A Fast Bit Allocation Method for MPEG Layer III

Chi-Min Liu, Chin-Ching Chen, Wen-Chieh Lee, and Szu-Wei Lee

Department and Institute of Computer Science and Information Engineering
National Chiao Tung University, Hsinchu, 30050, Taiwan
E-Mail: cmliu@csie.nctu.edu.tw

Abstract

The MPEG/Audio layer III has been widely used to store the CD-quality music sequences. The layer III is an asymmetric codec in the sense that the complexity of the audio encoder is much higher than that of the decoder. A key task leading to the higher complexity is the bit allocation. The MPEG/Audio layer III encoder has adopted the schemes such as the nonuniform quantization and Huffman coding to achieve high coding efficiency. However, all these schemes do not have the problem in quality and bit-rate controls. The usual method follows an iterative approach to perform bit allocation to achieve both quality and bit rate tuning. This method has immense complexity and will sacrifice the quality to fit bit rate. This paper presents an efficient allocation method, which have merits in both complexity and quality control.

1. Introduction

MPEG audio coding with its three compression layers has been widely used in consumer electronics, telecommunications, and the broadcasting. Among these three layers, layer III [1] has the highest compression ratio and has been widely used to compress the CD-quality music sequences. The MPEG/Audio layer III is an asymmetric coding in the sense that the complexity of the audio encoder is much higher than that of the decoder. The bit allocation is one of the main tasks leading to the high complexity in the encoder. This paper proposes an efficient bit allocation method for the encoding process.

Fig. 1 illustrates the block diagram of the encoding process in layer III. A hybrid transform transfers an audio sequence into the spectral lines frame-by-frame. The spectral lines are nonuniformly quantized, Huffman coded, and packed into a MPEG bit-stream. In the process, the nonuniform quantizer quantizes the spectral lines subject to the bit allocator, which decides the quantization manners with consideration to the resultant audio quality and the required bits. In aid of the allocator, the psychoacoustic model and the bit-reservoir provide the perceptual resolution and the available number of bits for each audio frame. For the encoding process in Fig. 1, the bit allocator is the key block deciding the final coding quality.

The objective of the bit allocator in layer III is to assign suitable parameters to quantizer to achieve the best audio quality under the restricted bit number. Hence the control over quality and the bit number through the quantization parameters are two fundamental requirements. However, both the quality and the bit-rate controls are difficult in MPEG layer III. The difficulty from the quality control can be examined from the quantizers. The quantizer in MPEG layer III is a nonuniform quantizer, which means that the noise introduced from the quantization depends on the quantized value instead of just the quantization parameters like the general uniform quantizers. In other words, the general quality-control manners, which assign quantization parameters according to the perceptually allowable noise does not work for nonuniform quantizers.

The difficulty on the bit-rate control can be examined from the Huffman coder in MPEG layer III. The Huffman coder assigns different bit-length to different values, which means that the bit consumed should be obtained from the quantization results, and can not be from the quantization parameters alone.

2. Related Researches

The above two difficulties lead to the problem in evaluating the quantization parameters. An iterative method has been proposed in [1],[2] to solve the problem. The method evaluates

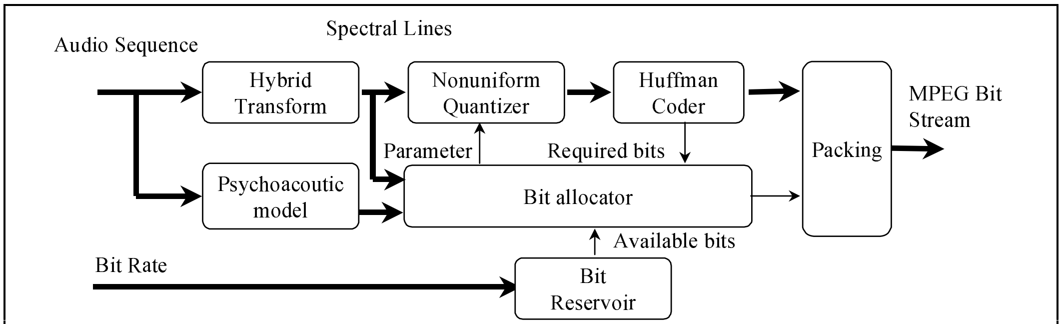

Fig. 1 Block diagram of the MPEG/Audio layer III

the quantization parameters through two iteration loops: the inner loop and the outer loop. The inner loop adjusts in iteration the parameter values to fit to the limited bits the consumed bits, which are obtained by performing quantization and Huffman coding for spectral lines. The outer loop adjusts in iteration the parameter values to fit to a perceptual criterion the quantization noise that needs to be evaluated by performing the inverse quantization. The method can be examined from the complexity and the induced audio quality. The complexity of the method for a frame with 576 spectral lines is $O(576 \cdot U \cdot I \cdot \eta + 576 \cdot U \cdot \kappa)$, where the U and I are respectively the numbers of outer iterations and inner iterations while the η and κ are the computation complexity to handle a spectral line in an inner loop and an outer loop, respectively. The inner loop complexity η will be from the quantization and the Huffman coding of a spectral line while the outer loop complexity κ from the inverse-quantization. Both the complexity η and κ are high. Also, the number of iterations U and I depends on the initial values of quantization parameters and the adjust manners. For the initial values and the adjusting manners in the draft [1], the magnitude of $U \cdot I$ maybe as large as forty. The complexity will be even larger than the total complexity of hybrid transform and the psychoacoustic model in Fig. 1. The second problem is on the quality of the coded audio. Since that there are two separate rules controlling the quality and bits consumed in two loops, there may lead to infinite loops, generally referred to as "dead-lock problem". A general method to handle the deadlock problem is to set the maximum number of iterations; however, the quality should be sacrificed to meet the bit rate constraint. This paper presents a new bit-allocation method, which has merits in both complexity and audio quality.

3. New Bit Allocation Method

As illustrated in Fig. 1, the hybrid transform transfers an audio sequence into the spectral lines frame-by-frame. These spectral lines are divided into several groups (referred to as scalefactor bands), and nonuniformly quantized. The formula of the quantization is

$$I_j(k) = nint\{[X_j(k) * scale_j / gain]^{3/4} - 0.0946\} \qquad (1)$$

where *j* and *k* are respectively the index of the scalefactor band and spectral lines in bands while $X_j(k)$ and $I_j(k)$ are respectively the spectral lines before and after quantization. In the above formula, the *nint* denotes the function of the nearest integer while the $scale_j$ and *gain* are two quantization parameters. The $scale_j$ controls the quantization noise of the associated band relative with the other bands. The *gain* controls the overall number of consumed bits. The $scale_j$ is the quality-related parameter and the *gain* is the bit-related parameter.

The new bit-allocation method is developed with two considerations: First, we calculate $scale_j$ without iteration through

$$scale_j \approx \left\{\left(27 * M_j * NMR_j\right) / \left(4 * E[X_j^{1/2}(k)]\right)\right\}^{2/3} * C \qquad (2)$$

where NMR_j is the noise-to-masking ratio of band *j*. The bit-related parameters can be calculated from the efficient iteration method. *E[*]* is the expectation function and *C* is an arbitrary constant. The formula in (2) eliminates the outer loop complexity κ since this complexity κ is from the inverse-quantization.

Secondly, this paper presents an efficient iterative method to fit the bit-related parameters to the bit rate. In this iteration method, the paper set the initial value of the *gain* as that from the previous frame and adjusts the gain according to the resultant bits.

4. Results

The complexity of the new method is $O(576 \cdot \Gamma \cdot \eta)$ where Γ is the number of iterations. The complexity in each iteration is the same as that in the inner loop of the method mentioned above. To compare the complexity with that in [1],[2] mentioned above, we can check $U \cdot I$ and Γ for the two methods.

Table. 1 lists the average number of iterations for both allocation methods. The results indicate that the new method has improved more than ten times the complexity. The music sequences, which have been used are the same as those used in our previous papers [4],[5]. Also, for quality consideration, this paper has adopted the quality criteria mentioned in [3] as the iteration objective. The criterion illustrates the graceful degradation criterion for audio coding; that is, the condition which best degrades the audio quality when the bit-rate is not high enough to have the coded quality the same as the original audio music. Our informal subjective quality measure indicates the better quality at bit rate 128 kbits/sec which is most widely bit rate for internet music transmission.

Music Type	The average iteration number ($U \cdot I$) in [1]	The average iteration number (Γ) in the new method
Flute	38.094096	3.409902
Violin	38.141144	3.70203
Harp	42.045818	3.627614
Drum	35.066728	3.955105
Female	40.575967	3.589319
Male	40.008610	3.810886

Table. 1 The average number of iteration

References

[1] ISO/IEC IS 11172-3, Coding of Moving Pictures and Associated Audio for Digital Storage Media up to About 1.5 Mbit/s. – Part 3: Audio, 1992.

[2] K. Brandenburg, "OCF – A New Coding Algorithm for High Quality Sound Signals," in Proc. ICASSP 1987. pp. 141-144.

[3] X. Wei, M. J. Shaw and M. R. Varley, "Optimum Bit Allocation and Decompositon for High Quality Audio Coding," in Proc. ICASSP 97, pp. 315-318.

[4] C.M. Liu and J.C. Liu, "A new intensity stereo coding scheme for MPEG1 Audio encoder-- Layer I & II," *IEEE Trans. on Consumer Electronics*, Aug., vol. 42, no. 3, pp.535-539, 1996.

[5] C.M. Liu, W.C. Lee and S.Y. Juang, "Design of the coupling schemes for the Dolby AC-3 coder in stereo coding," *IEEE Trans. on Consumer Electronics*, Aug., vol. 44, no. 3, pp. 878-882, 1998.