Optimal fixed point implementation of MPEG-4 AAC encoder

Ashok Magadum, Vinod Prakash
Ittiam Systems (P) Ltd.
Consulate 1, Richmond Road,
Bangalore, India-560025.
Tel: +91-80-2237660

AAC (Advanced Audio Coding) is a perceptual audio coding technique standardized by MPEG. At a bit rate of 64Kbps per channel MPEG-4 AAC provides transparent quality (indistinguishable from original) according to EBU (European Broadcasting Union) Measures, this makes AAC a suitable codec for audio applications involving high compression ratios while retaining the audio quality.

The reference ISO code and the technical details provided in the specification given by MPEG are very complex for real time fixed point implementation on any platform due to usage of complex trigonometric and non-linear functions. In this paper we present the limitations of the techniques specified in the standard and outline the algorithmic and implementation level optimizations carried out in order to get a high quality, low complexity audio encoder that can be easily ported onto multiple platforms. Similar optimization techniques can be used for AAC like perceptual coders (for example: MP3 encoders).

Initially, we describe the various blocks of the AAC encoder in brief. We have chosen the MPEG4 Low Complexity encoder profile with support for the LTP (Long Term Prediction) object type as a case study. This includes a short introduction to the psychoaccoustic, quantization and LTP modules

Next, we go into the details of the major MIPS consuming blocks while highlighting the difficulties in fixed-point implementations. The major blocks being: The standard available **Psychoacoustic** process involves many complex mathematical functions that are high on MIPS. Also the wide dynamic ranges involved make the fixed point implementation more difficult. The bit rate dependant iterative loops in the **Quantization** module contribute significantly to the complexity. The Long Term Prediction tool (part of the **LTP** object type) in the raw form consumes lot of processor resources due to high length auto and cross correlations routines.

Next we describe the optimizations carried out. We have split the optimizations into two categories, algorithmic and implementation level. Briefly, In the **Psychoaccoustic** block we have changed the tonality estimation measure making it more robust for distinguishing between noise and tones. Also the scheme employed is much more amenable for fixed point conversion. In the **Quantization** block a novel technique has been used to meet the bit rate /distortion criteria as opposed to the loops normally used. Use of FFT's to implement the auto correlations and cross correlations in the **LTP** module has considerably brought down the complexity.

Finally we present the results obtained due to the above-mentioned optimizations. The results provide a block-by-block breakup of the performance enhancement achieved. We also indicate the quality Vs complexity at different bit rates and the relative comparison with the existing ISO code.