COMPARISONS BETWEEN THE ONE-LOOP AND TWO-LOOP SOLUTIONS FOR IMPROVING THE CODING EFFICIENCY OF FGS

Feng Wu[1], Shipeng Li[1], Xiaoyan Sun[2], Rong Yan[3], Ya-Qin Zhang[1]
1 Microsoft Research China, Beijing, 100080, China
2 Harbin Institute of Technology, Harbin, 150001
3 Beijing Institute of Technology, Beijing, 100081, China

ABSTRACT

As the MPEG core experiments have shown, the current FGS video coding scheme suffers from a 2~3dB coding efficiency loss compared with the non-scalable single-layer coding for a wide range of bit rates. This is mainly due to the open-loop enhancement structure that ensures drifting free coding. Two categories of schemes are proposed to address this issue with the same basic idea, i.e., introducing some enhanced reference layers into the motion compensation loop. In the one-loop solution, several enhancement bit planes are directly introduced to the motion compensation loop of the base layer. On the other hand, the two-loop solution establishes another motion compensation loop at the enhancement layer. There are many different characteristics with these two solutions. This paper analyzes the one-loop and two-loop solutions in detail from different respects, including coding efficiency, drifting errors, flexibility, and complexity. In the end, the experimental results at the same conditions verify the differences between the one-loop and two-loop solutions.

1. INTRODUCTION

Fine Granularity Scalable (FGS) video coding scheme was already promoted to the FDAM of MPEG-4 Amendment 4 in Pisa meeting [1]. The key to the FGS scheme is the open-loop enhancement structure. All layers are only predicted and reconstructed from the base layer. This assures that the same references are used in both the encoder and the decoder even when the whole enhancement bitstream is somehow dropped. The bit-plane coding technique is employed to encode the quantization residue at the enhancement layer. The FGS scheme has many long desired features, such as that the enhancement bitstream can be truncated and decoded with fine granularity and can easily adapt to the channel bandwidth fluctuations, the errors that occur at the enhancement layer would not affect the frames followed, and so on. However, since only the lowest quality base layer is used as reference in the motion compensation, the coding efficiency of FGS is 2~3dB lower than that of the single layer coding for a wide range of bit rates [2].

The main reason for coding efficiency loss is because of the open-loop enhancement structure that ensures drift free coding. Although the bit-plane coding of DCT coefficients is more efficient than the run-level coding, it yet cannot compensate the coding efficiency loss due to the open-loop structure. Therefore, the technical challenge is how to include a certain amount of enhancement layer information into the prediction loop while minimizing the effect of drifting potentially caused by the prediction mismatch. There have been two categories of potential techniques proposed to address this problem.

The first one is the one-loop scheme proposed in [3], [4]. In this solution, several enhancement bit planes are directly introduced to the motion compensation loop of the base layer. Only limited extra cost is needed in the one-loop solution. However, if the enhancement bit-plane layers used for prediction are not transmitted to the decoder, the one-loop solution will inevitably cause severe drifting errors in the video reconstructed from the base layer and lower enhancement layers.

The second one is the two-loop solution (PFGS) proposed in [5]~[7]. In this solution, a second motion compensation loop is first established at the enhancement layer. Several bit planes of the enhancement layer plus the DCT coefficients of the base layer generate a high quality reference used in the second motion compensation loop. It is clear that some additional costs are needed in this solution, such as frame buffer, inverse DCT and motion compensation. However, there is no drifting error in the base layer of the two-loop solution. Moreover, the two-loop solution allows flexible encoding optimization, as well as some new functionalities, such as spatial scalability.

This paper analyzes the different characteristics of the one-loop and two-loop solutions in detail. The advantages and disadvantages of these two schemes are compared from coding efficiency gain, drifting errors, flexibility, and complexity aspects, respectively. The rest of this paper is arranged as follows. The one-loop and two-loop decoders are described in the Section 2. Section 3 gives the comparisons between the one-loop and two-loop solutions. Experimental results in Section 4 further verify the analyses and conclusions of this paper. Finally, Section 5 concludes this paper.

2. ONE-LOOP AND TWO-LOOP DECODERS

The decoder of the one-loop solution is depicted as in Figure 1. Compared with the FGS scheme, n(t) lower bit planes are introduced to the base layer for reconstruction of the reference. It is very clear that the reconstructed references have higher quality than that of the original FGS scheme. As a result, when the network bandwidth is broad enough to transmit the first n(t) lower bit planes to the decoder, one-loop solution can provide higher coding efficiency.

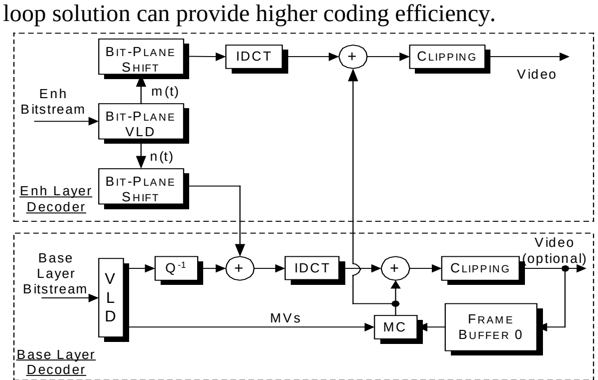

Figure 1 The decoder of the one-loop solution.

The decoder of the two-loop solution is depicted as in Figure 2. The decoder of base layer in the two-loop solution is exactly the same as that of FGS. No bit planes are introduced from the enhancement layer to the motion compensation loop of the base layer. Therefore, the two-loop solution can get the same video quality in the base layer as in FGS. In order to improve the coding efficiency of the FGS scheme, the two-loop solution establishes another set of motion compensation at the enhancement layer. The DCT coefficients in the base layer and the first n(t) lower bit planes in the enhancement layer plus the temporal prediction are used to reconstruct the high quality reference for the second motion compensation loop. Since the enhancement layer is predicted from the high quality reference, the two-loop solution can also provide the higher coding efficiency.

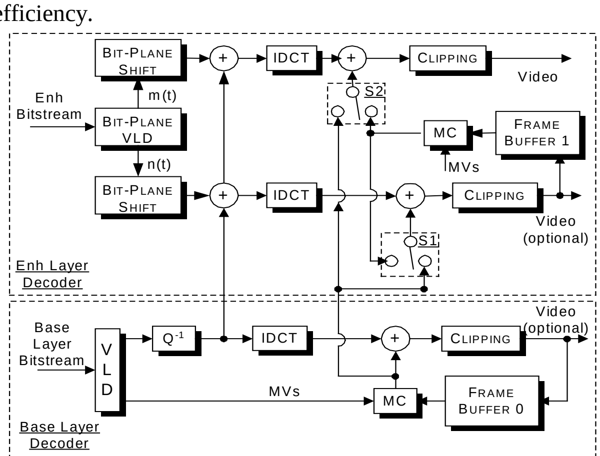

Figure 2 The decoder of the two-loop solution.

Since there are two different quality references in the two-loop solution, the enhancement macroblocks have more flexibility to choose the reference for the prediction and reconstruction. Three inter coding modes proposed in Figure 3 are used for the enhancement macroblock coding. Gray rectangular boxes in Figure 3 denote those bit plane layers to be reconstructed as references. Solid arrows with solid lines are for temporal predictions, hollow arrows with solid lines are for reconstruction of high quality references, and solid arrows with dashed lines are for predictions in DCT domain.

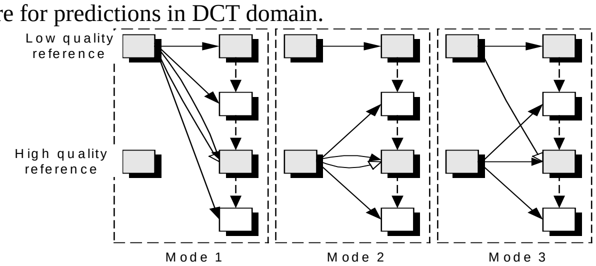

Figure 3 The coding mode for the enhancement macroblock in the two-loop solution.

In the mode 1, all layers are predicted and reconstructed from the low quality reference. If all macroblocks are encoded with this mode, the two-loop solution is exactly the same as FGS. In the mode 2, the enhancement layers are predicted and reconstructed from the high quality reference. This mode can provide high coding efficiency. In the mode 3, the enhancement layers are predicted from the low quality reference. However, the current high quality reference is reconstructed from the previous low quality reference. This mode can effectively terminate the error propagation. These coding modes are implemented by the switch S1 and S2 in the decoder of the two-loop solution.

3. COMPARISONS BETWEEN THE ONE-LOOP AND TWO-LOOP SOLUTIONS

In this section, we will analyze the performance of the one-loop and two-loop solutions from different viewpoints in detail. Firstly, comparison of the coding efficiency of four coding schemes is illustrated in Figure 4. The single layer scheme and the FGS scheme have the same quality at the base layer. As the bit rate increases, more and more bits are used for the reconstruction of references in the motion compensation loop in the single layer scheme, while the bit rate used for the reference in the FGS scheme always remains at the same level of the base layer. As a result, the gap of coding efficiency between the single layer scheme and the FGS scheme could be up to 3.0dB.

In the one-loop solution, several enhancement bit planes are introduced into the base layer for reconstruction of references. If the decoder cannot receive the enhancement layer bitstream used for reconstruction the references due to the limited network bandwidth, the difference between references used in the encoder and the decoder will inevitably cause drifting errors. Therefore, the coding efficiency of the base layer might suffer a loss of up to 2.0dB in the one-loop solution. As the bit rate increases, the coding efficiency of the one-loop scheme rapidly increases too. Once all bit planes that are used for reconstruction of the references are available in the decoder, the one-loop scheme could obtain up to 2.0dB coding efficiency gain.

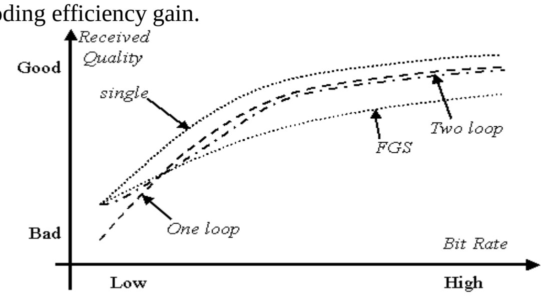

Figure 4 The illustration of coding efficiency of four coding schemes: single layer, FGS, the one-loop solution and the two-loop solution.

In the two-loop scheme, there is no change at the base layer compared with the FGS scheme. Consequently, the coding efficiency of the two-loop scheme at the base layer is the same as that of the single layer scheme and the FGS scheme. If all enhancement macroblocks are coded as Mode 2, the two-loop scheme would have some degree of quality loss at the lower enhancement layers because the bit planes needed to reconstruct the high quality reference may be partially and completely dropped due to bandwidth limitation at lower bit rates. Fortunately, the two-loop solution can effectively control the quality loss by optimally choosing the coding mode of each enhancement macroblock. The curve in Figure 4 is an example of the two-loop solution with limited quality loss at the lower bit rates.

The maximum coding efficiency gain of the two-loop solution can be as high as 2.0dB compared with the FGS scheme. However, the gain is still slightly lower than that achieved in the one-loop solution. There are three reasons that could reduce the coding efficiency gain of the two-loop solution. Firstly, the base layer in the two-loop solution is still encoded using the low quality reference. Secondly, since different references are used for the base layer and enhancement layer coding, there are fluctuations between the base layer and the enhancement layer in the DCT domain. Finally, although Mode 3 can effectively terminate the error propagation, it also affects the coding efficiency, because the high quality reference could not obtain the best quality it could get in this mode. However, the two-loop solution can provide a good trade-off between high coding efficiency and low drifting error by optimally choosing the coding mode of each enhancement macroblock.

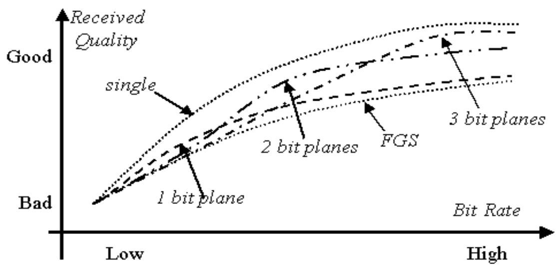

Figure 5 The illustration of different bit planes used in the motion compensation loop.

The one-loop and two-loop solutions have different flexibility at coding optimization. Both the one-loop solution and the two-loop solution can control the maximum coding efficiency gain by choosing how many bit planes entering the motion compensation loop. Figure 5 gives three cases of the two-loop solution. If only 1 bit plane is used to reconstruct the high quality reference, the maximum coding efficiency gain is relative small and appears only at the low bit rate end. As more and more bit planes are introduced to the motion compensation loop, the coding efficiency gain also becomes more and more significant. Furthermore, the maximum gain normally appears at higher bit rates.

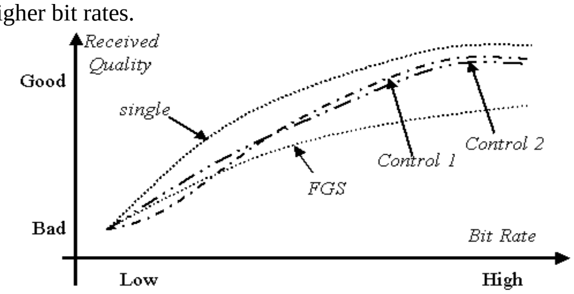

Figure 6 The illustration of two-loop scheme with different optimization.

Besides the flexibility of selecting the high quality reference, the two-loop solution can provide more flexibility by selecting the coding mode of each enhancement macroblock. This is very difficult to implement in the one-loop solution, because only one reference is used in this scheme. Figure 6 gives the PSNR curves of the two-loop solution with different optimization criteria. If all enhancement macroblocks are encoded in Mode 1, the two-loop scheme essentially is equal to FGS. If all enhancement macroblocks are encoded in Mode 2, there will be a significant quality loss at the low bit rates. However, the two-loop scheme has the greatest coding efficiency gain in this case. Two curves in Figure 6 provide a trade-off between low drifting error and high coding efficiency gain. The curve of Control 1 allows a limited quality loss at low bit rates with less macroblocks encoded in Mode 3, and the curve of Control 2 minimizes the quality loss at low bit rates with more macroblocks encoded in Mode 3.

Although both the one-loop and two-loop solutions support temporal scalability, only the two-loop solution can integrate the spatial scalability into fine granularity scalability. Since there are two references in the two-loop scheme, the base layer and lower enhancement layers can be down-sampled to low resolution and predicted from the low quality reference. Other enhancement layers with high resolution are predicted from the high quality reference. Therefore, the two-loop solution can readily support the new functionality of spatial scalability with small cost of coding efficiency.

At last, we will compare the complexity of the one-loop and two-loop solutions. We can clearly see the differences from Figure 1 and Figure 2. There is no additional buffer and computational complexity needed in the one-loop solution. However, the two-loop solution needs an additional buffer at the enhancement layer coding. The increased computational complexity in the two-loop decoder is an additional motion compensation module and another inverse DCT module.

4. EXPERIMENTAL RESULTS

The experiments are designed to verify the above comparisons in the last Section. Two sequences in QCIF format, Foreman and Coastguard, are used in our experiments. The coding frame rate is 10Hz and each GOP lasts 2 seconds. The coding mode is IPP… and there is no B frame in each GOP. The bit rate of base layer is 32kbits/s. The maximum total bit rate is 320kbits/s with an interval of 16kbit/s. These same experimental conditions are used in all the four coding schemes.

In Figure 7, the first two enhancement bit planes are introduced into the motion compensation loop. The coding efficiency gain of both the one-loop and two-loop solutions is close to 2.0dB. Since only a few bit planes are in the motion compensation loop, the maximum gain appears at the middle bit rates. In Figure 8, the first three enhancement bit planes are introduced into the motion compensation loop. Both of these two solutions can achieve more coding efficiency gain while the maximum gain appears at the high bit rate. In Figure 9, the Coastguard sequence is used in the experiment, and only first two enhancement bit planes are introduced into the motion compensation loop. Generally, we can see that the coding efficiency gain of the two-loop solution is slightly lower than that of the one-loop solution for moderate bit rates to high bit rates.

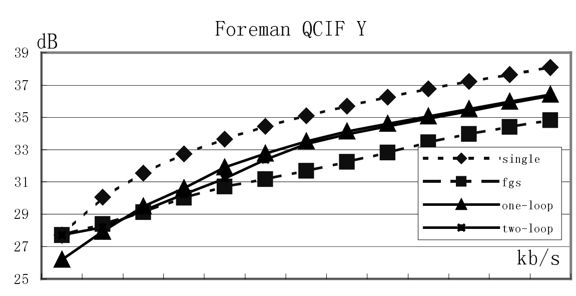

Figure 7 The PSNR curves with 2 bit planes into motion compensation loop for Foreman sequence.

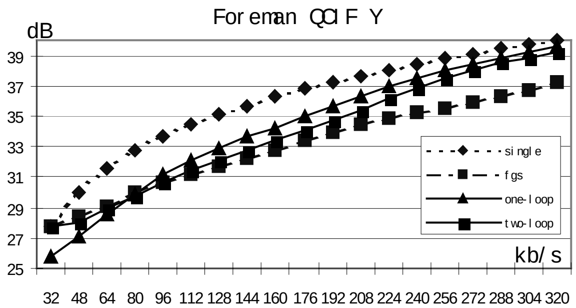

Figure 8 The PSNR curves with 3 bit planes into motion compensation loop for Foreman sequence.

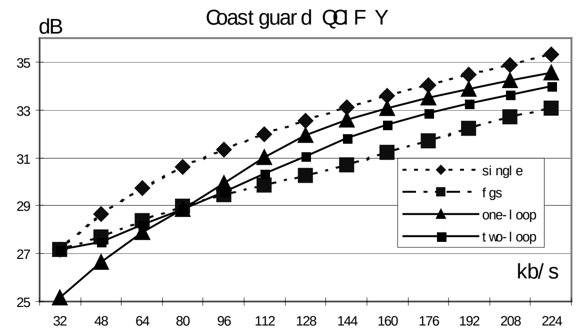

Figure 9 The PSNR curves with 2 bit planes into motion compensation loop for the Coastguard sequence.

However, for all three cases that there is about 2dB quality loss in the base layer of the one-loop solution from Figure 7, Figure 8, and Figure 9. The reason is that the decoder of the one-loop solution could not get the first few enhancement bit planes at the bit rate of the base layer. In fact, the drifting errors severely affect the PSNR and visual quality of base layer in the one-loop solution. Figure 10 gives the PSNR of each individual frame of the one-loop and two-loop solutions. We can see that the maximum quality loss is up to 4.0dB in the one-loop solution compared with two-loop solution. Moreover, the PSNR values in the one-loop solution fluctuate a lot with the period of a GOP. This results in the infamous "pumping" or "breathing" artifact commonly known in the video industry. The visual quality tests with D1 machine also show the severe quality loss at the base layer.

In the authors' point of view, such a base layer with severe pumping artifacts is useless in practice. Any scalable coding scheme should consistently perform across all the bit rates without severely degrading the quality of certain bit rates. The FGS scheme and the two-loop scheme are valid in practice since they do not introduce any other artifacts at any bit rate though the overall coding efficiency is not the best. Moreover, better coding efficiency gain can be achieved in the two-loop approach by simply increasing the base layer bit rate to that of these first few enhancement layers in the one-loop solution.

5. CONCLUSIONS

The one-loop solution provides a cheap and effective method to improve the coding efficiency of FGS at the moderate and high bit rates, but the quality loss and "pumping" artifact caused by drifting errors at the base layer or low enhancement bit rates are intolerable for any real applications. Although the two-loop solution need some additional costs in memory and computation, its performance is consistently higher than that of FGS. The maximum gain in coding efficiency of the two-loop solution is still close to 2.0dB. Meanwhile, the two-loop solution provides great flexibility in encoding optimization by choosing high quality reference and macroblock coding modes. The two-loop solution also allows for extending the current fine granularity techniques to full scalability, including SNR, temporal and spatial scalability.

6. REFERENCES

[1] Video group, "Text of ISO/IEC 14496-2/FDAM4", ISO /IEC JTC1/SC29/WG11, N3904, Pisa, Italy, Jan 2001.

[2] W. Li, "Fine granularity scalability in MPEG-4 for streaming video", ISCAS 2000, Geneva, Switzerland, pp299-302, May 28-31, 2000.

[3] M. Schaar, H. Radha, "Motion Compensation based Fine-Granular Scalability (MC-FGS)", ISO/IEC JTC1/SC29/ WG11, M6475, La Baule, France, October, 2000.

[4] R. Kalluri, M. Schaar, "Single-Loop Motion-Compensated based Fine-Granular Scalability (MC-FGS), with cross-checked results", ISO/IEC JTC1/SC29/WG11, M6831, Pisa, Italy, Jan, 2001.

[5] S. Li, F. Wu and Y. Q. Zhang, "Study of a new approach to improve FGS video coding efficiency', ISO/IEC JTC1/SC29/WG11, M5583, Dec, 1999.

[6] S. Li, F. Wu and Y. Q. Zhang, "Experimental Results with Progressive Fine Granularity Scalable (PFGS) Coding", ISO/IEC JTC1/SC29/WG11, M5742, Noordwijkihout, Netherlands, March, 1999.

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

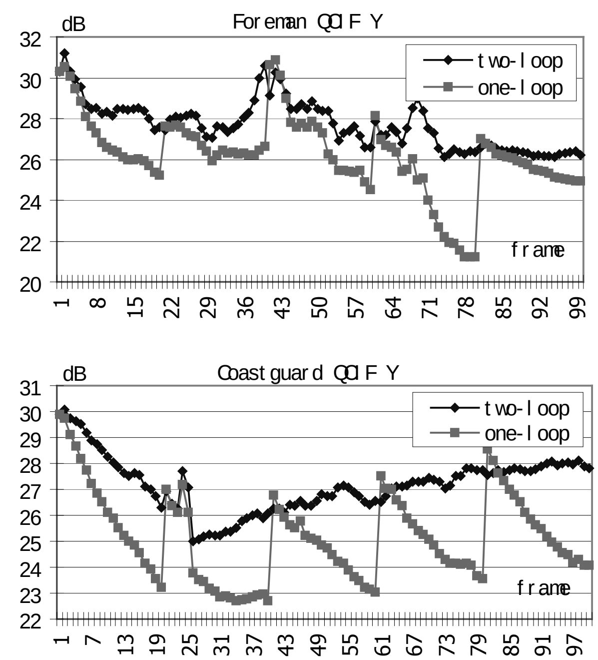

Figure 10 The curves of PSNR versus frame in the base layer