

Real Time MPEG1 Audio Encoder and Decoder Implemented on a 16-bit Fixed Point DSP

Seonjoo Kim, Yi Li, Heesu Kim, Hanmook Choi and Youngbum Jang

DSP Team, Micro Device Business, Semiconductor Division, Samsung Electronics Co., Ltd.
Suwon P.O. Box 105, Kyungki-Do, Korea 440-600

Abstract

In this paper, it is described a combined hardware and software solution for MPEG1 audio encoder and decoder system implemented on samsung 16 bit fixed-point DSP(Digital Signal Processor) and 2K gate FPGA(Field Programmable Gate Array) logics. The MPEG1 audio encoder and decoder (layer I & II) implemented on SSP1605¹ and a compact/cost effective board are running in real time. The developed code satisfies the MPEG1 standard requirement.

1. Introduction

The MPEG1 audio [1] becomes a popular algorithm for audio data compression due to its high quality with various compression rate, sampling rate and mode. In addition to these variety choices, the MPEG1 is classified into 3 different layers by its complexity. It can compress 1.5Mbit/sec CD quality audio data into 32 to 448 Kbit/sec for layer I and 32 to 384 Kbit/sec for layer II. It requires to support sampling rates of 32, 44.1 and 48Khz. Mode can be selected in one of single channel, dual channel, stereo and joint stereo.

The basic structure of MPEG1 audio can be viewed as 32 subband filter banks. In encoder, the input signal is partitioned into

32 different frequency region, and then decimated for Nyquist sampling rate. Due to the special choice of analysis filters, this filtering process can also be interpreted as windowing and MDCT (Modified Discrete Cosine Transform). For the implementation of subband filtering process on DSP, the overlap-and-add method is used. [2] After frequency samples in each subband are quantized, bits are adaptively allocated to quantized samples in each subband according to the result of psychoacoustic model. We can simply understand the decoding procedure as the reverse of encoding one except for psychoacoustic modelling part.

2. Encoder

The main algorithm of MPEG1 audio is subband filtering and psychoacoustic modelling parts. By using subband filter banks, we can give more bit to quantized samples in the subband with more information. This decision is made from psychoacoustic modelling part using masking effect. This effect is using the fact that weak signals in a subband can be masked by strong ones in neighboring subbands. In other words, this weak signals cannot be heard by human ear. This model is frequently used in high quality audio compression algorithm.

2.1 Encoding procedure

As shown in Fig. 1, audio PCM data is forced to both of the filter bank block and

1. SAMSUNG 16 bit fixed point DSP

the psychoacoustic model block. The processing of these two blocks can be independently accomplished.

In psychoacoustic modelling block, the input time samples are transformed into frequency domain using FFT to obtain the power spectral density. The final output of this block is SMR (signal to mask ratio).

In the filter bank block, each input samples are assigned into each subband. Then, by windowing and MDCT, they are transformed into frequency samples. These frequency samples are now adaptively quantized and bit-allocated using the output of psychoacoustic model, i.e., SMR.

In bitstream formatting block, header, bit allocation information, scalefactor and samplecodes are all combined to a bitstream.

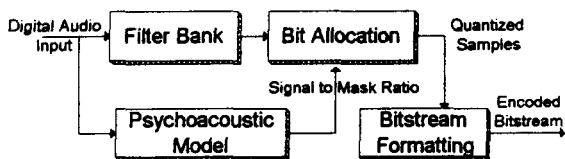


Fig. 1. Block Diagram of Encoding Process

2.2 Implementation of encoder

In implementation of encoder, we simplify the psychoacoustic model part as follows: The relative masking level in a subband is obtained by the convolution of the signal energy and the spreading function [1]. To obtain signal energy, we take 4 times 256 FFT for a frame instead of 2 times 512 FFT for the sake of computational savings. Between 256 point FFTs, 32 samples are spaced to cover a frame.

The SMR, which is the output of psychoacoustic model, is used for bit allocation to the quantized frequency samples resulting

from analysis subband filtering. The number of bit corresponding to bit allocation table in [1] is assigned to the subband having the minimum SMR. Then, the SMR of this subband is updated by a new SNR (signal to noise ratio) from the table in [1]. This procedure is repeated until total available bits are out of short.

Since the analysis subband filtering part is reverse process to the synthesis subband filtering part in decoder, we omit the explanation on it to avoid duplication with decoder.

3. Decoder

3.1 Decoding procedure

The encoded MPEG1 audio bitstream is decoded in frame by frame in this decoder as follows: First, the header of a frame is searched and decoded. From the information of the layer, bitrate and mode included in the header, the scalefactor, bit allocation and audio data are unpacked for all channels. Second, the subband samples are dequantized and descaled. Finally, the reconstructed samples are sent to the synthesis filter to produce the PCM audio signal. Then, the decoder continues to decode the next frame as these three steps until the end of the MPEG1 audio bitstream. The whole decoding flow is shown in Fig. 2.

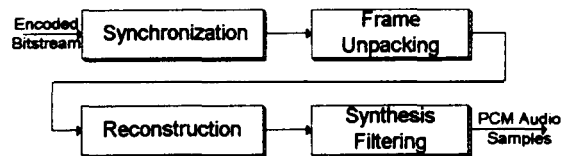


Fig. 2. Block Diagram of Decoding Process

3.2 Reconstruction

At the same time of unpacking audio

data, each 32 subband samples are dequantized and descaled. The combined dequantization and descaling are slightly different from the recommendation[1] so that higher accuracy is achieved on a 16-bit fixed point DSP. The following constants C and D are used for dequantization step[1].

$$C = 32768 \times \left(\frac{2.0}{nlevels} \right)$$

$$D = \frac{nlevels}{2}$$

However, this C cannot be accurately represented with 16-bit fixed-point arithmetic. Decomposing the coefficient C into c1 and c2 as in below, the dequantization is implemented as:

$$c1 = C \quad \text{integer part of } C$$

$$c2 = \frac{C}{c1} - 1.0 \quad \text{correction factor of } C$$

Then, we have

$$x = (1 + c2) \times (c1 \times (S - D))$$

Thus, the reconstructed sample can be obtained as:

$$x = (sf \times (1 + c2)) \times (c1 \times (S - D))$$

where sf represents the scalefactor. Let us denote

$$cs = sf \times (1 + c2)$$

then

$$x = cs \times (c1 \times (S - D))$$

Note that cs does not need to be calculated for each sample, since the scalefactor keeps unchanged for every 12-sample. In implementation on the SSP1605, c1 and c2 are properly scaled and compensated for the shifts in the multiplier. In this way, the RMS(root mean square) of the decoded audio data can be reduced about 50%.

3.3 Synthesis filter

The subband filtering is one of the most compute-intensive portion in the MPEG audio coding standard. However, by making the use of the symmetry of the IMDCT coefficients and windowing coefficients, the total calculation can be reduced in a great

percentage. For an example, in MPEG decoding, the original 2048 multiply-accumulate operations can be reduced to 384 multiply-accumulate operations. At the same time, the data memory size to store IMDCT coefficients is reduced from 2048 words to 248 words.

In MPEG audio decoding, the IMDCT in the subband filter is defined as[1]:

$$V(i) = \sum_{k=0}^{31} S(k) \cos \left[\frac{\pi}{64} (2k+1)(i+16) \right]$$

$$i = 0, 1, \dots, 63$$

The analysis of the IMDCT coefficients shows following facts

1. Not all the coefficients of IMDCT need to be stored. In fact, they have the symmetry of cosine function, so only 248 elements among 64x32 need to be stored in data memory
2. The 64 IMDCT output have the following symmetry [3], i.e., Thus, only half of

$$V(i) = -V(32 - i) \quad \text{for } i = 0, 1, 2, \dots, 15$$

$$V(i) = V(96 - i) \quad \text{for } i = 33, 34, \dots, 47$$

the IMDCT output need to be calculated.

3. Since the half of the IMDCT output can be obtained from the other half, the 1024 subband samples buffer[1] can be reduced to 512. As a result, the window of the synthesis filter has to be performed as even-windowing and odd-windowing by alternating 32 subband samples.

3.4 Decoding accuracy

Two kinds of accuracy are used to verify the MPEG1 audio decoder[4] as

$$\text{full accuracy:} \quad RMS = \frac{2^{-15}}{\sqrt{12}}$$

$$\text{limited accuracy:} \quad RMS = \frac{2^{-11}}{\sqrt{12}}$$

However, only a limited accuracy can be satisfied by such an implementation on 16-bit fixed point DSP. We tested the decoder

with test vectors generated by IRT (in Germany), and resulted in a 0.000027622 average RMS.

4. Hardware Description.

A block diagram of the MPEG1 audio system is shown in Fig. 3. Since SSP1605 is a general purpose DSP, several interface units are required to implement specific DSP application such as MPEG1 audio. Required interface units, which are designed in a FPGA of 2K gate, are MIU (Memory Interface Unit), CIU (Codec Interface Unit) and HIU (Host Interface Unit).

The SSP1605 is a 16-bit fixed point DSP which includes a 32-bit ALU/accumulator, a 16 by 16-bit multiplier with a 32-bit result, two 1K words data RAMs and 8 user definable external registers (EXT0-EXT7). Using these external registers, we can easily implement for various DSP applications. This DSP also includes some additional logics and pins to support the emulation. It supports the IEEE1149.1 standard JTAG interface. Its clock cycle is 33-40 Mhz at 3.3-5V.

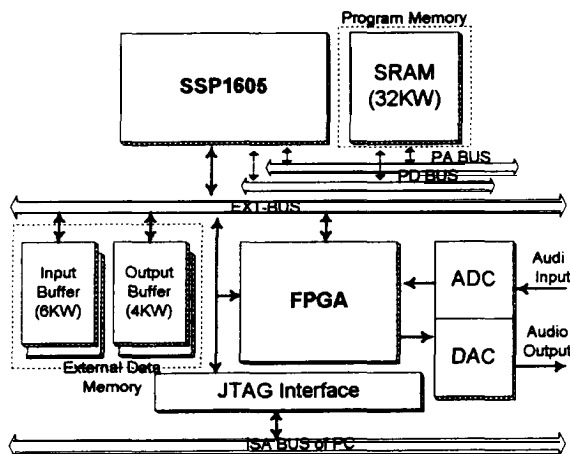


Fig. 3. Block Diagram of MPEG 1 Audio System

4.1 Proposed design of interface unit

The block diagram of proposed design for several interface units is shown in Fig. 4. The interface between DSP and these interface units for MPEG1 audio is done by 16-bit external data bus in SSP1605.

CIU(Codec Interface Unit):

The FPGA includes CIU for interfacing external 16-bit stereo audio codec¹. The CIU includes 32-bit buffer and 32-bit shift register for interfacing A/D and D/A part of codec chip. The double buffering in each of A/D and D/A part reduces the reaction time of the DSP process and can make audio data transfer more safely between DSP and codec chip. In the case of A/D, audio input data is serially transferred to the 32-bit shift register, and then, the content of this shift register is transmitted to 32-bit buffer at every frame sync signal. The content of this buffer is transferred to DSP every interrupt signal synchronized by sampling rates. The data transfer of D/A part is similar to A/D part.

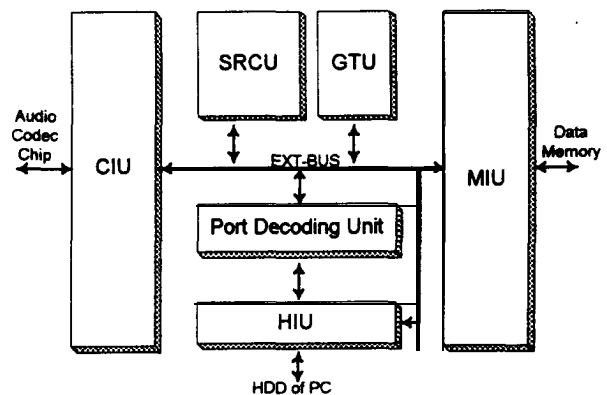


Fig. 4. Block Diagram of Proposed FPGA Design

MIU(Memory Interface Unit):

This unit supports a circular address mode and a block address mode. These

1. CS4216 (Crystal Inc.) is used in proposed design

functions are very useful for reducing code size in block-processing of MPEG1 audio algorithm porting. The loop size of block addressing is defined by setting start address and end address. The maximum access cycle of external memory interface is 4 cycles and the minimum access cycle is 1 cycle.

HIU(Host Interface Unit):

This unit for the interface with Host (PC) is used to transfer encoded bitstream between external data memory and host through external bus in SSP1605. This unit contains a memory-to-host and a host-to-memory buffer. Memory-to-host buffer is used to send encoded bitstream to HDD(Hard Disk Drive) in PC at encoder processing. Host-to-memory buffer is used to send the bitstream stored in HDD to memory at decoder processing. For these data interface, SSP1605 is monitoring flag-signal whenever every interrupt is generated.

SRCU(Sampling Rate Control Unit):

We can change sampling rate which is useful for various DSP applications. Control data is loaded to this block in decoding process when DSP write control data to this block after parsing the header information of encoded bitstream.

GTU(Global Test Unit):

The GTU is assigned to check the internal function units of FPGA. We can monitor the state of CIU, MIU, HIU and SRCU by using LED on/off state.

5. Concluding Remarks

So far, we have discussed the combined hardware and software solution for MPEG1

Audio Encoder and Decoder system using 16-bit fixed point DSP and 2K gate FPGA logics.

We have focused on the reduction of code size and hardware size without degrading audio quality. The main advantage of this system is compact and cost-effective compared to the other MPEG1 Audio Systems. Most of all, cost and performance are incomparably excellent. The developed MPEG1 audio code has the following specification.

	MIPS	Program Memory	Data Memory
Encoder	22	14.2 KWords	13.3 KWords
Decoder	12	6.3 KWords	6 KWords

Furthermore, we can easily obtain an ASIC(Application Specific Integrate Circuit) product which includes the DSP and additional peripheral logics. For the next step, we will develop DSP Embedded ASIC Chip for more compact and cost-effective MPEG1 audio system.

References

[1] Committee draft CD11172-3, Coding of moving pictures and associated audio for digital storage media at up to about 1.m Mbit/s, June, 1992.

[2] Rabiner, L.R., and Schafer, R. W., Digital processing of speech signals, Prentice Hall, Englewood Cliffs, NJ, 1978

[3] Konstantinides, K., Fast subband filtering in MPEG audio coding, IEEE signal proc. letters, v.1, no. 2, Feb., 1994.

[4] Committee draft CD11172-4, Compliance testing, June. 1994.