EFFICIENT BIT ASSIGNMET STRATEGY FOR PERCEPTUAL AUDIO CODING

*Cheng-Han Yang and Hsueh-Ming Hang***

Department of Electronics Engineering
National Chiao Tung University Hsinchu, Taiwan, R.O.C.
**hmhang@cc.nctu.edu.tw; Fax: (886)-3-5723283

ABSTRACT

For the purpose of efficient audio coding at low rates, a new bit allocation strategy is proposed in this paper. The basic idea behind this approach is **"Give bits to the band with the maximum NMR-Gain/bit"** or **"Retrieve bits from the band with the maximum bits/NMR-Loss"**. The notion of "bit-use efficiency" is suggested and it can be employed to construct a bit assignment algorithm operated at band-level as compared to the traditional frame-level bit assignment methods. Based on this strategy a new bit assignment scheme, called Max-BNLR, is designed for the MPEG-4 AAC. Simulation results show that the performance of the Max-BNLR scheme is significantly better than that of the MPEG-4 AAC Verification Model (VM) and is close to that of TB-ANMR [3], which is the (nearly) optimal solution. Moreover, the Max-BNLR scheme has the advantages of low computational complexity comparing to TB-ANMR.

1. INTRODUCTION

Many highly efficient and high quality audio coding schemes have been developed and proposed to meet the growing demand of multimedia applications. The MPEG-4 Advanced Audio Coding (AAC) is one of the most recent audio coder specified by the ISO/IEC MPEG standards committee [1]. It is a very efficient audio compression algorithm aiming at a wide variety of applications, such as Internet, wireless, and digital broadcast arenas [2]. For the applications where the bandwidth is very limited, the low rate audio coding with good quality becomes essential.

The procedure of bit assignment is one of the most important elements in audio coding. Particularly, when bits are scare, how to make the best use of the limited bits is critical in producing the best quality audio. Up to now, the popular strategies on bit assignment are as follows ([2][3][5]).

1. "Give bits to the band which has the largest value of NMR (perceptual distortion)."
2. "Give bits to the bands of which the distortion is larger than the masking threshold".

In these strategies, the bit-use (giving away bits) is considered at frame-level and only the value of distortion is taken into consideration at band-level. Hence, it is hard to control the bit-use efficiency (the NMR improvement due to adding one bit) at band level and thus results in a less efficient compression scheme.

In this paper, we suggest the notion of *bit-use efficiency* and propose a new strategy to improve the bit-use efficiency, which can be evaluated at band-level. Moreover, a new bit assignment scheme based on this new strategy is proposed for MPEG-4 AAC.

The organization of the paper is as follows. Section 2 describes the aforementioned new strategy. A new AAC bit assignment scheme is delineated in section 3. Finally, the complexity analysis and the simulation results are presented in section 4.

2. EFFICIENT BIT-USE STRATEGY

How to make use of the bits more efficiently is always the key issue in audio coding. The traditional strategies, "Giving bits to the band with the largest NMR" or "Giving bits to the bands of which the distortion is larger than masking threshold", do not necessarily provide the best bit-use efficiency. For example, there are two candidate bands, A and B, and their NMR characteristics are listed in the table below. Which band should the first available bit be assigned to? In this table, NMR-Gain/bit means the gain in NMR by allocating one bit to this particular band. A more precise definition of NMR-Gain/bit will be given in section 3.

Band	NMR (dB)	NMR-Gain/bit
A	3.5	0.5
B	3	1.5

Following the traditional strategy, we would assign this one bit to band A; however, considering the bit-use efficiency, this one bit should be assigned to band B so that the overall NMR reduction is maximized. The essence of this new strategy can be summarized by the following statements.

"Give bits to the band with the maximum NMR-Gain/bit" or ***"Retrieve bits from the band with the***

0-7803-7663-3/03/$17.00 ©2003 IEEE

maximum bits/NMR-Loss”, where bits/NMR-loss is the number bits we save if we give away one unit of NMR.

3. MAX BITS/NMR-LOSS BIT ASSIGNMENT SCHEME

In this section, a new bit assignment scheme designed for MPEG-4 AAC based our new strategy is described. First, we define NMR-Gain/bit and bits/NMR-Loss by the following equations.

$$NMR-Gain/bit = (NMR_{ref} - NMR_{new})/(bits_{new} - bits_{ref}) \quad (1)$$

and

$$bits/NMR-Loss = (bits_{new} - bits_{ref})/(NMR_{new} - NMR_{ref}) . \quad (2)$$

Figure 1 is the block diagram of the Max bits/NMR-Loss based bit assignment scheme. Each step in Figure 1 will be elaborated in the following sub-sections.

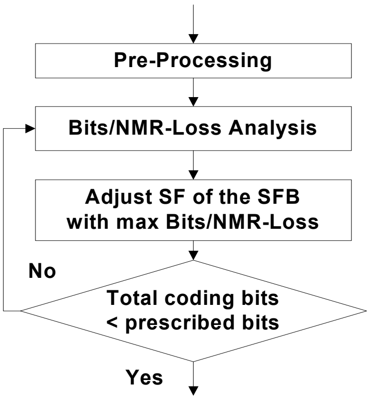

Figure 1. Max bits/NMR-Loss bit assignment scheme

3.1. Pre-Processing

The pre-processing step is to initialize two of the major parameters in the *bits/NMR-Loss* analysis: the reference NMR and the reference bits. There are no particular values associated with these parameters and thus the design of the pre-processing is case-dependent. In our implementation, we set the reference *NMR*=1 (0dB) for all the scale factor bands (SFB) at the beginning of processing a frame. After that, the reference scale factor (SF) for each SFB and the reference bits are calculated based on the input audio data.

3.2. Bits/NMR-Loss Analysis and SF Adjustment

In this scheme, only one SF value (of one SFB) is adjusted in one adjustment iteration. The detailed process is described below.

1. Initialization. Get the reference bits (B_{ref}), and the reference SFs (sf_{ref}) and NMRs (NMR_{ref}) for all SFBs (*N_SFB* SFB in total) from the pre-processing step. Start the max *bits/NMR-Loss* analysis from the first SFB and thus set the SFB index i=1.
2. Find the local max *bits/NMR-Loss* ratio of the ith SFB, $BNLR_i$, by computing
   $BNLR_i = \max_{sf}\{(B_{ref} - B_{sf})/(NMR_{sf,i} - NMR_{ref,i})\}$
   $\forall sf \text{ and } sf_{ref,i} < sf \le sf_{\max,i}$
   The B_{sf} is the new value of the total coding bits in the current frame if the SF value (of the ith SFB) is changed from $sf_{ref,i}$ to $sf_{sf,i}$. The $sf_{\max,i}$ is the SF value that quantizes all the spectral coefficients in the ith SFB to zero. The local optimal SF (of the ith SFB), $sf_{opt,i}$, is the SF with the maximum *BNLR*. The local optimal coding bits of the ith SFB, $B_{opt,i} = B_{sf_{opt,i}}$, is also recorded.
3. If $i<N\_SFB$, update i to i+1 and go to step 2.
4. Find the global maximum *bits/NMR-Loss* ratio, $BNLR_{global}$, by computing
   $BNLR_{globe} = \max_i\{BNLR_i\} \quad \forall i,\ 0 \le i < N\_SFB$
   The global optimal SFB, sfb_{global}, is the SFB that has the $BNLR_{global}$. Then, the global optimal SF, sf_{global}, is the local optimal SF of the sfb_{global}*-th* SFB. Similarly, the global optimal coding bits, B_{global}, is the coding bits of the sfb_{global}*-th* SFB.
5. Set the SF of the sfb_{global}*-th* SFB to sf_{global}. Update parameters for the sfb_{global}*-th* SFB; that is, $sf_{ref,sfb_{global}} = sf_{global}$ and $NMR_{ref,sfb_{global}} = NMR_{sf_{global},sfb_{global}}$.
6. Compare B_{global} to the prescribed rate, B. If $B_{global} > B$, update B_{ref} to B_{global} and go to step 2.

Note that, in performing the local maximum *bits/NMR-Loss* ratio analysis in step 2, only the SF of one SFB that is being examined is modified. The SF of the other SFBs are kept unchanged.

3.3. Trellis-Based Optimization on HCB

Total coding bits calculation in step 2 in the **Bits/NMR-Loss Analysis** (in sub-section 3.2) is one of the most computational-intensive processes. When the SF for each SFB is determined, the quantized spectral coefficients are also fixed. Before calculating the total coding bits, the HCB for each SFB has to be chosen first. The MPEG-4 AAC Verification Model (VM) has a simple algorithm for this purpose; however, a more efficient algorithm is needed for HCB decision. Thus, we adopt the Viterbi-based approach in this paper.

The problem for finding the optimal HCB can be reformulated as minimizing the following cost function:

$$C_{HCB} = \sum_i b_i + R(h_{i-1}, h_i)\ , \tag{3}$$

where b_i is the coding bits of the quantized spectral coefficients for the *i*th SFB, h_i is the HCB for the *i*th SFB, and R is the run-length coding function (bits needed) for coding HCB. We find that the contribution of h_i to C_{HCB} depends only on the *previous* choice, h_{i-1}. Therefore, the minimization of C_{HCB} can be achieved by finding the optimal path through the trellis using the Viterbi algorithm.

A trellis is thus constructed for minimizing C_{HCB}. Each stage in the trellis corresponds to an SFB and each state at the *i*th stage represents a HCB candidate for this scale factor band. In other words, for the *i*th stage, if a path passes through the *m*th state, the *m*th HCB is employed for encoding the *i*th SFB. The Viterbi search procedure is outlined below.

The *k*th state at the *i*th stage is denoted by $S_{k,i}$ and the minimum accumulative-partial cost ending at $S_{k,i}$ is denoted by $C_{k,i}$. The transition cost from $S_{n,i-1}$ to $S_{m,i}$ is $R(h_{n,i-1}, h_{m,i})$.

1. Initialize $C_{m,0} = 0, \forall m$. Initialize *i*=1.
2. *Search*. $\forall m$, the best path ending at $S_{m,i}$ is found by computing
$$C_{m,i} = \min_n \{C_{n,i-1} + b_{m,i} + R(h_{n,i-1}, h_{m,i})\}$$
3. If $i < N$, set $i = i+1$ (SFB) and go to step 2.

3.4. Fast algorithm for Bits/NMR-Loss Analysis

The most time-consuming computation in this bit assignment scheme is the trellis-based HCB optimization for coding bits calculation in step 2 (Search). For each SF modification in step 2, the new value of total coding bits needs to be recalculated. Therefore, for one SF adjustment iteration, we need to perform $(sf_{\max,i} - sf_{ref,i})$ times trellis-based HCB optimization processes for the local *bits/NMR-Loss* analysis. Hence, the total number of calculations for finding the global maximum *bits/NMR-Loss* is

$$\sum_{i=1}^{N\_SFB} (sf_{\max,i} - sf_{ref,i})\ . \tag{4}$$

There are at least two ways to reduce computations. One is to reduce the complexity of the trellis-based HCB optimization; the other is to reduce the number of trellis-based HCB optimization.

By analyzing the local optimal parameters, $sf_{opt,i}$ and $BNLR_i$, we find some interesting properties.

1. The average value of the difference between the local optimal SFs of the *m*th and the *(m+1)*th iterations, $sfdiff_{ave}$, is often close to zero.

$$sfdiff_{ave} = \frac{1}{(N\_SFB-3)} \times \sum_{i \notin S} abs(sf_{opt,i}^{m+1} - sf_{opt,i}^{m})\ ,$$

where $S = \{sfb_{global}^m - 1, sfb_{global}^m, sfb_{global}^m + 1\}$ and sfb_{global}^m is the global optimal SFB of the *m*th SF adjustment iteration.

2. The average value of the difference between the local max *bits/NMR-Loss* ratio of the *m*th and the *(m+1)*th iteration, $BNLRdiff_{ave}$, is typically quite small.

$$BNLRdiff_{ave} = \frac{1}{(N\_SFB-3)} \times \sum_{i \notin S} abs(BNLR_i^{m+1} - BNLR_i^{m})$$

Using these two properties, we can drastically reduce the number of *bits/NMR-Loss* analyses (trellis-based HCB optimizations). We only need to perform the *bits/NMR-Loss* analysis on three SFBs after the first SF adjustment iteration.

4. SIMULATION RESULTS

The computational complexity and objective quality based on our simulations are summarized in this section. The bits assignment schemes used in comparison are as follows.

(1) The MPEG-4 VM of AAC (VM-TLS) without modification.
(2) The modified MPEG-4 VM of AAC (VM-TLS-M), in which the HCB decision algorithm is replaced by the TB-HCB optimization procedure described in section 3.3.
(3) The trellis-based ANMR optimization (TB-ANMR) and the MNMR optimization (TB-MNRM), which are implemented as described in [3] and [4].
(4) The normal and fast max bits/NMR-Loss schemes (max-BNLR).

Ten audio files with sampling rate 44.1K are used as test sequences. Two of them are extracted from MPEG SQAM [6], and the others are from EBU [7].

4.1. Computational complexity

The storage and computational complexity of one iteration in various schemes are summarized in Table 1.

Table 1. Complexity Analysis

	Search complexity	**Storage**
VM-TLS	1	--
VM-TLS-M	$12^2 \times N\_SFB$	$12 \times N\_SFB$
TB-ANMR TB-MNMR	$(60 \times 2)^2 \times 12^2 \times N\_SFB$	$60 \times 2 \times 12 \times N\_SFB$
Max-BNLR	$N\_SFB \times \text{Ave\_SF} \times 12^2 \times N\_SFB$	$12 \times N\_SFB$
Fast Max-BNLR	※ (a) ($N\_SFB \times \text{Ave\_SF} \times 12^2 \times N\_SFB$) (b) $3 \times \text{Ave\_SF} \times 12^2 \times N\_SFB$	$12 \times N\_SFB$
※ (a) is only for the first iteration; all the rest are using (b)		

In this table, Ave_SF is the average number of SF tested for the max BNLR analysis for each SFB and its typical value is around 17 or so. Table 2 is the statistics collected from the simulations on audio sequences. It is clear that in terms of computational requirement:
Fast Max-BNLR<< Max-BNLR<< TB-ANMR(MNMR)

Table 2. Statistics on Computational Complexity

	Average iteration /frame	Average TB HCB optimi-zations/ frame	Average TB HCB optimi-zations/ iteration	Complexity ratio
TB-ANMR (MNMR)	12	14400*12	14400	1
Max-BNLR	50	10103	10103/12 = 842	1/17
Fast Max-BNLR	50	1153	1153/12 = 96	1/150

4.2. Objective results

Two common objective quality measurements, average noise to mask ratio (ANMR) and maximum noise to mask ratio (MNMR) [5], are adopted in the performance comparison. Note that, in evaluating distortion, the NMR is set to 0 dB if the original NMR value is less than 0 dB. The rate-distortion curves of six bit assignment schemes are shown in Figures 2 and 3. (Note: TB-ANMR and TB-MNMR are similar algorithms aiming at two different target NMRs.) We can find that the ANMR performance of the Max-BNLR scheme is almost as good as that of TB-ANMR. There is almost no loss of ANMR performance in using the fast algorithm for Max-BNLR either. The MNMR values of TB-ANMR, Max-BNLR and Fast Max-BNLR are also similar. The characteristic of the proposed Max-BNLR scheme is closer to that of TB-ANMR as compared to TB-MNMR. Again, TB-ANMR and TB-MNMR are the optimal solutions tuned for their target cost functions, ANMR and MNMR, respectively [3][4].

4. CONCLUSIONS

In this paper, we propose a new concept, bit-use efficiency, for improving audio coding performance. Furthermore, a new bits assignment scheme based on this new concept (strategy) is proposed for MPEG-4 AAC, named Max-BNLR. Simulation results show that the Max-BNLR scheme has a performance close to TB-ANMR and is much better than the MPEG VM. In addition, its computational complexity is much lower than that of TB-ANMR.

5. REFERENCES

[1] ISO/IEC JTC1/SC29, "Information technology – vary low bitrate audio-visual coding," *ISO/IEC IS-14496 (Part 3, Audio)*, 1998

[2] M. Bosi, *et al.*, "ISO/IEC MPEG-2 advanced audio coding," *Journal of Audio Engineering Society,* vol. 45, pp. 789-814, October 1997.

[3] A. Aggarwal, S.L. Regunathan, K. Rose, "Trellis-based optimization of MPEG-4 advanced audio coding," *Proc. IEEE Workshop on Speech Coding*, pp. 142-4 2000.

[4] A. Aggarwal, S.L. Regunathan, K. Rose, "Near-optimal selection of encoding parameters for audio coding," *Proc. of ICASSP*, vol. 5, pp. 3269-3272, Jun 2001.

[5] H. Najafzadeh and P. Kabal, "Perceptual bit allocation for low rate coding of narrowband audio," *Proc. of ICASSP*, vol. 2, pp. 893-896, 2000.

[6] "The MPEG audio web page." http://www.tnt.uni-hannover.de/project/mpeg/audio.

[7] European Broadcasting Union, *Sound Quality Assessment Material: Recordings for Subjective Tests,* Brussels, Belgium, Apr. 1988.

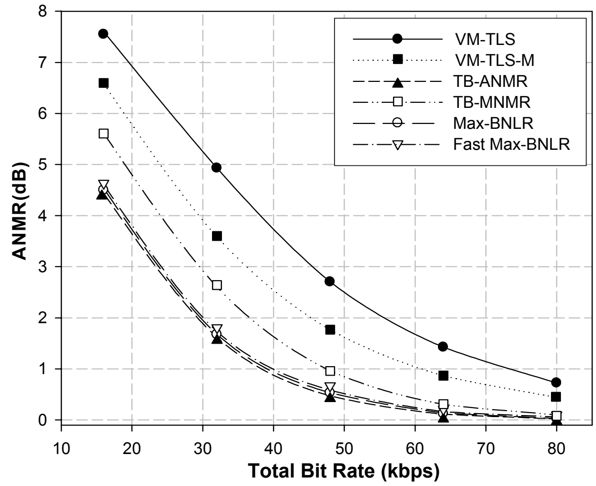

Figure 2. ANMR rate-distortion analyses

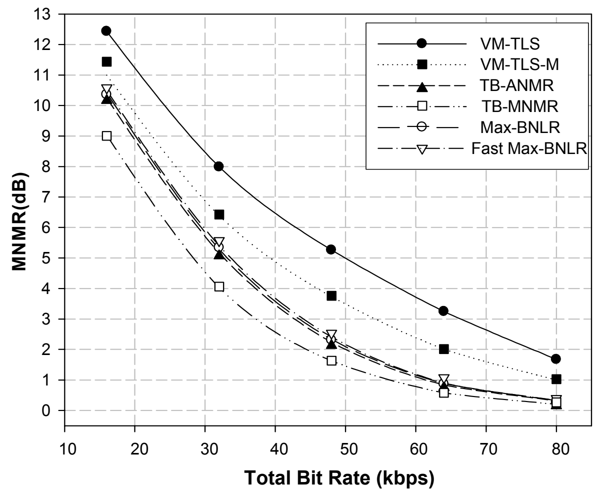

Figure 3. MNMR rate-distortion analyses