MPEG-4   DSP
MPEG-4 Multimedia Video Communication for
for Mobile Applications

ᴄ

NSC   89   2213   E   009   137

   88   8   1   89   7   31

   89   10   31

MPEG-4 與 DSP

MPEG-4 Multimedia Video Communication for Mobile Applications

MPEG4 的
PDAM4 中 Streaming Video Profile(SVP)
的 Fine Granularity Scalability (FGS)

Internet
FGS syntax
FGS

MPEG-4
FGS

ABSTRACT

In this paper, we use the Fine Granularity Scalability (FGS) as defined in the Streaming Video Profile (SVP) of MPEG4 PDAM4 [1] and demonstrate how FGS can match the requirement of the Internet environment and other applications with heterogeneous networks such as terrestrial broadcasting, Internet broadcasting, and wireless communications. We present two algorithms that utilize an FGS syntax tool, selective enhancement, to further improve the performance of FGS using Rate-Distortion optimization technique.

Keywords: MPEG-4, Multimedia Video Communication for Mobile Applications , FGS, National Science Council

1. INTRODUCTION

With the rapid advance in communication technology, more and more information with various styles can be transmitted through different forms of media, which depends on the applications, such as Internet, wireless channel, satellite broadcasting. The common feature for the digital video communications is that they need vast bandwidth for representing the information. Hence, the compression is needed before they are transmitted or stored. In the past, the issues for video data compression are mainly focused on how to achieve higher compression gain under reasonable complexity. However, the requirements for video compression today are different from that in the past. What the most compression algorithms today have to provide is not only good coding efficiency but also scalability. The scalability mentioned here means that the compressed bitstream can still be decoded with reasonable quality after the bitstream is extracted or truncated. The reason for bitstream truncation is due to the fact that the effective transmission bandwidth might varies with time. This phenomenon can be easily seen on Internet or wireless communications. In addition, the compressed bitstream might need to be stored in the server side and requested by different client users with different hardware complexity. For this situation, the bitstream can be extracted at the sever side to meet the hardware decoding capability of the client side. Many other applications that need scalability can be found. There are several approaches including SNR (Signal to Noise Ratio) scalability, spatial scalability and temporal

scalability. All of these scalabilities can be merged in a hybrid form. In the following discussion, we will consider the SNR scalability only.
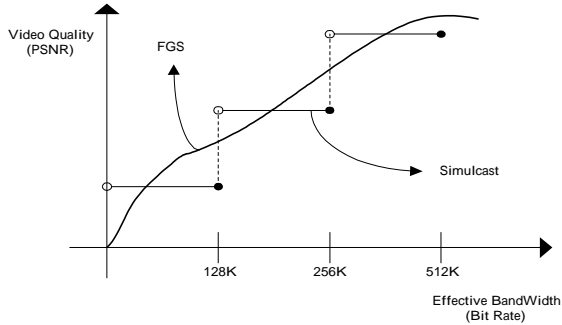

Fig. 1 SNR scalability comparison between FGS and Simulcast

For the SNR scalability of video compression, the simulcast is the most institutive way to achieve the goal. The simulcast for SNR scalability means that we can first predetermine some possible rates and then code the video signals with these distinct rates independently. However, as one can observe that it is not efficient from the compression point of view since there exits correlation between these independently coded bitstreams. In addition, the simulcast cannot best match the effective bandwidth provided. Hence, some bits might be wasted. In order to solve these problems, the FGS (Fine Granularity Scalability) scheme is proposed in MPEG4. The FGS can best match any effective bandwidth and support reasonable quality. In other words, the graceful degradation property can be obtained through FGS. The key point of the current adopted FGS algorithm in MPEG4 is to separate the coding bitstream into two layers, Base layer and Enhancement layer. The base layer is formed by coarsely quantizing the DCT coefficients or the prediction residual. And the enhancement layer is obtained by coding the difference between the original DCT coefficients and the coarsely quantized baser layer coefficients in bit plane by bit plane order.

The detail structure will be briefly given in section II. Fig.1 shows the performance of SNR scalability between the simulcast and FGS.

Several schemes had been proposed to code the FGS enhancement layer. Sen-ching Samson Cheung et al. [2] proposed a matching pursuit coding for FGS coding. Hayder Radha et al. [3] proposed a wavelet-coding scheme using the SPIHT [4], a variation of EZW algorithm, to code the FGS Enhancement layer. Weiping Li [5] proposed a DCT based bit plane coding scheme. The bit planes are formed from the residue between original DCT coefficients and the coarsely quantized DCT coefficients of base layer as mentioned before. These bit planes are further coded as FGS Enhancement Layer. The MPEG4 SVP standard has adopted this algorithm.

As one can expect that it is trade-off between scalability and compression efficiency. Without exception, the tradeoff also happens in FGS. In the current syntax, the prediction from the enhancement layer is not allowed. Thus, the coding efficiency is sub-optimal. Other proposals that allow prediction from enhancement layer at the cost of higher complexity can be found in [6][7]. Beside the SNR scalability mentioned, the temporal scalability is also proposed [8].

In the following, the FGS decoder and encoder architectures will be introduced in section II. Meanwhile, we will also describe the parameters that can be defined by the user in the current FGS syntax in that section. After that, two primitive algorithms are proposed in section III to further use these parameters for different purpose rather than the current way used in MPEG4. Finally, conclusion and possible future works will be given in Section IV.

2. Architectures of FGS encoder and decoder

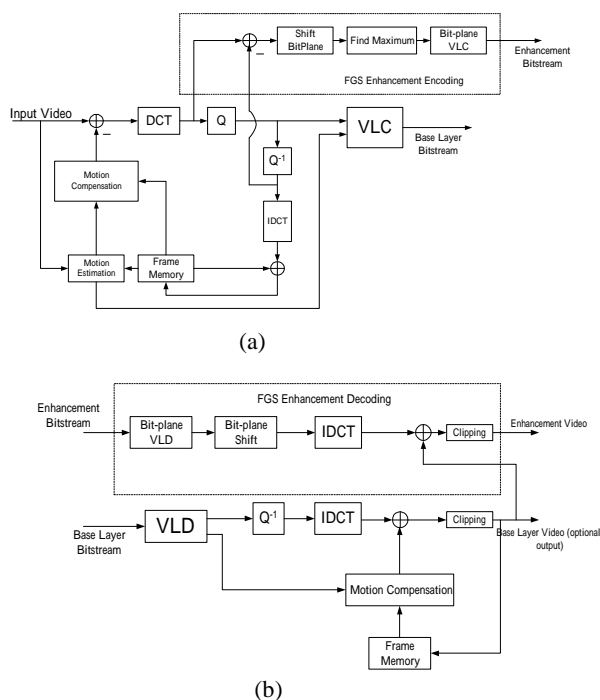## The FGS decoder and encoder architectures

(a)

(b)

Fig. 2 FGS (a) Encoder [9] (b) Decoder [1] architecture

are shown in Fig. 2 (a) and Fig. 2(b) respectively. The decoder architecture can be found in [1]. On the other hand, the encoder architecture is constructed based on the decoder structure and [9]. At the encoder side, the base layer encoder is similar to those defined in MPEG-1/2. Again, the DCT transform is used. The scalar quantizer of the encoder is used to determine the bit rate of the base layer. Usually, the stepsize is set to be large so that the encoder can provide coarse approximation with low bit rate. On the other hand, the data used for enhancement layer coding is the difference between the original DCT coefficients (or residual in P, B frames) and the coarse approximation obtained by coarsely quantizing the original DCT coefficients (or residual in P, B frames) as shown in Fig. 2 (a). The difference is then bitplane coded. At the decoder side, all of the operations are the inverse of those on the encoder side. However, as one take a close look at the encoder and decoder structures, it can be found that the encoder and decoder structures

are not exactly matched. The subtraction of base layer and enhancement layer on the encoder side should be performed between the input of DCT and output of IDCT (the decoder feedback loop in the encoder) respectively since the addition of base layer and enhancement layer is after IDCT. However, the mismatch does not lead to any error drift problem or what else since the DCT or IDCT operation is linear. Some analyses have been carried in [9][10]. It is proved that the encoder and decoder structures shown in Fig. 3 are better choice. The other notable thing is the "BitPlane Shifting" module on both encoder and decoder side. The module is used for the shifting operation of the MBs (Macro Blocks) bitplane and 8x8 block bitplane defined in current FGS syntax. The purpose of the 8x8 block bitplane is to reduce the flickering effect and the shifting of MBs is to facilitate the region of interest functionality. Since the enhancement layer bitstream is transmitted bitplane by bitplane within each FGS VOP, it is straightforward to implement the shifting operation to determine the priority for which MBs should be transmitted first. The shifting operation for MBs is called "selective enhancement" (SE). In the current syntax, the shifting factor is specified at MB level. More detail about the SE can be found in [11][12][13]. In addition to the SE, another shifting operation for 8x8 FGS block is specified in VOL (Video Object Layer) level that is used to reduce the flickering effect. The flickering effect is resulted from the fact that we will transmit the high frequency component first at the enhancement layer where we have larger error terms located at high frequency bands due larger stepsize of quantizer for high frequency component at the base layer. As expected larger stepsize of quantizer for high frequency component will result in larger magnitude of quantization noise for high frequency component. That is why we also got larger magnitude of FGS coefficients located at high frequency bands. However, the fact described
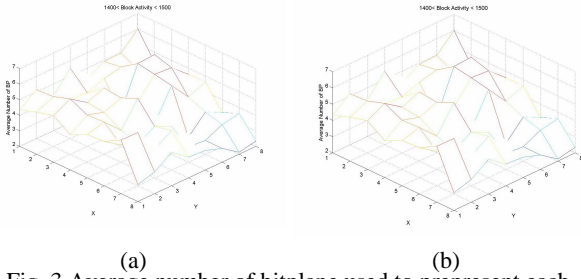
3

(a)   (b)

Fig. 3 Average number of bitplane used to prepresent each
coefficient within the 8x8 block with (a) low block activity
(b) high block activity

above is true only for blocks that have larger block activity that is defined as the absolute summation of all 64 coefficients. Based on our experiment, block with lower activity have most of larger magnitude of coefficients located at low frequency bands. In Fig. 3, we show the average number of bitplanes used by each coefficient within the 8x8 block versus the block activity. Note that only two extream cases are shown. Fig. 3 (a) shows the bitplane distributation for low blck activity and Fig. 3 (b) denotes that for high block activity. The left corner in each figure indicates the DC position. And the rightmost corner represents the highest frequency component. The shifting matrix specified in VOL level for reducing the flickering effect is named as "Frequency Weighting Matrix". Since it is given in VOL level, all 8x8 blocks located at the same layer use the same matrix. Hence, one has to determine the FW matrix that is suitable for most of blocks. Our experiments shows that it can be determined from the block activity which is the absolute sum of the 64 coefficients. Note that non proper FW matrix will let the advantage of better visual quality be nullfied by the poor overall compression efficiency since the total number of bitplanes need to be coded might increase after FW shifting. As mentioned in [14][15], the advantage of FW shifting can not be seen from the PSNR, it is mainly due to the fact that coefficients located at lower frequency bands with smaller magnitude might be shifted up so that the transmission order is prior to those coefficients located at higher

frequency bands with larger magnitude. Thus, one can observe that blocks after FW shifting have lower PSNR comparing to the non-shifted version under the same constraint rate. The transmission order has benefits only from the subject quality. More analyses can be found in [14][15].

In these user defined parameters, we find that the SE shifting factors specified for each MB can not only be used for the function of region of interest but also can be used to further change the R-D curve of the overall FGS VOP. Different combinations of shifting factors will lead to different rate consumed and different amount of distortion reduced for each bitplane of the overall FGS VOP. Form this point of view, we think that there exists potential to further improve the R-D performance of FGS under the current syntax constraint. In the following sections, we propose two algorithms for changing the R-D curve.

3. R-D improvement of FGS using the Selective Enhancement shifting factors

As mentioned, we will discuss how to use the SE shifting factors for R-D improvement purpose in this section. Before we further to illustrate the algorithm, let us first to have some observations on the behavior of the amount of distortion reduction contributed by each bitplane located at different levels. In the following, we list these observations without showing the derivation process. As a matter of fact, the amount of distortion reduction contributed by each bitplane can be derived precisely and we omit it here. Recall that we are on the encoder side. We can have all information needed. The calculation of distortion measure is based on the reconstruction behavior of the decoder that is specified in [1]. The distortion measure is MSE (Mean Square Error).

Observations:

1) The amount of distortion reduction by transmitting one whole bitplane of certain MB is mainly determined by those bits that are the MSBs (Most Significant bits) of the corresponding coefficients.
2) The MSB bits located at higher level can lead to more distortion reduction while they are transmitted.
3) The R-D curve of any MB within the FGS VOP is not necessary to be strictly decreasing.
4) The all zero bitplanes, before any MSB is reached, of any FGS MB contribute no distortion reduction. However, 1~2 bits are needed to represent this case. It happens because the maximum number of bitplanes to be coded for one VOP might be larger than necessary for representing certain FGS MB especially for those blocks which have lower block activity.

For the corresponding rate consumed by each bitplane, it is not straightforward to model the behavior. However, the trend is that more bits are needed for recording the lower bitplanes. Thus, one can deduce from the observations listed above and the trend of the corresponding rate consumed that the R-D slope is larger for higher bitplane. Intuitively, we will want to optimize the overall R-D performance by rearranging the transmission order of MBs by SE shifting factors so that the final overall R-D curve for one FGS VOP can have larger slopes at each bitplane than the original one. If we can achieve that, the R-D performance is improved. Based on this, we develop the following algorithm.

3.1 Algorithm 1

3.1.1 Shifting up or Shifting down
In the current FGS syntax, the shifting operation provided is to shift up the MBs. As a matter of fact, one can have two choices based on the
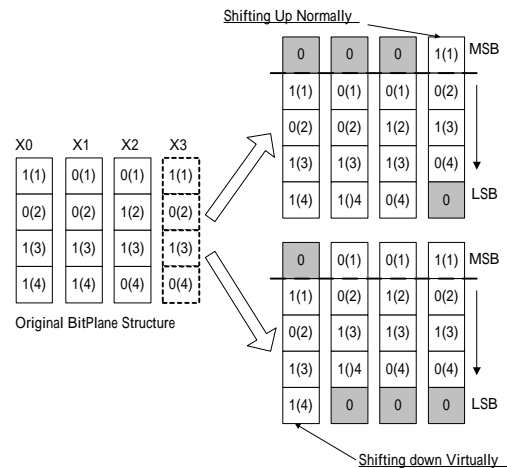

Fig. 4 The difference between physically shifting up and virtually shifting down operations

current FGS framework and the deductions introduced. Since the bitplanes of MBs that have larger R-D slopes should be transmitted first, one can directly shift up those MBs with bitplanes having larger R-D slopes or virtually shift down those MBs with bitplanes having lower R-D slopes. The virtually shifting down operation is to physically shift up all the other MBs except those that should be virtually shifted down. The main difference can be illustrated by means of Fig. 4. In Fig.4, each rectangle block represents one bitplane of MB Xi. The text enclosed in each block, **m(n)**, indicates that the current bitplane located at level **n** are all zero bitplane while **m**=0 otherwise the bitplane contains at least one nonzero bit. The level represents MSB of the whole VOP when **n**=1 and LSB (Least Significant Bit) while n=4 in the case of Fig. 4. Note that the MSB of the VOP is not necessary to be the MSB of each MB. Assuming that the MB, X3, is we want to shift up physically. In this case, one can virtually shift down the MB, X0. That is to physically shift up MBs, X1, X2 and X3. Or one can directly shift up the MB, X3. Note that shifting down the bitplanes with all zero before MSB, which is the highest level where at least one nonzero bit first appears for certain MB, is meaningless since this operation does not lead to any distortion reduction due to the fact that another all zero bitplane will

replace the original one. That is why X1 and X2 are not virtually shifted down. As one can observe in Fig.4, shifting up the MB will introduce more all zero bitplanes that consume bits but contribute nothing at all to distortion reduction. However, shifting down virtually has totally opposite situation. For both situations, the number of zero bitplanes introduced due to shifting is the same. The main difference between the two situations is that the positions of these all zero bitplanes introduced are different. In addition, it seems that each bitplane will have benefit by shifting down operation. For example, the R-D slope of the 2nd bitplane of the whole FGS VOP is increased since the first bitplane of the shifted down MB is merged into the 2nd bitplane of the overall FGS VOP. Generally, the R-D slope of the first bitplane is larger than that of the 2nd bitplane. Thus, it seems that the R-D slope of 2nd bitplane for the whole FGS VOP can be improved since MB bitplane with lower slopes are replaced with higher slopes. But, the improvement might be negligible. In the scheme 1, we will use the shifting down operation.

3.1.2 Algorithm 1

The key concept of scheme 1 is to shift down the MBs that have bitplanes with lower R-D slope. Note that the shifting down operation will replace the original R-D slope of certain bitplane to be 0 since the all zero bitplane is introduced. However, the main drawback to deal with slope problem is that it is not linear operation. For example, if we got N line segments and each has its own slope, the decision for which line segments should be replaced with 0 slope so that the overall final slope resulted by the other rest lines is the largest is not a linear problem. Intuitively, we would first discard the line segments that have lower slopes. However, the optimal solution cannot be obtained by solving the problem sequentially. That is to sequentially discard the

line segments with slopes in increasing order might not be the optimal solution. Fig. 5 shows a special case that violates the intuition. In Fig. 5, there are 3 vectors, vector 1, 2 and 3. If what we want to do is to discard a vector so that the final combined vector slope is the largest. From intuition, one should eliminate the line that has
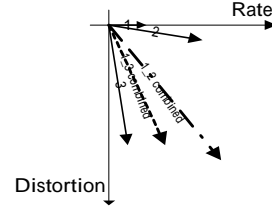

Fig. 5 Illustration of the nonlinear behavior while dealing with slopes

lower slopes first. In the case of Fig. 5, one might select vector 1 whose slope is zero. However, the better choice is to discard the vector 2 as shown in Fig. 5. In terms of the non-linear characteristic, we precisely formulate the relationship involved in the slope selection problem. Again, considering that we have three line segments with different slopes, S1, S2 and S3. And our goal is to find the vector and replace it with zero slope so that the final overall slope of vector composed by the other rest vectors is the largest. Assuming that vector S3 is to be discarded. The following derivation shows the sufficient conditions for S3 to be the better choice than the other two vectors. And the stop (or select) criterion is also given.

<u>Derivation I:</u> (Having a better choice): For S3 to be the better choice, the final slope by replacing the S3 should be greater than that by replacing S1 or S2.

$$\frac{S_1\Delta R_1 + S_2\Delta R_2 + 0}{\Delta R_1 + \Delta R_2 + 1} > \frac{S_1\Delta R_1 + 0 + S_3\Delta R_3}{\Delta R_1 + 1 + \Delta R_3}$$
$$\Rightarrow (S_2 - S_1)\Delta R_2\Delta R_1 + (S_1 - S_3)\Delta R_3\Delta R_1 + (S_2 - S_3)\Delta R_2\Delta R_3 + C(\Delta D_2 - \Delta D_3) > 0 \qquad C = 1 \qquad here \quad Eq(1)$$
$$\Rightarrow (\Delta D_2\Delta R_1 - \Delta D_1\Delta R_2) + (\Delta D_1\Delta R_3 - \Delta D_3\Delta R_1) + (\Delta D_2\Delta R_3 - \Delta D_3\Delta R_2) + C(\Delta D_2 - \Delta D_3) > 0$$

$$\frac{S_1\Delta R_1 + S_2\Delta R_2 + 0}{\Delta R_1 + \Delta R_2 + 1} > \frac{0 + S_2\Delta R_2 + S_3\Delta R_3}{1 + \Delta R_2 + \Delta R_3}$$

$$\Rightarrow (S_1 - S_3)\Delta R_1\Delta R_3 + (S_1 - S_3)\Delta R_3\Delta R_1 + (S_2 - S_3)\Delta R_2\Delta R_3 + C(\Delta D_1 - \Delta D_3) > 0 \quad where \ C=1 \quad here \quad Eq(2)$$

$$\Rightarrow (\Delta D_1\Delta R_2 - \Delta D_2\Delta R_1) + (\Delta D_1\Delta R_3 - \Delta D_3\Delta R_1) + (\Delta D_2\Delta R_3 - \Delta D_3\Delta R_2) + C(\Delta D_1 - \Delta D_3) > 0$$

Derivation II (Stop or Select Criterion):

Replacing S3 with zero slope should lead to higher final overall slope. If the final overall slope is not increased while the vector S3 is replaced, one should find another possible solution or just stop to proceed.

$$\frac{S_1\Delta R_1 + S_2\Delta R_2 + 0}{\Delta R_1 + \Delta R_2 + 1} > \frac{S_1\Delta R_1 + S_2\Delta R_2 + S_3\Delta R_3}{\Delta R_1 + \Delta R_2 + \Delta R_3}$$

$$\Rightarrow \Delta D_1(\Delta R_3 - 1) + \Delta D_2(\Delta R_3 - 1) - \Delta D_3(\Delta R_1 + \Delta R_2 + 1) > 0$$

$$\Rightarrow (\Delta D_1 + \Delta D_2)(\Delta R_3 - C) > \Delta D_3(\Delta R_1 + \Delta R_2 + C) \qquad C = 1 \qquad here$$

$$\Rightarrow \frac{(\Delta D_1 + \Delta D_2)}{(\Delta R_1 + \Delta R_2 + C)} > \frac{\Delta D_3}{(\Delta R_3 - C)} \qquad Eq\ (3)$$

Deduction:

A. Sufficient condition for better choice

From Eq. (1), S3 should satisfy the sufficient condition, $S_2 > S_1 > S_3 \quad , \Delta D_2 > \Delta D_3$, so that it is the better choice comparing to S2. In addition, S3, based on Eq. (2), still have to satisfy the sufficient condition, $S_1 > S_2 > S_3 \quad , \Delta D_1 > \Delta D_3$, so that S3 is the b

etter choice comparing to S1. Note that these are sufficient conditions not necessary conditions. Hence, one might violate these sufficient conditions and still have the best choice.

B. Stop (or select) criterion

In order to have the correct choice, Eq (3) have to be satisfied anyway. It is sufficient and necessary condition for the correct choice. The correct choice means that discarding the selected vectors will lead to higher final overall slope. Eq (3) also tells us when we should stop to further discard vectors. In order to avoid losing generality, the zero slope used to replace those vectors so that the final overall slope is the largest is assumed to have zero distortion reduction and **C** bits are consumed to represent the information. To find the optimal solution, one has to first decide how many MBs should be shifted down, i.e., discarding the

corresponding slopes. As mentioned before, one has to exhaustively try all possible solutions so that the optimal can be exactly found. For example, if we are dealing with MBs at the MSB level of the FGS VOP and have totally N MBs for selecting, then the number of MBs needed to be shifted down can be 1, 2, 3,…….N-1 and the maximum number of total possible solutions are $\displaystyle\sum_{i=1}^{N-1}\frac{N!}{(N-i)!*i!}$

where N!=1*2*3…..N. One has to check all the possible solutions and eliminate those that do not satisfy Eq (3). If one can find a solution among those which have the same number of discarding MBs that satisfies the sufficient condition of better choice, then the solution is the best with the current discarding number of MBs. Finally, the overall best solution is to select the one that lead to the largest final slope among all the legal solutions. As we can find that such kind of searching operation is exhaustive and not feasible, a heuristic solution is proposed as follows:

Procedures of Algorithm 1:

1) First of all, calculating the average slope of the current bitplane from those MBs that are not all zero bitplane.
2) Define the threshold slopes as 0.25*Average slope, 0.5*Average slope, 0.75*Average slope (This can be arbitrarily defined as needed).
3) The possible solutions for the current bitplane is to shift down the non-all zero MBs with slopes lower than the threshold slopes. Note that the MBs with shifting factors to be fixed cannot be further shifted down. In the above case, we can obtain maximum 3 solutions.
4) Discard the solutions that don't satisfy Eq (3).
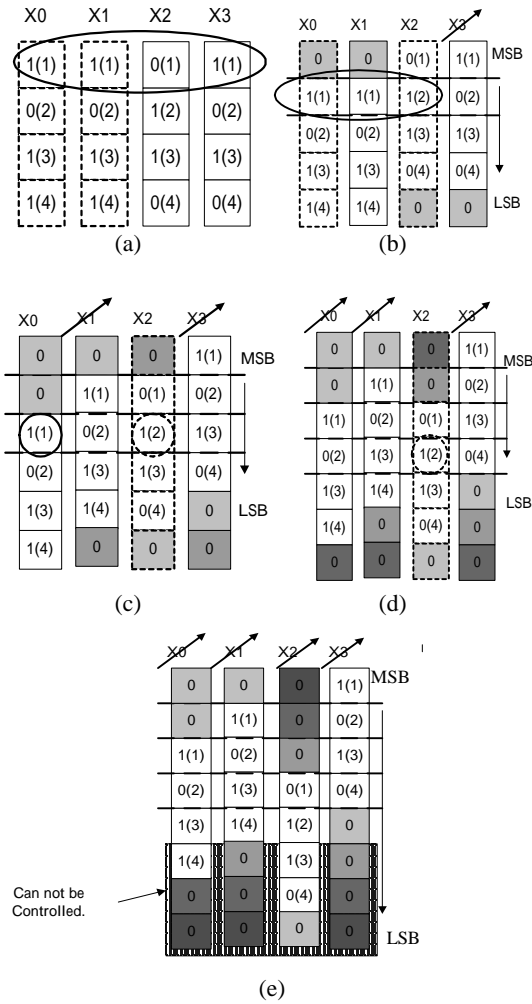5) Keep the best solution that satisfies Eq (3) and results in the largest final overall slope.

(a)

(b)

(c)

(d)

Can not be
Controlled.

(e)

Fig. 6 An example of Algorithm 1

6) The shifting statuses are fixed for those MBs that are not shifted down and not all zero bitplane.

7) Go back to the next bitplane and repeat Step 1-6.

To illustrate the algorithm, we use an example shown in Fig. 6 for illustration. The respective illustrations are listed as follows:

a) We are at the MSB level and X0, X1 need to be shifted down after calculation. Since X2 is all zero bitplane, the only MB that should be fixed is X3.

b) Now, we go the 2nd bitplane. The MBs X0, X1 and X2 are candidates that can be

shifted down. After calculation, X0 and X2 are shifted. Thus, X1 is fixed after this step.

c) For the 3rd bitplane, the MBs still remain movable are X0 and X2. Again, calculating the average slopes and shifting down the MBs according to procedures (3)-(6). Assuming that X2 should be shifted down. Then, X0 is now fixed.

d) For the 4th bitplane, the only MB that can be shifted down is X1. Assuming that X1 is shifted down.

e) For the 5th bitplane, X1 cannot be further shifted down since shifting down X1 will lead to bitplane 5 cannot satisfy Eq (3).

The last few bitplanes might be not controlled at all as shown in Fig. 6 (e) since all the shifting factors are determined at the first few bitplanes. The simulation results will be given in section 4.

3.2 Algorithm 2

Although the slopes of the R-D line segments correspond to the first few bitplanes are improved with Algorithm 1, the length of
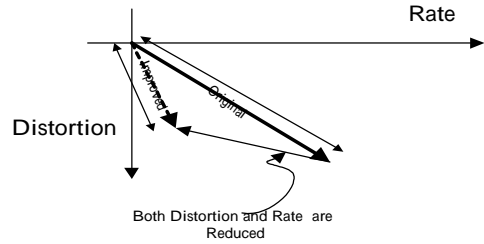

Fig. 7 The shorten effect of shifting down the MB

these R-D line segments is shorten. The shorten effect will limit the improvement of our
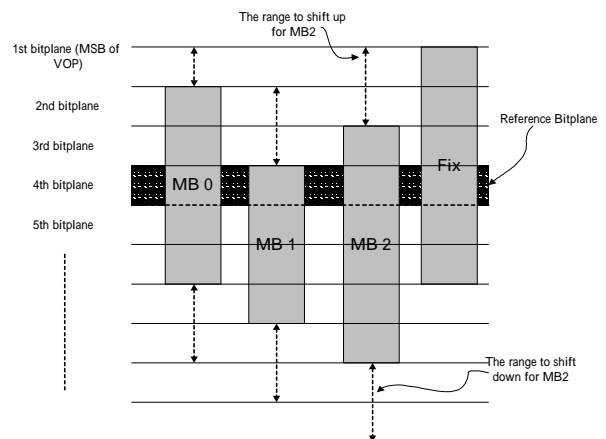

Fig. 8. The 2nd approach that involves the shifting up operation.

8

original idea. The improved R-D curve might quickly cross over the original one such that the area enclosed by the improved R-D curve and the original one is reduced. The smaller size of the area, the less gain we will have. The reason for this comes from the fact that the slope improvement is obtained by shifting down the MBs that have lower slopes and replace with zero slope by introducing all zero bitplane. In general, the average rate used to represent all zero bitplane is about 1~2 bits and it increases when the all zero bitplane appears at lower level. On the other hand, the distortion reduction of the introduced all zero bitplane is 0. Thus, the amount of distortion reduction and corresponding rate of certain bitplane for FSG VOP are both reduced while we shifting down some MBs. Fig. 7 shows the shorten effect. To conquer the shorten problem, we think that the shifting up operation should also be involved in the improvement of R-D curve. In scheme 2, the primitive idea is to first find a bitplane for reference and use Lagrange optimization method to optimize the R-D curve before the reference bitplane. Note that the MBs can be shifted up or down while comparing to the reference bitplane. Fig. 8 shows an example. In Fig 8, the reference bitplane is 4th bitplane. For each MB in Fig 8, shifting up will need more bits to represent the extra data while the transmission is stop at the reference bitplane. However, the amounts of distortion reduction for these MBs are increased. We will get an opposite situation while the MBs are shifted down. Fig. 9 shows the R-D curve of MB2 in Fig. 8 versus shifting operations. Note that the rate mentioned here are the total bits needed for representing before the reference bitplane. As one can find that, we might have R-D dependence problem among these MBs. For instance, if MB2 is shifted up to surpass the maximum bitplane of FGS VOP, all of the MBs will generate all zero bitplane. The R-D curve will changed for each MB. Thus, one cannot guarantee that the R-D curve for the
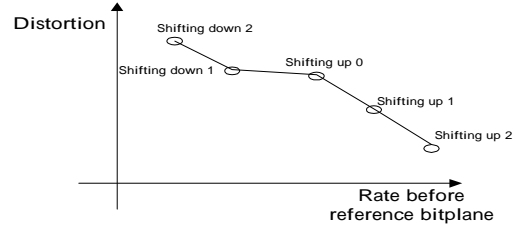

Fig. 9 The R-D curve of MB2 in Fig. 7 versus shifting operation

current MB is fixed since the later MBs might change the R-D curve by introducing the all zero bitplane. Our strategy for this is chose an upper bound. No MBs can be shifted up to surpass the upper bound. Since the bitplane information of different MBs are independent under the proposed framework, the typical Lagrange optimization problem can then focus on the optimization of each individual MB with the constraint rate. The problem is formulated as follows:

$$\min{mize}(\sum_{i=1}^{N} D_i) \textbf{ Subject to } \sum_{i=1}^{N} R_{i,j} < Rc$$

$$\Rightarrow \qquad \min{mize}(\sum_{i=1}^{N} D_i + \lambda \sum_{i=1}^{N} R_{i,j}) \qquad \Rightarrow$$

$\min{mize}(D_i + \lambda R_i)$ **where N is the number of MBs within the current VOP.**

Now, the typical Lagurange solving approach can be applied. The parameters that can be adjusted in this algorithm is listed below:

1) The determination of the reference bitplane.
2) The corresponding Constraint Rate, Rc.
3) Which MB should be fixed?
4) The shifting factor of the fixed MB.

In our algorithm, the Lagurange multiplier might not find the optimal solution since the operational points (maximum 5 points for each MB) on the R-D curve is sparse and is not necessary convex [Ortega]. The dynamic programming such as Viterbi Algorithm may help.

4. Simulation Results
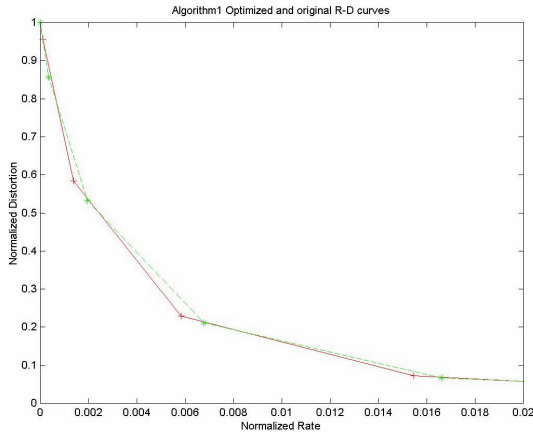
The simulation results are presented in the

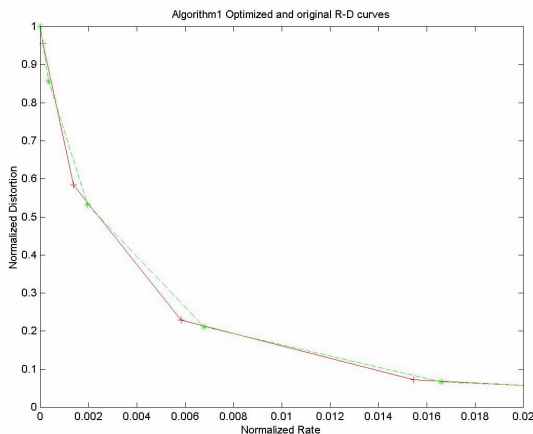Fig. 10 Original R-D curve and that of Algorithm1

Fig. 11 The original R-D curve and that of Algorithm II

following two Figures. The simulation results use single I frame of "foreman" with base layer coded at 256k. Fig. 10 shows the original R-D curve and that of Algorithm 1 and Fig. 11 shows the results of Algorithm 2. The solid lines in both figures are the original R-D curves and the dash lines are the R-D curves of the proposed algorithms. In Fig. 11, we have multiple dash lines that correspond to different reference bitplanes in algorithm 2. In order to zoom in the difference, the dynamic range of X-axis is changed. For Fig. 10, we can have slight gain at the low bit rate. However, as

expected there is a cross point that allows important coefficients being sent first. Thus, it provides better performance at lower bitrate at the cost of inferior performance at higher bit rate.

5. Summaries and Conclusion

In this paper, we briefly describe the FGS architecture and algorithm adopted by MPEG4. The feature and advantage of FGS over heterogeneous networks environment is also pointed out. In addition, we also suggest that the determination of user definable parameter, Frequency Weighting matrix, used to reduce the flickering effect should depend on the block activity of most 8x8 blocks that is strongly related to the bit rate that the base layer is coded and the complexity of input sequence. For the Selective Enhancement shifting factors, we treat it from different point of view and explore the possibility to use it in that way. We believe that the SE shifting factor for each MB can be used not only for region of interest function but also for facilitating the improvement of the R-D curve of FGS VOP. In this paper, two algorithms are proposed to demonstrate how to achieve the R-D improvement within the current syntax constraints. Our algorithms are attractive since all the operations follow the FGS standard.

6. Reference

[1] "Streaming video profile-- Proposed Draft Amendment (PDAM 4)", MPEG00/N3315

[2] Sen-ching Samson Cheung and Avideh Zakhor, "Matching Pursuit Coding for Fine Granular Video Scalability", MPEG98/m3991.

[3] Hayder Radha*, Yingwei Chen, Kavitha Parthasarathy and Robert Cohen, "Scalable Internet video using MPEG-4", Signal

Processing: Image Communication 15 (1999) 95}126

[4] Said, Pearlman, A new fast and efficient image codec based on set partitioning in hierarchical trees, IEEE Trans. Circuits Systems Video Technol. 6 (June 1996).

[5] Weiping Li, "Fine Granularity Scalability Using Bit-Plane Coding of DCT Coefficients", MPEG98/m4204.

[6] Shipeng Li, Feng Wu and Ya-Qin Zhang ,"Study of a new approach to improve FGS video coding efficiency", MPEG99/m5583.

[7] Shipeng Li, Feng Wu, Ya-Qin Zhang, "Experimental Results with Progressive Fine Granularity Scalable (PFGS) Coding", MPEG2000/m5742.

[8] Mihaela van der Schaar, Hayder Radha, Yingwei Chen, "An all FGS solution for hybrid Temporal-SNR scalability", MPEG99/m5552.

[9] Weiping Li, "Complexity Analysis of Two Residue Computation Methods for FGS", MPEG98/m5082

[10] Hong Jiang, "Experiments on Using Post-Clip Addition in MPEG-4 FGS Video Coding", MPEG2000/m5826

[11] Weiping Li, "Experiment Results on Syntax Improvement of Selective Enhancement in FGS", MPEG2000/m5812

[12] Yingwei Chen and Bertrand Thebault, "Results on Selective Enhancement for Fine Granularity Scalability", MPEG99/m5091

[13] Weiping Li, "Improvement of Selective Enhancement Syntax for FGS", MPEG98/m5633

[14] Weiping Li, "Frequency Weighting for FGS", MPEG98/m5589

[15] Hong Jiang, "Using Frequency Weighting in Fine-Granularity-Scalability Bit-plane Coding for Natural Video", MPEG99/m5489