A Multi-Resolution Sinusoidal Model

final assignment report

1. For this assignment I chose 2 real world polyphonic recordings from "freesound.org" that have both a relevant melodic and percussion components. They are orchestrial sounds. The links from "freesound.org" are the following ones:

   http://freesound.org/people/xserra/sounds/217543/

   http://freesound.org/people/dshoot85/sounds/331071/

2. For both sounds I performed an analysis with "Sonic Visualizer" and I was able to decide the following parameters by their frequency content:

   Sound #1: --> frequency bands --> [0Hz-500Hz], (500Hz-3000Hz], [3000Hz-22050Hz]
   --> window sizes --> 4096, 2048, 1024 (samples)

   Sound #2: --> frequency bands --> [0Hz-300Hz], (300Hz-1000Hz], [1000Hz-22050Hz]
   --> window sizes --> 4096, 2048, 1024 (samples)

   The basic melody of the first sound was in the middle frequencies (500Hz-3000Hz) and that's why I dedicated a whole frequency band there. The same goes for sound #2 for frequency band (0Hz-300Hz). As for the percussive part, sound #1 had low frequency content (kick → 0Hz-500Hz) and sound #2 middle frequency content (snares → 300Hz-1000Hz). Finally sound #1 had a lot of high frequency content compared to sound #2 so I dedicated a narrower band to sound #2 for better definition.

3. The results, as you can hear, are very good. Through the whole process of the multi-resolution analysis and synthesis with the sinModel we have maintaned sharp attacks and also very well acoustically defined frequency content, even in the low frequencies. Therefore, we can say at a point that we have constrained the time/frequency analysis tradeoff.
   Moreover, i should mention that we have to pay attention at the coding technics for this kind of multi-resolution implementation because the computational complexity of the algorithm has increased a lot.
   Finally, in future work we cound perform simirarly a multi-resolution approach to the HPR and HPS models. With this technique the harmonic content will be defined better than the original models and therefore the residual will be less correlated with harmonic content and it could be described better with stochastic approaches.

4. If I was to extent it to HPR and HPS models I admit that there will be some challenges that I will have to face. Speaking, especially, for sinusoid tracking and F0 estimation algorithms, there will be a discontinuity between different frame sizes. Therefore, a more efficient and robust implementation must be done in order to

perform 3 different analysis and then combine the results in a way that all sinusoids can be tracked correctly without misses and estimated f0's also.

5. I can think of a further method to improve the time-frequency resolution trade-off. I have studied techniques where the algorithm estimates the onsets of the sound (as a function of time) and performs a masking technique in which all the analysis and synthesis is done to the whole sound except from the transients (onsets).