Music 422 - Final project proposal

Prateek Murgai, Orchisama Das

1 Features to be implemented

1.1 Stereo Coding

We are interested in implementing stereo coding as most of the audio available to us is present in stereo format and provides us with an opportunity to exploit the redundancies that occur between two correlated data streams. The main advantage of stereo coding comes when there is high correlation between the two channels and we can employ the difference signal (Left Channel - Right Channel) which is automatically a very small signal thus requiring low bit allocation for representing this difference signal.

For implementing this features following are the steps that we need to figure out :

- Come up/Decide a method to estimate the correlation between the data of the channels. This is important as this correlation factor will tell us when to use the mid-side encoding/decoding and when to use Left-Right encoding/decoding. We plan to combine for L/R (independent) and M/S stereo coding as done in MPEG-2 [1].
- Incorporating a psychoacoustics model that takes into account binaural masking (and binaural masking thresholds) at different frequencies (for example binaural separation at low frequencies is difficult so figuring out a way to incorporate this would be interesting) and generating different masking curves when Mid-Side is being employed.

1.2 Block Switching

To prevent pre-echo and transient smearing, we plan to implement block switching the way it is done in MPEG-2 AAC [1]. This involves detection of transients and sub-division of blocks into eight 256 sample sub-blocks (for a block length of 2048). For proper overlap add, two transition windows are used before and after the shorter blocks. For transient detection, we plan to use the traditional perceptual entropy based classification, or spectral flatness based classification as given in [2], which performs more accurate detection for low-energy transients.

1.3 Variable bit rate/Bit reservoir

Adaptive block size also calls for variable bit rate allocation. The number of bits allocated to each block will depend on the block size. Details of this algorithm is also given in [1]. Another alternative is using the algorithm in [3], depending on how much time we have.

1.4 Other Features

We will implement two small other features , they are :

- Kaiser Bessel Derived Windows
- Generating masking curves for noise maskers.

2 Timeline

- Implement Stereo Coding and Block Switching by 3rd March.
- Implement variable bit rate by 7th March.
- Test Coder and Listening Tests by 13th March.

References

[1] Marina Bosi, Karlheinz Brandenburg, Schuyler Quackenbush, Louis Fielder, Kenzo Akagiri, Hendrik Fuchs, and Martin Dietz. Iso/iec mpeg-2 advanced audio coding. *Journal of the Audio engineering society*, 45(10):789–814, 1997.

[2] Xuemin Zhang, Changqing Cai, and Jianhong Zhang. A transient signal detection technique based on flatness measure. In *Computer Science & Education (ICCSE), 2011 6th International Conference on*, pages 310–312. IEEE, 2011.

[3] A Szwabe and C Jedrzejek. Perceptually transparent audio compression based on a variable bit rate aac coder. In *Video/Image Processing and Multimedia Communications, 2003. 4th EURASIP Conference focused on*, volume 2, pages 685–690. IEEE, 2003.