A Frame Loss Concealment Solution for Spatial Scalable HEVC using Base Layer Motion

Thuc Nguyen Huu[1], Thuong Nguyen Canh[1], Xiem HoangVan[2], and Byeungwoo Jeon[1]
[1]Department of Electrical and Computer Engineering, Sungkyunkwan University, South Korea
[2]Vietnam National University - University of Engineering and Technology, Vietnam

***Abstract*—Scalable High Efficiency Video Coding (SHVC) is the most recent video coding solution designed mainly for network adaptive or device adaptive applications. It follows a layered coding structure with one base layer (BL) and one or several enhancement layers (ELs) which can be unequally protected. SHVC is often sensitive to packet loss in unreliable networks, especially in case of ELs. In this paper, we propose a novel error concealment method for the SHVC EL under an assumption that the BL is well protected. First, we recover the partitioning and resample motion data from collocated BL frame. Following, we remove outliers of motion field by a motion vector refinement algorithm. Lastly, we conceal loss frame by using motion compensation and deblocking filtering. Experiments conducted with a rich set of test sequences using the spatial-scalable SHVC have shown that our proposed method significantly outperforms the existing error concealment methods, e.g., BL Reconstruction Up-sampling (RU) and BL-SKIP in both subjective and objective quality assessments.**

Keywords—frame loss, error concealment, scalable video coding, spatial SHVC, unequal protection.

I. INTRODUCTION

Error resilience (ER) and error concealment (EC) are important for real time video transmission and storage over unreliable networks and environments [1]. For error resilience, the techniques of Forward Error Correction [2] and Unequal Error Protection [3] have been widely employed to effectively protect the video bitstream. The bitstream is classified into different levels of importance, or so-called layers. The important layer, that is, the base layer, is assigned more redundant parity bits to ensure no data loss during the transmission. This technique can be applied in video coding following the layered structure, especially in the temporal layer of High Efficiency Video Coding (HEVC), or in the Scalable HEVC (SHVC).

TABLE I. A survey on related works

SHVC scalability type	Resolution variation between layers?	Existing EC methods
Temporal (from HEVC)	No	Any EC methods for HEVC, such as [3]
SNR (quality)		RU, BL-SKIP [4], [5]
Color gamut or Bit depth		
Spatial	Yes	RU, BL-SKIP

This research was supported in part by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017R1A2B2006518), and partly supported by VNU University of Engineering and Technology under project number CN18.13.

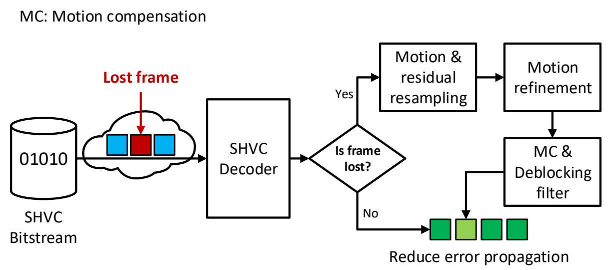

Fig. 1. Proposed BMR-EC framework.

The SHVC standard which was finalized in 2015 follows a layered coding structure with one base layer (BL) and one or several enhancement layers (ELs) [4]. It is noted that while BL is well-protected, the ELs are vulnerable in error-prone environments since they contain less parity bits than BL. Thus, they call for designing an efficient error concealment (EC) method. In the burst loss environments, multiple slices of frame are usually lost together [5], leading to loss of the whole frame. Therefore, we only consider frame loss in this paper.

SHVC straightforwardly inherits the temporal scalability of HEVC, so one can directly employ those EC methods developed for HEVC such as spatial error concealment (SEC) or temporal error concealment (TEC) [5]. For other types of scalabilities (SNR, Color gamut, Bit depth, Spatial), the BL Reconstruction Up-sampling (RU) and BL Motion Up-sampling (BL-SKIP) are the two conventional EC methods [1]. While RU uses the up-scaled version of reconstructed BL frame, the BL-SKIP re-samples the BL motion and then performs the motion compensation. Beside these two conventional EC methods, several researches have investigated the EC methods for the SHVC scalabilities in case of the same resolution among layers (SNR, Color gamut, or Bit depth scalability). For instance, the work [6] has proposed a hybrid method to adaptively select between RU and BL-SKIP candidates in block-based context. More recently, Xiem *et al.* [7] have developed a Joint-layer model between BL and EL, which can be used to create the EC frame. Throughout recent works [6] and [7], their model only works if resolution between layers remain equal. Up to now, far too little attention has been paid to solve the EC problem on different resolutions of layers, particularly on Spatial Scalability. The recent works are summarized in TABLE I.

In this paper, we propose an EL frame loss concealment method for the Spatial SHVC under the assumption of well-protected BL. Because the proposed method exploits BL motion and residual energy, for convenience, we name our proposed method as Base layer Motion and Residual based Error Concealment (**BMR-EC).**

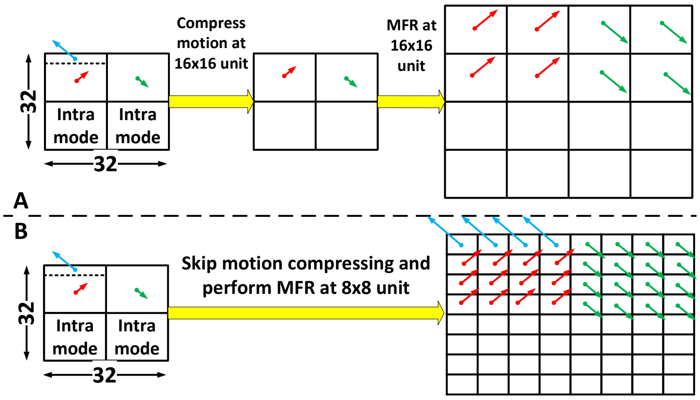

Fig. 2. Motion field resampling comparison between:
(A). SHVC MFR.
(B). **Proposed MMFR.**

Here the resolution ratio $\boldsymbol{\alpha = 2}$. Some output blocks do not have motion because their collocated blocks are coded by Intra mode. The collocated position is determine by eq. (2).

The rest of the paper is organized as follows. Section II presents the BMR-EC method including motion & residual resampling as well as motion refinement. Section III gives the experimental results. Lastly, we conclude the paper in section IV.

II. Proposed Error Concealment Method

Our key idea in this paper is to recover motion of current lost frame in EL by resampling the BL collocated motion, and then to perform motion compensation to retrieve the frame loss. To further enhance the EC frame quality, the resampled BL residual energy is used as an indicator to determine the motion reliability. If motion is not reliable, a refinement process is applied. Fig. 1 illustrates our BMR-EC framework which is implemented at the decoder side. The proposed techniques will be presented in the following sub-sections.

A. *Motion resampling*

SHVC offers a motion field resampling to create motion vector prediction for EL. At first, we are going to discuss how SHVC deals with motion resampling and its disadvantages in EC. After that, we will address our proposed motion resampling process to overcome those issues.

1) *Motion field resampling in SHVC standard*

SHVC employs the *motion field resampling* (MFR) technique to map the motion information in EL to BL. If all layers have the same resolution, EL motion can be inherited directly from the collocated BL sample position. However, in *Spatial Scalability* which has different spatial resolution over layers, one needs to identify new *collocated position* as well as the *motion amplitude difference*.

Let's denote α be the resolution ratio between EL and BL, then one can compute EL motion at position $(x,\ y)$ as follows:

$$mv^{EL}_{(x,y)} = \alpha \times mv^{BL}_{pos\_mapping(x,y)} \tag{1}$$

where mv^{L}_{p} denotes a motion vector at position $p \in R^2$ at layer L (here L can be BL or EL); and $pos\_mapping(x,y)$ is a function, $R^2 \rightarrow R^2$, which maps EL position (x,y) to BL collocated position (u,v). The $pos\_mapping$ function can be determined by:

$$\begin{aligned} u &= ((x - offsetX_{EL})/\alpha - ((phaseX/\alpha + 8)/16 + 2^{11}) \gg 12 + offsetX_{BL} \\ v &= ((y - offsetY_{EL})/\alpha - ((phaseY/\alpha + 8)/16 + 2^{11}) \gg 12 + offsetY_{BL} \end{aligned} \tag{2}$$

Where $phaseX$, $phaseY$ are the signaled horizontal and vertical resampling phases, respectively; and $offsetX$, $offsetY$ are the signaled left, top offset, respectively; and $\gg$ indicates the right bit shifting operator [4]. Those parameters are all related to down-scaling process between EL and BL.

By using (1), one can compute motion at every pixel in EL. However, SHVC executes MFR in a unit of 16×16 blocks. This decision makes sense because of following two reasons. Firstly, SHVC is a scalable extension of HEVC which does block-based motion coding. Secondly, due to memory restriction, once a picture is decoded, SHVC/HEVC compresses motion information into units of 16x16 blocks (by computing motion of central sample), thus making sense to perform MFR in block-based context. Fig. 2(A) demonstrates the MFR mechanism in SHVC. The figure shows that MFR does not take input from exact BL motion but from the compressed version. In short, the motion output from SHVC MFR is processed in two phases: (i) motion compressing; and (ii) motion mapping and resampling. This observation strongly motivates us to design a new technique for MFR which is more suitable in EC.

MMFR algorithm

1	$F_{EL} \leftarrow$ the current EL lost frame
2	$F_{BL} \leftarrow$ the collocated BL frame
3	$\alpha =$ resolution(EL)/resolution(BL)
4	Skip motion compressing in F_{BL}
5	For each 8×8 block in F_{EL}:
	$(x,y) \leftarrow$ central position of current block
	$(u,v) = pos\_mapping(x,y)$
	if (sample (u,v) in F_{BL} is coded by Inter mode):
	$cur\_mv = \alpha \times mv^{F_{BL}}_{(u,v)}$
	else: //*Intra mode*
	$cur\_mv \leftarrow None$
	Set motion of current block to cur_mv

2) *Proposed motion field resampling for EC*

As discussed above, motions resulting from SHVC MFR are likely to be distorted due to the two-pass process. We visualize the distortion in Fig. 2(A). After motion compression, the motions in blue color is dominated by other motions (in red and green color) and eliminated after MFR process.

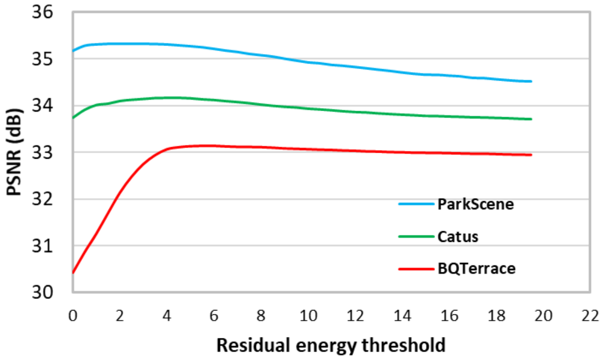

Fig. 3. EC quality with respect to residual energy (frame number 5 of ParkScene, Cactus, BQTerrace sequences).

SHVC MFR was actually designed to generate motion vector predictions (MVPs) at ELs. In fact, when MVPs are not correct, the signaled motion vector differences (MVD) is ready to compensate that error [4]. Unfortunately, in a frame loss scheme, we do not have any chance to correct the error since every data is completely lost. While MFR keeps the balance between memory restriction and motion accuracy, it turns out that MFR is not suitable in frame loss scheme where we want to achieve best motion accuracy as possible.

To address this problem, we propose a Modified version of MFR which is toward the Error Concealment scheme (**MMFR**). The algorithmic detail is presented above. Here, we eliminate the overlapping problem by skipping motion compressing process and increase the motion sampling rate. That is, when EL picture is detected as lost, the collocated BL postpones the motion compressing until MFR is completely finished. Moreover, the overlapping problem still persists if we perform MFR at a large block size which is 16×16 block unit in SHVC. Therefore, we increase the motion sampling rate to 8×8 block unit to provide denser and more accurate motion results. In summary, Fig. 2(B) demonstrates the difference between the original MFR and our proposed MFR-EC. From this figure, one can observe that the blue motions are still preserved in our MMFR whereas they are not seen in the SHVC MFR.

B. *Residual resampling*

Apart from MFR, the residual resampling is very straightforward to understand. At this point, the residual of BL collocated picture is up-sampled by a resolution ratio α to match the EL picture size. One of well-known interpolation methods, such as bilinear, bicubic, or Lanczos can be used without significantly affecting the final result. In this paper, we use the bilinear interpolation method for simplification.

C. *Motion refinement*

Until now, we have finished recovering the motion parameters for EL frame loss using the proposed MMFR. At the first thought, the Motion Compensation can be applied to retrieve EC frame. However, that approach has some serious problems. On the one hand, MMFR cannot resample Intra Coding Block, which indicates that this approach is not a complete solution. On the other hands, even MMFR resamples correct motion parameters from BL collocated frame, we are not totally sure whether motion parameters describe object movement perfectly. If not correct, it might lead to unexpected artifacts in the motion-compensated frame.

To solve those problems, we propose *a motion refinement algorithm* which works for each 8×8 block and employs the reliability degree of motion information. This algorithm can be described as follows. We compute residual energy for each 8x8 block by calculating the average of absolute values in corresponding residual block. For each 8x8 block, its motion is marked as unreliable if one of the following conditions occurs: (1) this motion is not available due to Intra mode, or (2) the residual energy is larger than a certain threshold. If motion is marked as unreliable, we replace it by zero motion with respect to up-scaled BL reference index.

At the final step, motion compensation process is applied to retrieve to EL frame loss. Furthermore, as the block basis is the key element in video coding, the blocking artifact naturally occurs even with the correct motion information; hence, we apply a de-blocking filter for the final EC frame.

III. Experimental Results

A. *Test conditions*

To evaluate the performance of the proposed BMR-EC method, we conducted an extensive experiment using five common test sequences suggested in [8]. For generating BL input, we use the built-in downscaling software included in reference software SHM 12.3 [9]. According to eq. (2), we specify parameters related to downscaling process, like PhaseX, PhaseY, OffsetX, and OffsetY to all zero. The resolution ratio here is set to 2.0. Additionally, the spatial SHVC with Random Access configuration is examined in this assessment, and the packet loss rate of 5% is also considered to reflect the network transmission issue. Two well-known existing EC solutions, namely, BL Reconstruction Up-sampling (RU) and BL-SKIP [1], are used as benchmarks. For fair comparison, we also apply MMFR for the BL SKIP method. Furthermore, we also include the “No loss” case as an upper-bound for EC.

TABLE II. Summary of test conditions

Software	SHM 12.3 [9]
Scalability	Spatial scalability 2.0×
Coding scheme configuration	Random Access, GOP size = 16, Intra period = 32
Sequence, EL resolution, frame rate	BQTerrace, 1920×1080, 60Hz BasketballDrive, 1920×1080, 50Hz Cactus, 1920×1080, 50Hz Kimono1, 1920×1080, 24Hz ParkScene, 1920×1080, 24Hz
Down-scaling filter parameters	PhaseX = PhaseY = 0 OffsetX = OffsetY = 0 Resolution ratio $\alpha = 2$
Packet loss rate	5%

B. *Results and discussion*

In this section, we show the subjective quality assessment accounted for lost frames only and objective quality measurement in PSNR (dB) in Table II in comparison with various methods. In the objective quality, it is easy to observe that our proposed BMR-EC method significantly outperforms both the RU and BL-SKIP based EC solutions, notably with nearly 2dB and 14.5dB higher, respectively on average.

No loss RU BL-SKIP **BMR-EC**

Fig. 4. Subjective quality comparison of various concealment methods applied to the sequence Cactus.

Especially, we achieved up to 3.3 dB gain comparing to the RU method in sequence BQTerrace.

The smallest gain comes in with Kimono1 as expected, since this sequence comprises of a lot of low frequency areas which help up-scaling behavior in RU method. The objective performance gain is consistent from low rate to high rate (that is, small QP to large QP values). Especially, our results are close to the upper-bound case of "No loss", proving the effectiveness of the proposed method.

TABLE III. PSNR [dB] comparison of EC methods

Sequences	**Method**	**Quantization parameters (BL/EL)**			
		26/26	**30/30**	**34/34**	**38/38**
BQTerrace	No loss	35.83	34.55	33.18	31.64
	BL-SKIP	18.27	19.03	19.64	21.35
	RU	30.45	29.83	29.01	27.97
	BMR-EC	**33.75**	**33.05**	**31.98**	**30.74**
Basketball Drive	No loss	37.77	36.28	34.66	32.97
	BL-SKIP	14.52	14.83	14.47	15.40
	RU	32.58	31.77	30.74	29.58
	BMR-EC	**34.46**	**33.69**	**32.45**	**31.06**
Cactus	No loss	36.97	35.48	33.74	31.90
	BL-SKIP	16.93	17.30	17.28	18.24
	RU	32.64	31.65	30.39	28.99
	BMR-EC	**34.94**	**33.84**	**32.32**	**30.69**
Kimono1	No loss	39.92	37.92	35.79	33.74
	BL-SKIP	19.96	21.75	20.10	21.46
	RU	37.27	35.19	33.09	31.17
	BMR-EC	**38.14**	**36.15**	**34.14**	**32.25**
ParkScene	No loss	37.58	35.41	33.30	31.38
	BL-SKIP	20.55	21.21	21.44	21.71
	RU	33.14	31.85	30.39	28.88
	BMR-EC	**35.51**	**33.90**	**32.11**	**30.40**

Surprisingly, quality of BL-SKIP is seen to be decreasing along with QP values, which reflects the opposite trend with other methods. However, we can still find the reason since there are more intra coding blocks with a low QP value, compared to the higher QP case. Because BL-SKIP cannot resample intra coding blocks, going from low QP to high QP makes BL-SKIP quality even worse.

The relation between the threshold of residual energy and EC frame quality is shown in Fig. 3 which shows that the EC frame quality is seen to increase along with the threshold, but it will start decreasing beyond a certain point. In this paper, we fix the residual energy threshold at 2.0. Still, a study on choosing optimal threshold is necessary in our future work.

In Fig. 4, the proposed method shows visually more pleasing result compared to relevant methods. The RU method typically blurs the whole picture due to up-scaling, while the BL-SKIP method creates artifacts at the bottom part. The artifacts can be explained by the fact that: first, BL-SKIP cannot resample Intra Coding Block, which makes the EC frame has some green holes; second, some motions resampled by BL-SKIP is not refined, leading to serious blocking problem. Both RU and BL-SKIP methods can degrade subjective quality in the spatial scalable SHVC. In contrast, our proposed BMR-EC can still preserve fine details for the whole picture.

IV. Conclusion

In this paper, we proposed a novel BMR-EC method for spatial scalable HEVC. Throughout the paper, we have introduced the new MMFR method and Motion refinement algorithm to enhance the EC frame quality. Our experimental results have shown superiority of the proposed method compared with other state-of-art methods. Our future work could focus on studying the optimal threshold in Motion refinement algorithm.

References

[1] Chen, et. al, "Frame loss error concealment for SVC," in *Journal of Zhejiang University-Science*, vol.7, no. 5, pp. 677-683, 2006.

[2] Yao Wang, S. Wenger, Jiantao Wen, and A. K. Katsaggelos, "Error resilient video coding techniques," in *IEEE Signal Processing Magazine*, vol. 17, no. 4, pp. 61-82, 2000.

[3] E. Maani and A. K. Katsaggelos, "Unequal Error Protection for Robust Streaming of Scalable Video Over Packet Lossy Networks," in *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no.3, pp. 407-416, 2010.

[4] J. M. Boyce et. al, "Overview of SHVC: Scalable Extensions of the High Efficiency Video Coding Standard," in *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 26, no. 1, pp. 20-34, 2016.

[5] Liu C., Ma R., and Zhang Z. "Error Concealment for Whole Frame Loss in HEVC," in *Advances on Digital Television and Wireless Multimedia Communications. Communications in Computer and Information Science*, vol 331. 2012.

[6] T. N. Huu, et. al, "Base layer constrained error concealment solutions for robust SHVC video transmission," in *Proc. 2018 Int. Workshop on Advanced Image Technology (IWAIT)*, Chiang Mai, pp. 1-4, 2018.

[7] X. HoangVan and B. Jeon, "Joint Layer Prediction for Improving SHVC Compression Performance and Error Concealment," in *IEEE Trans. on Broadcasting, 2018* (in press).

[8] Frank Bossen, "Common test conditions and software reference configurations.", in Joint Collaborative Team on Video Coding (JCT-VC) of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, 5th meeting, 2011.

[9] SHVC reference software, https://hevc.hhi.fraunhofer.de/shvc.