MP2 CODEC Using Software Configurable Processor

Chetan. P. Kamath and U. C. Niranjan

Abstract-- **Design of audio coders necessitates retaining high perceptual quality of the reconstructed signal with robustness to variations in the spectra and levels which can be achieved using compression. Mpeg-1 layer II (MP2) is a lossy compression scheme in which the nature of the human perception is taken into account to selectively remove the parts of audio data files, without the loss being noticed. This paper highlights the implementation of MP2 codec on a software configurable processor.**

Index terms-- **MP2, codecs, software configurable processor, subband, filter bank, psychoacoustic, fast fourier transform, signal to mask ratio.**

I. INTRODUCTION

The purpose of any data reduction system is to reduce the data rate, the product of sampling frequency and word length. This can be accomplished by decreasing sampling frequency. Audio compression coder/decoders (codecs) are designed so that all of the frequencies in the range of human hearing (20 Hz – 20 kHz) are preserved. MPEG codecs use perceptual coding techniques to exploit masking properties allowing them to greatly compress a file while maintaining good quality. MPEG is a standard compression method used in the audio and video industry.

Codecs are implemented as either software program or as dedicated hardware. Typically the user loads PCM audio file into the encoder and then transmits and stores its output (encoded) file. The input and output files are not in the same format so a decoder is required to reconstruct the original format audio file. There are two basic types of audio compression codecs: Time domain and Frequency domain. Time domain codecs include M-law and A-law companders, ADPCM, vocoders and linear predictive coders. Frequency domain codecs include two more types, transform and subband. Here we have implemented a subband based MP2 codec [2].

II. SOFTWARE CONFIGURABLE PROCESSOR (SCP)

A software configurable processor combines a traditional RISC processor with a field-programmable instruction extension unit that lets the system designer tailor the processor to a particular application. Figure 1 below gives a pictorial view of the SCP architecture. To add application specific instructions to the processor, the programmer adds a C or C++ function declaration, and the compiler then turns the function into a single instruction. This ability of the software configuration to modify hardware is beneficial to be used for computationally intensive algorithms. In our work we have implemented the MP2 codec on the Stretch's SCP. The major components of the architecture include Tensilica Xtensa ISA (Fixed instruction set), ISEF (Instruction set extension fabric for executing extension instructions), 128 bit WR register files (Wide registers to hold data).

Chetan. P. Kamath is with Manipal Dot Net Pvt. Ltd, Manipal-576104, India. (chetan.kamath@manipal.net)
U. C. Niranjan is with the Manipal Dot Net Pvt. Ltd., Manipal -576104, India-(email: niranjan@manipal..net).

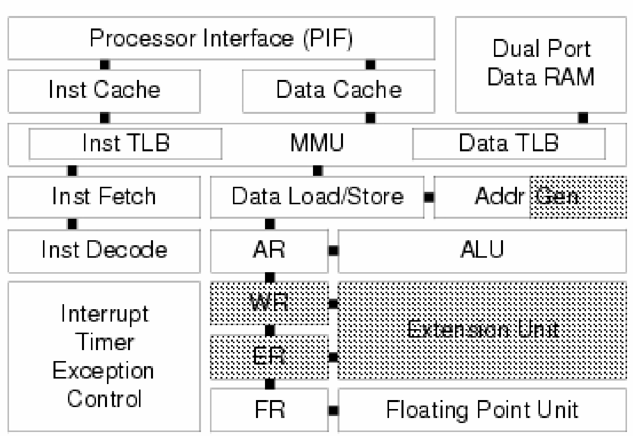


Figure 1.Software Configurable Processor Architecture

StretchBIOS (SBIOS) provides the foundation for applications running on the S5 family of processors. It includes run-time libraries for C and C++, reset, exception, and interrupt handlers and drivers for the peripherals available on S5 processors. SBIOS is included through the linker support packages (LSPs) provided with the Stretch development tools [6].

III. IMPLEMENTATION DETAILS

The MP2 codec has four major parts namely subband filter bank, psychoacoustic model, bit allocation, quantization and framing. The figure 3 and figure 4 below shows the basic structure of MP2 encoder and decoder. The implementation was done keeping in view the MP2 standard and various algorithms specified in the standard.

The advantage of using SCP is that all programming including the on-chip FPGA can be done in C language. We used an audio file (.wav format) which was encoded using MP2 standard and again the same was decoded. To analyze the spectral components of the audio signal a filter bank was used. The output of analysis is referred as subband signal with as many subbands as there are filters in filter bank. The filter bank serves to isolate different frequency components in a signal. In this project 32-subband polyphase filter bank is used. In analysis filter bank the input audio signal is divided into 32 subbands and which uses perceptual coding models of minimum hearing threshold and masking to achieve data reduction.

The MP2 codec consists of an encoder part and a decoder part.

A. ENCODER: The encoder analyzes the spectral components of the audio signal by calculating the Fourier transform and applying a psychoacoustic model to estimate noticeable noise level. In the quantization and coding stage, the encoder tries to allocate the available number of data bits in a way such that both the bit rate and masking requirements are met. Figure 3 gives a basic structure of an MP2 encoder.

B. DECODER: The decoder’s task is to synthesize an audio signal out of the coded spectral components. Decoder searches for the header frame and decodes it. The decoded frame is dequantized and descaled, which are finally reconstructed and frames are sent to synthesis filter to produce PCM audio signal. Figure 4 gives a basic structure of an MP2 decoder.

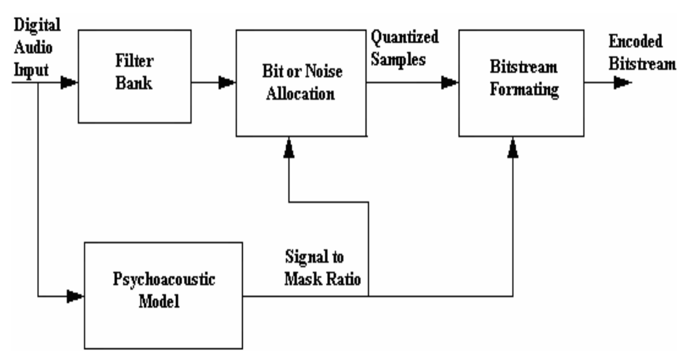


Figure 3. Basic structure of MP2 frequency domain encoder

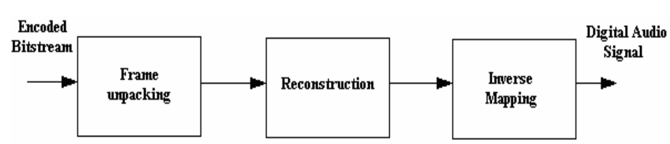


Figure 4. Basic structure of MP2 frequency domain decoder

IV. RESULTS

An input PCM signal which is used to compress using the MP2 standard is shown in Figure 5. The various plots obtained from psychoacoustic model in finding the Signal to mask ratios (SMR) are given below. 1024 samples were taken and FFT was calculated. Figure 6 below gives the list of local maxima. Next the tonal components which are above the absolute threshold are considered and the components (both tonal and non-tonal) which are below the absolute threshold curve are discarded. Figure 7 gives the graph containing tonal components versus absolute threshold. The tonal components which are too close get decimated. Figure 8 gives a plot where the tonal components had been decimated. A curve with components (tonal and non-tonal) above the absolute threshold is calculated which is shown in figure 9 called as global masking threshold curve. Finally, the 32 subbands obtained after calculation is taken and the curve obtained from global masking threshold is used to plot minimum masking threshold curve which is shown in figure 10.

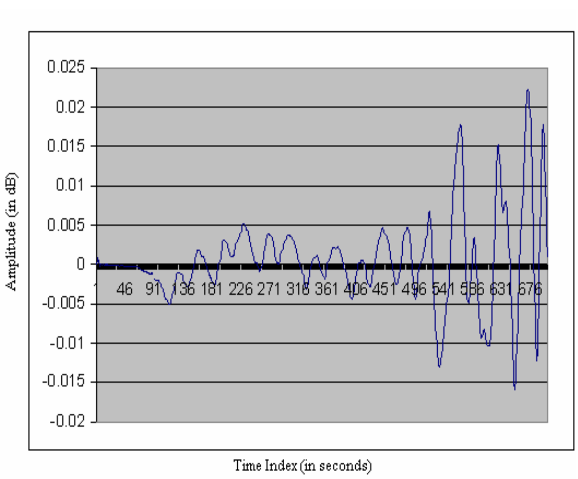


Figure 5. Plot showing the PCM data input signal

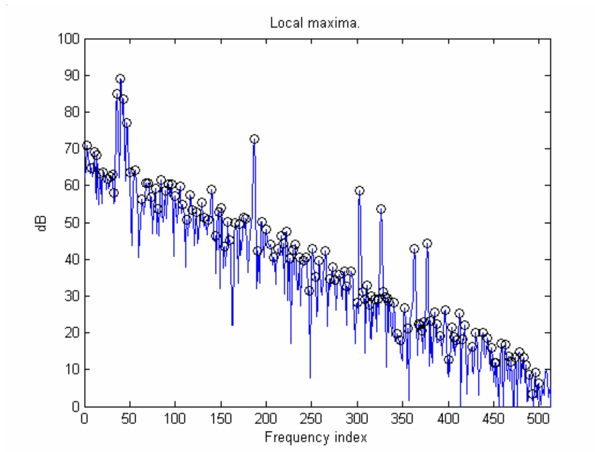


Figure 6. Plot showing List of local maxima obtained from FFT lines

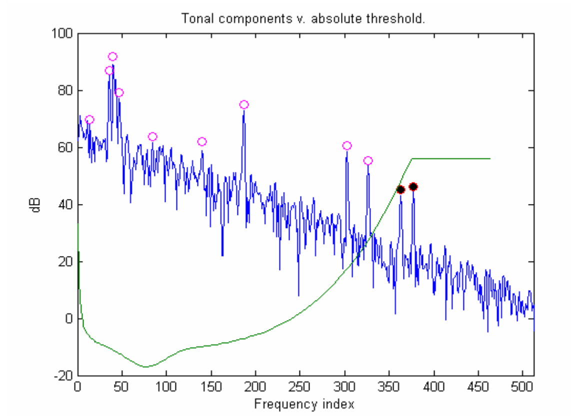


Figure 7. Plot showing Tonal Components Vs Absolute Threshold

Figure 11 and figure 12 shows the decompressed output PCM signal at lower bit rate and higher bit rate respectively. From the graph it is clear that as the bitrate of decoder increases the reconstructed signal is much similar to the input data PCM signal.

Note: ○ Represents tonal components in the signal
● Represents non-tonal components in the signal
__ Absolute threshold
--- Individual masking threshold of tonal components

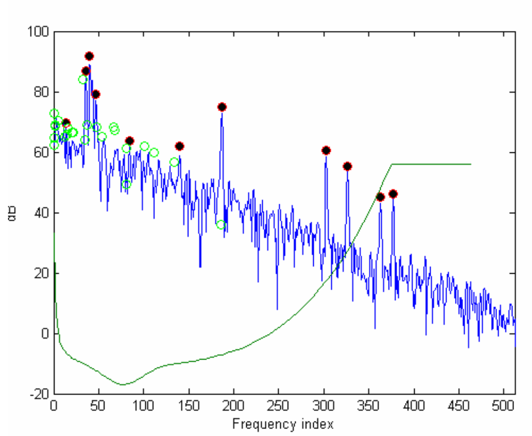


Figure 8. Plot showing Tonal Components are too close to decimate

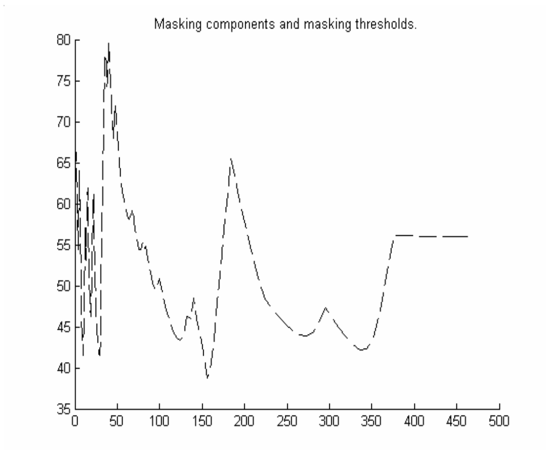
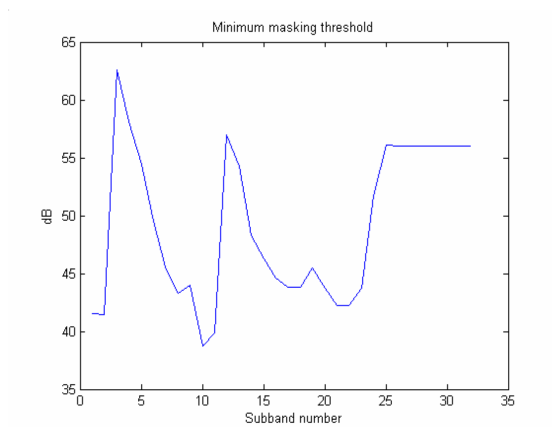


Figure 9. Plot showing Global Masking Thresholds



Figuere 10. Plot showing Minimum Masking Threshold

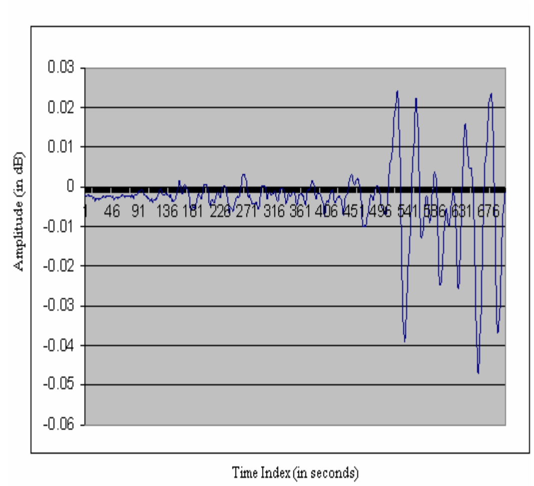


Figure: 11 Plot showing output decompressed PCM data signal at a lower bit rate

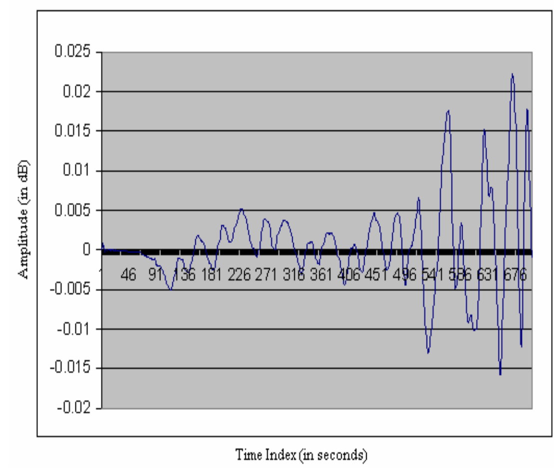


Figure: 12 Plot showing output decompressed PCM data signal at a higher bit rate

---

Note: --- Global Masking Threshold curve

V. CONCLUDING REMARKS

The implementation of the MPEG-1 Layer II (MP2) audio codec for mono recordings attempts to model audio data by mimicking human auditory system using subband filters. This psychoacoustic modeling of the input audio data leads to significant data compression. Though this results in a significantly complex encoder, the decoder is quite simple.

The advantages of the SCP include the fact that it can be programmed in C and coding of compute-intensive part of the program in to on-chip special hardware. The outputs of the implemented Codec are compared with the outputs of other Codec algorithms and found to be acceptable. In future, the algorithm needs to be extended to stereo recordings.

VII. ACKNOWLEDGEMENTS

We are grateful to Dr. Narsimha B. Bhat for providing us the opportunity to carry out this work at Manipal Dot Net Pvt. Ltd. We are also thankful to Krishnaraj and Harishiva for their help throughout the project. We also thank Stretch Inc., CA, USA.

VII. REFERENCES

[1] Charles D. Murphy, *Student Member*, *IEEE*, and K. Anandakumar, *Student Member*, *IEEE*, “Real-time MPEG-1 Audio Coding and Decoding on a DSP chip”, Vol. 43, No. 1, Feb. 1997, pp. 40-47.

[2] Ken C. Pohlmann, “Principles of Digital Audio”, McGraw-Hill, Fourth Edition, 2000, pp. 303-348.

[3] William H. Press, Saul A. Teukolsky, William T. Vellerling, Brain P. Flannery, Numerical Recipes in C The Art of Scientific Computing, Cambridge India: Cambridge University Press, pp. 504-508.

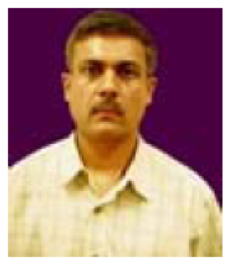
[4] Seonjoo Kim, Yi Li, Heesu Kim, Hanmook Choi and Youngbum Jang, “Real Time MPEG1 Audio Encoder and Decoder Implemented on a 16-bit Fixed Point DSP”, DSP Team, Micro Device Business, Semiconductor Division, Samsung Electronics Co., Ltd.

[5] ISO Standards, “CD 11172-3 coding of moving pictures and associated audio for digital storage media at up to about 1.5 mbit/s”, pp. 1-39.

[6] http://www.stretchinc.com/

About the Authors:

Chetan. P. Kamath is a final year M. Tech student at NMAMIT, Nitte. He is a student member of IEEE.



U. C. Niranjan, is a PhD in Biomedical Engineering from the Indian Institute of Science. He is currently member of the Technical Staff at Manipal Dot Net Pvt. Ltd. He also holds positions of adjunct faculty in the BME and ECE Departments at Manipal Institute of Technology, Manipal.