ERROR DRIFTING REDUCTION IN ENHANCED FINE GRANUALITY SCALABILITY

Wen-Hsiao Peng and Yen-Kuang Chen

Intel Corporation

ABSTRACT

We incorporate fading and reset mechanisms in an enhanced fine granularity scalability algorithm to reduce the drifting error at low bit rate while still maintaining 1.5dB PSNR gain at high bit rate over the current MPEG-4 fine granularity scalability. Many of previous works use enhancement layers to predict enhancement layers so as to increase the compression efficiency. Drifting error occurs because enhancement layer, the predictor, is not received as expected. Our fading mechanism linearly combines the current reconstructed base layer and previously reconstructed enhancement layer with fading factors between 0 and 1. Our reset mechanism sets the reference frame for prediction to be the base layer periodically. Our theoretic formulation and experiment results show that drifting error can be distributed more uniformly and maximum accumulated mismatch error is significantly reduced while our mechanisms are turned on. Around 1dB can be improved at low bit rate comparing to the one without any drifting reduction mechanism.

1. INTRODUCTION

In fine granularity scalability (FGS) defined in MPEG-4 Streaming Video Profile [1], the enhancement layer is predictively encoded with the reconstructed base layer as the predictor. Poor coding efficiency of MPEG-4 FGS is expected because only the base layer with poor visual quality is used as the predictor [5].

Using the enhancement layer, which has better visual quality, for prediction can improve the coding efficiency. Our previous work, mode adaptive fine granularity scalability (MFGS) [5], introduced three macroblock based prediction modes, Type B, Type E, and Type BE. Type B mode is the same as MPEG-4 FGS. Type E exploits the previously reconstructed enhancement layer frame as the predictor and Type BE takes the average of predictors from Type E and Type B. Other approaches exploiting enhancement layer for prediction can be found in [2,3,4,6]. All these schemes try to find a better predictor.

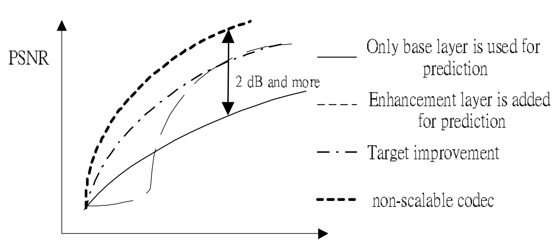


Figure 1: Trade-off between coding efficiency and drifting error.

Approaches using enhancement layer for prediction, however, will encounter drifting errors because of predictor mismatches. To reduce the drifting error, in [7], advance prediction biplane coding artificially sets the reference frame for prediction to be different from that for display. A similar concept is used in our previous MFGS, which has a reset mechanism at the macroblock level.

Additionally, in this paper, we proposed a fading mechanism from our generalized Type BE prediction mode. The generalized Type BE mode linearly combines the current reconstructed base layer and previously reconstructed enhancement layer with fading factors $1-\alpha$ and α . Our algorithm can simultaneously combine the proposed fading and reset mechanisms.

Our goal is to preserve all the scalability features as MPEG-4 FGS while offering better coding efficiency with limited drifting errors. Figure 1 characterizes our target from the rate-distortion perspective. Experiment results show that drifting error can be effectively reduced. Around 1dB can be improved at low bit rate compared to the one without any drifting error reduction mechanism.

2. REDUCTING DRIFTING AND ACCUMULATION ERROR

Predictor mismatches between encoders and decoders cause drifting and accumulation errors. In scalable coding, the predictor of enhancement layer at the decoder is not guaranteed to be the same as the one at the encoder. The perceptual quality could get worse and worse because of drifting and accumulation errors. In this work, we solve the problem in the prediction structure without referring to other error correction techniques.

2.1. Drifting and Accumulation Error

In motion compensated predictive coding, prediction residue $\varepsilon(t)$ at time t can be written the following:

$$\varepsilon(t) = I_s(t) - m.c._t < I_s(t-1) > \quad (1)$$

where $I_s(t)$ denotes the source frame.

Via relating $I_s(t)$ and $\varepsilon(t)$ at different time instances, Eq. (1) can be written as Eq. (2), assuming zero motion vectors. (Such assumption is for notation simplicity and does not affect our final inferences.)

$$I_s(t) = \sum_{k=1}^{t} \varepsilon(k) + I_s(0) \quad (2)$$

Furthermore, taking the mismatch error into account, we modify Eq. (2) as the following:

Table 1: Definitions of predictors for error drifting reduction.

Modes	Equivalent mathematical representation
Generalized Type BE	$P(t) \equiv \alpha * IQ.Q\langle I_B(t)\rangle$ $+ (1-\alpha) * m.c.\langle I_E(t-1)\rangle$
Generalized Type BE+R	$\varepsilon(t) = I_{source}(t) - P(t)$ $\tilde{P}(t) \equiv \alpha * IQ.Q\langle I_B(t)\rangle$ $+ (1-\alpha) * m.c.\left\langle IQ.Q\langle I_B(t-1)\rangle\right\rangle$ $P(t+1) \equiv Trun.\langle \varepsilon(t)\rangle + \tilde{P}(t)$

$IQ.Q\langle x\rangle$: Quantization and inverse quantization.

$m.c.$: Motion compensation

α : A value between 0 and 1

$Trun.\langle x\rangle$: keep certain number of bitplans.

$$\tilde{I}(t) = \sum_{k=1}^{t}\tilde{\varepsilon}(k) + I_s(0)$$
$$= \sum_{k=1}^{t}(\varepsilon(k)+d(k)) + I_s(0) = I_s(t) + \sum_{k=1}^{t} d(k) \quad (3)$$

where $d(k)$ denotes the mismatch error at time k.

Drifting and *accumulation error* can be observed from the implications of Eq. (3) as the following:

1. Drifting: Any mismatch error at time instance t_1 will drift to the reconstruction frames after time t_1, i.e., $\{I_s(t) \mid t \geq t_1\}$.
2. Accumulation: The total error collected by reconstruction frame at time t_1 is the accumulation of the mismatch error before time t_1, i.e., $\{d(t) \mid t < t_1\}$.

2.2. Fading Mechanism

Our fading mechanism is to exponentially fade out the mismatch error term $d(k)$ along time axis. From Table 1, prediction residue of generalized Type BE can be written as Eq. (4) where α is less than 1, $I_B(t)$ and $I_E(t-1)$ are the current reconstructed base-layer frame and previously reconstructed enhancement-layer frame.

$$\varepsilon(t) = I_S(t) - \alpha * m.c. < I_E(t-1) > -(1-\alpha) * I_B(t) \quad (4)$$

Applying the same substitution and assumption in Eq. (3), we can rewrite Eq. (4) as the following:

$$\tilde{I}(t) = I_s(t) + \sum_{k=1}^{t} d(k)\alpha^{t-k} \quad (5)$$

In Eq. (5), the mismatch error occurring at time k can be neglected if the current time instance t is far away from k, i.e. $t >> k$. To enable our fading mechanism, we only require α be less than 1.

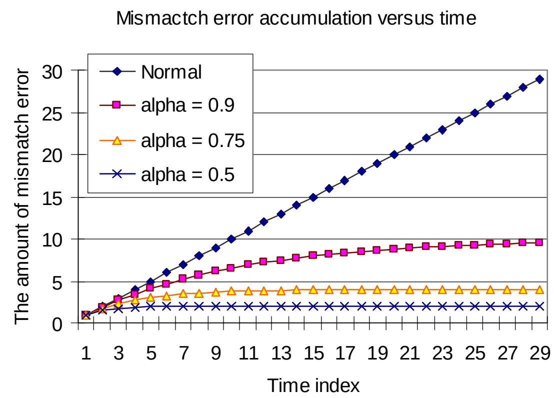


Figure 2: The accumulated mismatch error versus time. Te first frame is I-frame and the rest are P-frame.

Figure 2 illustrates the accumulated mismatch error with the premise that $d(t)$ equals to a constant, **delta**. Smaller α is with more robustness against drifting error. The drifting error is more uniformly distributed and the amount of accumulation error is significantly reduced. However, smaller α has less coding efficiency. In the extreme case, MPEG-4 FGS is with $\alpha = 0$.

2.3. Reset Mechanism

Our reset mechanism is to periodically force the prediction from base layer. This is, we periodically break the dependence of enhancement layer frames. Thus, our reset mechanism is equivalently to let the last term of Eq. (3) equipped with the form of $\sum_{k=1}^{t} d(k)u(k - \left\lfloor \frac{t}{T} \right\rfloor * T)$ where T is the reset period and $u(t)$ is the unit step function. Table 1 defines our reset predictor denoted as "Type BE+R".

When reset is on, the predictor for time instance $t+1$, i.e., $P(t+1)$, is calculated by adding the modified predictor of generalized Type BE, $\tilde{P}(t)$ with the truncated prediction residue $\varepsilon(t)$. The modified $P(t)$ replaces $I_E(t-1)$ by $IQ.Q\langle I_B(t-1)\rangle$ in motion compensation term and it is the same as the advanced prediction bitplane coding in [7].

3. MODIFIED MFGS ENCODER WITH FADING AND RESET MECHANISM

Our fading mechanism only requires the α value of Type E be less than 1. To implement the reset mechanism requires two extra switches.

Different from our previous work, we generalize the definition of Type E, Type BE and Type B by α value. Now, each prediction mode can flexibly have its own α. The only restriction is that α of Type E, Type BE and Type B are defined to be in descending order. While there could be infinite combinations of α, experiment results show that three prediction modes can bring reasonable improvement at acceptable complexity increasing.

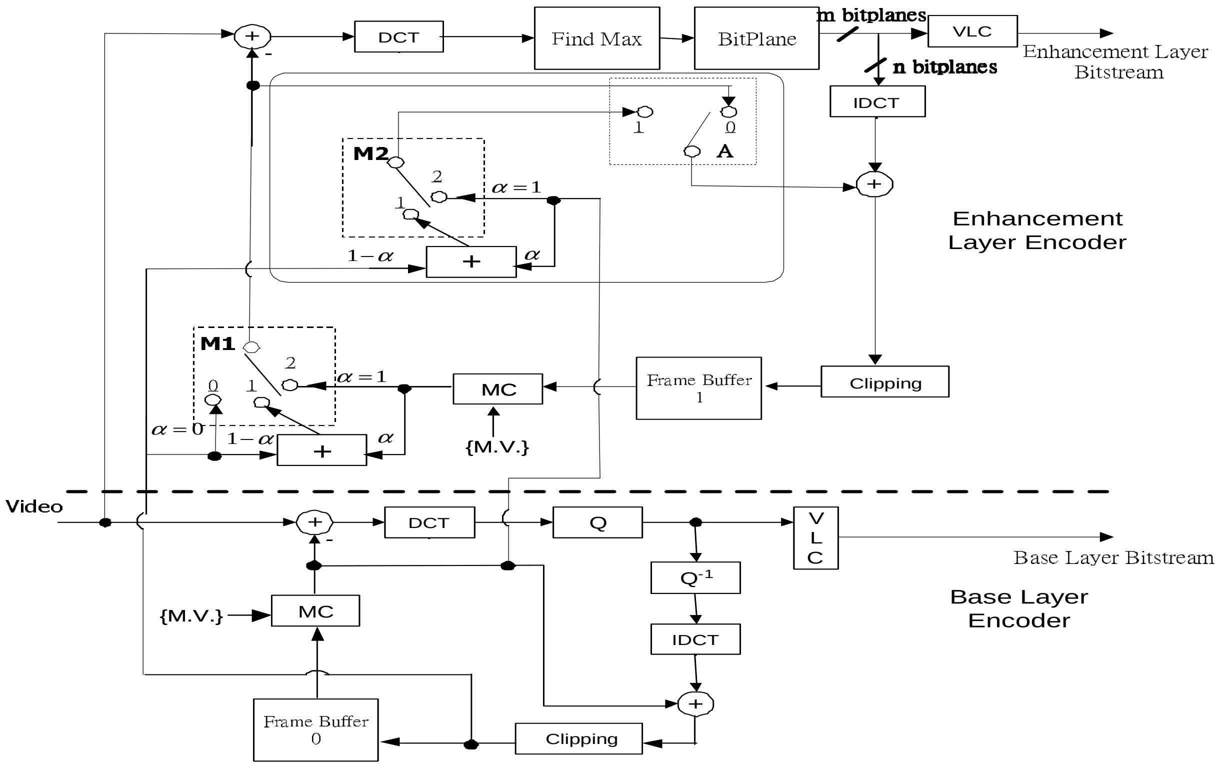


Figure 3: Modified MFGS encoder structure with the reset mechanism.

The modified MFGS encoder with the reset mechanism is depicted in Figure 3. Setting switch **A** at "0" position leads to the original MFGS encoder. Two extra switches, **M2** and **A**, are added to enhancement layer encoder of MFGS for the reset mechanism. Switch **M1** is used to adapt and choose the prediction mode at the macroblock level for coding efficiency improvement. Switches, **M2** and **A**, are for implementation of generalized Type BE+R mode. Switch **M2** is synchronized with **M1** and the mode decision is performed in **M1** module only, i.e. the adaptation result of **M1** is input to **M2** once reset is required. The corresponding decoder can be derived following the same way. Table 2 lists the configurations of different switches and their corresponding prediction modes implemented.

The side information required to our decoder are (1) prediction modes for each macroblock at enhancement layer, (2) the number of bitplanes used for enhancement-layer prediction and (3) the α of our fading mechanism. Specifically, we have (2) and (3) be transmitted at frame and video object layer level respectively. Their overheads are minor. However, (1) is required at the macroblock level. Entropy coding is applied.

Table 2: Configurations of encoder switches and their corresponding prediction modes

Configuration (M1,M2,A)	Corresponding prediction modes
(0,x,0)	Type B
(1,x,0)	Type BE
(2,x,0)	Type E
(2,2,1)	Type E+R
(1,1,1)	Type BE+R

x: Don't care

4. EXPERIMENT RESULTS

In the experiments, there are 59 P frames between 2 I frames. The base layer is coded by fixing the quantization parameter to be 31. To the fading mechanism, we only changed α of Type E. α of Type BE and Type B remains 0.5 and 0 respectively. When α of these three modes are less than 1, fading mechanism is enabled. Our reset predictor is inserted periodically at frame level. In the figures and tables, MFGS_F**m** represents our fading mechanism with α of Type E being **m**. MFGS_R**n** denotes the reset mechanism with reset period being every **n** frames. MFGS_R**n**_F**m** uses two mechanisms simultaneously. The rest of test conditions are the same as in [5].

Figure 4 gives the time-based profiling of PSNR variation using different configurations. The MPEG-4 FGS is used as our basis. Figure 4 illustrates that our fading and reset mechanisms can uniformly distribute and significantly reduce the drifting and accumulation error compared to the one without any drifting reduction mechanism. In our configurations, the fading mechanism can offer more uniform visual quality over the reset mechanism. To have constant visual quality with our reset mechanism, we need additional algorithm to uniformly insert the reset predictor at the macroblock level.

Figure 5 shows the rate-distortion curves of various configurations. At low bit rate, while the original MFGS has the worse performance due to drifting error, configurations with fading and reset mechanisms can have competitive performance to MPEG-4 FGS. Without any drifting reduction, PSNR drops 1~1.5dB at low bit rate, compared to base layer. On the other hand, our fading and reset mechanisms can be significantly improved with ignorable degradation at higher bit rate.

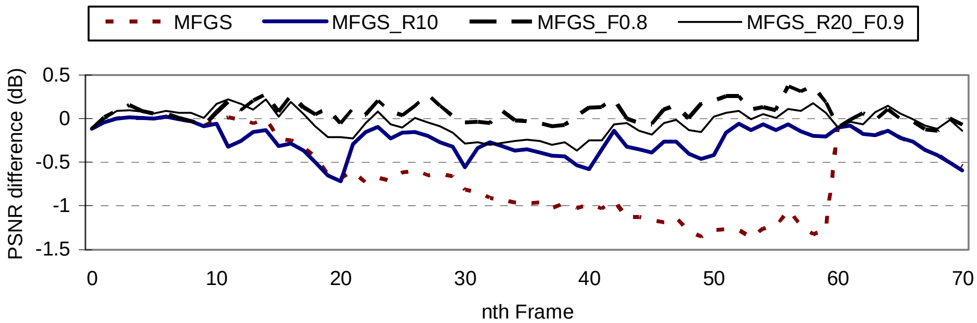


Figure 4: PSNR profiling for drifting error analysis. Y axis = PSNR of proposed codec – PSNR of MPEG-4 FGS. The input is Akiyo of CIF format and 30 frames/s. Base-layer is coded at 80kbits/s. Enhancement layer is truncated at 128kbits/s.

5. CONCLUSIONS

To reduce error drifting within MFGS framework, reset and fading mechanisms are introduced. Our reset mechanism requires two extra switches, compared to our previous work. The fading mechanism is a generalized Type BE prediction mode and requires no extra switches for implementation. Experiment results show that the fading mechanism can more uniformly distribute the mismatch error and significantly reduce the maximum accumulated error. While still maintaining 1.5dB PSNR gain at high bit rate over the current MPEG-4 fine granularity scalability, around 1dB can be improved at low bit rate comparing to the one without any drifting reduction mechanism.

The improvement is at the cost of 30%~40% additional complexities over the original MPEG-4 FGS. The MFGS without drifting reduction mechanisms is about 17%-23% more complex than MPEG-4 FGS (encoder and decoder). Our fading and reset mechanisms further increase the encoder and decoder complexity by 10%-16%.

The performance of our fading and reset mechanisms are affected by the motion characteristic. Less efficient motion estimation generally produces more prediction residue. The prediction residue is the mismatched amount of enhancement layer used for prediction between the encoder and the decoder when the decoder receives no enhancement layer in the worst case. Since the drifting error is proportional to the mismatched amount, fast-motion sequences normally have more drifting error than slow-motion sequences.

Combining the fading and reset mechanisms offers more optimization opportunities. For instance, α of Type E can be higher while the reset frequency is higher. This can cause better performance at higher bit rate without losing performance at lower bit rate. Better performance can be expected with further optimization on these parameters.

6. REFERENCES

[1] "Proposed Draft Amendment 4", ISO/IEC JTC1/SC29/ WG11, MPEG00/N3315, Mar 2000.

[2] J. F. Amold, M. R. Frater, and Y. Wang, "Efficient Drift-Free Signal-to-Noise Ratio Scalability," IEEE Trans. on CSVT, vol. 11, pp. 332-344, Feb. 2000.

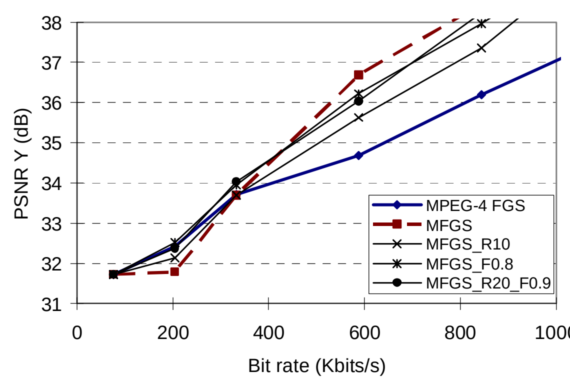
[3] Y. He, R. Yan, F. Wu, and S. Li, "H.26L-Based Fine Granularity Scalable Video Coding," ISO/IEC JTC1/SC29/ WG11, MPEG01/m7788, Dec. 2001

[4] R. Kalluri and M. Schaar, "Single Loop Motion-Compensated Based Fine-Granular Scalability (MC-FGS), with Cross-Checked Results," ISO/IEC JTC1/SC29/ WG11, MPEG00/M6831, Jan. 2001.

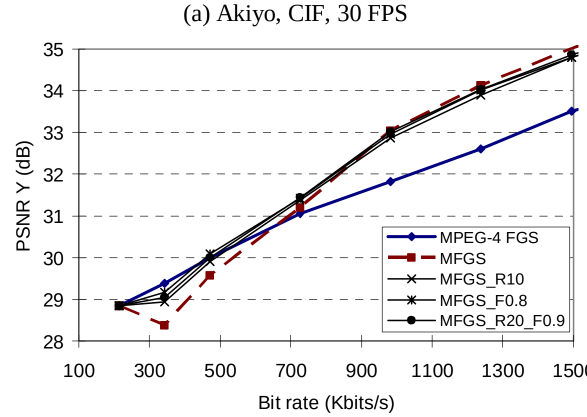
[5] W. H. Peng and Y.-K. Chen, "Mode-Adaptive Fine Granularity Scalability," ICIP, pp. 993-996, Oct. 2001.

[6] K. Rose and S. L. Regunathan, "Toward Optimality in Scalable Predictive Coding," IEEE Trans. on Image Processing, vol. 10, pp. 965-976, July 2001.

[7] F. Wu, S. Li, and Y. Q. Zhang, "A Framework for Efficient Progressive Fine Granularity Scalable Video Coding," IEEE Trans. on CSVT, vol. 11, pp. 332-344, Mar. 2001.



(a) Akiyo, CIF, 30 FPS



(b) Foreman, CIF, 30 FPS

Figure 5: Rate distortion performance of modified MFGS with drifting error reduction.