Baharak Zali Homework # 3
8679928115 Feb. 3, 2005

Tempo and beat analysis of acoustic musical signals

By Eric Scheirer

The paper explains a method for analyzing the tempo of music signals, as well as extracting their beat, by use of band-pass filters and banks of parallel comb filters. The goal of this research was to find a computational algorithm to extract what human listeners call “beat” or “pulse”, from a piece of music. For this purpose, beat of a music piece is considered as the sequence of equally spaced phenomenal impulses which define a tempo for the music. The concern of this research is on the pieces and style of music that have a complex texture and timber but a simple rhythm.

Most of the previous works on this subject was based on a two phase process: first the music is segmented into notes, onsets, timbers, and so on; then post-processing algorithms grouped the previous information, or rhythms, in order to track the beats. These methods are called transcriptive methods. Doing such grouping is very difficult though!

Psychoacoustic studies showed that certain kind of signal manipulations and simplifications can be performed without affecting the perceived pulse content of a musical signal. In this kind of transformation the only thing preserved is the amplitude envelopes of the filter-bank outputs, not the notes. It proved that only this much information is enough to extract pulse and meter from a musical signal which results in huge reduction in the size of input data.

The algorithm that is used for beat-tracking works on the acoustic data and uses a bank of resonators to phase-lock with the beat of the signal and also to determine the frequency of the pulse. When the signal comes in, a filter-bank is used to divide the signal into six bands. Then for each sub-band, the amplitude envelope is calculated and derived. Each of the envelope derivations is passed on to another filter-bank of tuned resonators; in each resonator filter-bank, one of the resonators will phase-lock, the one for which the resonant frequency matches the rate of periodic modulation of the envelope derivative. Next, outputs of the resonators are examined and the information about the ones showing phase-lock behavior is tabulated for band-pass channels. The frequency of the resonator with maximum energy output is selected as the tempo of the signal. For extracting the phase, all of the tabulations are summed across the frequency filter-bank to arrive at the frequency estimate for the signal, and reference back to the peak phase points in the phase-locked resonators determines the phase of the signal.

The system is implemented in C++, and as input, it can accept audio files from a CD or other audio recordings, as well as live music coming from a microphone. It is possible to change some parameters to control the speed vs. accuracy of the program; these parameters include: frequency filter-bank, envelop sampling rate, number of resonators per frequency channel, and analysis frame rate. It is also possible to control behavior of the system by changing parameters in comb filters, in a way that it locks on to a beat and follow that tempo regardless of the new information (it can be used for rock or pop music), or change them very fast (it may be used in classical music).

Engineering Approaches to Music Perception & Cognition

Baharak Zali
8679928115

Homework # 3
Feb. 3, 2005

The results of the beat-tracking are validated then. In summary, among 60 trials conducted using different genres of music, they showed 41 (68%) correct detections, 11 (18%) partial detections and 8 (14%) wrong detections. For classical music correct detection was 4/9, and 4/9 partial detection; and for jazz they had only 3/8 correct detections and these two genres were the most problematic ones. For the next step, five experienced adult musicians were asked to extract beat and tempo of some pieces and their results were compared to the algorithm's extraction. The results of the human subjects and the system were comparable and in 4/7 cases the algorithm was the most or second most accurate tapper.