Yu Shiu
ISE 575
Feb 3, 2005

Review of "Tempo and beat analysis of acoustic musical signals"
by Eric Scheirer

The basic idea of the author's tempo extracting algorithm is, first of all, to calculate the onset from each subband, second, to calculate the energy of each tempo using comb-filter resonant filterbanks and, finally, sum the energy of each tempo from different subbands and choose the tempo with the largest energy as the tempo of the input signal.

The frequency division of subband, in his design, corresponds to human perception of octave. Along with other choice of parameters, the author tries hard to map the audio signals and their derived features to correspond to human perception. On the other hand, the use of comb-filter in computer music is one of the author's major contributions. Comb-filter basically is a periodic impulse train function with attenuation (decay). The advantage of comb filter is that it will resonate dramatically (it will has large energy) if the audio signal has similar tempo component as the comb filter and will not if the audio signal has different tempo component as the comb filter. The period T corresponds to the value of tempo. Therefore, the author needs to define the possible values of the tempo in advance and assign each one of them to a comb filter. Since the number of comb filters is limited, the frequency resolution of tempo is limited too. In order to achieve higher frequency resolution, the author needs to have large number of comb filters in each subband. Conversely speaking, one of the reasons that the author uses limited number of comb filters is that the algorithm needs to sum the energy output of comb filters from different subbands. If the number of comb filters is large, the frequency resolution is so high that the signals of the same tempo with variation or noise might be assigned to different tempos. In this case, the author needs further processing to group them together or quantize so that he could sum over 6 subbands. Or, the author could add another stage before resonant filterbanks to calculate the possible tempo in advance. It seems to me that the author uses the former strategy.