AN EFFECTIVE MODE DECISION SCHEME IN MACROBLOCK-BASED PFGS

Tiantian Sun[1], Jizheng Xu[2], Feng Wu[2], Wen Gao[1]

[1]Institute of Computing Technology, Chinese Academy of Sciences, P. R. China
[2]Microsoft Research Asia, Beijing, P. R. China

ABSTRACT

The Progressive Fine Granularity Scalable (PFGS) coding is a promising technique for streaming video applications. The macroblock-based PFGS coding proposes several INTER modes for the enhancement macroblock coding and provides a simple method to determine a coding mode for each macroblock. Because it does not consider the drifting error possibly occurs in decoder, the mode decision method is not effective. In this paper, we present a mode decision scheme to make a trade-off between the low drifting error and high coding efficiency, by adding a middle bit-rate loop in encoder to estimate the possible drifting error in decoder and using a R-D model to estimate the improvement of coding efficiency by selecting high quality reference. Then a nearly optimal coding mode for each enhancement macroblock can be easily determined. The experimental results show that the proposed scheme can effectively reduce the drifting error at low bit rates, while maintaining improved coding efficiency at high bit rates.

1. INTRODUCTION

With the steady increase in the access bandwidth, more and more Internet applications start to use the streaming audio and video contents. In these applications, the servers normally have to serve a large amount of users with different screen resolutions, frame rates and network bandwidth. Several scalable video coding schemes have been proposed to satisfy this requirement, such as MPEG-4 FGS [3] [4], PFGS [5] and MCTF-based scheme [1]. Recent progress in scalable video coding shows that Motion Compensated Temporal Filtering (MCTF) [1] can provide high coding performance while maintaining scalabilities. But there are still some problems in MCTF such as delay, complexity issues, which make it inapplicable for some applications. So some other techniques, such as FGS, PFGS, are still investigated in MPEG [2].

In response to the increasing demand on streaming video applications over the best-effort Internet, the coding objective for streaming video is to optimize the video quality for a wide range of bit rates so that[1] it can be adapted to the channel conditions. The Fine Granularity Scalable (FGS) video coding adopted in MPEG-4 standard is just such a technique [3] [4]. Because of the fine granularity scalability provided by the bit-plane coding technique in the enhancement layer, the FGS scheme can easily be adapted to the channel bandwidth fluctuation. One major feature of FGS coding scheme is that the base layer and all enhancement layers in a frame are always predicted from the reconstructed version of the base layer in the reference frame. Therefore, the FGS coding scheme provides good error recovery from occasional data losses or errors in enhancement layers. However, since the prediction is always based on the lowest quality base layer, the coding efficiency of the FGS scheme is not so good.

In order to improve the coding efficiency of the FGS, a basic framework, named Progressive Fine Granularity Scalable (PFGS) video coding was proposed [5]. By using the high quality reference for prediction of the enhancement layer, PFGS outperforms FGS by more than 1.0dB in average PSNR because of more accurate motion compensation. On the other hand, drifting error will also be introduced and propagated at low bit rates if this high quality reference bitstream is not received by the decoder. However, it is very difficult for the PFGS scheme to provide a good trade-off between low drifting error and high coding efficiency at the frame level. Then a flexible macroblock-based PFGS coding was proposed [6]. Several modes are provided for the enhancement macroblock coding, and a mode decision is made to choose an appropriate coding mode for each enhancement macroblock. However, because it only considers the improvement of coding performance by selecting the high quality reference, this mode decision cannot effectively reduce the drifting error at low bit rates.

In this paper, we address the mode decision issue in macroblock-based PFGS. A new scheme is proposed to make a good trade-off between drifting error and coding efficiency. There are two key points in our scheme. First, a middle bit-rate loop is added at the encoder to estimate the

---

possible drifting error. Second, we use a simple R-D model to calculate the improvement of coding efficiency by selecting the high quality reference. With these tools, a nearly optimal coding mode can be determined for each enhancement macroblock. This scheme is expected to reduce the drifting error at low bit rates and maintain the high coding performance at high bit rates.

The rest of the paper is organized as follows. The basic macroblock-based PFGS framework is introduced in Section 2. Section 3 describes the whole proposed mode decision scheme in detail. Experimental results are shown in Section 4 and Section 5 concludes the paper.

2. MACROBLOCK-BASED PFGS FRAMEWORK

Figure 1 is the diagram of macroblock-based PFGS encoder. There are two sets of predicted DCT coefficients. (1) A set of predicted DCT coefficients X_b are prediction errors formed by referencing the previous low quality (base layer) reference, and (2) a set of predicted DCT coefficients X_e are prediction errors formed by referencing the previous high quality reference. The DCT coefficients X_b are encoded in the base layer. In the enhancement layer, mode decision module decides which reference will be used, and the differences between X_e and the reconstructed $\tilde{X}_b$ form the enhancement bitstream. Obviously, if base layer is selected as reference, $X_e = X_b$.
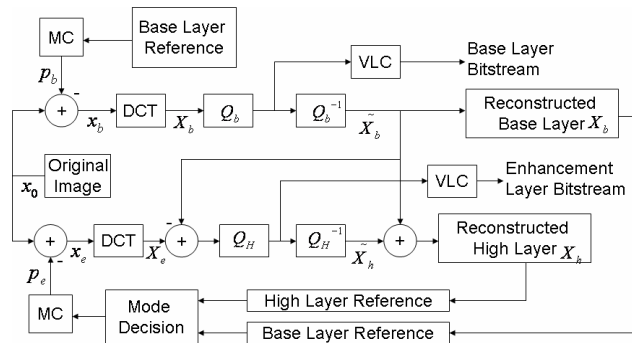

Figure 1.   The Diagram of a macroblock-based PFGS encoder

So there are two INTER modes in the framework. In the LPLR mode, the enhancement layer macroblock is both predicted and reconstructed from the previous low quality reference. Hence the coding efficiency is low. However, there is no drifting error in this mode. In the HPHR mode, an enhancement layer macroblock is both predicted and reconstructed from the previous high quality reference. In this mode, the PFGS scheme can provide the highest coding efficiency but the drifting error will be introduced. The mode decision module selects the one whose residual energy encoded in enhancement layer is the minimum.

3. PROPOSED MODE DECISION SCHEME

Obviously, if the encoder uses the high quality frame to predict the next frame, while the decoder cannot receive the whole high quality reference because of the bandwidth fluctuation, it can only use a reconstructed low quality frame as reference. Then a drifting error will be introduced and propagated. So the total distortion of each enhancement macroblock is composed of two aspects. One is the distortion produced by its own residual quantization and another is the mismatch of the reference between encoder and decoder. It can be formulated as follows:

$$D = D_{quantization} + D_{mismatch} \qquad (1)$$

So our objective is selecting a INTER mode to minimize the total distortion D.

In LPLR mode, because the base layer is always received by the decoder, there is no drifting error and (1) will become:

$$D_{LPLR} = D_{quantization\_BL}$$

And in HPHR mode, (1) will be:

$$D_{HPHR} = D_{quantization\_EL} + D_{mismatch}$$

Comparing the two formulas, we will see that if $D_{quantization\_BL} - D_{quantization\_EL} > D_{mismatch}$, which means $D_{LPLR} > D_{HPHR}$, the HPHR mode should be selected to achieve better performance. Otherwise, the LPLR mode is used to prevent the drifting error.

A.   Drifting Error

First, we will estimate the possible mismatch between encoder and decoder which causes drifting error. Suppose that the decoder receives only part of the bitstream, it will reconstruct a middle quality reference using these bits. If we simulate the same procedure in the encoder side, the drifting error will be the difference between the high quality reference and the middle quality reference.
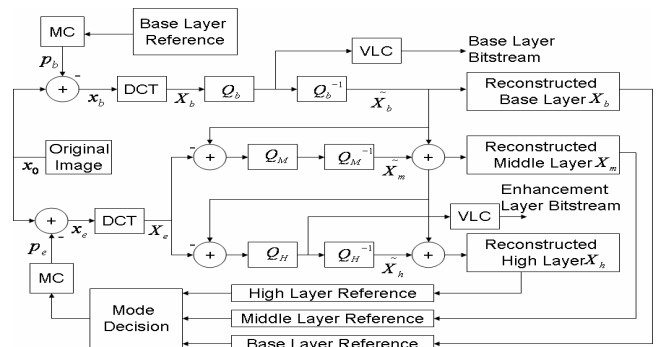

Figure 2.   The Diagram of a macroblock-based PFGS encoder with a middle bit-rate loop

A middle bit-rate loop is added to the macroblock-based PFGS encoder as depicted in Figure 2. During the

encoding of enhancement layer, a middle quality reference will be reconstructed to simulate the reference used in decoder. We use half of bits to reconstruct the middle quality reference because it gives attention to both the low bit rates and the high bit rates.

As discussed above, the distortion caused by mismatch can be estimated by the difference between high quality reference and middle quality reference, that is:

$$D_{mismatch} = D(R_{EL} - R_{ML})$$

where R_{EL}, R_{ML} are the high quality reference and the middle quality reference respectively.

B. Coding Efficiency

In Figure 1, the input signal encoded in the enhancement layer will be:

$$s = \begin{cases} X_b - \tilde{X_b}, & LPLR \\ X_e - \tilde{X_b}, & HPHR \end{cases}$$

We use a simple R-D model [7] to describe the relationship between distortion and bit rate in the enhancement layer encoding. That is:

$$D_{quantization} = D(s - \tilde{s}) = D_0 e^{-\alpha R}$$

where D is the distortion for a given bits R, $\tilde{s}$ is the reconstruction of signal s, α is a constant number and D_0 is the distortion when $R = 0$, that is the energy of input signal s.

So the improvement of coding performance at half bit-rate can be formulated as follows:

$$D_{quantization\_BL} - D_{quantization\_EL} = (D_{0\_BL} - D_{0\_EL})e^{-\alpha R}$$

where D_{0_BL}, D_{0_EL} is the input signal energy using base layer (LPLR) or enhancement layer (HPHR) frame as reference respectively and R is half of bits encoded at the enhancement layer.

Now the next problem is how to determine the parameter α. In temporal domain, because the statistic is highly correlated among neighboring frames, previous frame's information can be used to estimate α. During the enhancement layer coding of previous frame, a set of points (D_i, R_i) and D_0 are tested. Thus the R-D curve can be simulated for the current frame.

C. Mode Decision

Besides the INTER modes, INTRA mode is also allowed in the enhancement layer coding. INTRA mode or INTER mode are determined by the motion estimation module. If a macroblock is encoded with INTRA mode in the base layer, the corresponding enhancement macroblock is also encoded with INTRA mode. If a macroblock in the base layer is

encoded with INTER mode, an INTER coding mode should be decided in the corresponding enhancement macroblock by comparing the total distortion of each mode.

Because each macroblock has only one bit to indicate which reference it uses, the overhead bits can be negligible.

4. EXPERIMENTAL RESULTS

Extensive experiments have been performed to verify the proposed scheme of mode decision. The coding efficiency of the proposed scheme is compared with the original mode decision in macroblock-based PFGS and the FGS coding scheme of MPEG-4 standard. The software is JSVM2.0 [8] and for low delay application, only the first frame is coded as I frame, while the others are coded as P frames. Four sequences are used in this experiment: Bus and Mobile with QCIF and CIF format respectively. In base layer, QP=30 and only one bit-plane is used in enhancement layer.

The experimental results are depicted in Figure 2. The FGS coding scheme only use low quality frame as reference for base layer and enhancement layer of next frame. The original macroblock-based PFGS compares the input signal energy of LPLR and HPHR mode to make a decision for each enhancement macroblock. We can see that the coding performance losses from 0.1 to 0.5dB at low bit rates because of drifting error. However, the proposed mode decision scheme avoid the losses at low bit rates and nearly achieve the same performance as FGS, while at high bit rates, the improved PSNR is up to 1.5~2.0dB. Because more LPLR modes are selected to reduce the drifting error, the performance of our scheme is a little lower than that of the original macroblock-based PFGS at high bit rates. That also proves the middle quality reference at half bit-rate works effectively and makes a good trade-off between the low drifting error and high coding efficiency.
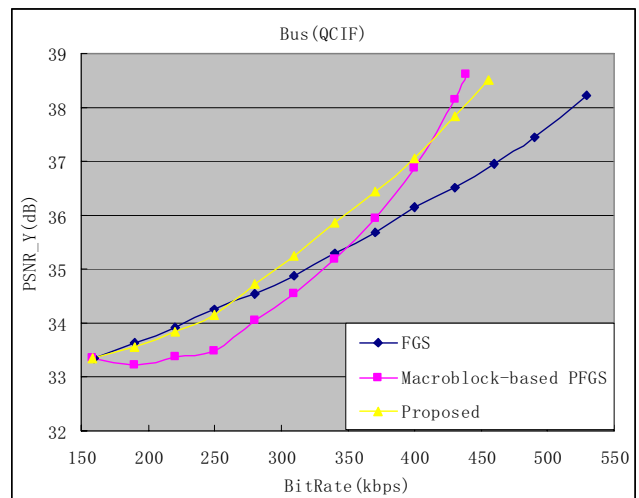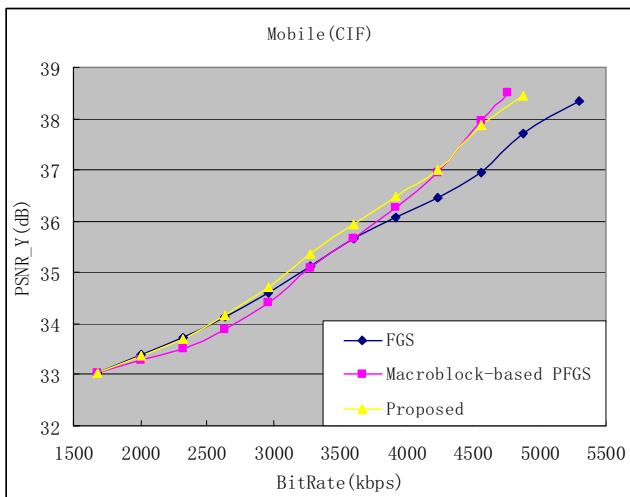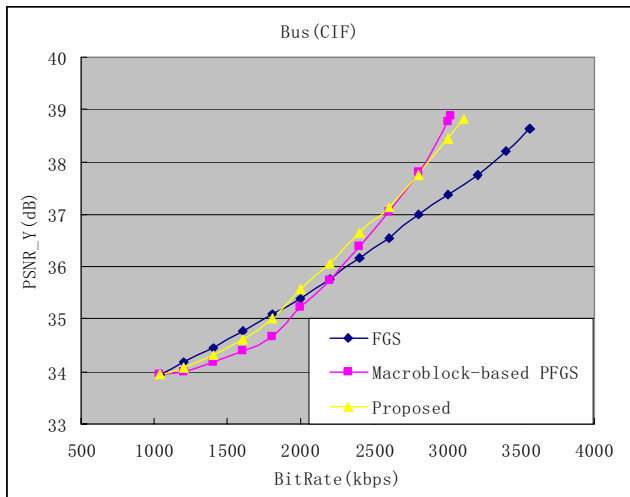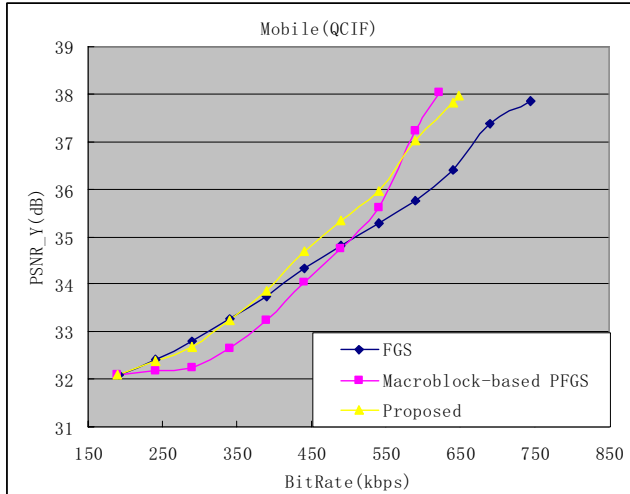
Figure 3.　PSNR versus bit rate comparison among FGS, original macroblock-based PFGS and proposed mode decision scheme

5. CONCLUSION

This paper proposed an effective scheme for mode decision in macroblock-based PFGS, which can achieve a good trade-off between coding efficiency and drifting error. By adding a middle bit-rate loop in encoder, a possible drifting error can be estimated and compared with the improvement of coding performance using high quality reference. Then a INTER coding mode is selected for each enhancement macroblock. The experimental results show that this scheme can reduce the drifting error at low bit rates and maintain high coding efficiency at high bit rates.

6. REFERENCES

[1] J. Reichel, H. Schwarz, M. Wien, "Draft of Joint Scalable Video Model (JSVM) 3.0 Reference Encoding Algorithm Description", ISO/IEC JTC1/SC29/WG11 N7311, Poznan, Poland, July 2005.

[2] "Core Experiment on FGS coding for low-delay applications (CE-7)," JVT-P307, Poznan, Poland, July 2005.

[3] W. Li, "Fine granularity scalability in MPEG-4 for streaming video", ISCAS 2000, Geneva, Switzerland, May 28-31, 2000.

[4] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard", IEEE trans. on Circuits and Systems for Video Technology, Vol.11, No.3, March 2001.

[5] F. Wu, S. Li and Y.-Q. Zhang, "A framework for efficient progressive fine granularity scalable video coding", IEEE trans. Circuit and Systems for video technology, Vol.11, No.3, March 2001.

[6] F. Wu, S. Li, X. Y. Sun, R. Yan and Y –Q. Zhang, "Macroblock-based progressive fine granularity scalable video coding", ISO/IEC JTC1/SC29/WG11, M6779, Pisa, Italy, Jan, 2001.

[7] H. Hang and J. Chen, "Source model for transform video coder and its application ----- Part I: Fundamental Theory", IEEE trans. Circuit and Systems for video technology, Vol.7, No.2, April 1997.

[8] "JSVM 2 Software", ISO/IEC JTC1/SC29/WG11 N7085, Busan, Korea, April 2005.