ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/228744106

Tempo extraction using beat histograms

Article · January 2005

CITATIONS
6

READS
248

1 author:

George Tzanetakis
University of Victoria

220 PUBLICATIONS 7,199 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project The Orchive View project

Project Display Wall immersive imagery research View project

All content following this page was uploaded by George Tzanetakis on 27 May 2014.

The user has requested enhancement of the downloaded file.

Tempo Extraction using Beat Histograms

George Tzanetakis
University of Victoria
Computer Science Department (also in Music)
gtzan@cs.uvic.ca

ABSTRACT

This abstract describes the tempo extraction algorithm used for the University of Victoria submission to the MIREX (Music Information Retrieval Exchange) 2005. The algorithm is mostly based on self-similarity rather than onset detection. However, an onset detection component is used to calculate the phase of the dominant periodicities. Multiple frequency bands are calculated using a Discrete Wavelet Transform. Subsequently the envelope of each band is extracted and autocorrelation is used to find the dominant periodicities of the audio signal. These dominant periodicities are accumulated into a *Beat Histogram* which is used to detect the primary and secondary tempo and their relative strength.

Keywords: tempo extraction, wavelet analysis

1 DESCRIPTION

The approach used by the University of Victoria submission to the MIREX 2005 tempo extraction contest is based on the calculation of the Beat Histogram described in Tzanetakis and Cook (2002). Most tempo extraction algorithms fall into two categories. Algorithms that do onset detection and then work directly from the onset positions and algorithms that use self-similarity at either single or multiple bands Gouyon and Dixon (2005). The main difference from the published algorithm is that the peak picking algorithm has been optimized to work better for tempo detection. In addition more heuristics regarding harmonic beat ratios are utilized. We are still currently experimenting with alternative front-ends to the wavelet approach described in Tzanetakis and Cook (2002). More information about other approaches to automatic tempo detection and other types of rhythm description can be found in Gouyon and Dixon (2005).

In our approach, the signal is separated into different frequency bands using a Discrete Wavelet Transform (DWT). Essentially each bands corresponds to the energy in an octave. The envelope of each band is calculated using Full Wave Rectification, Low Pass Filtering and Normalization. The purpose of the envelope extraction step is to retain events in that particular octave without being affected by particular pitches. The normalized envelopes are added with the effect of enhancing common periodicities with the same phase. The sum of the normalized envelopes is then processed through an autocorrelation function to detect the dominant periodicities of the signal. Figure 2 shows a schematic diagram of the process.

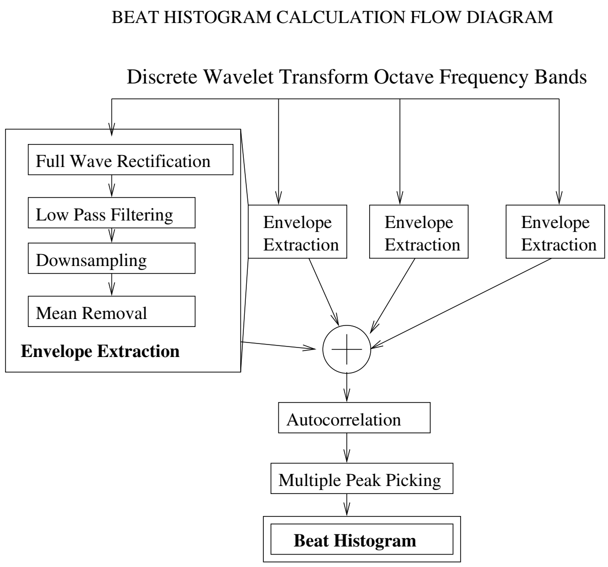

Figure 1: Beat Histogram Calculation Diagram

This front-end is based on the method for the calculation of Beat Histograms described in Tzanetakis and Cook (2002). The Beat Histogram (BH) shows the distribution of various beat periodicities of the signal. For example a piece with tempo 60 Beats-per-Minute (BPM) would probably exhibit BH peaks at 60 and 120 BPM (quarter and eight notes respectively). It is important to note that the BH is a representation (essentially an empirical distribution) of the various periodicities in the signal and can be potentially calculated using other methods.

Figure 1 shows a schematic diagram of the this calculation. The spread around each peak is an indication of tempo variability and the BH can accomodate pieces with complete changes of tempo by showing separate peaks for each one. Figure 2 shows the BH for a HipHop piece (notice the peaks at 96 BPM (main tempo) and 192 BPM). A BH of a country music piece at similar tempo would also exhibit pronounced peaks but with lower height. Although this information is not used for tempo estimation, it is valuable for other tasks such as genre classification.

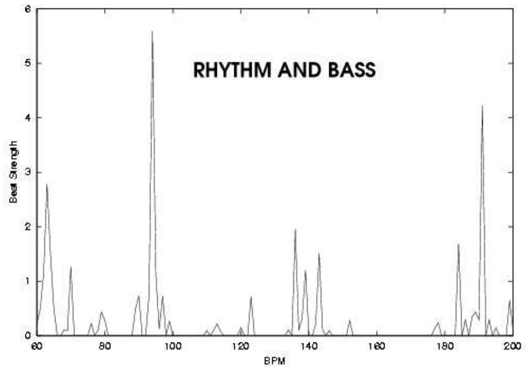

Figure 2: Beat Histogram

The two main peaks of the BH (subject to some heuristics) are selected as the two dominant tempos. The histogram is constructed by adding the autocorrelation values for each particular periodicity therefore the height of the peaks correpond to the "strength" of that particular beat and is used to calculate the relative strength of each tempo. Peak picking in the subband envelopes is used to calculate the phase of each tempo.

The main idea behind the heuristics is to enhance integer ratios of the detected periodicity. Basically for each peak of the Beat Histogram we form hypothesis such as it is is quarter note at 2/4 or a quarter note at 2/4 and use the information at the corresponding peaks (of these ratios) of the BH to score each peak. The peak with the highest score is selected. Therefore it is possible that the largest peak of the BH will not be selected if there is a smaller (still large) peak that has corresponding high peaks at integer ratios.

2 DISCUSSION OF EVALUATION RESULTS

The algorithm ranked 13th out of 13th participants and it is not particularly fast so it's clear that there are better alternatives. The BH was mainly developed for the purpose of representing rhythm-related audio content rather than explicitely tempo detection. Last year the algorithm was tested with only 6 samples and this year with the 20 development samples. There has been no optimization of parameters based on larger datasets so we believe that the performance of the algorithm can be improved with more work. The choice of the wavelet front-end is also relatively arbitrary and other front-ends could be used while retaining the Beat Histogram representation.

The main difference from the submission to last year's contest was the addition of various heuristics to enhance integer ratio of tempo and detect the secondary tempo and other parameters requested for the contest this year. The main direction for future work is to optimize and understand the errors over larger datasets and use ideas from other algorithms to improve the performance.

3 IMPLEMENTATION

The entire system is implemented in *Marsysas-0.2* a free cross-platform software framework for audio analysis and synthesis [1]. The C++ source code can be found in the tempo.cpp file which also contains code for other types of rhythm analysis. Please contact the author for any questions/problems/ideas/suggestions you would have about *Marsyas-0.2*.

ACKNOWLEDGEMENTS

Many thanks to Stuart Bray, Ajay Kapur, Jennifer Murdoch, and Adam Tindale and Antonin Stefanutti for help with the development of Marsyas.

References

F. Gouyon and S. Dixon. A review of automatic rhythm description schemes. *Computer Music Journal*, 24(3), 2005.

G. Tzanetakis and P. Cook. Musical Genre Classification of Audio Signals. *IEEE Trans. on Speech and Audio Processing*, 10(5), July 2002.

[1] `http://marsyas.sourceforge.net`

View publication stats