Decoder Motion Vector Estimation for Spatial Scalable Video Error Concealment

Ruiduo Yang and Michael S.Brown

Department of Computer Science
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
{yangrd,brown}@cs.ust.hk

Abstract

In this paper, we propose to recover the lost enhancement layer information in spatial scalable video based on the information from the current base layer, previous base layer and previous enhancement layer using the Decoder Motion Vector Estimation Method [1]. An average of 1 dB improvement is reported for the enhancement layer. The accelerated method is also demonstrated to make the concealment real time while with insignificant PSNR loss.

1. Introduction

Scalable video has been widely researched in recent years. In scalable video, the base layers will be coded independently and transmitted with higher priorities. Enhancement layers can be coded based current based layer or previous enhancement layer or a linear combination of these two. Enhancement layer can provide video with higher quality but often suffers from a higher loss rate due to network congestion.

Error concealment is one common technique to deal with the package lost situation. Concealment approaches do not need any extra feedback channel and can be done simply on the decoder side, which make this technique more attractive to use. In scalable video coding, since we generally consider that the base layer is transmitted under a error free channel, error concealment will be done only in the enhancement layer provided the information from the current base layer are correctly received.

Previous work on scalable video error concealment consider only about using either using the base layer information (Upward Error Concealment: UEC) [1] or using the previous enhancement layer information (Forward Error Concealment: FEC) [2]. In [2], lost areas are substituted with the motion compensated areas of the previous decoded frame provided the base layers' motion vector. While in [1] the erroneous parts of the picture will be simply concealed with the upsampled base layer, by which the picture quality will be decreased (the sharpness), but with no shift. A switch per-pixel error concealment method (SPEC) has been proposed by Zhang et al. [3] to select a better one among those two utilizing the packet loss history and the base layer quantized residual. Another approach is to use frequency domain error concealment to conceal the error. Zhang et al. [4] proposed an optimal method to do this based on a statistical model for the evolution of transform coefficients from frame to frame. However, the optimal approach suffers from the requirement of large storage of previous frames to be the training data.

In this paper, we tried to recover the lost enhancement layer information in scalable video based on the information from the current base layer, previous base layer and previous enhancement layer using the Decoder Motion Vector Estimation Method [1]. The paper is organized as: In section 2, we review the Decoder Motion Vector Estimation Method and proposed our reformation to the scalable video. Section 4 we proposed fast motion estimation method to cooperate with our concealment. Section 3 covers the proposed algorithm to do error recovery. Simulation Results can be seen in section 4. Section 5 gives the conclusion.

2. Decoder motion vector estimation for scalable video error concealment

In this section, we proposed a method to do error concealment for scalable video. The new method is based on decoder motion vector estimation (DMVE) approach proposed in [1]. In their work, Zhang et al. performs the motion estimation in the decoder that is similar to that done in the encoder. The difference is that they try to do the motion vector search based on the pixels that is surrounding the lost block instead of a rectangular block. The best matched motion vector is used to recover the lost block after the decoder motion vector estimation. The disadvantage is the motion vector search is computational expensive on the decoder and hence cannot be used in a real time application while the advantage is a better motion vector can be found if there is less correlation between the motion vector in the lost block and adjacent blocks.

In our approach, we reform the method of the DMVE for scalable video. Since in scalable video the base layer is generally well protected, we assume the decoder has the correct copy of both previous and current base layer's information. In the case where the previous enhancement layer is correctly received, we could perform our error concealment method for the current enhancement layer. Instead of using the surrounding pixels to do the motion estimation like in [1], the motion search could be done based on the base layers to induce a better motion estimation for the enhancement layer. After this, the obtained motion vector will be projected into enhancement layer to conceal the current lost blocks.

The difference between our method and FEC is the additional search for a better motion vector but with more computation. Meanwhile, the motion vector in the base layer can be used as an initial motion vector, and much smaller block, instead of 16x16 macroblock, will be used to achieve better granularity

We assume the original MPEG sequence is coded using a Macroblock size of 16x16 with a Spatial Scalable Coding Scheme. Our algorithm is depicted below:

1. *For each 16x16 lost MacroBlock, Find the corresponding 8x8 Block m in the base layer, Upsample the 8x8 Block to get a 16x16 Block M.*
2. *Split M into 4 8x8 Block, denoted as Mi (i=1..4).*
3. *For each Mi, search at the previous base layer (upsampled) for a better motion vector Mvi.*
4. *Project Mvi to Enhancement Layer to locate a 8x8 block in the previous enhancement layer. Using this block to conceal the lost block.*

3. Fast method for DMVEEC

As discussed before, though generally not all of the macroblock needs to be concealed, the computational time is still huge to do the decoder motion estimation especially for a higher frame rate which is desired to achieve a better playback quality, hence a faster algorithm is desired.

Fortunately, due to the similarity between the DMVE and the encoder block based motion estimation, we found many faster algorithms could be ported to DMVE from the encoder side. In our approach, we used the Diamond Search (DS) [6] to replace the full search in motion estimation. In particular we used the 5-point small diamond search pattern (SDSP). For the initial search point x, SDSP is applied and the smallest SAD position among the 5 point will be chosen as the next initial position. The process will continue until the best search result is in the central point or the motion vector goes out of the search window. By using this SDSP search strategy, we reduce the computational complexity from quadratic to linear.

At the same time, we notice that the residual block between the base layers can be naturally used as a criterion to make early determination in our algorithm. The idea is to choose FEC directly when the Base Layer's prediction is good enough. In this case, no additional search will be applied for subblock. The new fast DMVEEC is designed as follows:

1. *For each 16x16 lost Macroblock E: If the corresponding base layer residual is small enough, For example, if the residual is smaller than a threshold1, go directly to FEC, else continue.*
2. *For E, find the corresponding 8x8 block m in the base layer, upsample the 8x8 block to get a 16x16 block M, split M into 4 8x8 block, denoted as Mi (i=1..4)*
3. *For each Mi, fast search at the previous base layer (upsampled) for a better motion vector Mvi. If Mvi is not good enough. For example, if the residual between the base layers predicted from Mvi is larger than a threshold2. Go directly to UEC to do the recovery, otherwise continue.*
4. *Project Mvi to enhancement layer to locate an 8x8 block. Using this block to conceal the lost block.*

4. Simulation results

Simulation have been done based on the UBC H.263+ codec[7]. Previous method is used to compare the results, which are UEC, FEC and SPEC. In our simulation, the sequences are coded using a GOP pattern of IPPPP …, the channel of base layer is assumed to be error free, and the current enhancement layer lost its motion vector as well as its residual.

The propagation of errors between frames are also neglected in our simulation, which means we tends to conceal a frame separately provided all the adjacent frames are correctly received. This assumption is based on the fact that error concealment is always used together with other techniques like Frame Retransmission or Intra Block Refresh which can stop the propagation of errors; this can help us focus on the result of error concealment. Although the assumption is not always true, in our experiment we found we can always use surrounding blocks to do the prediction which can be within the same frame or in earlier frames because of the spatial and temporal consistency in the video, thus our algorithm will not get stuck because of the propagation errors.

In UEC, the base layer is interpolated bilinearly to get the enhancement layer.

In FEC, since we assume the motion vectors are lost as well, the enhancement layer will be recovered using base layer's motion vector. SPEC is conducted with the assumption that previous base and enhancement layer are correctly encoded. Substantial gain in PSNR has been found in our experiments especially under the situation of complex motion. Table gives us results for different kind of videos. 5 sequences, foreman, akiyo, news, hall and coastguard are coded in 2 layers using spatial scalable technique, base layers are coded given a bitrate of 25.6, 51.2, 76.8, 102.4 kbps at 10fps, which is QCIF format. Enhancement layers are coded in a CIF format with a constant QP that is 16. Each GOB (Group of Blocks) in the enhancement layer consist 22 marcoblocks which is one row of the frame under the CIF format. One packet loss will result in one whole GOB cannot be decoded correctly, and we assume the position of the error has been detected in the decoder part. In our experiments, we use a threshold of 256 to switch between our method and FEC.

The PSNR is plotted against frame number in Figure 1 and Figure 2 for each method including Original (No Loss), UEC, FEC, SPEC and our method. For fast approach, Table 2 gives us the average value of PSNR in these frames. The overhead of the DMVEEC and the Fast DMVEEC is computed in terms of operation taken on average for each macroblock (see Table 1). By using the fast search method and early determination scheme, we can see the computation has been lowered dramatically while maintaining a good PSNR.

Subjective results are also performed in Figure 3 for comparing FEC, UEC and our approach. From the figure we can see our method protects the resolution from the loss that occur with UEC, while we suffer with less block effect which is found in FEC.

		25.6kbps		51.2kbps		76.8kbps		102.4kbps	
		EC	Fast	EC	Fast	EC	Fast	EC	Fast
Akiyo	PSNR	33.53	33.47	34.09	33.94	34.27	34.09	34.38	34.10
	Overhead	142	0.42	142	0.41	142	0.39	142	0.40
Hall	PSNR	30.65	30.65	31.10	31.09	31.26	31.23	31.31	31.28
	Overhead	142	0.65	142	0.63	142	0.62	142	0.62
Foreman	PSNR	26.49	26.58	27.60	27.71	28.27	28.35	28.63	28.67
	Overhead	142	2.22	142	2.42	142	2.52	142	2.52
News	PSNR	29.60	29.63	30.32	30.29	30.63	30.58	30.83	30.74
	Overhead	142	1.11	142	1.03	142	1.01	142	1.00
Coastguard	PSNR	23.96	24.02	24.64	24.69	24.93	24.98	25.12	25.1
	Overhead	142	1.61	142	2.06	142	2.03	142	2.08

Table1: result of DMVEEC (column EC) and Fast DMVEEC (column Fast). The compare is based on the PSNR(dB) and the overhead in terms of the additional thousand integer operation taken by CPU

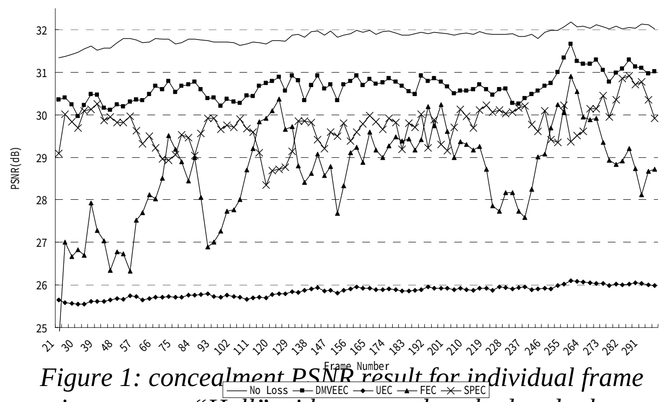

Figure 1: concealment PSNR result for individual frame in sequence "Hall" with proposed method and other methods (base layer bitrate=25.6kbps).

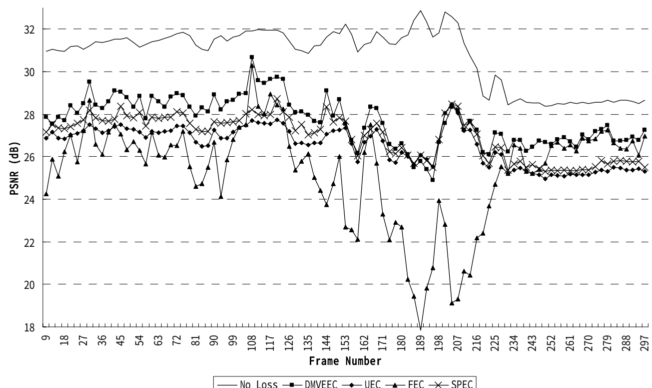

Figure 2: concealment PSNR result for individual frame in sequence "Foreman" with proposed method and other methods (base layer bitrate=51.2kbps).

		No loss	DMVEEC	UEC	FEC	SPEC
Akiyo	25.6kbps	34.38	33.47	31.55	32.86	32.42
	51.2kbps	34.85	33.94	32.18	33.26	32.84
Hall	25.6kbps	31.84	30.65	25.84	28.64	29.73
	51.2kbps	32.05	31.10	26.25	28.87	29.30
Foreman	25.6kbps	30.57	26.58	25.39	24.95	26.00
	51.2kbps	30.73	27.71	26.54	25.31	26.95
News	25.6kbps	31.34	29.63	26.02	27.43	28.62
	51.2kbps	31.49	30.29	26.74	27.69	28.95
coastguard	25.6kbps	28.79	24.02	22.69	23.91	23.27
	51.2kbps	28.81	24.69	23.36	24.27	23.68

Table 2: average PSNR result. Column "DMVEEC" provides the proposed method's result. Average 1dB gain is observed.

5. Conclusion

In this paper, we propose an error concealment method for transmitted video. Our approach is based on a layered video coding scheme, in which the base layer can be guaranteed to provide the basic service. During the loss of current enhancement layer, the previous relationship between base and enhancement layer is learned and used to recover it with the current base layers information. To reduce the computational complexity which is not desired for real time video streaming, fast search and early determination approaches are taken.

6. References

[1] J. Zhang, J.F. Arnold., “Video error concealment using decoder motion vector estimation”, *IEEE Region 10th Annual Conference. Speech and Image Technologies for Computing and Telecommunications, Volume: 2,* 2-4 Dec 1997 Page(s): 777 - 780 vol.2

[2] S.Aign and K.Fazel, “Temporal & spatial error concealment techniques for hierarchical MPEG-2 video codec,” *IEEE International Conference on Communications,* Seattle WA, USA, June 1995.

[3] M.Ghanbari and V.Seferidis, “Cell-loss concealment in ATM video codecs,” *IEEE Transactions on Circuits and Systems for Video Technology, vol.3, no. 3,* pp. 238-247, June 1993.

[4] R. Zhang, S. L. Regunathan and K. Rose, “Switched Error Concealment and Robust Coding Decisions in Scalable Video Coding,” *International Conference on Image Processing,* Vancouver, Canada, Sept. 2000.

[5] R. Zhang, S.L. Regunathan and K. Rose, “Optimal Estimation for Error Concealment in Scalable Video Coding,” *International Conference on Image Processing,* Vancouver, Canada, Sept. 2000.

[6]Wing Cheong Chan, ”Fast global motion estimation and color interpolation.” 2002. M.Phil Thesis, Electrical and Electronic Engineering, Hong Kong University of Science and Technology

[7] H.263+ codec, http://spmg.ece.ubc.ca/

A. Original Result (No Loss)

B. DMVEEC Result

C. UEC Result

D. FEC Result

Figure 3: the Subjective Results. B. is the DMVEEC. C. is the UEC result (suffer from the resolution loss); D. is the FEC Result (Suffer from the artifacts);