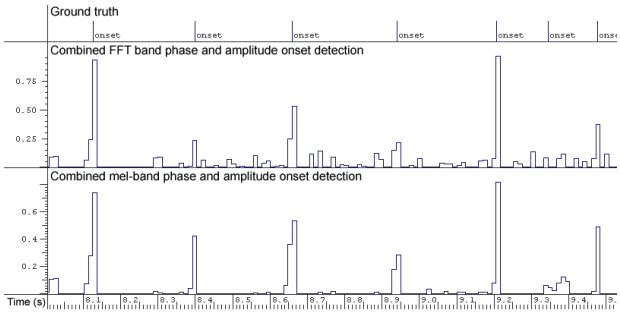

Figure 1. An example of the effect of Mel-scale filtering of onset detection functions

REFERENCES

[1] Duxbury, C., Bello, J. P., Davies, M., and Sandler, M. "A combined phase and amplitude based approach to onset detection for audio segmentation", Proceedings of the 6th Int.

Conference on Digital Audio Effects (DAFx-03), London, UK.