NEW IMPLEMENTATION TECHNIQUES OF A REAL-TIME MPEG-2 AUDIO ENCODING SYSTEM

Hyen-O OH, Sung-Youn KIM, Dae-Hee YOUN, Il-Whan CHA

ASSP Lab., Dept. of Electronic Eng., Yonsei Univ., Seoul 120-749, Korea
E-mail: oho@cyclon.yonsei.ac.kr

ABSTRACT

In this study, new implementation techniques of a real-time MPEG-2 audio encoding system are presented. The system is developed using general-purpose DSP's. It consists of one master unit and five slave units, and its structure is basically based on our early work [1]. Two fast algorithms are developed and applied to the most compute-intensive routines of the encoding process. These algorithms play a key role to improve the entire system performance. The implemented system is designed to encode audio signal into MPEG-2 layer II bitstream with full configurations up to 5.1 channels and 640Kbps, and intended to support state-of-the-art quality. Generated bitstream can be stored in hard disk on PC or sent to integration system to be multiplexed with corresponding video-bitstream.

1. INTRODUCTION

MPEG-2 multichannel audio (ISO/IEC 13818-3 [2][3]) is a very high-quality subband coding standard. It is a backward compatible multichannel extension to MPEG-1 [4] up to 5 main channels plus a 'low frequency enhancement' (LFE) channel. The bit-rate range is extended up to about 1Mbps.

Many have tried to develop MPEG-2 decoding system using the ASIC technology. But the development of encoding system has been delayed yet because of two reasons. One is that encoder has much higher complexity than decoder. The other is that while the commercial value of the decoding system is expected to grow rapidly, the use of the encoding system is limited in such areas as broadcasting. Nevertheless, the utilization of the MPEG-2 audio in multimedia applications depends on MPEG-2 encoding system running in real-time.

In this paper, some implementation issues of the real-time MPEG-2 layer-II audio encoding system are presented. The professional broadcasting-purpose encoder requires high-end performances and qualities, so it is encouraged to develop an encoder by using general-purpose floating-point DSP's. In our implementation, a floating-point DSP processor (TMS320C30 [5]) was chosen. The system is composed of five slave units and one master unit. The system hardware is based on our early work [1]. However, some parts of the hardware are upgraded and modified with extended-manner to support higher bit-rate and, thus, higher quality.

In the software implementation issues, we focus on two fast algorithms; fast subband analysis and efficient bit-allocation technique. Both methods reduce computational load dramatically. Fast subband analysis algorithm describes that the matrixing operation in the MPEG subband analysis can be efficiently computed by a fast inverse discrete cosine transform. Fast bit-allocation algorithm is developed from statistical characteristics of the encoded bitstream. This fast bit-allocation algorithm gives a solution to extend the system to run at 640Kbps. In addition, the job scheduling to each DSP and the synchronization between multiple DSP's are presented here.

2. SYSTEM OVERVIEW

MPEG-2 layer II is a block processing algorithm. It is processed frame by frame. Each frame is composed of 1152 samples. An analysis subband filterbank is used to split input signal into 32 subband signals. Each subband signal consists of 36 subband samples. The calculation of the scalefactor for each subband is performed every 12 subband samples. The maximum of the absolute value of these 12 samples is determined. In parallel to the analysis filter bank operation, PAM (Psychoacoustic Model) is calculated. The result of the PAM calculation is the SMR (Signal-to-Mask) ratio of the each subband. Bit-allocation of the 32 subbands is calculated on the basis of the SMR ratios of all the subbands. Subband samples are quantized according to the bit-allocation. For more detail, refer [4].

Fig. 1 illustrates a block diagram of the real-time MPEG-2 audio encoding system developed in this work. It consists of six units; one master and five slaves. Each unit comprises two DSP's (they are referred to as DSP-a and DSP-b, respectively), memories and peripheral circuits.
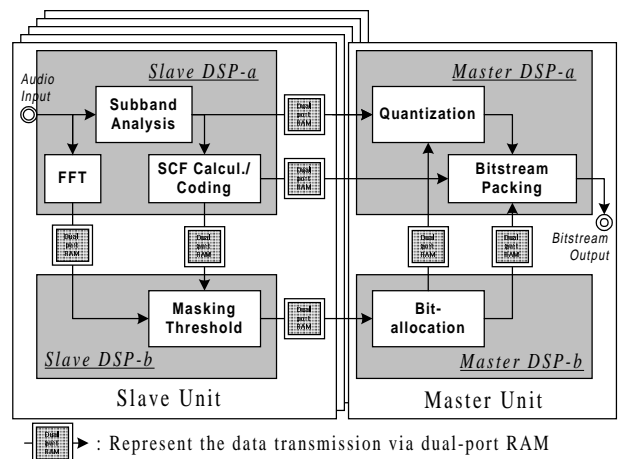
Figure 1. Block diagram of the MPEG-2 audio encoding system.

Each slave unit receives one channel audio samples from external ADC module and converts them into 36 subband samples, and performs the FFT (DSP-a). Simultaneously, it computes the SMR using the results of the FFT (DSP-b). The master unit receives the results from five slave units and allocates the available bits to the subband samples based on the SMR (master DSP-b). The subband samples are quantized to the number of allocated bits and packed into a bitstream (master DSP-a). The generated bitsream can be sent to a hard disk on PC or to video/audio multiplexer of the MPEG-2 broadcasting system.

To increase the sound quality of the real-time system, new specifications of the master unit has been designed. In new version of the master unit, another DSP has been added. As a result, the master and slave units in the real-time encoding system have almost identical structures. Fig. 2 illustrates the core architecture of each unit (phrases in parenthesis are for master unit).

Both DSP's have their own local 32K × 32 bits ROM (only DSP-a in the master unit has 64K × 32 bits ROM) containing the program codes, constant tables, and boot-loader. To avoid unnecessary waiting cycles for fetching from ROM, they also have 32K × 32 bits SRAM. During boot-loading, the necessary program and tables corresponding to the operating mode are fetched and stored into the SRAM. A 2K × 32bits dual-port RAM is used to exchange the data between DSP-a and DSP-b.
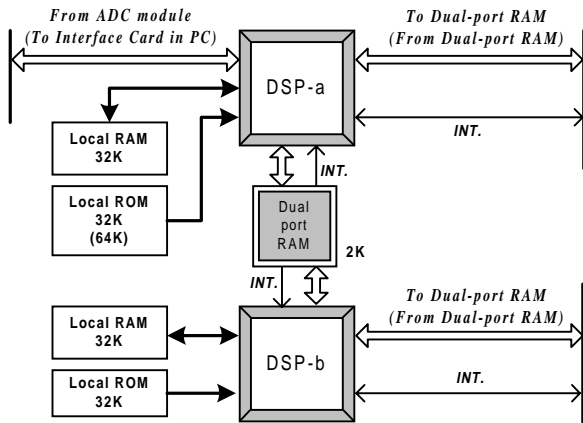
Figure 2. The architecture of the slave (master) unit.

The optimized timing diagram of the overall encoding process is illustrated in Fig. 3. The encoding system starts its process with the slave DSP-a when the 1st FFT-buffer (F#1) is full. And all routines for one frame encoding are finished in the master DSP-a. The details of the timing, input sample control method, buffer management, and data transfer between DSP's can be found in our early paper [1].

3. NEW TECHNIQUES FOR EFFICIENT IMPLEMENTATION

There are several issues associated with the real-time software to implement 5channels MPEG-2 audio encoding system. In this section, we will introduce two fast algorithms for the real-time implementation and show the effectiveness of the algorithms.
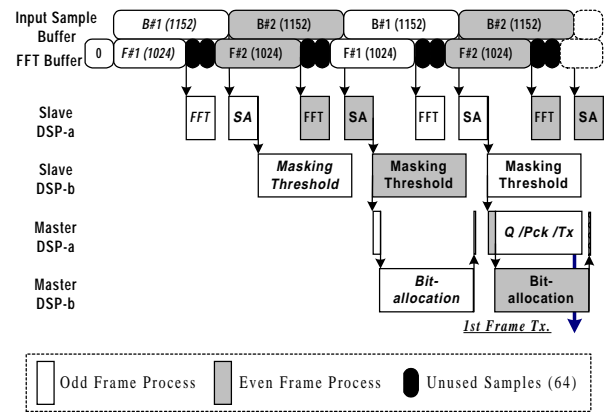
Figure 3. The optimized timing diagram of the processes including buffer management.

The first algorithm is concerned with a fast subband analysis which is the most computationally demanding process among the jobs in the slave DSP-a. The second algorithm is concerned with a fast bit-allocation which has been developed from statistical characteristics of the encoded bitstream. This fast bit-allocation algorithm plays a key role to realize the real-time MPEG-2 encoding system at high bit-rate modes.

3.1 Fast Subband Analysis

The subband analysis is the most computationally demanding process among the jobs allocated to the slave DSP-a. Hence, fast subband analysis algorithms are of prime importance in low-delay encoding and decoding systems. As an effort to obtain this objective, the algorithm proposed by *K. Konstantinides* [6] was applied to the encoding system. Using this algorithm, the matrixing operation in the MPEG subband analysis can be efficiently computed by a fast inverse discrete cosine transform (IDCT). There have been many efficient algorithms to compute the IDCT.

The fast IDCT algorithms can be divided into two catagories. The first one is related to the computation of the IDCT using the FFT algorithm, and the second one is the consideration of the factorization of the IDCT matrix. In order to choose the most efficient algorithm, three IDCT algorithms proposed by *B.D. Tseng* [7], *M.J. Narasimha* [8], and *B.G. Lee* [9] were tested. The first two algorithms in [7] and [8] are based on the FFT, while the third algorithm uses the decimation-in-time scheme.

The test was conducted by programming the three IDCT algorithms using TMS320C30 Assembly language, and the execution time for each algorithm was measured. Results are summarized in Table 1. The results show that the Lee's fast IDCT method is the most efficient among the three methods considered. However, it should be mentioned that, if the MPEG-2 encoding system was implemented via ASIC designing, the methods by Tseng or Narasimha would have been better choice, because the FFT may offer better architecture to be designed by logics.

By using the 32-point Lee's fast IDCT method, rather than the matrixing operation, the encoding system could save the computations by around 43% compared to the original method.

Table 1. Comparison among three IDCT algorithms.
*The execution times (msec) are experimentally acquired for the whole executions of 36 times 32-point IDCT including pre/post processes.

	Standard	Tseng [8]	Narasimha[9]	Lee [10]
Exec. Time*	6.35	5.18	4.95	4.38

3.2 Fast Bit-allocation Algorithm

MPEG bit-allocation algorithm has the following characteristics:

- In each iteration, bits are given only to one band. Then, we are obliged to perform that operation many times to control bit usage precisely, because allocation varies very slowly.
- Only one error (MNR) curve is used for all channels to be quantized. So, if the number of channel is increased, the number of subband to be compared to find the minimum MNR value is also increased.
- The bit-allocation chosen in previous frames is not taken into consideration in the current frame.

The computational demand of the bit-allocation process grows as the channel number increases. Also, the bit-allocation process consists of iterative operation for finding minimum values. In the real-time encoding system, TMS320C30 does not support efficient operations to find a minimum value. Instead, a series of 'compare' and 'conditional branch' instructions should be used. When the entire bit-allocation process is implemented in DSP-a of the master unit, around 10.3msec is required for 2channels-192Kbps mode, and 62.5msec is needed for 5channels-640Kbps mode. Considering that the encoding process for one frame data must be ended within 24 msec, the encoding system fails to perform the 5channels-640Kbps MPEG-2 encoding process in real-time.

In this study, we developed a fast bit-allocation algorithm in order to implement the 5channels-640Kbps MPEG-2 encoding system using the hardware described in the previous section.

The fast bit-allocation algorithm developed in this study is based on the statistical characteristics of the final bit-allocation indices. To analyze the statistical characteristics of SMR and the amount of the allocated bits to each subband, intensive experiments were conducted. The experimental results of the final bit-allocation indices for three kinds of test materials are shown in Fig. 4. The test materials having spectral distributions different from one another were used for the experiments, and the final bit-allocation indices used for this analysis were obtained from 3000 frames' encoding/decoding procedure for each test material. From the results, it could be found that there were certain thresholds of indices, and bits were always allocated above the thresholds for all frames and for all test materials. This implies that the allocation iterations required to reach the thresholds are unnecessary, so that they can be eliminated by allocating bits up to the thresholds in prior to the bit-allocation iteration, which is referred to as pre-allocation. Another useful result of the analysis was that there was also a group of subbands, in high frequency bands, in particular, having always zero indices for all frames.

Thus, via a pre-exclusion, it is also possible to exclude those subbands from candidates of the minimum MNR value search.
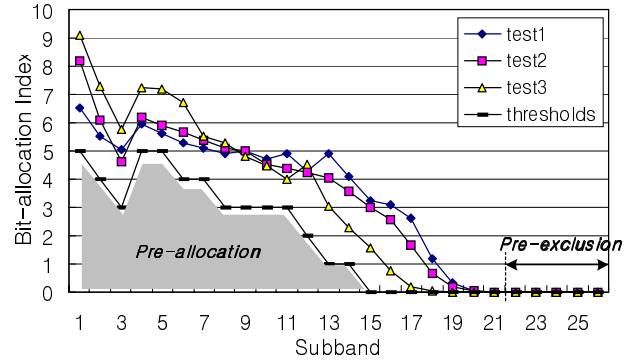
Figure 4. Results of statistical analysis for bit-allocation.

By using the two approaches described above, *i.e.,* the pre-allocation and the pre-exclusion, we could successfully implement the real-time bit-allocation for all configurations up to 5 channels-640Kbps. Moreover, the number of bits allocated to each subband is transmitted to the decoder as a side information. So, the encoder may have a little degree of freedom to perform bit-allocation at given bit-rate. As a result, it will also be possible to control the execution time by managing the level of threshold indices and the range of exclusion.

Using the proposed fast algorithms, the bit-allocation process takes just 17.5msec for 5 channels-640Kbps which is corresponding to only about 28% of the conventional MPEG bit-allocation operations.

4. EXPERIMENTAL RESULTS

To verify the performance of the encoding system, we compared each subroutine's output with that of MPEG standard encoder [10] written in c language. During the comparison, we confirmed that the standard encoder and the encoding system built in this study produced exactly the same outputs at each step to the final bitstream. An example of the verification test for bit-allocation process is presented in Fig. 5. Here, it is found that the final bit-allocation result using proposed pre-allocation/pre-exclusion method is identical to the original bit-allocation result acquired from the execution of [10].

After the step-by-step verifications, program codes corresponding to the jobs for each DSP were written on ROM's to build a stand-alone system. We tested the real-time encoding system for various materials recorded in multichannel. The generated bit-streams were stored in hard disk. Later, the bit-streams were decoded by a standard decoder [10]. The quality of the decoded signals was evaluated via subjective tests. However, it should be emphasized that the subjective test for the encoding system can not hold its importance any more, because the output bitstreams are identical to those of the standard encoder's.

The execution times needed by each routine are summarized in Table 2. Both results obtained before and after optimization are shown. To calculate the execution time, we counted the number

of input samples stored in the input buffers during the execution period of each subroutine. Each slave DSP-a receives input sample exactly every 20.833micro-seconds (that is, 1/48000 KHz). Results in Table 2 indicate that the most significant reduction in execution time is made in the bit-allocation procedure.

Overall encoding delay, the total execution time from the first sample in to first frame bitstream out, can be calculated from Fig. 3 and Table 2 using the following equation:

Encoding Delay

  = (Time for 3-Frames in) + (S.B. Analysis) + (SCF coding)

  + (Data Fetching) + (Quantization) + (Packing)

  = 24.0 × 3 + 6.22 + 1.81 + 0.93 + 3.42 + 6.21

  = 90.59 (msec)

Table 2. The execution time (msec) results for 48KHz, 5channels-640Kbps mode**.**

DSP'S		ROUTINES	EXEC. TIME	
			BEFORE	AFTER
Slave	a	*Subband Analysis*	*10.82*	*6.22*
		SCF Coding	2.06	1.81
		FFT/Power Spectrum	4.97	3.65
	b	Masking Threshold	18.67	18.28
Master	a	Data Fetch from Slaves	-	1.81
		Quantization	-	3.42
		Bit-stream Packing	-	6.21
		Bit-stream Transmitting	-	7.52
	b	*Bit-allocation*	*62.53*	*17.52*

5. CONCLUSION

In this paper, we described new implementation techniques of a real-time MPEG-2 audio encoder. The implementation system was developed using general-purpose floating-point DSP's. The system employs twelve highly-parallel DSP's that are distributed onto five slave units and one mater unit. Two noble techniques are developed in the software implementation. The fast subband filtering using Lee's fast DCT was implemented in 57% of the original method. The fast bit-allocation via pre-allocation and pre-exclusion is executed only in about 28% of conventional method. And it enables higher bit-rate multichannel encoding.

The implemented system covers up to 5.1channels-640Kbps encoding. Since the implemented encoding system uses floating-point DSP's, the decoded sound quality is exactly the same that can be obtained by using the MPEG standard encoder written in c language. The system can be used in such application as the audio codec for digital TV (also HDTV) broadcasting.
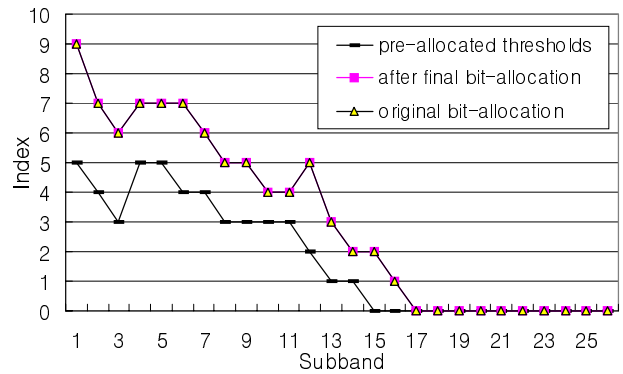
Figure 5. Example of the experimental verification test for bit-allocation.

6. REFERENCES

[1] S.Y. Kim, H.O. Oh, K.S. Kim, D.H. Youn and J.Y. Lee, "A Real-time Implementation of the MPEG-2 Audio Encoder", *IEEE Transactions on Consumer Electronics*, Vol.43, No.3, Aug. 1997.

[2] ISO/IEC JTC1/SC29/WG11 No.703 "Generic Coding of Moving Pictures and Associated Audio - CD 13818-3(Part3. MPEG-Audio)", Mar., 1994.

[3] ISO/IEC JTC1/SC29/WG11 No.1519 "Generic Coding of Moving Pictures and Audio - CD 13818-3(Part3. MPEG-Audio)", 2nd Edition, Feb., 1997.

[4] ISO/IEC JTC1/SC29/WG11 No.71 "Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1.5 Mbit/s - CD 11172-3(Part3. MPEG-Audio)"

[5] Texas Instruments, *TMS320C30 User's Guide*, 1992.

[6] K. Konstantinides, "Fast Subband Filtering in MPEG AUDIO Coding.", *IEEE Signal Processing Letters*, vol. I, no. 2, pp. 26-28, Feb., 1994.

[7] B. D. Tseng, W. C. Miller, "On Computing The Discrete Cosine Transform", *IEEE Trans. on Comput.*, vol. C-27, pp. 966-968, Oct. 1978.

[8] M. J. Narasimha, A. M. Peterson, "On The Computation of The Discrete Cosine Transform", *IEEE Trans. on Commun.*, vol. COM-26, pp. 934-946, Jun. 1978.

[9] B.G. Lee, "A New Algorithm to Compute the Discrete Cosine Transform", *IEEE Trans. on Acoust., Speech, Signal Processing*, pp. 1243-1245, Dec. 1984.

[10] ftp://ftp.tnt.uni-annover.de/pub/MPEG/audio/mpeg2/