Multimed Tools Appl (2017) 76:25511–25527
DOI 10.1007/s11042-017-4835-2

Robust real-time UHD video streaming system using scalable high efficiency video coding

Eun-Seok Ryu[1] · **SunJung Ryu**[2]

Received: 18 August 2016 / Revised: 26 March 2017 / Accepted: 17 May 2017 /

Abstract With a new video coding standard high efficiency video coding (HEVC), the ultra high definition (UHD) TV service with robust video streaming technology is emerging in the TV industry. This paper addresses the system architecture for the UHD video streaming and proposes three main ideas: (i) picture prioritization method, (ii) error concealment mode signaling (ECMS), and (iii) Tile complexity-based parallel video processing. In the experiments using HEVC reference model conducted, the proposed picture prioritization method shows the gains in video quality from 2.2 to 7.5 dB in Y-PSNR, and the error concealment mode signaling gains from 0.2 to 2.5 dB in Y-PSNR, with corresponding subjective improvements. In addition, proposed parallel processing method for real-time decoding shows around 20% decoding speed up gain.

Keywords Robust video streaming · Picture priority · Error concealment · Scalable high efficiency video coding

---

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A02037743).

✉ SunJung Ryu
ryuhappy@dankook.ac.kr

Eun-Seok Ryu
esryu@gachon.ac.kr

[1] Department of Computer Engineering, Gachon University, 1342, Seongnam-daero, Sujeong-gu, Seongnam-si, Gyeonggi-do, 13120, Republic of Korea

[2] Department of French, Dankook University, 29 Anseodong, Dongnamgu, Cheonan 330-714, Republic of Korea

1 Introduction

Among the emerging technologies in the TV industry, ultra high definition TV with robust video streaming is one of the most emerging technologies. Korea's cable TV companies opened the world's first UHD (4K resolution, 3840 × 2160) channel in April 2014, and Japan started the UHD service around 11 months later. Korea plans to activate UHD market with PyeongChang 2018 Winter Olympics, and Japan also has a plan to open 8K (7680 × 4320) UHD TV service in time for the Tokyo 2020 Olympic games. In addition, France broadcasted the EURO 2016 with UHD resolution, and *Technicolor* and *ATEME* developed the real-time SHVC encoder and decoder, and conducted world's first field test for UHD-HD switching technology with *ETRI* in Korea. The UHD video (4K and 8K) has the resolution of 4 times and 16 times larger than FHD (Full High Definition) Video. To support the UHD resolution, JCT-VC (Joint Collaborative Team on Video Coding) has developed the HEVC (High Efficiency Video Coding) standard with better video compression technologies in January 2013. However, the UHD resolution makes some technical difficulties in the area of real-time video streaming as well. For example, the bitrate of UHD bitstream is much higher than previous video standards such as H.264 AVC, and the error resilient (robust) video streaming over error-prone wireless network is very difficult. Besides, due to the resolution of the UHD, the real-time picture decoding is difficult as well. Thus, this paper introduces four main technologies for robust real-time UHD video streaming.

First, to provide the UHD video service over error prone networks, robust video streaming technologies including video packet error protection and EC are essential. However, the current video coding standards, HEVC and scalable HEVC (SHVC), are only focusing on the video compression without careful consideration of video transmission. Besides, the MPEG-H part 1 system standard, MPEG media transport (MMT), that is considering the transmission issue also does not have any syntax and semantics for the picture priority in the same temporal level of hierarchical-B-structure and error concealment (EC) at all. Regarding the EC technology, it is very difficult to find the best EC mode among multiple EC methods provided by decoder without original pictures. This is the limitation of the EC method that only works at a video decoder side.

By extending our previous researches [10, 12–16], this paper proposes two methods for the robust video streaming over error-prone networks; (i) a new picture prioritization method in the hierarchical B structure of HEVC, and (ii) a new EC mode signaling method that signals best EC mode(s) which is calculated and determined at an encoder side to a decoder. Second, regarding the real-time of UHD video decoding, this paper proposes one more method; (iii) HEVC Tile-based video parallel processing using asymmetric CPU multicores, which reduces a decoding delay as well as a transmission delay. Table 1 shows difficulties for robust real-time UHD video streaming and three proposed methods. This

Table 1 Difficulties from UHD resolution and three proposed methods

Difficulties and goal	Error resilience / robust UHD video streaming	Real-time decoding
Proposed methods	(1) Picture prioritization method and unequal error protection (2) Error concealment (EC) mode signaling method that signals best EC mode(s)	(3) Tile-based parallel processing using asymmetric CPU multicores

paper provides several robust video streaming technologies with proposed two methods ((i) picture prioritization and (ii) error concealment mode signaling). Thus, it explains the expecting performance gains of each method for fair comparison.

In the general architecture of the video streaming system, the video server consists of multiple modules such as video encoder, error protection, selective scheduler, and quality of service (QoS) controller for streaming. The video client includes an EC module. From a network point of view, the video packet could be transmitted over error-prone network. Thus, the transmission has to consider the packet loss condition that could happen in wireless connection by signal interference or by dropping packets for congestion control. The network may use the methods such as automatic repeat request (ARQ) and forward error correction (FEC) to recover the packets from the network error, but extra transmission delay and jitter may occur unpredictably. Because of the undesirable delay and jitter, the cross-layer optimization is avoiding to use the retransmission (e.g. ARQ) and error protection (e.g. FEC) in the link and physical layers, instead, the technologies such as video content-aware error protection (e.g. unequal error protection (UEP)) and EC methods are preferred in the application layer. Consequently, the video server and client need to provide error resilient streaming methods and EC methods along with flow control and congestion control technologies [7]. In Fig. 1, the server and client exchange control messages (signal) to control the QoS metrics, and the signaling effort may enhance the overall video quality significantly.

The paper is organized as follows: Section 2 provides several related work including HEVC and SHVC standards, selective video streaming with picture priorities, FEC, general EC methods, and algorithms for estimating picture complexity. Section 3 proposes the UEP-based robust video streaming and HEVC Tile allocation algorithm for real-time decoding using asymmetric multicores. Lastly, Section 4 summarizes our conclusions and outlines future work.

2 Related work

This section details the background information necessary to understanding the UEP-based video streaming over error-prone networks and the proposed Tile allocation algorithm for asymmetric multicores.

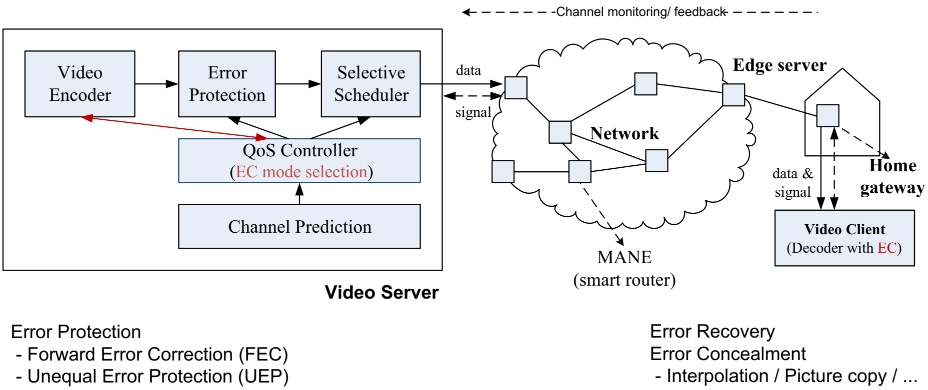

Fig. 1 General architecture of video streaming system

2.1 JCT-VC HEVC video coding standard with parallel processing tools

After successfully standardizing H.264/AVC (Advanced Video Coding) [18], ISO/IEC MPEG and ITU-T VCEG have been jointly developing next generation video standard called HEVC. This new standard targets next-generation HDTV displays and IPTV services, addressing the concern of error resilient streaming in HEVC-based IPTV. Comparing to H.264/AVC, the HEVC includes new features such as extended prediction block sizes (up to 64 × 64), large transform block sizes (up to 32 × 32), sample adaptive offset (SAO), and so on [19].

H.264 scalable video coding (SVC) is an H.264 AVC scalable extension that combines spatial, temporal, and quality scalabilities simultaneously. That is, the SVC can support multiple resolutions, frame rates, and video qualities within a single bitstream because the bitstream consists of multiple layers as shown in Fig. 2.

In the layered structure of SVC encoder. The original high quality video input is spatially down-sampled for multiple layers, and each layer encodes the input video with interlayer prediction. Because of the layered feature, the SVC has several advantages. First, by supporting many clients with a single video content file (bitstream) the SVC enables video service providers to reduce overall network bandwidth (BW) [9], disk storage for video contents, and computational complexity for transcoding. Second, the SVC is applicable to many UEP methods using the priorities of each layer [11]. For example, a base layer (BL) can be provided at a level of error protection higher than that of the other enhancement layers (ELs) because the decoder cannot reconstruct a video sequence without the BL [6, 17], suggesting a higher priority for it. Third, SVC can support the diverse screen sizes and resolutions of user devices, and also diverse network BWs.

The SHVC standard is designed to have low complexity for ELs by adding the reconstructed BL picture to the reference picture lists in EL [20, 21]. In addition, SHVC uses multiple loops decoding to make a decoder chipset simple while the SVC uses single loop decoding. SHVC also provides a standard scalability by supporting AVC with BL and HEVC with EL. Thus, UHD TV services that supporting legacy HD TVs as well as simple bitstream level rate control (layer switching) need the SHVC.

2.2 Picture prioritization and error concealment methods

In video compression and transmission, picture prioritization is of utmost importance for the role it plays in UEP, picture dropping for bandwidth adaptation, as well as quantization

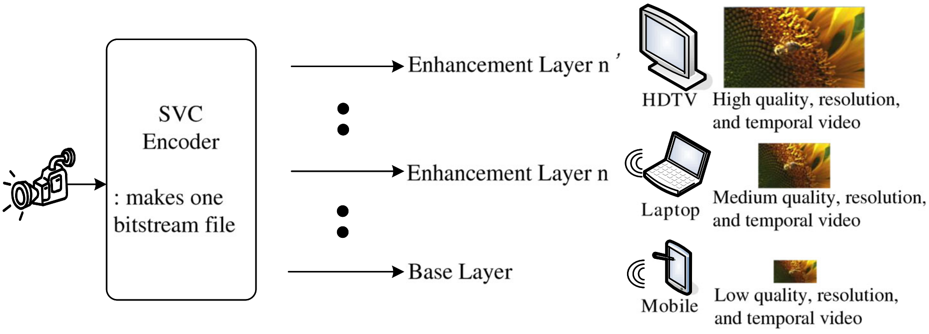

Fig. 2 The layered structure of SVC and SHVC technology

parameter (QP) control for enhanced video quality, to name a few. There have been many studies to prioritize individual video pictures and slices with precision and reliability. Layer information of video packets is widely used. For example, in the encoded bitstream of H.264 SVC, the reconstruction pictures of BL is used to decode the pictures of the ELs, and the video packet of BL must be processed with the highest priority and transmitted with greater reliability or lower packet loss rates. Otherwise, losing a single BL packet could result in severe error propagation in both layers.

Figure 3 shows four different methods of picture prioritization based on picture characteristics. (a) Use picture type information which is related to temporal reference dependency for picture prioritization; (b) Use temporal level information in hierarchical B structure, and higher layer will not be referenced by lower layer; (c) Use location information of slice groups (SGs) (SG-level prioritization); and (d) Use the layer information of the SVC/SHVC. In most cases, I-pictures, pictures in low temporal level, slice group of region of interest (ROI), and pictures in BL of the SVC/SHVC have higher priority than the others. (e) Regarding the ROI, the flexible macroblock ordering (FMO) method in H.264 or the Tile method in HEVC/SHVC could be used.

Picture prioritization can also be used for QoS handling in video streaming: (i) dropping less important pictures in the transmitter or scheduler of the server for bandwidth adaptation; (ii) allocating more important pictures to more stable channels (or antennas) in multi-channel networks or MIMO; (iii) protecting more important pictures with larger

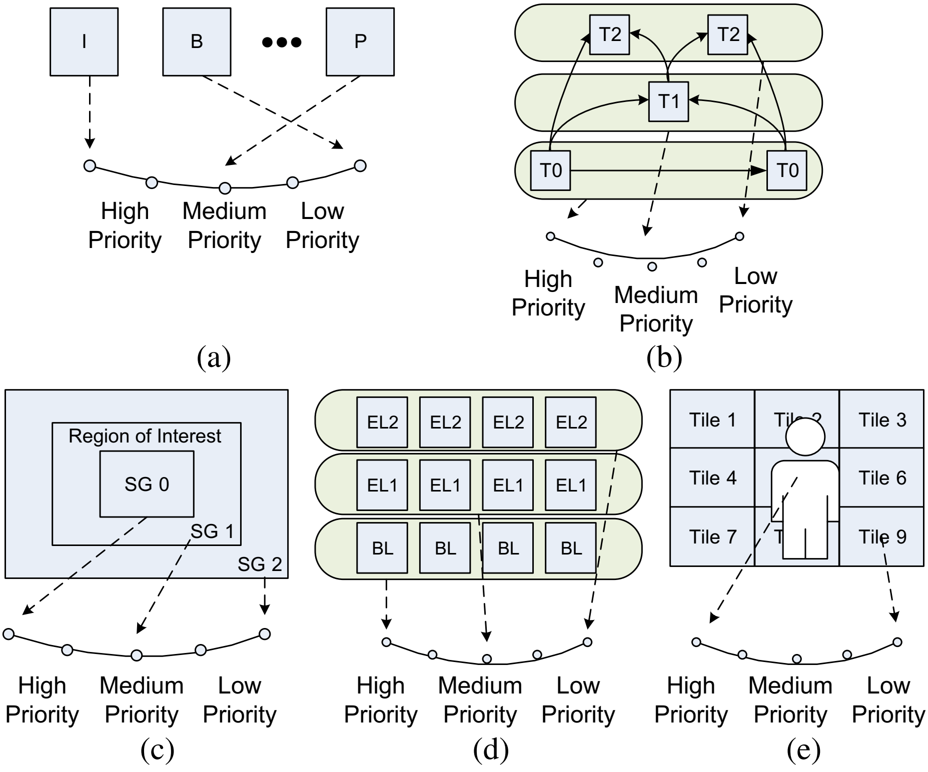

Fig. 3 Examples of picture prioritization methods

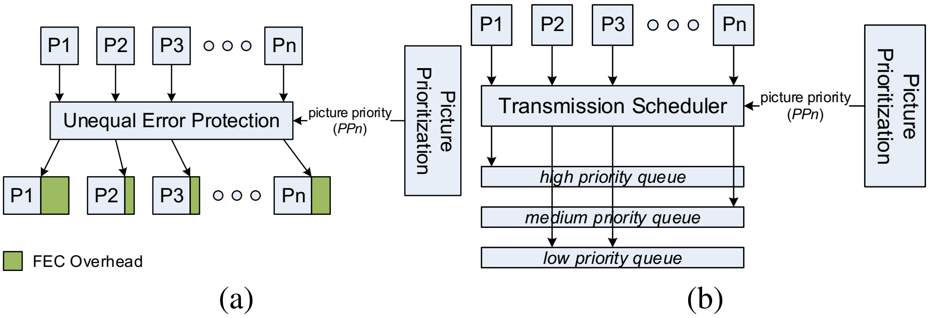

Fig. 4 Two use cases of the picture prioritization method: **a** UEP and **b** transmission scheduler

overhead of FEC code in application or physical layers; (iv) scheduling more important pictures first in application or MAC layers; and (v) differentiating services in the media aware network element (MANE), edge server, or home gateway.

Among them, Fig. 4 shows two use cases in detail. Once the encoder has decided the priority of a picture, the UEP and/or transmission scheduler can use the priority in both robust streaming and QoS handling. Figure 4a applies different FEC overheads to pictures according to picture priority (PP_n), and Fig. 4b allocates pictures to different prioritized queues according to picture priority, the high priority queue has higher throughput. Therefore, picture priority is essential for optimal QoS handling in video streaming and communication applications. Other standardization working groups such as MMT and IETF H.264 over RTP consider picture priority at the system level, which can enhance video server (scheduler) and MANE (smart router) for QoS improvement by differentiating among packets with various priorities when congestion occurs in networks.

The layering feature of SHVC could be very useful for error resilient video transmission. For example, when a picture in EL is damaged during transmission, a decoder could use the picture in BL to make up the lost EL picture. In addition, the upsampling method using lower layer pictures and motion compensation method using same layer pictures could be applied for EC as well. Here, the upsampled lower layer picture is prepared at inter-layer picture (ILP) buffer in decoding step, and this picture level upsampling and picture copy methods are easy to implement and have low computational complexity. On the other hand, there could be more sophisticated coding unit (CU) level error concealment methods because the lost or damaged video packets would make partial picture errors. In the case, the other EC methods could use motion vector (MV) and CU level motion compensation and copying. Table 2 shows general EC methods for SVC.SHVC as [5].

2.3 Parallel processing and picture complexity prediction algorithms

The HEVC parallel processing tools support different picture partition strategies such as Tiles and wavefront parallel processing (WPP) as shown in Fig. 5. Tiles partitions a picture

Table 2 General EC methods for SHVC

Picture Copy (PC)
Temporal Direct (TD)
Motion Copy (MC)
Base Layer Skip (BLSkip; Motion & Residual upsampling)
Reconstructed BL upsampling (RU)

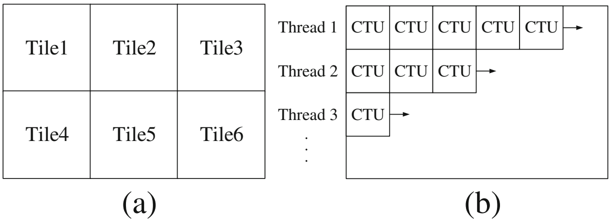

Fig. 5 Two parallel processing tools of HEVC; **a** example of a frame divided into Tiles, **b** example of WPP

with horizontal and vertical boundaries so that it provides better coding gains compared to multiple slices. WPP is used when a slice is divided into rows of coding tree units (CTUs). In WPP, the first row of CTUs is decoded normally, but each additional row requires decisions from the previous row. WPP has the entropy encoder that uses information from the preceding row of CTUs, and allows the parallel processing that has a better compression than Tiles.

Though the Tile can segment each picture into multiple rectangular regions and allocate them to CPU cores, it does not consider the computational ability of asymmetric CPU cores and just allocates video Tiles to CPU cores equally, which causes video processing (encoding/decoding) delays. These days, many CPU manufacturing companies release the device with asymmetric multicores because the asymmetric multicores have features that saving energy and the performance similar to symmetric cores as shown in Fig. 6a. For example, recent ARM processors incorporate the *big.LITTLE* architecture that uses asymmetric multicores. big.LITTLE allocates light-load threads into little cores (slower core) and allocates heavy-load threads into big cores (faster core) for improved energy efficiency and better performance [22].

Decoding complexity of video frames is influenced by many explicit factors including encoding parameters such as resolution, QP, and objects in picture. Recent researches on predicting encoding/decoding complexity are centered on optimizing power efficiency and encoding/decoding time.

One of those researches aims at adjusting clock speed of CPU to save power [1]: it sets low clock speed when a frame has less computational complexity and high clock speed for high computational complexity frames. This research is meaningful in improving power efficiency but it does not consider parallelism on multicore systems and optimization of decoding time.

The other research proposed a Tile partitioning algorithm based on the number of bits of CTUs [2]. The logic proposes a method to equalize the total number of bits in each Tiles in order to minimize the decoding time between Tiles that have a lot of bits or a few bits. It has many similarities with our research, but it uses a different method to predict complexity and does not consider asymmetric multicore environments.

3 Robust real-time UHD video streaming systems using SHVC

3.1 Proposed method 1: picture prioritization and unequal error protection

This section explains the picture prioritization and UEP method with our previous researches [10, 13, 15] including the JCT-VC standard meeting proposals. Figure 7 denotes

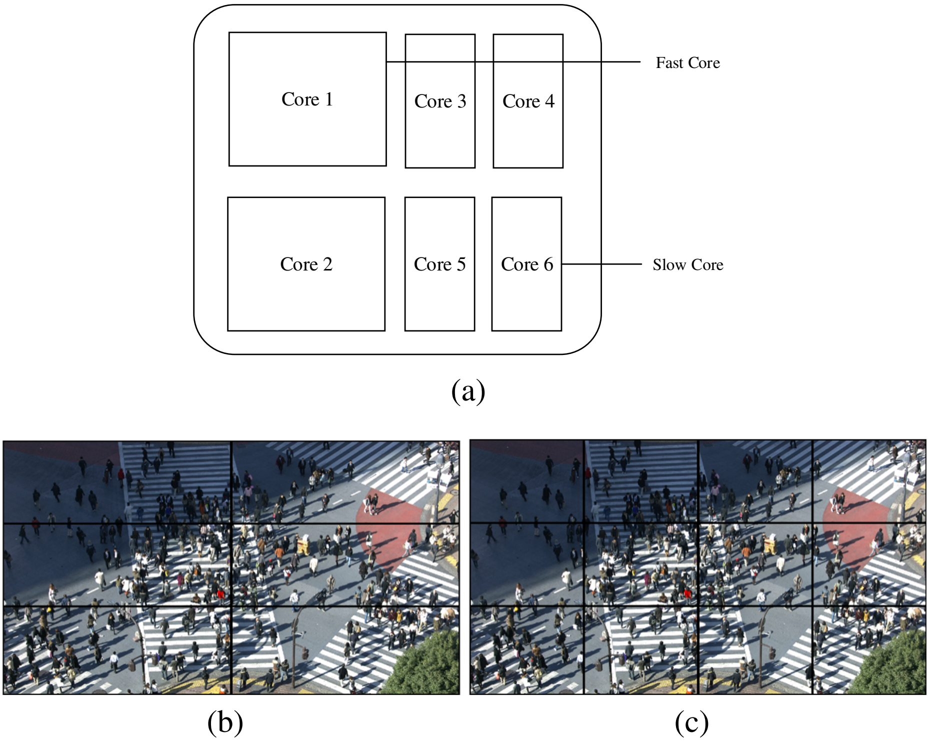

Fig. 6 Asymmetric multicores and HEVC Tiles; **a** asymmetric CPU cores, test sequence *PeopleOnStreet* with **b** 6 Tiles and **c** 12 Tiles

the current uniform prioritization method (applies same priority to pictures in the same temporal level of hierarchical B structure) with four dyadic stages in temporal domain. Although the temporal levels could tell the priority of a picture, the current HEVC standard provides no additional methods for assigning priorities to pictures at the same temporal level. Figure 7 shows the random access (RA) setting in the common test condition of HEVC, with picture order count (POC) 2 and 6 having a same priority. However, the uniform prioritization at the same temporal level presents a problem when the importance of two (or more) POCs in each group of picture (GOP) varies according to both the reference picture

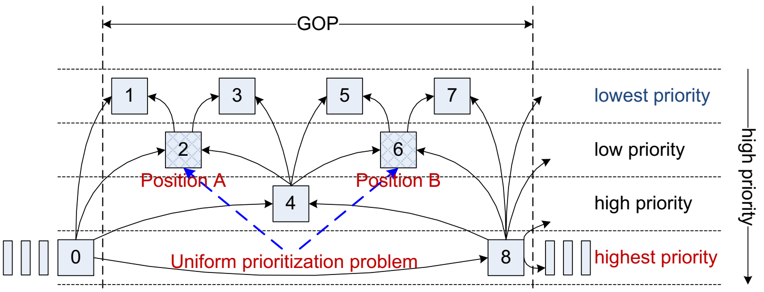

Fig. 7 Current uniform prioritization in hierarchical B pictures

set (RPS) and the size of the reference picture lists (RPL). In order to illustrate the problem, this paper uses two pictures at the same temporal level as an example, and defines *Pos.A* as pictures with a POC equal to $2 + N \times$ GOP, and *Pos.B* as pictures with a POC equal to $6 + N \times$ GOP, where GOP is 8 and N represents the number of GOP(s).

Figure 8 shows corresponding rate-distortion (RD) curves, indicating that the picture in *Pos.B* is more important than the picture in *Pos.A*; red curve is from original HEVC reference software (HM) 6.1 EC, and black curve is from modified HM 6.1 EC. The average BD-rate differences between the blue and black curves were 23.4% (*Kimono*) and 20.2% (*ParkScene*) when test sequences were encoded with the same TID (all TID is 0); error propagation effect. When TIDs from 0 to 3 were used according to their temporal levels, the average differences were 9.9% (*Kimono*) and 10.2% (*ParkScene*) respectively. The PSNR degradation caused by dropping a picture per intraperiod (=32 in this example) in *Pos.A* was less than the PSNR degradation caused by dropping pictures in *Pos.B* indicating that pictures even in the same temporal level in hierarchical B pictures should have different priorities in accordance with their prediction information.

To solve the uniform picture prioritization problem, this section proposes an implicit picture prioritization method that the encoder assigns priorities to pictures according to the RPS and the size of RPL of the encoding option without any additional delays. If a POC number is observed more often in the RPL, then the corresponding picture will earn a higher priority; this is because the number of observations implies the opportunity of being referenced in motion estimation. In case of a POC consists of multiple slices, the priority of the POC is assigned to those slices. The proposed picture prioritization method was combined with the implemented FEC code, Raptor codes, to show its UEP usage and performance gain. Each picture was encoded with a NAL packet and was protected with selected FEC redundancies. For example, when combined UEP, the proposed picture prioritization method protected pictures in *Pos.A* with 28% FEC redundancies (medium-low priority), and protected pictures in *Pos.B* with 32% FEC redundancies (medium-high priority), In contrast, when uniform UEP is used, pictures at *Pos.A* and at *Pos.B* were both protected with 30% FEC redundancies (medium priority). The other redundancies were as these: highest = 44%, high = 37%, low = 24%. Because the hierarchical B pictures with GOP 8 has 4 temporal levels, pictures in the lowest temporal level (e.g. POC 0 and 8) were protected with the highest priority, picture in temporal level 1 (e.g. POC 4) was protected with high priority, and pictures in the highest temporal level (e.g. POC 1, 3, 5, 7) were protected with low priority.

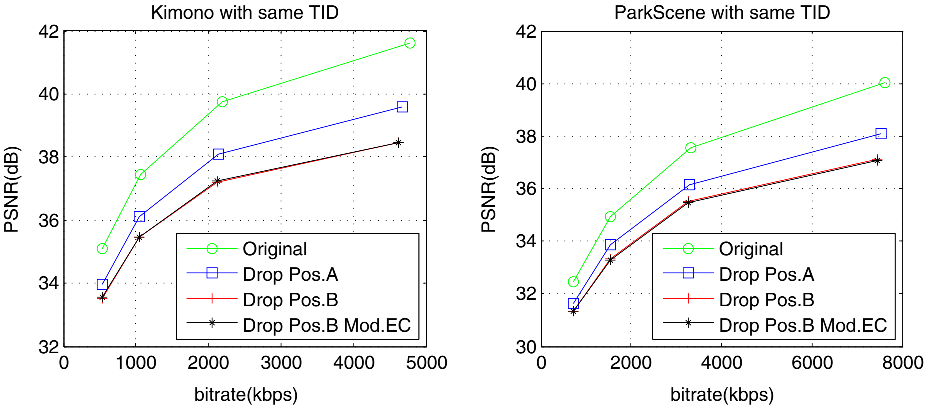

Fig. 8 RD curves for dropped packets in *Pos.A* and *Pos.B*

Finally, the gain of proposed picture prioritization method was from 2.2 dB to 7.5 dB in PSNR.

3.2 Proposed method 2: error concealment with mode signaling

This section explains the EC mode signaling method with our previous researches [12, 14]. The EC method is important for a scalable video coding transmission system over error-prone network [8]. Figure 9 shows an example of scalable coding with two layers (BL and EL, following number represents POC), where the picture EL2 in EL is lost. In the example, decoder can copy one of EL0, EL4, or BL2 to conceal the lost EL2 as a simple EC method (picture copy). Because EL2 could be referenced by EL1, EL3, and EL6, losing EL2 can cause error propagation in EL1, EL3, EL5, EL6, and EL7 (marked with red wave). Thus, applying the best EC method can improve not only the quality of the lost picture EL2, but also the quality of the other pictures such as EL1, EL3, EL5, EL6, and EL7 that are affected due to error propagation. The proposed EC mode signaling method works as the diagram in Fig. 10.

Though general video decoder supports some EC methods [5], it is difficult to find the best EC method among the supported EC methods at the decoder side without original pictures. The proposed EC mode signaling method enables the video encoder to (i) simulate various EC methods on a damaged picture, (ii) determine the best EC method that provides minimal disparity between an original image and a reconstructed image, and (iii) signal the best EC mode to the video decoder at the client.

This message-based EC mode signaling is designed to provide a general communication channel between multimedia server and client. Thus, any EC method that is developed by application developer can be used by using ‘user defined’. It could be in the HEVC SEI

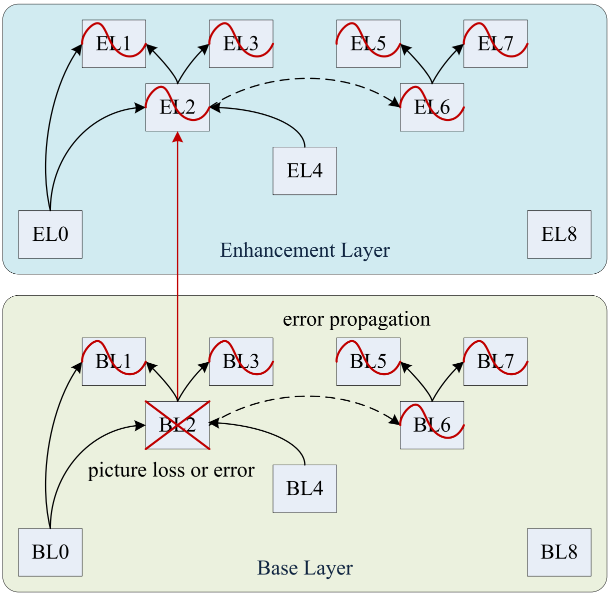

Fig. 9 The effect of the error propagation in SHVC

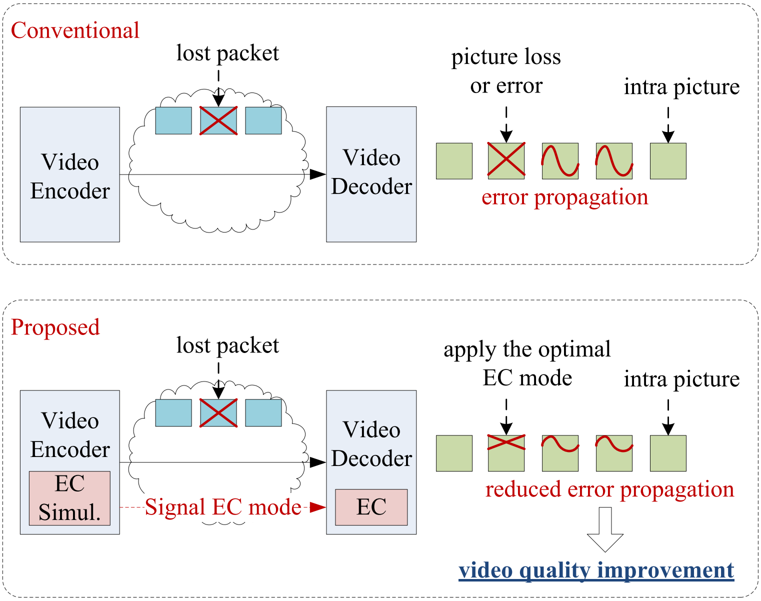

Fig. 10 EC mode signaling to reduce the error propagation (upper: conventional method, lower: proposed method)

message, MMT Transport Packet (MMTP) syntax, or MMT EC mode message as international standard documents provide by first author (ISO/IEC JTC1/SC29/WG11 M31189 and M32347.)

In the proposed EC mode signaling algorithm, the decoder sets EC mode to dealt EC mode 0, and it means the decoder copies previous reference picture if the picture is lost. In assumption, this study does not consider the picture loss of first intra-picture (I picture) of BL. Normally, in video streaming systems, the first intra-picture is guaranteed to transmit by using retransmission and FEC. There are some ways to determine the picture loss in the decoder side. If the whole one picture is lost during transmission, the picture number POC cannot be continued, and the decoder easily knows the lost picture number. If the partial one picture is lost during transmission, the decoder internally faces decoding errors in coding unit level, and its error handler lets the decoder conceal the damage by using the signaled EC mode. If the error handler just fills the damaged coding unit (in H.264, macro-block) with average Y/U/V values of neighboring coding units, severe blocky artifacts can be observed. If there is no EC method in a decoder, the decoder normally crashes with the packet loss.

Regarding the complexity of the proposed EC mode signaling method, it does not increase the computational complexity at encoder side. Comparing the disparities between the original picture and the reconstructed reference pictures is already included in the default encoding processes such as a motion estimation. Thus, the proposed method simply compares the disparities only and signals the best EC mode that has minimal disparities. In addition, the bitrate increasing for the EC mode signaling is very little because the method just sends a few bits per one picture. For example, if the encoder and decoder support 4 types of EC methods, the method needs only 2 bits to indicate the best EC mode. To verify the benefit of the proposed method, this study implements the optimal EC mode determination

Table 3 PSNR gain between EC modes for referenced pictures

Scalability	QP	Avg. Y-PSNR gain (dB)		
	BL 38	EC4 – EC0	EC4 – EC1	EC4 – EC3
Spatial 2×	EL 32	1.83	1.98	1.79
	EL 33	1.77	1.91	1.74
	EL 34	1.77	1.74	1.72
Spatial 1.5×	EL 32	1.91	2.10	0.20
	EL 33	1.91	1.89	1.87
	EL 34	1.86	1.89	1.82
SNR	EL 26	2.38	2.50	2.34
	EL 28	2.29	2.45	2.26
	EL 30	2.22	2.33	2.18

module in SHVC reference software (SHM) encoder and decoder version 2.0 [4]. For the fair performance comparison, this study implements some simple EC methods that use the picture copy method with multiple options as follows.

(a) EC0 (EC mode 0): picture copy from previous reference picture (1st picture in RPL0),
(b) EC1 (EC mode 1): picture copy from next reference picture (1st picture in RPL1),
(c) EC2 (EC mode 2): picture copy from the upsampled BL picture (picture in inter-layer picture buffer),
(d) EC3 (EC mode 3): picture copy from the reference picture that has lower QP (among 1st pictures in RPL0 and RPL1),
(e) EC4 (EC mode 4): (proposed method) signaling best EC mode.

The test bitstream (sequence named *Han* provided by Vidyo Inc.) consists of two layers (BL and EL) with three types of scalability, and their resolutions of the *spatial* 2× are 1080p and 540p. For *spatial* 1.5×, 1080p and 720p sequences are used, and SNR (quality scalability) uses same 1080p sequences with different QPs.

Tables 3 and 4 show the experimental results. The gains in video quality vary from 0.2 to 2.5 dB in PSNR. Thus, the advantage of the proposed EC mode signaling method in this section is also verified.

Table 4 PSNR gain between EC4 and EC2

	Scalability	QP		Y-PSNR gain (dB) in
		BL	EL	GOP (POC 65-72)
EC 4 - EC 2	Spatial 2×	38	32	1.03
			33	0.83
			34	0.81
	Spatial 1.5×	38	32	0.38
			33	0.35
			34	0.27
	SNR	38	26	0.37
			28	0.34
			30	0.20

3.3 Proposed method 3: tile-based video parallel processing for real-time decoding

This section proposes a new HEVC Tile-based parallel processing method for real-time requriement. The Tile allocation method considers the computational ability of asymmetric multicores as well as the computational complexity of each Tile. The computational ability of asymmetric multicores can be measured or provided by chip maker, and the computational complexity of each Tile can be measured by the amount of HEVC prediction unit (PU) partitioning.

The concept of CU in HEVC is similar to the macro block (MB) in H.264 AVC standard. For example, one partitioning mode can split a 64 × 64 CU into 32 × 32, 16 × 16, and 8 × 8 CUs, and each CU can be partitioned into one, two, or four PUs. The PU is a basic unit for prediction in CU.

Because the HEVC profiling results explain a motion compensation in video coding standard, it requires significant computing power [3]. This study assumes that the amount of PU partitions of a CU has a high correlation with computational complexity. There have been researches to estimate picture complexity; however, the PU partition-based complexity estimation is possibly one of the simplest methods because it is measured and given information by encoder while it encodes each CU.

The (implemented) modified HM encoder calculates the sum of the number of PU partitions of each CU in a Tile: if a Tile has many PU partitions, the processing complexity of the Tile is high. Therefore to verify our assumption, we implemented the counting module

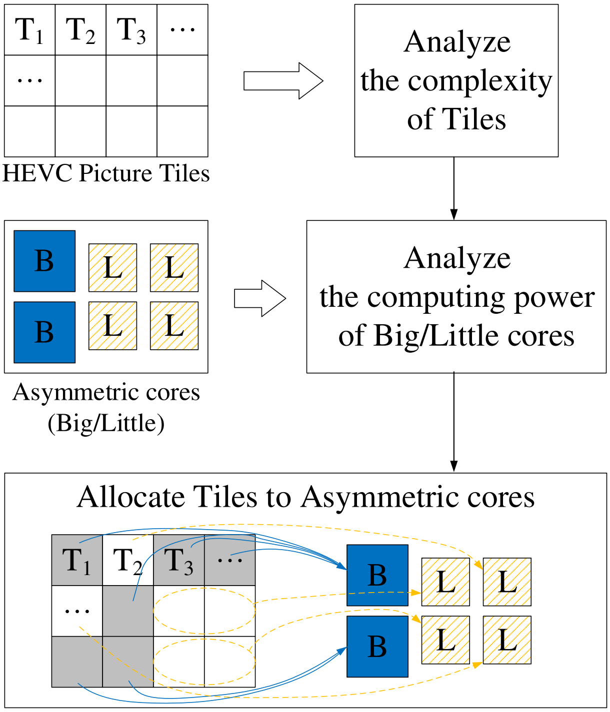

Fig. 11 Conceptual diagram of the proposed tile allocation method

of PU partitions in HM reference encoder, and the actual decoding time of each Tile is measured and compared with the amount of PU partitions. In our experiment, the relationship between the sum of PU partitions and actual decoding time has high correlation between with RA configuration. The experiment uses the test sequence, *PeopleOnStreet* (3840 × 2160), defined in HEVC common test condition (CTC). The sequence is encoded with RA configuration, and QP 22 is used.

For the real-time video decoding, this section proposes a new Tile allocation method by considering the computational ability of asymmetric multicores and the computational complexity of each Tile. The computational ability of asymmetric multicores can be measured or provided by the manufacturer while the computational complexity of each Tile can be measured by the amount of HEVC PU partitioning. Then, based on the measured/analyzed complexity of tiles and computing power of asymmetric cores, the proposed method allocates the set of Tiles to the Big and Little cores as shown in Fig. 11.

To verify the proposed parallel processing method for real-time decoding, this experiment uses HM15.0 (HEVC reference model) to encode/decode test sequences on JUNO ARM development platform that supports asymmetric big.LITTLE multicores. For encoding and decoding test sequence, *PeopleOnStreet* (3840 × 2160, 150 frames) is used with RA configuration and QP 22. The video picture is divided into 6 and 12 Tiles. The experimental JUNO ARM platform has 6 cores including 2 big cores and 4 little cores. Thus, when the video picture is divided into 6 Tiles, each Tile is allocated to a core, and when the video picture is divided into 12 Tiles, multiple Tiles are allocated into a core.

Table 5 shows the gains from conventional and proposed methods. When the performance was measured with Tile 6, and Tile 12, the gains with allocating fixed number of Tiles to cores were 5.24%, 8.44%, respectively. Then, the proposed method with allocating number of Tiles to cores adaptively (different number of Tiles) was experimented, and result show 18.03% decoding time speed up. The proposed method is now extending to a new HEVC Tile allocation method considering the computational ability of asymmetric multicores as well as the computational complexity of each Tile. Those two methods

Table 5 The gains of decoding time speed up using uniform Tile allocation and proposed adaptive Tile allocation

	6 Tiles		12 Tiles		
	Fixed		Fixed		Adapted
Loop	Control	Proposed	Control	Proposed	Proposed
1	68.52	64.64	68.22	62.65	56.13
2	67.96	64.55	68.14	62.63	55.96
3	68.13	64.50	68.33	62.31	55.86
4	68.00	64.64	68.21	62.46	56.11
5	67.98	64.62	68.06	62.63	55.87
6	68.47	64.58	68.22	62.37	56.03
7	68.36	64.58	68.40	62.72	55.86
8	68.20	64.66	68.21	62.44	55.71
9	68.10	64.57	68.17	62.15	55.88
10	67.91	64.56	68.36	62.42	55.90
Average	68.16	64.59	68.23	62.47	55.93
Gain	5.24%		8.44%		18.03%

Table 6 The performance gains of three proposed methods

Features	Proposed methods	Performance gain
Robust video streaming	(1) Picture prioritization method in hierarchical B structure of HEVC	2.2–7.5 dB
	(2) EC mode signaling method that signals best EC mode(s)	0.2–2.5 dB
Real-time decoding	(3) HEVC Tile-based parallel processing using asymmetric CPU multicores	18.03% decoding time speed up

are implementing on Samsung Galaxy S7 Edge smart phone, and their primitive performance gains were more than 30% decoding time speed up. Thus, the third proposed method described in this section was also verified with improved decoding time performance.

4 Conclusion

This paper proposes a new system architecture for real-time UHD video streaming and proposes three main ideas: (i) picture prioritization method, (ii) error concealment mode signaling, and (iii) Tile-based video parallel processing. In the experiments using HEVC and SHVC reference models conducted, (i) the proposed picture prioritization method shows the gains in video quality from 2.2 to 7.5 dB in Y-PSNR, (ii) the error concealment mode signaling gains from 0.2 to 2.5 dB in Y-PSNR, and (iii) the Tile-based video parallel processing for real-time decoding shows 18.03% decoding time speed up. Thus, three proposed methods described in this paper was verified as described in Table 6, and this paper proposes the architecture of our robust real-time UHD video streaming system using SHVC.

References

1. Ahn H, Jeong S (2013) Power-minimizing dvfs algorithm using estimation of video frame decoding complexity. The Journal of Korean Institute of Communications and Information Sciences 38(1):46–53
2. Baik H, Song H (2015) A complexity-based adaptive tile partitioning algorithm for hevc decoder parallelization. In: 2015 IEEE international conference on image processing (ICIP). IEEE, pp 4298–4302
3. Bossen F, Bross B, Suhring K, Flynn D (2012) Hevc complexity and implementation analysis. IEEE Trans Circuits Syst Video Technol 22(12):1685–1696
4. Chen J, Boyce J, Ye Y, Hannuksela MM (2013) Shvc test model 2 (shm 2). In: JCTVC-M1007, 13th joint collaborative team on video coding (JCT-VC) meeting, Incheon, South Korea
5. Guo Y, Chen Y, Wang YK, Li H, Hannuksela MM, Gabbouj M (2009) Error resilient coding and error concealment in scalable video coding. IEEE Trans Circuits Syst Video Technol 19(6):781–795
6. Hellge C, Schierl T, Wiegand T (2008) Receiver driven layered multicast with layer-aware forward error correction. In: ICIP 2008. 15th IEEE international conference on image processing, 2008, pp 2304–2307
7. Kim HS, Ryu ES, Jayant N (2010) Channel-adaptive video transmission using h.264 svc over mobile wimax network. In: 2010 digest of technical papers international conference on consumer electronics (ICCE), pp 441–442
8. Mansour H, Nasiopoulos P, Krishnamurthy V (2008) Joint media-channel aware unequal error protection for wireless scalable video streaming. In: Proceedings of the IEEE international conference acoustics, speech and signal processing ICASSP 2008, pp 1129–1132
9. Martini MG, Mazzotti M, Lamy-Bergot C, Huusko J, Amon P (2007) Content adaptive network aware joint optimization of wireless video transmission. IEEE Commun Mag 45(1):84–90
10. Ryu ES (2013) Prediction-based picture prioritisation method for hierarchical b-structure of high efficiency video coding. Electron Lett 49(20):1268–1270

11. Ryu ES, Jayant N (2011) Home gateway for three-screen tv using h.264 svc and raptor fec. IEEE Trans Consum Electron 57(4):1652–1660
12. Ryu ES, Kim J (2015) Error concealment mode signaling for robust mobile video transmission. AEU-Int J Electron C 69(7):1070–1073
13. Ryu ES, Ye Y, He Y, He Y (2012) Frame prioritization method based on prediction information. In: JCTVC-J0063 for Stockholm meeting on HEVC standard
14. Ryu ES, He Y, Ye Y, He Y (2013) On error concealment mode signaling. In: ISO/IEC JTC1/SC29/WG11 MPEG2013/M31189 for Geneva meeting
15. Ryu E, Ye Y, He Y, He Y (2014) Frame prioritization based on prediction information. In: U.S. Patent Application 13/931,362
16. Ryu ES, Ryu Y, Roh HJ, Kim J, Lee BG (2015) Towards robust uhd video streaming systems using scalable high efficiency video coding. In: 2015 international conference on information and communication technology convergence (ICTC). IEEE, pp 1356–1361
17. Schierl T, Schwarz H, Marpe D, Wiegand T (2005) Wireless broadcasting using the scalable extension of h. 264/avc. In: IEEE international conference on multimedia and expo, pp 884–887
18. Sullivan GJ, Topiwala PN, Luthra A (2004) The h.264/avc advanced video coding standard: overview and introduction to the fidelity range extensions
19. Sullivan GJ, Ohm JR, Han WJ, Wiegand T (2012) Overview of the high efficiency video coding (hevc) standard. IEEE Trans Circuits Syst Video Technol 22(12):1649–1668
20. Ye Y, Andrivon P (2014) The scalable extensions of hevc for ultra-high-definition video delivery. Multimedia, IEEE 21(3):58–64
21. Ye Y, McClellan GW, He Y, Xiu X, He Y, Dong J, Bal C, Ryu E (2014) Codec architecture for multiple layer video coding. In: U.S. Patent Application 13/937,645
22. Yoo S, Shim Y, Lee S, Lee SA, Kim J (2015) A case for bad big. little switching: how to scale power-performance in si-hmp. In: Proceedings of the workshop on power-aware computing and systems. ACM, pp 1–5

Eun-Seok Ryu is an Assistant Professor at the Department of Computer Engineering in Gachon University, Seongnam, Korea. Prior to joining the University in 2015, he was a Principal Engineer at Samsung Electronics, Suwon, Korea, where he led a multimedia team. He was a Staff Engineer at InterDigital Labs, San Diego, California, USA, from Jan. 2011 to Feb. 2014, where he researched and contributed to next generation video coding standards such as HEVC and SHVC. From Sep. 2008 to Dec. 2010, he was a Postdoctoral Research Fellow at GCATT in the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia USA. In 2008, he was a Research Professor at the research institute for information and communication technology in Korea University, Seoul, Korea. His research interests are in the area of multimedia communications that includes video source coding and wireless mobile systems. He received his B.S., M.S., and Ph.D. in computer science from Korea University in 1999, 2001, and 2008, respectively. He is a Senior Member of the IEEE, IEEE Computer Society, and IEEE Communications Society.

SunJung Ryu is an Associate Professor at the Department of French in Dankook University, Korea since 2009. Her research interests are in the area of French literature, Cultural contents (film, animation, musical comedy, game, festival, advertisement), Communications and Digital storytelling. She received her Ph.D. in French literature from Sorbonne University (Paris-IV) in 2002. She was a Postdoctoral Research Fellow in Cultural contents with the support of National Research Foundation of Korea.