Efficient Bit Allocation Algorithm For MPEG-4 Advanced Audio Coding

Cheng-Han Yang and Hsueh-Ming Hang*
Department of Electronics Engineering
National Chiao Tung University, Hsinchu, Taiwan, R.O.C.
*hmhang@mail.nctu.edu.tw; Fax: (886)-3-5731791

Abstract-In this paper, a new bit allocation algorithm, called EBFOS (Enhanced BFOS), is proposed for the MPEG-4 Advanced Audio Coding (AAC) scheme. The notion of "bit-use efficiency" is suggested and the basic idea behind our approach is "give bits to the band with the maximum NMR gain per bit" or "retrieve bits from the band with the maximum bits per NMR loss", which is similar to the basic concept of the generalized BFOS (Breiman, Friedman, Olshen, and Stone) algorithm. However, the inter-band dependency of coding process in AAC is also taken into consideration in our approach. Simulation results show that the performance of the EBFOS scheme is significantly better than that of the MPEG-4 AAC Verification Model and is close to that of the trellis-based scheme, which achieves the minimum of a predefined NMR cost function. Also proposed in this paper is a fast version of our algorithm to reduce its computational complexity.

I. INTRODUCTION

In the last decade, analog audio has been gradually replaced by high-fidelity digital audio. Moreover, to meet the demand of efficient transmission and storage of digital audio for diversified multimedia applications, many high-efficient perceptual audio coding schemes have been developed, such as MPEG-1/2/4 audio coding standards and Dolby AC-3 [1]. The MPEG-4 Advanced Audio Coding (AAC) is one of the most recent-generation audio coders specified by the ISO/IEC MPEG standards committee [2]. The core part of the MPEG-4 AAC is based on the MPEG-2 AAC technology. The MPEG-4 AAC features a number of additional coding tools and coder configurations comparing to MPEG-2 AAC [3][4]. Consequently, the MPEG-4 AAC is a very efficient audio compression algorithm aiming at a wide variety of different applications, such as internet, wireless, and digital broadcast arenas.

In perceptual audio coding, the coding performance is highly dependent on the efficiency of the bit allocation scheme at the encoder. Particularly, when the bits are scare, how to make the best use of the limited number of available bits is critical in producing the best achievable audio quality. A typical bit allocation scheme in the MPEG audio encoder has two nested iteration loops, the outer loop and the inner loop. Thus, it is often called the *two-loop search* (TLS). The outer loop is the *distortion control loop* that handles the distortion associated with each band. The inner loop, also called the *rate control loop*, adjusts coding bits to fit the target bit budget for a frame. For such a bit allocation scheme, it is difficult to control the bit-use efficiency at band-level; thus, it

produces a not-very-high coding performance. The "bit-use efficiency" in this paper is referred to the amount of distortion reduction owing to the used bits.

The BFOS (Breiman, Friedman, Olshen, and Stone) algorithm was an efficient algorithm developed for pruning tree-structured classifiers. It was extended for designing vector quantizer and then adopted for sub-band coding [5][6]. The research in [7] shows that the generalized BFOS algorithm is a near optimal bit allocation scheme for MPEG-1 Layer I /II audio coding. However, our study in this paper shows that the generalized BFOS algorithm becomes less efficient for MPEG-4 AAC when the inter-band dependency of coding process exists.

Two trellis-based high performance bit allocation algorithms for AAC were proposed by [8][9]. One distinct feature of these bit allocation algorithms, comparing to TLS, is that both bit rate and distortion are controlled simultaneously and the inter-band dependency of coding process is also considered. As discussed in [8][9], the objective and subjective quality of the trellis-based schemes is significantly better than that of TLS. However, its computational complexity is extremely high.

In this paper, an improved bit allocation algorithm is proposed for MPEG-4 AAC. It is developed based on the generalized BFOS scheme. We suggest the notion of bit-use efficiency at band-level and also take the inter-band dependency into account in our bit allocation algorithm. Moreover, a fast version of this bit allocation scheme is also proposed for reducing calculations while the coding performance degradation is small.

The organization of the paper is as follows. In Section II, a brief overview of the typical MPEG-4 AAC encoder is provided. The proposed new bit allocation algorithm and its fast version are described in Section III. For comparison purpose, we also integrate the conventional BFOS algorithm into the AAC system in Section III. The complexity analysis and the simulation results are presented in Section IV.

II. OVERVIEW ON AAC ENCODER

The block diagram of a typical MPEG-4 AAC encoder is shown in Fig. 1. The time-domain audio signals are first converted to their frequency-domain representation (spectral coefficients) by the modified discrete cosine transform (MDCT). Motivated by the human auditory system, these spectral coefficients are grouped into a number of bands,

called *scale factor bands* (SFB). The pre-processing modules, which are optional tools, can help enhancing the coding performance and allowing AAC to process a wide frequency range of signals. The pre-processing modules in the MPEG-4 AAC include Temporal Noise Shaping (TNS), Long Term Prediction, Intensity/Coupling, Prediction, Perceptual Noise Substitution, and Mid/Side (M/S) Stereo Coding. The psychoacoustic model calculates the hearing masking threshold, which is the base for deciding coding parameters in the Rate-Distortion (R-D) controller. Note that the psychoacoustic model and the R-D controller are not specified by the AAC standard. The design engineers have the freedom to insert their own preferred elements into the encoder.

The spectral coefficients in one SFB are quantized by a non-uniform quantizer. The step size of the quantizer, which determines the quantization distortion (noise-to-masking ratio, NMR), is controlled by the parameter, scale factor (SF). The quantized coefficients in one band are then entropy-coded by one of the twelve pre-designed Huffman codebooks (HCB). Each SFB can choose its own quantization step size and HCB. In addition, the indices of SFs and HCBs have to be coded and transmitted as side information. In AAC, the SFs are differentially coded relative to the previous SF and then Huffman coded using a pre-designed codebook [2]. Taking Fig. 2 as example, instead of encoding the SF value of the 2nd SFB, 65, the difference between the 2nd SFB and the 1st SFB, 5, is coded. The indices of HCBs are coded by run-length codes [10]. A run-length code in AAC is 9 bits long, which is made of a 4-bits codebook index and a 5-bits run index. For example, as shown in Fig. 2, the 3rd HCB is used from the 1st SFB to the 10th SFB; therefore, these 10 HCB indices (same value) are coded together by one run-length code, in which the codebook index is 3 and run index is 10. The bit allocation scheme (referred as the R-D controller in Fig.1), our focus in this paper, is to determine two critical parameters, the values of SF and HCB, for each SFB to minimize the selected distortion criterion under the given bit rate constraint.
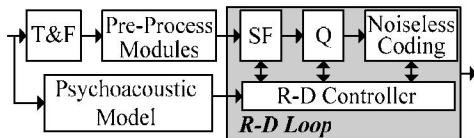

Fig. 1. Block diagram of a typical MPEG-4 AAC encoder.
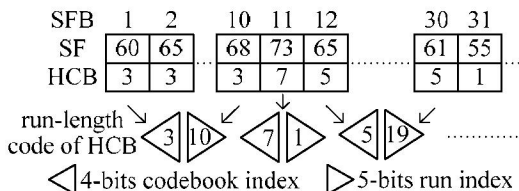T&F: Transform/Filter Bank, SF: Scale Factor, Q: Quantizer.

Fig. 2. An example of values of SF and HCB.

III. ENHANCED BFOS (EBFOS) SCHEME FOR MPEG-4 AAC

How to make use of the bits more efficiently is always the key issue in the perceptual audio coding. In this section, the proposed bit allocation scheme for MPEG-4 AAC is described. The basic idea behind our approach is "give bits to the band with the maximum NMR gain per bit" or "retrieve bits from the band with the maximum bits per NMR loss". "NMR gain per bit (ratio)" (NGPB) is the gain in NMR when one bit is given to a particular band and is defined mathematically by (1). "Bits per NMR loss (ratio)" (BPNL) is the number bits we save if we increase one unit NMR and is defined by (2).

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

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

The subscripts, *ref* and *new*, in (1) refer to the NMR or the bits before and after adding in one more bit in the case of NGPB. Similar definitions applied to (2) when the NMR is increased by one unit. In principle, our proposed scheme tries to reduce the total NMR of all bands. Hence, it has a performance close to the algorithm that minimizes the averaged NMR criterion.

In Section III-A, we first describe the mechanism of generating bits and distortion in AAC. The proposed EBFOS scheme and its fast version are described in Sections III-B and III-C, respectively. In addition, for comparison purpose, we integrate the BFOS algorithm with AAC and the scheme is described in Section III-D.

A. Rate-Distortion Relationship in AAC

As discussed in Section I, the perceptual distortion (or NMR, more precisely) of the spectral coefficients in each AAC band is controlled by the SF value. In general, a large SF value (leads to a larger quantization step size) will result in a large NMR value. After been quantized, the quantized spectral coefficients (in each band) are entropy-coded by a proper choice of HCB. The indices of SFs and HCBs for all bands are coded using the differential method and the run-length codes, individually. Let the values of SF and HCB for the ith SFB be denoted by s_i and h_i, respectively. The total coding bits, *TB*, for a frame can be expressed as follows.

$$TB = \sum_i B_i = \sum_i \left(b_i + D(s_i - s_{i-1}) + R(h_{i-1}, h_i)\right). \quad (3)$$

In the preceding equation, symbol $D()$ is a function of SF, representing the bits produced by the differential coding of SF indices. Symbol $R()$ is a function of HCB, representing the bits produced by the run-length coding of HCB indices. The returned function values in both cases are the numbers of bits to encode the arguments. Parameter b_i is the number of bits for coding the quantized spectral coefficients and parameter B_i is the total coding bits for the ith SFB. It is obvious that the differential and run-length codes induce the inter-band dependency in the coding process.

B. Bit Allocation Procedure of EBFOS Scheme

The block diagram of the Enhanced BFOS (EBFOS) scheme is shown in Fig. 3. Every step in Fig. 3 is elaborated below.

1. *Initialization.* This step is to initialize the reference NMR for each band, $NMR_{ref,i}$, at the beginning of maximum NGPB/BPNL analysis. Then, we determine the reference SF for each band, $s_{ref,i}$, and calculate the value of the total coding bits of a frame, TB_{ref}, based on the adopted reference NMR value. In general, larger $NMR_{ref,i}$ values will result in smaller TB_{ref} values. There seems to be no theoretically optimal choice for these values. Based on our experiences, we set the reference NMR to 1 (0 dB) for all the bands in our implementation; that is, $NMR_{ref,i}=1$, $\forall i$. In other words, we are targeting at perceptually lossless coding at the beginning of processing a frame.

2. *Local Maximum NGPB/BPNL analysis.* This step is to find the local maximum NGPB and BPNL values for all bands. We determine the local maximum NGPB and BPNL of the ith SFB, denoted by $NGPB_{L,i}$ and $BPNL_{L,i}$, by computing:

$$NGPB_{L,i} = \max_{s_{new,i}}\left\{(NMR_{ref,i} - NMR_{new,i})/(TB_{new} - TB_{ref})\right\} \quad (4)$$

$$\forall s_{new,i}, (s_{ref,i} - n_i) \le s_{new,i} < s_{ref,i}$$

$$BPNL_{L,i} = \max_{s_{new,i}}\left\{(TB_{ref} - TB_{new})/(NMR_{new,i} - NMR_{ref,i})\right\} \quad (5)$$

$$\forall s_{new,i}, s_{ref,i} < s_{new,i} \le (s_{ref,i} + n_i)$$

TB_{new} and $NMR_{new,i}$ are the new value of total coding bits for a frame and the new value of distortion for the ith SFB, respectively, when its SF value is changed from $s_{ref,i}$ to $s_{new,i}$. The local optimal SF value of the ith SFB, $s_{opt,i}$, is the SF value associated with the local maximum NGPB or BPNL. The parameter n_i in (4) (or (5)) determines the candidate number of the $s_{new,i}$ values, which is approximately 12 on the average based on the statistics of the coded data.

   Note that, in performing the local maximum NGPB or BPNL analysis for the ith SFB, only the SF value of the ith SFB is changed from $s_{ref,i}$ to $s_{new,i}$. The SF values of the other SFBs are kept unchanged ($s_j = s_{ref,j}$, $\forall j$, $j \ne i$).

3. *Global Maximum NGPB/BPNL analysis.* We find the global maximum NGPB and BPNL value, $NGPB_G$ and $BPNL_G$, for a frame by calculating:

$$NGPB_G = \max_i\left\{NGPB_{L,i}\right\} \quad \forall i, 1 \le i \le 49 \quad (6)$$

$$BPNL_G = \max_i\left\{BPNL_{L,i}\right\} \quad \forall i, 1 \le i \le 49 \quad (7)$$

   The global optimal SFB, sfb_G, is the SFB that has $NGPB_G$ (or $BPNL_G$). Then, we adjust the SF value of the sfb_G-th SFB only to its local optimal SF value (determined in Step 2).

4. Update $NMR_{ref,i}$ (as well as $s_{ref,i}$) of the sfb_G-th SFB and TB_{ref}. Go to Step 2 if the bit budget constraint is not met.

As described in the Step 2 of the preceding procedure, after changing the SF value from $s_{ref,i}$ to $s_{new,i}$, we need to calculate TB_{new} and $NMR_{new,i}$. The value of $NMR_{new,i}$ depends only on the value of $s_{new,i}$. However, (3) tells us that the value of TB_{new} depends not only on the value of $s_{new,i}$; it also depends on the choice of HCB. In our bit allocation scheme, we adopt the trellis-based optimization algorithm for deciding HCB as proposed in [11].

Because the coding process in AAC is inter-band dependent, the SF value change of the sfb_G-th SFB in Step 3 has an impact on the local maximum NGPB/BPNL analyses of the other SFBs in Step 2. Therefore, we have to perform the local maximum NGPB/BPNL analyses for all the bands in each iteration.

In general, either the maximum NGPB analysis or the maximum BPNL analysis (but not both) has to be performed for each iteration. The maximum NGPB analysis is used for spending more bits (when the bit budget is positive) and the maximum BPNL analysis is used for recovering the bits (when the bit budget is negative).
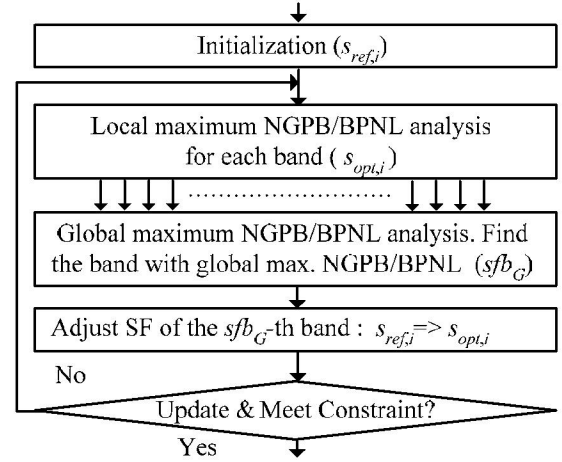

Fig. 3. EBFOS bit allocation scheme

C. Fast Algorithm for EBFOS scheme

The complexity of our EBFOS bit allocation scheme highly depends on the iterations the NGPBR/BPNL calculation in Step 2 (the local maximum NGPB/BPNL loop in Section III-B). Taking the maximum NGPB analysis as example, we need to perform n_i times NGPB calculation for locating the local maximum NGPB of the ith SFB. Hence, the total number of calculations for finding the global maximum NGPB is $\sum_i n_i$.

It is obvious that the most effective way for reducing computation is to reduce the number of NGPB/BPNL calculations. From the statistics of the local optimal parameters, $s_{opt,i}$ and $NGPB_{L,i}$ (or $BPNL_{L,i}$) collected from the coded data, we find some interesting properties, which are summarized in Table I.

In Table I, i is the SFB index and m is the number of SF

adjustment iteration. And sfb_G^m is the global optimal SFB of the mth SF adjustment iteration and $S = \left\{ sfb_G^m - 1, sfb_G^m + 1 \right\}$, the set of two SFBs. The second column in Table I is the probability that $s_{opt,i}^{m+1}$ differs from $s_{opt,i}^m$ and is denoted by $P_C(s_{opt,i}^{m+1} \neq s_{opt,i}^m)$, where $s_{opt,i}^m$ is the local optimal SF value of the ith SFB in the mth SF adjustment iteration. Table I shows that $P_{i \in S}(s_{opt,i}^{m+1} \neq s_{opt,i}^m)$ reaches 28.72%, which is approximately 18 times of $P_{i \notin (S \cup sfb_G^m)}(s_{opt,i}^{m+1} \neq s_{opt,i}^m)$. In other words, typically the optimal SF value does not change much from the pervious iteration to the current one. The second row is related to the NMR values. Taking the maximum NGPB analysis as an example, $AD_C(NGPB_{L,i}^m, NGPB_{L,i}^{m+1})$ is the average value of the normalized differences between $NGPB_{L,i}^m$ and $NGPB_{L,i}^{m+1}$, where $NGPB_{L,i}^m$ is the local maximum NGPB value of the ith SFB in the mth SF adjustment iteration. More precisely, $AD_C(NGPB_{L,i}^m, NGPB_{L,i}^{m+1})$ is defined by (8), where M is 2 for $C = (i \in S)$ and is (49-3) for $C = (i \notin (S \cup sfb_G^m))$. As we can see, $AD_{i \notin (S \cup sfb_G^m)}(NGPB_{L,i}^m, NGPB_{L,i}^{m+1})$ is typically quite small, but $AD_{i \in S}(NGPB_{L,i}^m, NGPB_{L,i}^{m+1})$ is much larger.

$$AD_C(NGPB_{L,i}^m, NGPB_{L,i}^{m+1}) = \frac{\sum\limits_C \left(\dfrac{\left| NGPB_{L,i}^{m+1} - NGPB_{L,i}^m \right|}{NGPB_{L,i}^{m+1}} \right)}{M} \quad (8)$$

It is clear that the differences of the local maximum NGPB/BPNL analyses between each SF adjustment iteration mainly locate at the SFBs belonging to S. Using these properties, we can drastically reduce the number of iterations in locating the maximum NGPB/BPNL. We only need to perform the maximum NGPB/BPNL search on three SFBs ($SFB = \left\{ sfb_G^m - 1, sfb_G^m, sfb_G^m + 1 \right\}$) after the first SF adjustment iteration. This is the *fast version* of our EBFOS algorithm.

TABLE I
<small>STATISTICS OF LOCAL OPTIMAL PARAMETERS IN MAXIMUM NGPB/BPNL ANALYSIS</small>

Condition (C)	$i \in S$	$i \notin (S \cup sfb_G^m)$
$P_C(s_{opt,i}^{m+1} \neq s_{opt,i}^m)$	28.72%	1.60%
$AD_C(NGPB_{L,i}^m, NGPB_{L,i}^{m+1})$ or $AD_C(BPNL_{L,i}^m, BPNL_{L,i}^{m+1})$	0.7428	0.0211

D. Generalized BFOS Bit Allocation scheme for AAC

The generalized BFOS algorithm is an efficient bit allocation algorithm for sub-band coding. For the purpose of analysis and comparison, we also implement a scheme that integrates the generalized BFOS algorithm into AAC based on the concepts described in [6]. The BFOS-based bit allocation scheme is described below.

1. *Initialization.* Same as the initialization step in Section III-B, we set the reference NMR to 1 (0 dB) for all bands. Then, we determine the $s_{ref,i}$ value and calculate the total coding bits for each band, $B_{ref,i}$ based on the reference NMR value, that is, $NMR_{ref,i} = 1, \forall i$.
2. *Local Maximum NGPB/BPNL analysis.* Differing from the EBFOS scheme, the local maximum NGPB and BPNL of the ith SFB for the BFOS scheme are determined by the formulas (9) and (10) respectively.

$$NGPB_{L,i} = \max_{s_{new,i}} \left\{ (NMR_{ref,i} - NMR_{new,i}) / (B_{new,i} - B_{ref,i}) \right\} \quad (9)$$
$$\forall s_{new,i}, (s_{ref,i} - n_i) \leq s_{new,i} < s_{ref,i}$$

$$BPNL_{L,i} = \max_{s_{new,i}} \left\{ (B_{ref,i} - B_{new,i}) / (NMR_{new,i} - NMR_{ref,i}) \right\} \quad (10)$$
$$\forall s_{new,i}, s_{ref,i} < s_{new,i} \leq (s_{ref,i} + n_i)$$

$B_{new,i}$ and $NMR_{new,i}$ are the new value of total coding bit and the new value of distortion for the ith SFB, respectively, when its SF value is changed from $s_{ref,i}$ to $s_{new,i}$. The local optimal SF value of the ith SFB, $s_{opt,i}$, is the SF value associated with the local maximum NGPB or BPNL.
3. *Global Maximum NGPB/BPNL analysis.* Same as Step 3 in Section III-B, we first find the $NGPB_G$ (or $BPNL_G$) for a frame by the (9) (or (10)) and thus sfb_G is determined. Then we adjust the SF value of the sfb_G-th SFB only to the its local optimal SF value (determined in Step 2).
4. Update $NMR_{ref,i}$ (as well as $s_{ref,i}$) and $B_{ref,i}$ of the sfb_G-th SFB. Go to Step 2 if the bit budget constraint is not met.

In the BFOS scheme for AAC, we also adopt the trellis-based optimization algorithm for HCB decision. However, differing from the EBFOS scheme, we only perform the local maximum NGPB/BPNL analysis for the sfb_G-th SFB after its SF value changes in Step 3.

As described in [6], the generalized BFOS scheme can be performed with or without the convexity assumption. When it is performed with convexity assumption, n_i in (9) (or (10)) is equal to 1. When it is performed without convexity assumption, n is approximately 14 on the average based the coded data statistics.

IV. SIMULATION RESULTS

In this section, we will examine the computational complexity and the coded audio quality of the aforementioned various schemes. Four types of bit allocation algorithms are

simulated and compared as described below using the MPEG-4 AAC Verification Model (VM) [12] as the testing platform. VM is the encoder software developed by the MPEG committee to verify the coding syntax.

(1) The TLS algorithm used by the MPEG-4 AAC VM (denoted as VM-TLS).

(2) The generalized BFOS algorithm with convexity assumption, denoted as BFOS-C; and the generalized BFOS algorithm without convexity assumption, denoted as BFOS-NC. Both are described in Section III-D.

(3) The trellis-based algorithm aiming at minimizing average NMR, denoted as JTB-ANMR; and the trellis-based algorithm aiming at minimizing maximum NMR, denoted as JTB-MNMR. Their details are described in [8] and [9].

(4) The EBFOS bit allocation scheme and its fast version, which are described in Sections III-B and III-C.

To focus only on the bit allocation performance, all the optional tools in AAC, such as TNS and M/S stereo coding, are not used in our simulations. Ten two-channel audio sequences with a sampling rate at 44.1 kHz are tested. Two of them are extracted from MPEG SQAM [12], and the rest are from EBU [13].

A. Complexity Analysis

The complexity analysis for the aforementioned several bit allocation algorithms is summarized in Table II. The "Computation" column is the average iterations of NGPB (or BPNL) calculation for a frame. For convenience of comparison, the BFOS-NC scheme is chosen to be the reference (ratio=1) and all the other schemes are rated based on this reference.

The experimental data indicate that the computation of the fast EBFOS scheme is approximately 2.6 times higher than that of the BFOS-NC scheme. Moreover, the fast EBFOS scheme is approximately 10 times faster than that of the EBFOS scheme.

TABLE II

COMPLEXITY ANALYSIS

	Computation	Ratio
BFOS-C	119	0.27
BFOS-NC	444	1
Fast EBFOS	1145	2.58
EBFOS	11848	26.68

B. Objective Quality

The rate-distortion curves of these 7 bit allocation schemes are shown in Figs. 4 and 5. Two common objective quality measurements, average NMR (ANMR) and maximum NMR (MNMR) [14], are adopted in the objective performance comparison.

$$ANMR = 10 \times \log 10 \left(\frac{1}{N} \sum_n \left(\frac{1}{Z} \sum_{i=0}^{Z-1} NMR_{i,n} \right) \right) \qquad (11)$$

$$MNMR = 10 \times \log 10 \left(\frac{1}{N} \sum_n \left(\underset{i}{Max} \left(NMR_{i,n} \right) \right) \right), \qquad (12)$$

where $NMR_{i,n}$ is the noise-to-masking ratio of the ith SFB in the nth frame.

The research in [7] shows that the BFOS-C scheme is a near optimal bit allocation scheme for MPEG-1 Layer I / Layer II audio coding, but the simulation results show that the BFOS-C scheme is less efficient for AAC. The performance of the BFOS-NC scheme is much better than that of the BFOS-C scheme particularly for the MNMR criterion, which means that the convex assumption is not suitable for AAC. However, both the ANMR and MNMR performances of the BFOS-NC scheme are approximately 1dB worse than that of the JTB-ANMR scheme. Clearly, the performance of the EBFOS scheme is much better than that of VM-TLS. If we look at the ANMR plot (Fig. 4), the performance of the EBFOS scheme is slightly worse than that of JTB-ANMR but they are very close. It is somewhat better than the JTB-MNMR scheme since the latter is not optimized for the ANMR criterion. If we look at the MNMR plot (Fig. 5), the EBFOS scheme is somewhat worse than JTB-MNMR but it is slightly better than the JTB-ANMR scheme. As stated earlier, the EBFOS scheme is aiming at reducing the overall NMR, which pretty much leads to minimizing ANMR. As for the fast version, there is almost no loss of performance (less than 0.06dB loss).

C. Subjective Quality

Listening test by human ears is the traditional way to subjectively evaluate the audio quality and is also the most recognized subjective quality measure. However, the exact subjective test is expensive, time consuming, and difficult to reproduce. As described in Section 6.2 in [9], the subjective quality (mean opinion score, MOS [15]) of the JTB-ANMR (or JTB-MNMR) scheme is significantly better than that of the VM-TLS. The informal listening tests on the aforementioned schemes show that it is hard to tell the difference between JTB-ANMR and the EBFOS scheme. Also, we conduct a "simulated" subjective measure, Objective Difference Grade (ODG). ODG is generated by a procedure designed to be comparable to the Subjective Difference Grade (SDG) judged by human ears. It is calculated based on the difference between the quality rating of the "reference" signal and the "test" signal. The ODG has a range of [-4, 0], in which –4 stands for a very significant difference and 0 stands for an imperceptible difference between the reference and the test signal [16][17].

The ODG results of the aforementioned bit allocation schemes are shown in Fig. 6, in which the reference signal is the original audio sequence. Interestingly, JTB-ANMR is the best algorithm judged by ODG. According to the collected test data (Fig. 6), the EBFOS scheme is better than that of the BFOS-NC and BFOS-C schemes. Moreover, the difference between the EBFOS and the JTB-ANMR schemes is rather small.

V. Conclusions

In this paper, we propose an efficient bit allocation algorithm, which is an enhanced version of the generalized BFOS algorithm for MPEG-4 AAC, called EBFOS. For reducing calculations, a fast version of the EBFOS scheme is also proposed. Simulation results show that the original generalized BFOS algorithms become less efficient for AAC. The performance of the proposed EBFOS scheme is better than that of the VM-TLS and BFOS algorithms. Moreover, the EBFOS scheme has a performance close to the trellis-search based algorithm (optimized for the averaged NMR, JTB-ANMR). As for the fast version, there is almost no loss of performance in adopting the fast algorithm for the EBFOS scheme.
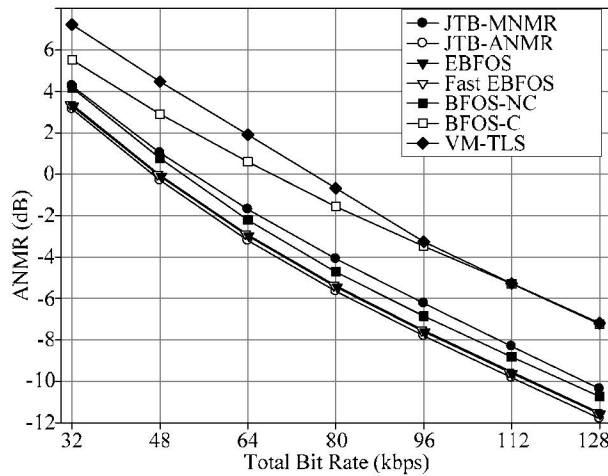
Acknowledgment

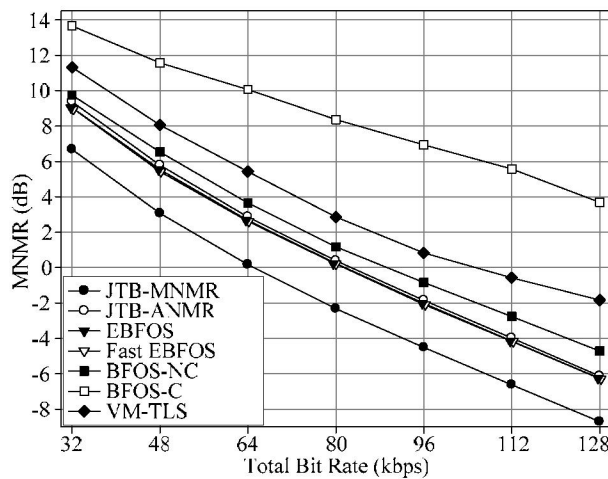Fig. 4. ANMR rate-distortion comparison of various bit allocation schemes

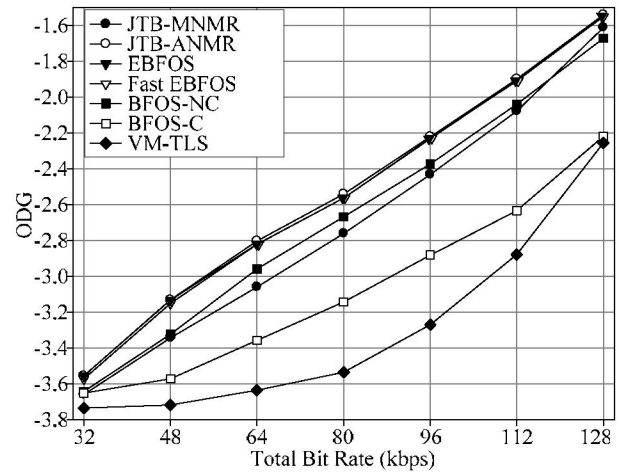Fig. 5. MNMR rate-distortion comparison of various bit allocation schemes

Fig. 6. ODG performance of various bit allocation schemes

References

[1] T. Painter and A. Spanias, "Perceptual coding of digital audio," *Proc. of IEEE*, vol. 88, pp. 451- 515, Apr. 2000.

[2] ISO/IEC JTC1/SC29, "Information technology – Coding of audio-visual objects," *ISO/IEC IS-14496 (Part 3, Audio)*, 1999.

[3] J. Herre, B. Grill, "Overview of MPEG-4 audio and its applications in mobile communications," *Proc. of WCCC-ICSP*, vol. 1, pp. 11-20, Aug. 2000.

[4] H. Purnhagen, "An overview of MPEG-4 audio version 2," *AES 17th International Conference on High-Quality Audio Coding*, Firenze, Sept. 1999.

[5] P. H. Westerink, *et al.*, "An optimal bit allocation algorithm for sub-band coding," *Proc. of ICASSP*, pp. 757-760, 1988.

[6] E. A. Riskin, "Optimal bit allocation via the generalized BFOS algorithm," *IEEE Trans. On Information Theory*, vol. 37, No. 2, Mar. 1991.

[7] G. Diego and R. Sudhakar: "Optimal bit allocation for MPEG audio standard using the generalized BFOS algorithm", *98th AES Convention*, Paris, Feb. 1995

[8] A. Aggarwal, *et al.*, "Trellis-based optimization of MPEG-4 advanced audio coding," *Proc. IEEE Workshop on Speech Coding*, pp. 142-4 2000.

[9] A. Aggarwal, *et al.*, "Near-optimal selection of encoding parameters for audio coding," *Proc. of ICASSP*, vol. 5, pp. 3269-3272, Jun 2001.

[10] S. Golomb, "Run-length encodings," *IEEE Transactions on Information Theory*, vol. 12, pp. 399-401, July 1966.

[11] C. H. Yang, H. M. Hang, "Cascaded trellis-based optimization for MPEG-4 Advanced Audio Coding", *AES 115th Convention*, Oct. 2003, New York, USA.

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

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

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

[15] ITU-R BS.1116, "Method for subjective Assessment of Small Impairments in Audio Systems Including Multichannel Sound Systems," 1994.

[16] Draft ITU-T Recommendation BS.1387: "Method for Objective Measurements of Perceived Audio Quality," July 2001.

[17] A. Lerchs, "EAQUAL software", Version 0.1.3 alpha, [Online]. Available: http://mitiok.free.fr/c.htm