LOW-COMPLEXITY FGS CODING FOR INTERLACED SVC IN LOW-DELAY APPLICATIONS

Yanyan Zheng[1], Xiangyang Ji[1], Feng Wu[2], Debin Zhao[1], Wen Gao[1]

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

ABSTRACT

JVT SVC has adopted AR-FGS (Fine Granularity Scalability with Adaptive Reference) to efficiently encode the progressive sequences for low-delay applications. Generally, low-delay applications demand low complexity. However, to some extent, the significant gain of AR-FGS is achieved at the expense of much higher coding complexity. In this paper, we propose an efficient FGS coding scheme for low-delay applications that can more smartly compress the interlaced video sources by fully utilizing the special temporal correlation of interlaced sequences. Compared to AR-FGS in coding interlaced sequences, the new approach is able to significantly reduce both the encoding and decoding complexity while maintaining comparable compression performance. Moreover, benefited from the introduced open-loop MC, the proposed scheme can also support additional temporal scalability.

Index Terms— SVC, interlace, AR-FGS, low complexity, low delay, open-loop.

1. INTRODUCTION

As an extension of H.264/AVC, Scalable Video Coding (SVC) is well known for its efficient support for temporal, spatial and SNR scalabilities with high coding performance [1]. Specifically, SVC can enhance multimedia compressing and transporting with combinations of diverse capabilities, including different video quality (SNR), frame rate (temporal), display resolution (spatial), error resilience, and so on. In practice, each combination corresponds to a particular application. When it comes to the environment that demands low delay, for instance the real-time video communication, the primary requirement of such kind of application is to minimize the end-to-end delay, and accordingly the video sequences are encoded as either I-slice or P-slice without hieratical B coding structure embedded which will incur a long delay. Nevertheless, the compression performance of conventional close-loop P coding has been proved to be quite far from ideal without the aid of hieratical B coding structure.

Currently in the case of progressive coding for low-delay applications, SVC has accepted an enhanced FGS coding scheme customarily called AR-FGS [2] which employs an adaptive reference and cyclical sub-band/block coding method. It is able to efficiently improve the coding performance (up to 4-5 dB) for SNR enhancement layer, due to the utilization of more accurate temporal prediction which is adaptively formed from both the enhancement layer reference and base layer reference on the basis of the motion and mode information coded in the base layer. Moreover, the leaky prediction used in AR-FGS plays a key role to keep balance between the coding efficiency and the error robustness. However, compared to the conventional MPEG-4 FGS, AR-FGS has to incorporate some extra modules which take up a quite considerable amount of computation operations at both encoder and decoder. Consequently, the resulted increasing complexity will inevitably, to some extent, hinder its extensive application, particularly for the mobile terminals with low computation capability and limited power.

As to the case of interlaced coding, since the interlaced scanning technology has persisted in various camera and display designs many years, there are still a large amount of interlaced video contents widely used in digital TV broadcast and storage applications, which also need to be delivered and displayed in real time by the orders of the clients. Thereupon, based on our previous work [6], in this paper, we mainly focus on the low-delay coding of interlaced video source with relatively much lower encoding/decoding complexity as well as still high compression efficiency.

The rest of this paper is organized as follows: Section 2 briefly introduces the current FGS coding method for low-delay applications adopted in JVT SVC standard. Section 3 presents and analyzes our proposed FGS coding approach that can more efficiently compress the interlaced video sequences. In Section 4, experimental conditions and results are displayed to verify the advantages of our technique.

2. FGS CODING FOR LOW-DELAY APPLICATIONS

The basic idea of AR-FGS is to employ a leaky based prediction solution to improve the FGS coding of the close-loop P frames by using temporal prediction signal which is adaptively formed form both of the enhancement layer and base layer reference frame based on the information coded in the base layer. Currently in AR-FGS, all the frames are coded as key pictures, and the base layer is coded either as I-slice or P-slice in close-loop motion compensation, while the SNR enhancement layer is coded as F-slice trying to make full use of the temporal correlation among the enhancement layer. Fig.1 shows SVC encoder scheme with AR-FGS embedded (in green).

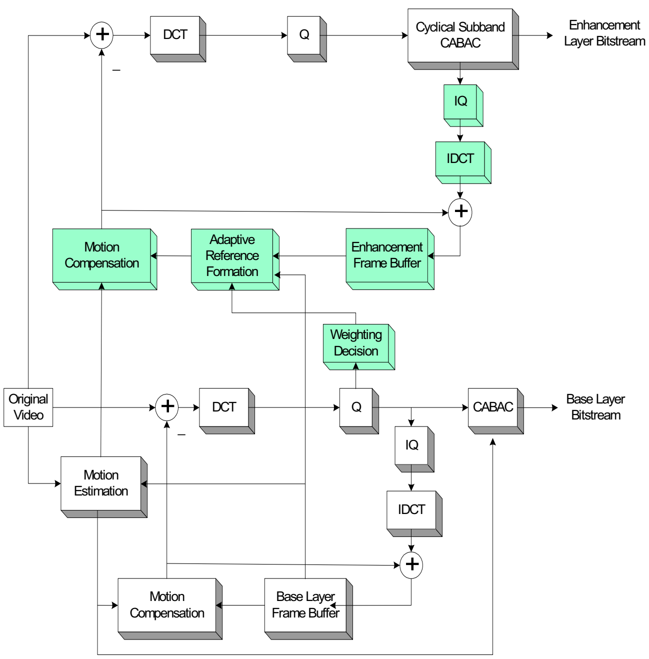


Fig.1. SVC encoder architecture with AR-FGS coding.

3. INTERLACED FGS CODING

Although the FGS coding efficiency of is significantly improved by AR-FGS for low-delay applications, it is worth noting that the compression performance gain is achieved at the expense of much more complicated encoding/decoding process with several additional coding modules. As shown in fig.1, in order to efficiently encode the enhancement layer, AR-FGS has to integrate some addition coding modules (highlighted in fig.1), including motion compensation (MC), inverse-transform (IDCT), inverse-quantization (IQ) and an enhancement layer frame buffer.

Moreover, other new function modules are also needed such as "adaptive reference formation" which actually consists of another series of DCT－>Q－>IQ－>IDCT operations to be able to more accurately set the leaky parameters in either frequency domain or spatial domain based on the judgments of "*weighting decision*". All theses added modules introduced by AR-FGS, especially the two-loop motion compensations (MC) which are always required for constructing the base layer and SNR enhancement layer at both encoder and decoder, are not a neglectable burden, especially for those power-limited decoding devices.

For the interlaced video sequences in low-delay applications, here we present a novel FGS coding method that can decrease half of the amount of computations of AR-FGS and meanwhile keep similar coding efficiency. Considering the special features of interlaced sequences, different from progressive sequences, there should be more flexible FGS coding methods for compressing interlaced video sources. To compare clearly, a straightforward way to implement interlaced FGS coding in low-delay SVC is first introduced below.

3.1. Normal AR-FGS Interlaced Coding

As the anchor algorithm, normal AR-FGS coding with close-loop motion compensation is conceptually shown in Fig.3. When coding the base layer pictures, the current bottom field is predicted from the top field, with which it belongs to same frame; and the following top field in the next frame will predict from the current bottom field.

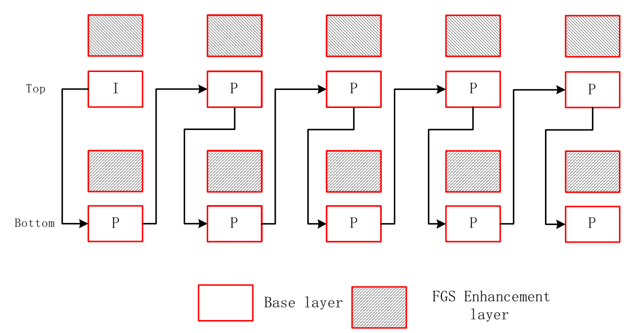


Fig.2. Normal AR-FGS with close-loop MC

All these prediction references are base-layer quality, which means that no enhancement layer information is incorporated to form the references when coding the base layer pictures. While for encoding the FGS enhancement layer, the adaptive reference for the current FGS top or bottom field is formed with the base layer reference and the enhancement layer reference in the previous reconstructed bottom or top filed picture. The advantage of the coding scenario in fig.2 is that once the enhancement layer streaming is truncated or transmitted with errors, the decoder can still correctly decode the base layer pictures. Nevertheless, the overall coding efficiency is sacrificed due to the limit to the utilization of the enhancement layer data.

3.1. Improved AR-FGS Interlaced Coding

Considering the fixed temporal distance and the strong temporal correlation between the top and bottom fields which belong to the same frame, it should be suitable to perform open-loop motion compensation between them to efficiently remove the temporal redundancy in interlaced sequences.

For normal AR-FGS interlaced coding scenario, the compression ratio of both the base layer and enhancement layer are negatively influenced by always using the low quality references for coding the base layer. It is because there are still a lot of temporal correlations existing among adjacent filed pictures, which have not been fully explored due to the inefficient prediction of the base layer.

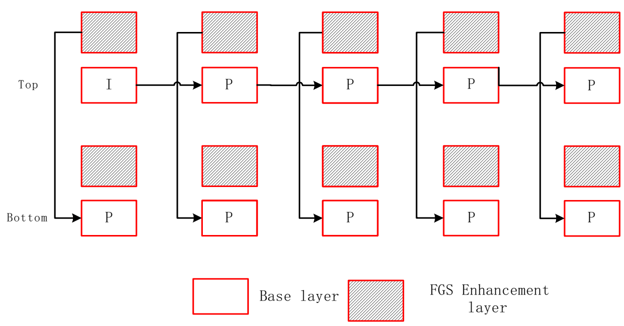


Fig.3. Improved AR-FGS with open-loop MC.

For the proposed AR-FGS interlaced coding method, as shown in fig.3, in order to more efficiently decrease the temporal redundancy, when encoding the bottom fields of the base layer, higher quality pictures instead of the base-layer quality pictures are employed as the prediction references. In other words, the prediction signals of bottom fields in base layer are achieved using open-loop motion compensation, with the enhancement layer reconstruction of top fields as their reference.

Thereupon for the bottom fields, the residuals of the base layer calculated by using improved AR-FGS will be much closer to the residuals of the enhancement layer coded with normal AR-FGS. That is because both employ the high quality references; except that before actually coding the enhancement layer residuals, they should have subtracted the base layer residuals.

Then for bottom fields, no further temporal prediction among the enhancement layer is really needed, since with the proposed method the base layer has predicted good enough to reduce the temporal redundancy between the adjacent top and bottom fields. Therefore, in our proposed improved AR-FGS (fig.3) coding scheme, those complicated operations as done in AR-FGS can be omitted without sacrificing coding performance, because the compression efficiency of the base layer with higher quality references and the enhancement layer with basic FGS coding method should be, more or less, as efficient as that of normal AR-FGS (fig.2).

As illustrated in fig.3, the motion-compensated prediction signals for base layer of P-slice in bottom fields are formed from the weighted combination of the base layer and the highest-quality representation of the corresponding top fields (e.g., with leaky parameter γ set to 0.2 for base layer and 0.8 for enhancement layer), in which way much more accurate prediction signal can be achieved. And no further temporal prediction occurs in enhancement layer of bottom fields. Contrarily, when coding the base layer of top fields, the base layer quality references from their previous top fields are employed. In this way, the error propagation will be able to be effectively prevented. And the enhancement layer of top fields is predicted using the adaptive references formed by leaky combination of both the base layer and enhancement layer references.

As far as the coding complexity is concerned, one major difference from normal AR-FGS is that, when encoding/decoding the base layer and SNR enhancement layer of bottom fields, only one motion compensation loop instead of two is required. As illustrated in fig.2, the highlighted coding modules which are introduced by AR-FGS can all be saved for bottom fields coding with our proposed method. Correspondingly, the computation complexity introduced by AR-FGS can be half reduced.

Another benefit is that this coding strategy can support further temporal scalability by dropping the bottom fields coded with open-loop motion compensation without influencing the coding efficiency of the preserved top fields that are coded with close-loop motion compensation. Therefore, the proposed interlaced FGS coding scheme should be more suitable for low-delay and low-complexity applications in fluctuating transmission channel.

4. EXPERIMENTAL RESULTS AND ANALYSIS

To evaluate the coding efficiency of the proposed improved AR-FGS coding scheme for interlaced SVC in low-delay applications, it has been integrated into interlaced SVC software [3]. And the results of improved FGS [5], normal AR-FGS and normal FGS (both are the original methods as in JSVM) are also tested and shown together with improved AR-FGS under the same coding conditions for overall comparison.

Simulations were carried out for various interlaced video sequences as suggested in [4] and here we display two results of interlaced SD frames (720x576) Mobile and Canoe @25Hz. Only the first key field is coded as intra picture and the remaining key fields are coded as inter-P pictures. The initial QP for base layer is set to 42, and two SNR enhancement layers are appended on top of the base layer, with each enhancement layer QP decreased by 6.

In fact, different proper combinations of leaky parameters (α, β) in normal AR-FGS are simulated and the pair with the best coding performance is selected and used in both normal AR-FGS and the proposed improved AR-FGS. And for improved AR-FGS, γ is also adjusted with suitable value to help keep tradeoff between coding efficiency and error drifting.

The experimental results in fig.4 have shown that for SNR enhancement layer, the proposed coding scheme improved AR-FGS is able to achieve comparable coding performance as normal AR_FGS for interlaced sequences with slight loss at base-layer bit-rate and some gain at middle and high bit-rate. Both the loss and the gain are mainly resulted from the adoption of open-loop MC in bottom fields, which technique is inclined to produce drifting error at low bit-rate and achieve more accurate prediction at high bit-rate. This tradeoff can be flexibly controlled by adjusting the weight parameter γ.

As shown in fig.4, the coding performance gain of normal AR-FGS over normal FGS is obvious; however, as mentioned in section 3, this compression gap between them is achieved at the cost of AR-FGS's high complicated computation. In contrast, our improved AR-FGS can keep similar coding performance as normal AR-FGS, but with medium complexity between those of normal AR-FGS and normal FGS.

5. CONCLUSIONS

In this paper, an efficient FGS coding scheme with low complexity for interlaced SVC in low-delay applications is presented. Compared with the normal AR-FGS in interlaced SVC software, the main advantages of our proposed coding scheme is that it is able to significantly reduce both of the encoding and decoding complexity by saving a series of coding modules including high complicated motion compensation, DCT/IDCT and Q/IQ when coding bottom fields, and keep the similar high coding efficiency as normal AR-FGS. Moreover, the proposed interlaced FGS coding scheme for low-delay applications can further support another temporal scalability by abandoning the bottom fields coded in open-loop MC without influencing the coding efficiency of the top fields, which means better adaptive capacity to fluctuant channel. Therefore, it can be expected that our proposed approach with features of low complexity, enhanced temporal scalability and equally high compression performance is more suitable to offer better quality of interlaced video service for low-delay applications.

6. REFERENCES

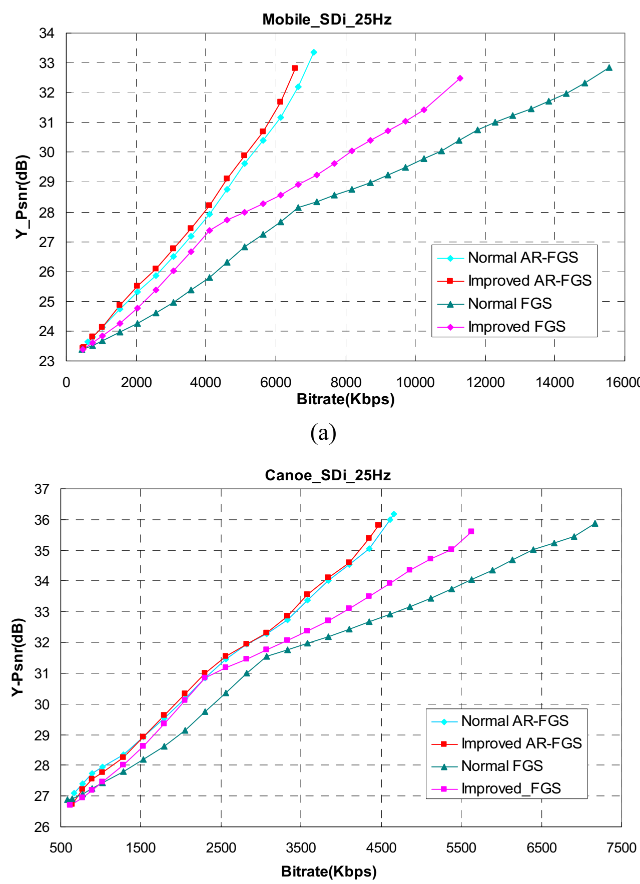
[1] J. Reichel, H. Schwarz and M. Wien, "Joint Scalable Video Model JSVM-5," Joint Video Team, Doc. JVT-R202, Bangkok, Thailand, Jan. 2006

[2] Y. Bao and M. Karczewicz, "CE7 Report: FGS coding for low-delay applications," Joint Video Team, Doc. JVT-O039, Nice, FR, Oct. 2005.

[3] V. Bottreau, J. Vieron, and E. François, "Thomson contribution to the AHG SVC interlaced coding," Joint Video Team, Doc. JVT-R015, Bangkok, Thailand, Jan. 2006.

[4] ITU-Video Coding Experts Group (VCEG), "Core Experiment on Interlaced Video Coding," Doc. VCEG-O59, Pattaya, Thailand, Dec. 2001.

[5] Y. Zheng, X. Ji, and F. Wu, "An Efficient FGS Coding Scheme for Interlaced Scalable Video Coding," Proc. ICIP 2006, Atlanta, Oct. 2006.



(b)

Fig.4. RD curves of improved AR-FGS VS normal AR-FGS.