CrossMark

Design of Efficient Key Video Frame Protection Scheme for Multimedia Internet of Things (IoT) in Converged 5G Network

Jong-Hyeok Lee[1] · Gwang-Soo Hong[1] · Young-Woon Lee[2] · Chang-Ki Kim[3] · Noik Park[3] · Byung-Gyu Kim[1]

Abstract

To guarantee the quality of video data into fast-responding transmission and high resolution output video using cost effective video processing is desirable in many services including Internet of Things (IoT) applications. The goal of this study is to develop a transmission method to improve a quality of service (QoS) to support for various multimedia contents with high quality on 5 generation (5G) convergence network. The main motivation is based on video feature and dependency between frames and blocks in coding structure. First, we investigate the existing methods and analyze them into some classes. From the analyzed result, we propose a priority-based key frame protection method for improving QoS of in 5G convergence network.

Keywords Multimedia-centric IoT · Quality of service · Ultra high quality video · 5 generation (5G) convergence network · Key frame protection

1 Introduction

While the 2-nd generation and 3-rd generation mobile communication support voice-centric service, the 4-th generation mobile communication technology took big step to data-centric service domain. With the change of service pattern, the demand on multimedia and social network service is widely growing up. It causes tremendous increase of data traffic with various IoT devices in [18, 22]. This trend accelerates to emerge the converged mobile technology and business with big data and artificial intelligence technologies.

To cover this change of mobile convergence, the current mobile technology should be reformed and improved due to the available frequency band, the limitation of antenna technology, cell coverage and so on. Now, many organizations such as 5G Forum (Korea), ITU-R, 3GPP (Europe), 5GPP (Europe), IMT-2020 (China), ARIB (Japan) are trying to lead 5G mobile communication standard which can solve the above problems. In Korea, the government of Korea is preparing a pre-service of 5G mobile communication system in Pyungchang Winter Olympic of 2018 and has plan to make international standard promptly. Figure 1 shows various organizations for 5G mobile communication standard.

IoT is an umbrella term includes the network of electronic devices like smart phones, wearable electronic devices or connected appliances that are capable of communicating wirelessly with each other by [3]. From applicability standpoint, these artificially intelligent networks interact with respective physical environment; allowing IoT to influence the present civilization from home automation to security, health monitoring and managing the daily task.

✉ Byung-Gyu Kim
  bg.kim@sookmyung.ac.kr

  Jong-Hyeok Lee
  jh.Lee@ivpl.sookmyung.ac.kr

  Gwang-Soo Hong
  gs.Hong@ivpl.sookmyung.ac.kr

  Young-Woon Lee
  yw.Lee@ivpl.sookmyung.ac.kr

  Chang-Ki Kim
  ckkim1@etri.re.kr

  Noik Park
  nipark@etri.re.kr

[1] Department of IT Engineering, Sookmyung Women's University, Seoul, Republic of Korea

[2] Department of Computer Engineering, Sunmoon University, Asan, Republic of Korea
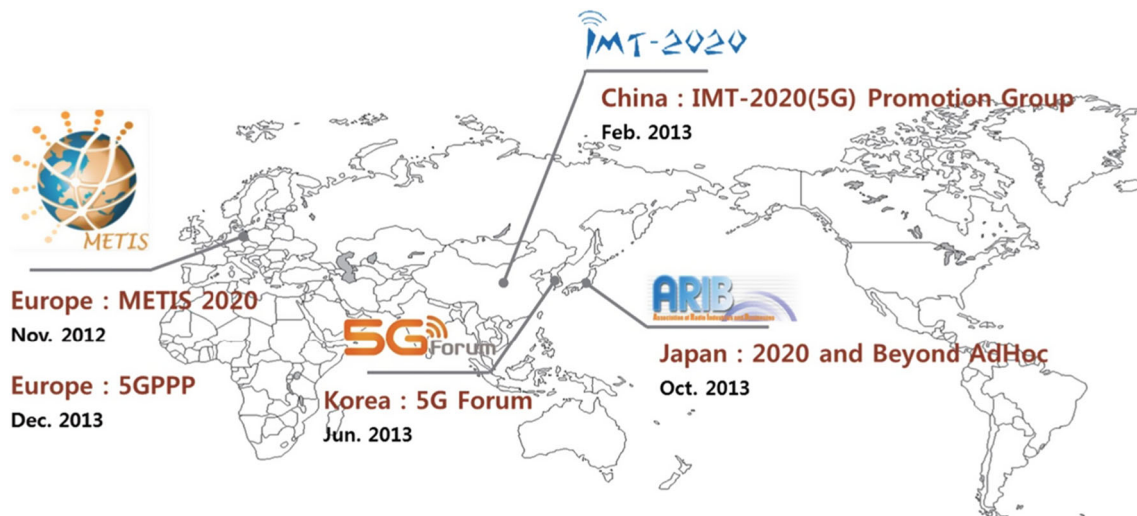
[3] ETRI, Daejeon, Republic of Korea

Fig. 1 The organizations for 5G mobile communication standard in the world

The growing applicability of IoT is hypothesized to reach over 50 billion devices by 2020 in [9].

With the advancement, considerable amount of data traffic is generated from IoT based devices, which are major multimedia data. The enormous size of raw multimedia data followed by its storage and transmission pushes the urge of its compression using video compression technology.

As the demand on the ultra-high quality media in mobile multimedia service extensively, the mobile service providers are extending their service infra-structure to meet consumer's demand. In Korea, the subscribers on the long-term-evolution (LTE) service are consuming the bandwidth of 3.3GB at October 2014 [23]. Also as Global Mobile Data Traffic Forecast by [8], The mobile traffic amount will be expected up to 15.6 Exa Bytes (EB)/month at 2018, from 1.5 EB/month in 2013. The current mobile multimedia system provides the 720p HD (High Definition) or 1080p full high definition (FHD) video service as high quality service standards [4].

In 2014, a new technology for the ultra high definition (UHD) video has been standardized for various industrial applications such as home appliance, broadcasting service, and mobile multimedia system by [28]. It is called as the high efficiency video coding (HEVC) standard. Many industries are studying how to support the real-time services with the UHD (4K:3,840 × 2,160 and 8K:7,680 × 4,320) video. The HEVC standard for the UHD video contents basically supports from 4 to 8 times of resolution, more bit depths (up to 12 bits), and various color formats. So its amount of data has been dramatically increased over 96 times of FHD video in [28].

Compared to FHD video service, the UHD is an essential part of the next generation realistic-media broadcasting service. The high efficiency video coding (HEVC) standard

now provides the creation of the UHD video content and define how to deliver it through networks even though the emerging IoT applications. To deliver UHD video content via mobile convergence network, the main problem is that the existing delivery techniques are weak to support the UHD video service because of the excessive fading of channel, very high traffic, deficiency of frequency resource, and variation of channel characteristics. To solve this situation, various works have been conducted by [10, 11, 16, 19, 24–26, 29]. There are several methods for allocating resource and scheduling them in mobile multimedia service in [11, 26, 29]. On the other hand, some authors proposed approaches based on the cross-layer optimization using packet loss, to improve the QoS in mobile environment by [10, 16, 19, 24, 25].

When delivering media data through network, an important issue is how to protect the media data. It means the communication system has to prevent the loss of data. We can classify into three approaches in physical layer: Video encoding/decoding part, channel coding in transmission, and packet error concealment at receiver part. For video encoding/decoding part, H.264/AVC [4] is being employed to support various media service. If some packets or frames are lost when delivering through network, it causes error propagation to other frames (data packets) where can be degraded significantly. Especially if the loss occurs in important frame (or key frame), then it can generate very serious quality degradation. Figure 2 shows an example of error propagation as frame goes. For providing the very high quality of video including audio signal, HEVC standard by [28] becomes popular in UHD contents and displays. However, there is no case study for the effect of content delivery through networks. Therefore, we need research on key frame protection in terms of the converged

Fig. 2 Error propagation between frames at decoder side
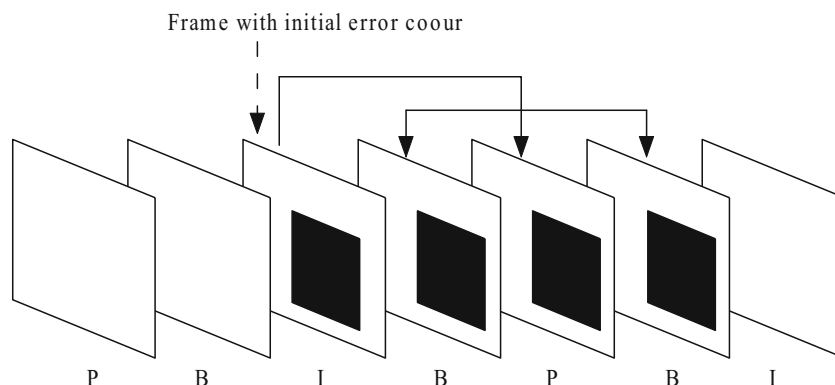
5G network layer, to improve the QoS. This study is aimed to suggest new method for protecting a key frame when UHD contents are transmitted with very high data traffic.

This paper is organized as follows: Section 2 will introduce the existing methods for key frame protection. In Section 3, we analyze the issue of the previous approaches and propose an alternate approach. Finally, we will give a concluding remark in Section 4.

2 Related works

2.1 Cross layer optimization approach

To achieve very high quality of video service, we can categorize main approaches into the followings: real-time video service optimization based on channel condition, wireless resource allocation using video characteristics, cross layer optimization of video service considering both of video characteristics and channel condition in [26–29]. In general, video stream is transmitted by the unit of IP packet and IP packets are delivered through Medium Access Control (MAC) and Physical Layer (PL). In physical layer, there are 15 Channel Quality Indicator (CQI) levels according to modulation and target code rate (TCR). Also, Modulation and Coding scheme (MCS) can be possible up to 29 levels. In this structure, we select the MCS level based on CQI level and the assigned Resource Block (RB). For example, for 16 Quadrature Amplitude Modulation (QAM), we have CQI = 9 and RB = 3. Then we can select 16 as the MCS level, which is corresponding to 9 of the CQI. The level of CQI is selected under condition of less than 10% Block Error Rate (BLER). In physical layer, there is Hybrid Automatic Repeat Request (HARQ) with MAC layer. At this time, an efficient bit rate to provide to application layer is a function of signal-to-noise ratio (SNR) and CQI as the following:

$$B(SNR, CQI) = \frac{TBA(CQI)}{t_{TBS} \times \eta_{HARQ}(SNR, CQI)}, \quad (1)$$

where t_{TBA} is the transmission period of Transport Block Size (TBS), η_{HARQ} is the average number of re-transmission of HARQ.

In application layer, one IP packet is composed of many Transport Block (TB). Based on information of the CQI, we can assign the number of TB. To guarantee QoS, the forward error correction (FEC) is applied on each IP packet of unit of Group of Pictures (GoP) as different video stream. We calla Packet Loss Rate (PRL) as "residual PLR" after doing FEC. Finally, we compute the bit rate by applying Reed-Solomon (RS) coding. However, this structure has a critical problem when LTE downlink system environment as:

1) when the SNR is given, it is very difficult to consider all CQI level by adopting Equi-PSNR (Peak Signal-to Noise Ratio) graphs which are from (Bit-rate, PLR) planes.
2) In physical layer, for computing efficient bit rate, there is no consideration of the number of e-transmission of HARQ.
3) Also, this is not suitable for solving the problem of optimization because this structure measures in actual, not prediction of the video quality using the average value of BLER.

To overcome this situation, there are several cross-layer approaches by [1, 5, 17, 21]. In [21], Lung-Jen et al. proposed a method using Dynamic Bandwidth Measurement and Dynamic Queue Adaption. Figure 3 shows the overall structure of Wang's method [21]. They used the existing video coding scheme to make packets in application layer and managed the bandwidth dynamically in NAT layer. In MAC layer, they provided 4 AC queue buffers to improve the congestion by access control of channel. Khalek et al. [1] proposed a new algorithm using Qos-to-QoE mapping and Unequal Error protection (EUP). Usually, the change of video signal is slower than that of channel, in physical and application layers. Thus they suggested an approach to design the physical layer to change its characteristics rapidly than that of video
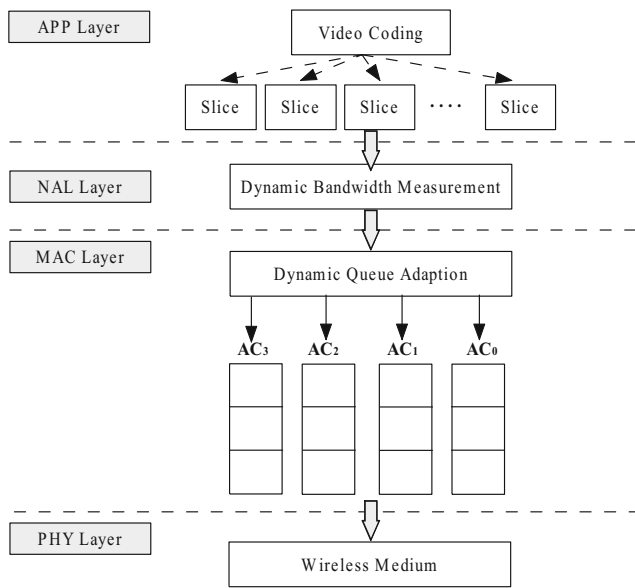
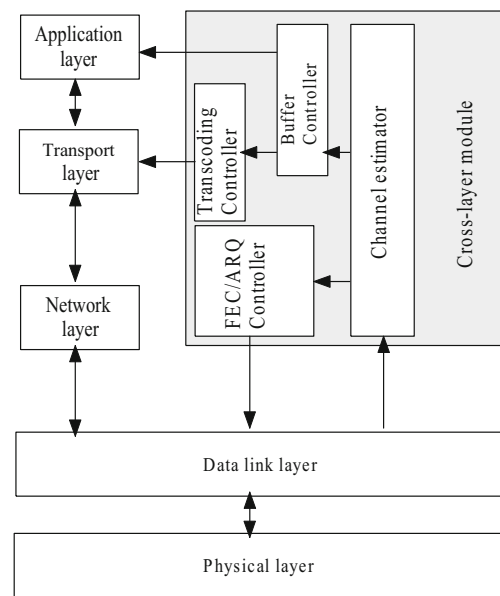**Fig. 3** A method by [21]

Fig. 5 Cross-layer Module (CLM) by [5]

signal in application layer (Fig. 4). In [5], Moid et al. developed a method using Channel Estimator. By channel estimator, they controlled the buffer amount and adopted time limitation to decode the lost packets as shown in Fig. 5.

From this approach, the consumed time for packet processing is a little increased but, it achieved the improvement of 5 dB in visual quality. Jassal et al. [17] have proposed a new algorithm to improve the 3D-HEVC video service quality in LTE network. It used ratio of packet loss and average access number of user. Also they made a packet

priority-based and cross-layer structure, and provided a quality-of-experience (QoE) to users. Like HEVC standard, the basic processing unit of 3D-HEVC is NAL. This NAL unit is classified into Video Coding Layer (VCL) NAL and Non-Video Coding Layer (Non VCL). Except information of slice and CTU, other parameters is sent via Non-VCL NAL. Jassal et al. [17] analyzed the header information of NAL units and gave priority scheme to improve the QoS in LTE network.
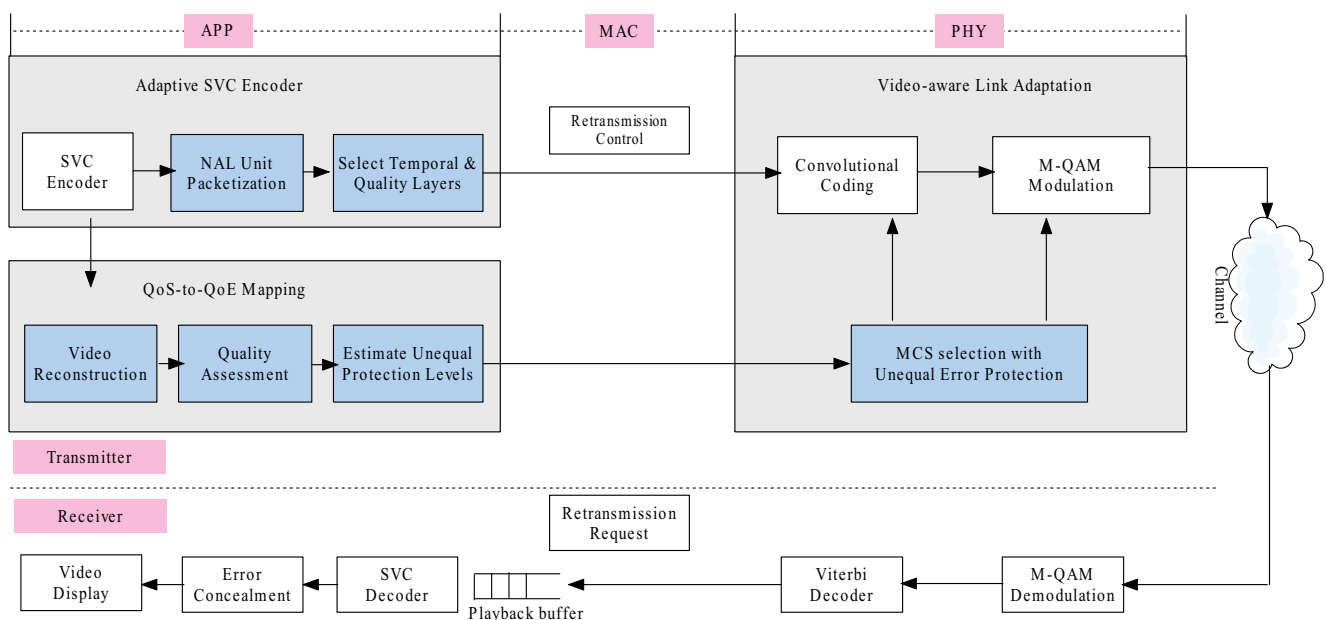

Fig. 4 Cross-layer QoE video delivery system by [1]

2.2 Priority-based data partitioning approach

In this subsection, we describe some schemes for partitioning the bitstream into segments according to its importance in terms of transmission loss. If the loss occurs in transmission, it causes serious error propagation of decoder side from frame to frame. The quality of video is degraded significantly. Especially, if motion vector, mode of macroblock are lost, the decoded video is highly degraded at decoder side because of the information loss. To cope with this problem, data partitioning technology based on degree of importance of the compressed bitstream, has been proposed in video standards [13, 30]. In H.264/AVC, it was extended to H.264/AVC bit stream from MPEG-4 Part2 (the existing data segmentation scheme) by [30]. It can support to transmit bit stream through various networks and is developed to keep the coding efficiency and network adaptation.

Zhang et al. [32] proposed a H.264/AVC video transmission system using the lost packet network. They made new data partition structure based on Unequal loss protection (ULP). In this structure, RS, XOR parity code, ULP, and partition A protection (PAP) are used to protect three types of loss. Through the experiments, they proved to make higher PSNR with smaller PAP. In [20],

Ksentini et al. suggested a cross-layer approach based on data partitioning scheme as shown Fig. 6. In this method, they designed a scheme for making adaptable to channel characteristics, limitation of QoS resource, and content characteristics. This method was implemented on IEEE 802.11 WLAN service with H.264/AVC video. In average performance, packet loss and the delay of End-to-End were minimized.

Go [12] developed SAVE (Systematic reAllocation and prioritization of Video packets for Error-resilient transmission). This approach make a difference in terms of followings. First, packet re-assignment and priority assignment are made after finishing encoding. So the coding efficiency of encoding is not affected by SAVE scheme. Secondly, this mechanism send the detailed packet-based priority assignment rather than video frame and data partition unit. Therefore, meta data of the encoded frame can be assigned the priority of PAT (Program Association Table), SEI (Supplemental Enhancement Information), P frame header and others in MPEG-2 TS. The third thing is that the SAVE is not dependant on the H.264/AVC profile because it uses cross-information of the header of the encoding frame and transmission protocol. In final, the SAVE can give better recovery performance with smaller

Fig. 6 Cross-layer structure based on data partitioning
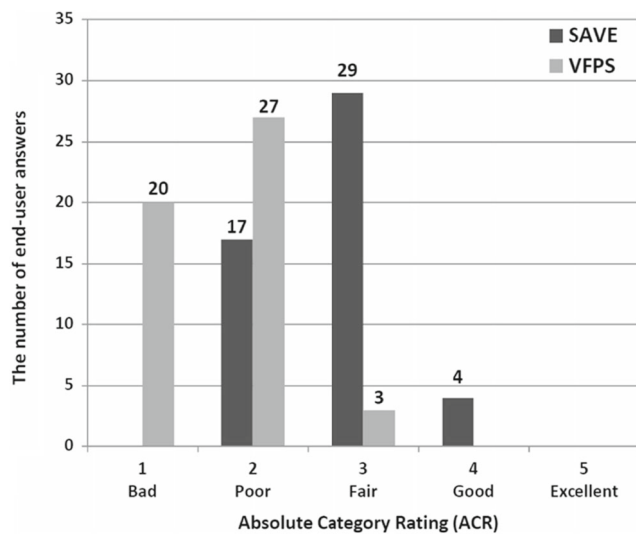

[htbp]
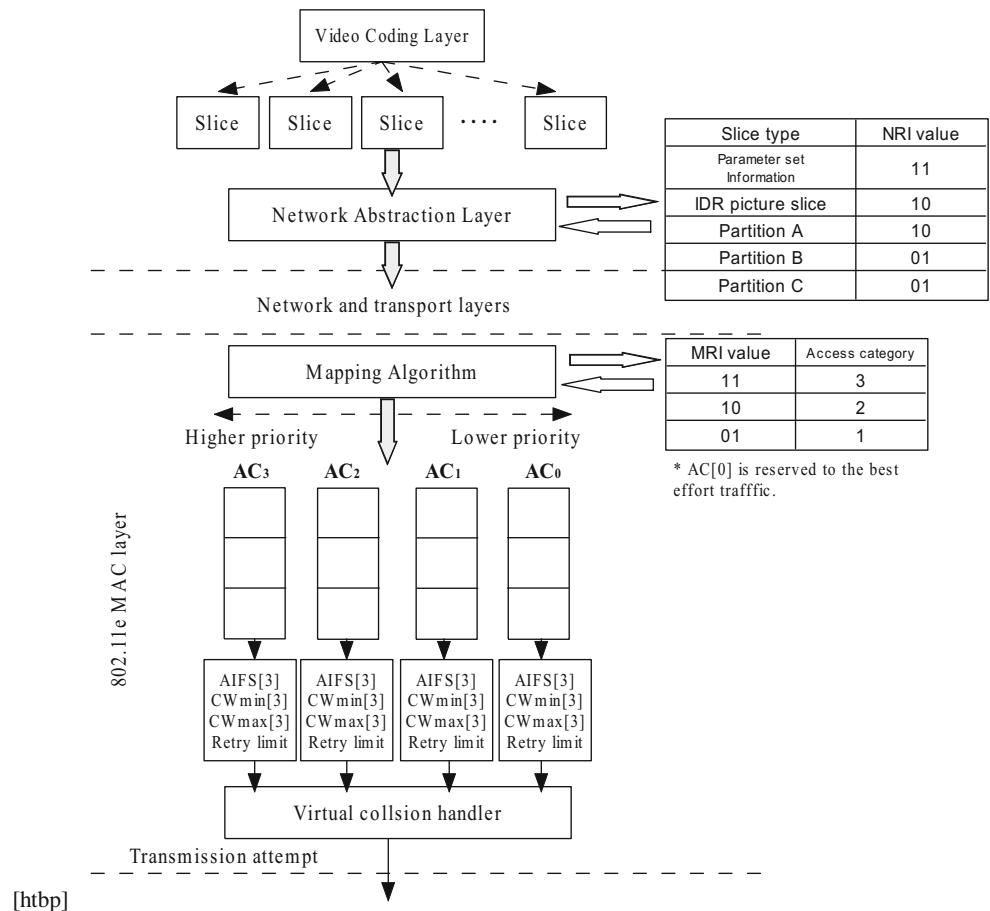
Fig. 7 Result of subjective quality test in SAVE algorithm

overhead rather than the previous methods although the size of video enlarges.

Figure 7 shows a result of subjective video quality test with SAVE algorithm. From the result, it achieved 137 ACR and 83 VFPS in total. Also more than a half of users among all 33 users voted to "so so" or "good" quality of the delivered video. As a result, the SAVE got a score of 2.74 in MOS.

2.3 Error robustness encoding approaches

In the previous subsections, we described several approaches in terms of network delivery, not video

encoding and decoding side. In this subsection, we would like to introduce some algorithms to recover the error in encoder/decoder side. To prevent the degradation of the quality, the corrupted image can be recovered by using the image patch with no error. This is called as "error concealment." However, we need to provide some useful information in encoding process for keep enough good quality, even though some errors occur in network delivery. This approach is defined as "error-robustness encoding." There are some existing methods by [7]-[6].

Jian et al. [15] proposed a method to recover the corrupted video data by using neighboring blocks and a co-located block. Figure 8 shows the recovered result by the suggested method. It gave a reliable result.

3 Proposed scheme

In this section, we describe a new scheme which is a priority-based key frame protection method for improving QoS of in 5G convergence network. In the existing methods, they used frame unit or data partition unit to make a robust transmission on packet errors. However to provide more robustness and coding efficiency, the protection scheme should reflect the relationship between frames. Also, it should be suitable for ultra high quality video service using the next generation video compression technique.

As the next generation video coding standard, the HEVC is specialized for ultra high quality video such as UHD contents by [28]. This also can provide various multimedia service when comparing to H.264/AVC standard, in terms of coding gain and network environment. The HEVC

Fig. 8 Result of subjective quality in "Foreman" sequence: **a** the original **b** the corrupted frame, **c** temproal replacement [2]: 26.16 dB, **d** JM (Joint video team Model): 29.12 dB, and **e** a result by [15]: 30.08 dB
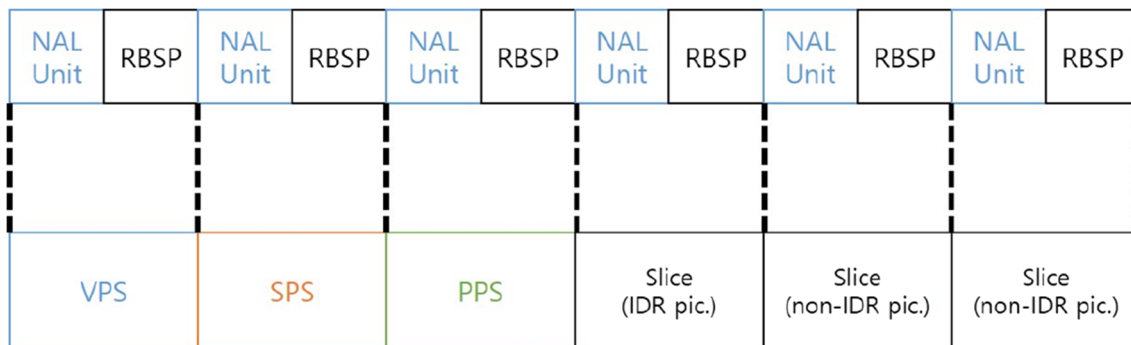

(a)　　　　　　(b)　　　　　　(c)

(d)　　　　　　(e)
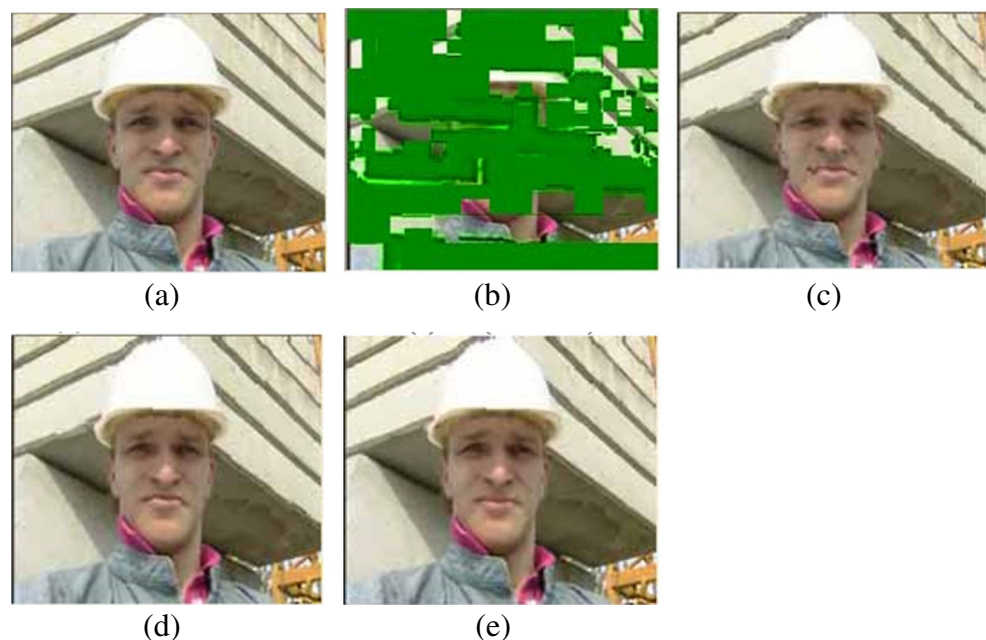
Fig. 9 NAL unit structure

give a structure of NAL layer and Video Coding Layer (VCL). VCL has the compressed MBs and various header information. NAL layer provide adaptable information on network transmission. In the standard, the combined form of both is defined as NAL unit.

Figure 9 shows the structure of NAL unit in H.264/AVC and HEVC. This NAL unit has 2 byte headers and Raw Byte Sequence Payload (RBSP) where actual video data is loaded. Table 1 describes kinds of NAL unit types and their contents. When we decode the video data, header information on meta data is very important. If this is lost, the decoder cannot recover the lost part. So we should protect this information strongly.

The suggested technique maintains the efficiency of the video coder by taking into account the reallocation

and priority of video packets after the encoding, thereby increasing the error resiliency in decoding by applying the reallocation techniques in the packet unit. The priority allocation for the subsequent image packet after the coding process and the packet configuration for effective transmission are achieved through the following two courses: The first step identifies important video data to be decoded based on the information recognition technology using the header information of image encoding and transfer protocols as illustrated in Fig. 10. Since it is unable to provide error restoration transfer to all encoded image data, since the available bandwidth is limited. Thus, it is important to identify important image information of codified image data to assign differentiated transmission priorities. Depending on the suggested techniques, the

Table 1 NAL unit: types and contents

| Type	Meaning	Class
0	Unspecified	non-VLC
1,2	non-TSA, non-STASA trailing picture (reference and non-reference)	VCL
3,4	Temporal sub-layer access (TSA) picture (reference and non-reference)	VCL
5,6	Step-wise TSA (STSA) picture (reference and non-reference)	VCL
7,8,9	Broken link access (BLA) picture (with DLP and TFD, with DLP only, and without LP)	VCL
10,11	Instantaneous decoder refresh (IDR) picture (with and without DLP)	VCL
12	Clean random access (CRA) picture	VCL
13	Decodable leading picture (DLP)	VCL
14	Tagged for discard (TFD) picture	VCL
15-20	Reserved	VCL
21-24	Reserved	non-VCL
25	Video parameter set (VPS)	non-VCL
26	Sequence parameter set (SPS)	non-VCL
27	Picture parameter set (PPS)	non-VCL
28	Access unit delimiter	non-VCL
29	End of sequence	non-VCL
30	End of bitstream	non-VCL
31	Filter data	non-VCL
32	SEI message	non-VCL
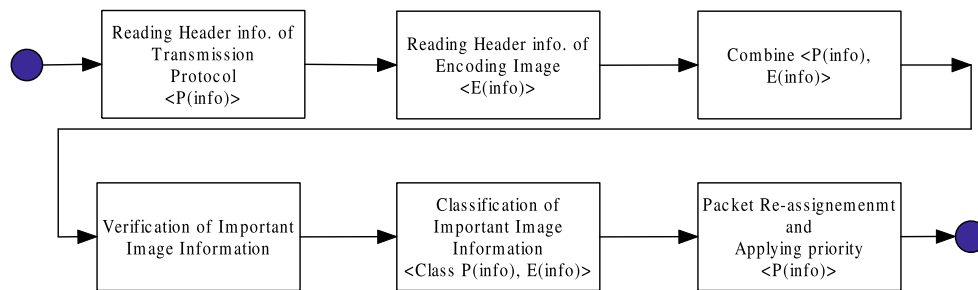
Fig. 10 Flow of image information as degree of importance

types of data identified are subdivided into the following categories:

- Data (HD^{t-1}) containing information to control image streaming such as PAT and PMT, SEI of HEVC, MPD of MPEG-DASH, etc.
- Data (HD^{t-2}) containing information of encoded images such as VPS, SPS, PPS, image frame headers, etc.
- Video Frame Data. Image Frame Data is divided into three classes according to importance.

    – Frame image data containing complete scene information excluding header (VD^{t-1})
    – P or B image data referenced by other scenes to decode (VD^{t-2})
    – B frame data (VD^{t-3}) not referenced by other scenes to be decoded

Since encoded images are sent to support the transport protocol, specific information of encoded images is included in the header information of the transmitted image.Also, the type of current image frame can be determined by reading the header information of the encoded image. Therefore, you can identify important image information by finding the header information of the header information in the header of the Transport Protocol $< P (info) >$ and encoded video $< E (info) >$. For example, to determine whether the MPEG-2 TS currently contains the header information of the I TS, the Payload TS Start indicator (PUSI) header information and stream field values must be parsed beforehand. When a parsing value is set to '1' and 'video', the specific header information in the image frame indicates that the image is currently included in the TS - TS packet : P (info). The NRI value in the HEVC NAL header is then used to determine whether the present MPEG packet contains the I, B (referenced), frame (referenced) header. When NRI is set to '3', it indicates that the current MPEG packet contains the I header. When NRI is set to '0' or '2', the current MPEG packet indicates that there is a reference B frame or non-referenced frame header : E (info).

Important information is also grouped into <Class> in three domains, taking into account the importance of dependencies between decoding and frames. First, the header data (HD^{t-1}), which contains information metadata used to control image streaming, is high priority because it presents a serious problem for each scene and program channel in the image. For example, if the PAT and Program Map tables (PMT), which contain information from a single TS containing the information of a single transmission stream, including the program number and the basic stream list, are lost during transmission, the client can not identify the channel information in the input image stream.

As a result, the image clients of the video data can not be known if the metadata is lost during the transfer of Streaming (MPD), i.e., HTTP (DASH), through HTTP (DASH) [14]. The MPD contains the location of the image data (i.e. URL).Second, the loss of encoding parameters and image frame header data (HD^{t-2}) has a significant effect on the decoding of transmitted image data. Therefore, a high priority is given to that metadata. For example, if the sequential parameter set parameter Set and Parameter parameter set in the image data is lost during transmission, the client can not decode subsequent image data. As a result, if the P header loses the transmission, the subsequent image data is discarded from the decoder, even though the subsequent image data is successfully passed in [14].

Finally, when image encoding is performed based on gan dependencies, the loss of image data, including the complete information of the scene, is significantly degraded because the image of the image is referred to for decoding in other scenes. Thus, I frame image data (VD^{t-1}) has a higher priority than referenced B (VD^{t-2}) and referenced (VD^{t-2} or VD^{t-3}) image data. By applying classification as a combination of classes, P (info), and E (info), you can identify sensitive image information in encoded images. It is possible to apply to all standards based on frame dependencies such as H.264/AVC, HEVC, and so on, since the unclassified results are not limited to specific encoding parameters and profiles.

As a result, the encoded image can be expressed as described in Eq. 2, considering the dependencies between frames.

$$GOP_j = \sum_{i=1}^{T_{pkt}} (HD_{(i)}^{t=1}||HD_{(i)}^{t=2}||VD_{(i)}^{t=1}||VD_{(i)}^{t=2}||VD_{(i)}^{t=3}),$$

(2)

where GOP_j is the j-th GOP of the recorded image, T_{pkt} is the number of packets in the GOP, and i is the order of the received packet in the GOP.

The second step is to systematically reassign video data that has the same priority in neighboring locations within dependency between frames. As illustrated in Figs. 11 and 12, each packet is placed in the adjacent position within the dependency of the frame.

Figures 11a and 12a are codified into IPPB structures with P structure, and Figs. 11b and 12b represent IBBB of Random Access in the HEVC structure. Because the size of the I-frame is typically larger than the IP packet unit, many IP packets are used to transmit I-frame. Because IP packets are minimal in the priority image transmission, all IP packets used to transmit a single frame must share the same transmission priority. Also, the same priority should be given in order to transmit adjacent image data as often as possible to the nearest IP packet. However, some of the critical image information, as shown in Fig. 11, is distributed according to the location of each frame (e.g., the header of P and the B – frame).

As a result, critical network packets are reassigned to the end of the I-frame, as per the following formula:

$$GOP_j = \sum_{i=1}^{T_{pkt}} (HD_{(a)}^{t=1}||HD_{(a)}^{t=2}||VD_{(a)}^{t=1})$$

$$+ \sum_{x=1}^{T_{fr}} \left(\sum_{b=1}^{P_{pkt}^{x}} \left(HD_{(b)}^{t=1}||HD_{(b)}^{t=2}||VD_{(b)}^{t=2} \right) \right.$$

$$\left. || \sum_{c=1}^{B_{pkt}^{x}} \left(HD_{(c)}^{t=1}||HD_{(c)}^{t=2}||VD_{(c)}^{t=3} \right) \right),$$

(3)

where I_{pkt} is total packets of the given I-frame, T_{fr} is the number of inter frames in the GOP, P_{pkt} and B_{pkt} are the total packet numbers of P-frame and B-frame respectively. Also, a, b and c are orders of packets in I-, P- and B-frames. Each frame is composed of two header data (HD^{t-1} and HD^{t-2}) and the corresponding image data (VD^{t-x}), since it (HD^{t-1}) can be placed anywhere in the encoded image.

Also, the re-assignment can be expressed as differently:

$$GOP_j = \sum_{a=1}^{I_{pkt}} \left(HD_{(a)}^{t=1}||HD_{(a)}^{t=2}||VD_{(a)}^{t=1} \right)$$

$$+ \left(\sum_{y=1}^{F_{pkt}^{high}} \left(HD_{(y)}^{t=1}||HD_{(y)}^{t=2} \right) + \sum_{z=1}^{F_{pkt}^{low}} \left(VD_{(z)}^{t=2}||VD_{(z)}^{t=3} \right) \right),$$
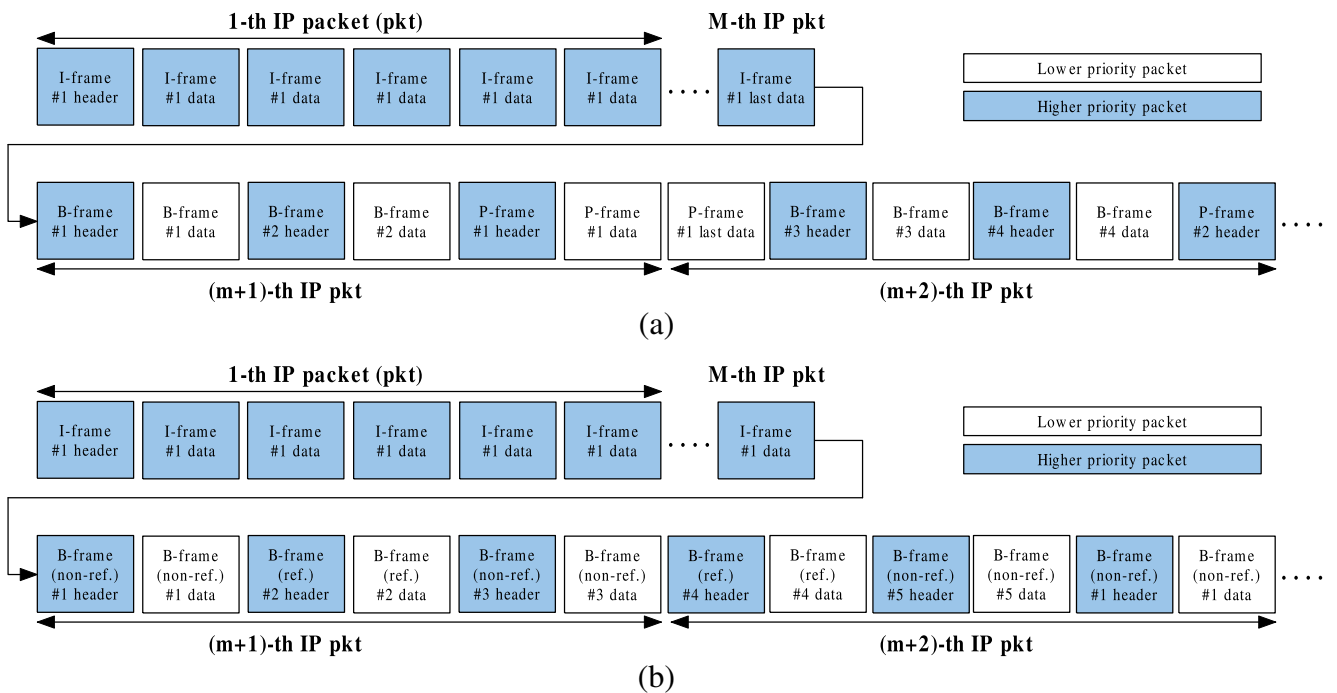
(4)

Fig. 11 Packet structure (before re-assigning) **a** IPPB structure, **b** IBBB structure
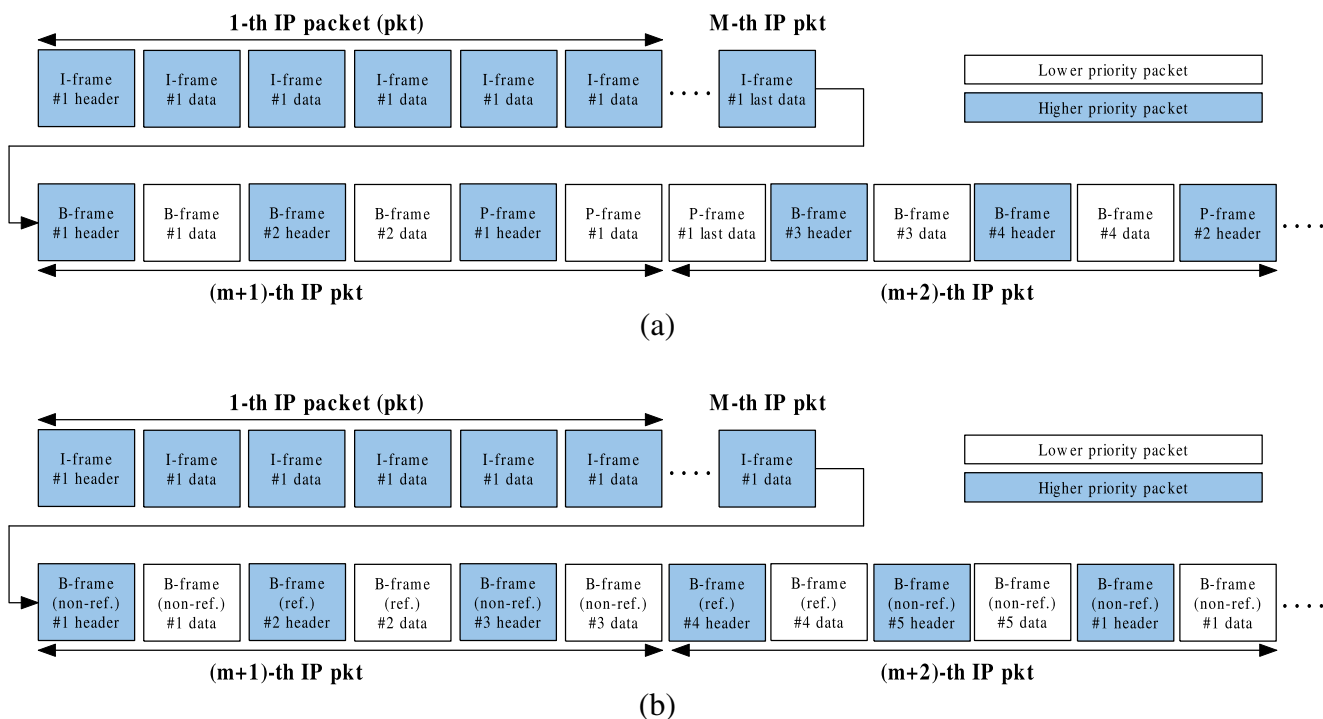
Fig. 12 Packet structure (after re-assigning) **a** IPPB structure, **b** IBBB structure

where F_{pkt}^{high} is total packet number including header data (HD^{t-1} and HD^{t-2}) of the referenced P-, B-frames, and non-referenced frame. F_{pkt}^{low} is total number of including image data (VD^{t-2} and VD^{t-3}) of P-frame, the referenced B-frames, and non-referenced frame. y is an order of image packet including metadata, and z is the order of packet of image data.

The reallocated image data shall be transmitted over the network according to the priority, and the transmit priority of the image packet to the IP packet code DSCP (Differentiated Services Code Point) of the IP packet shall be indicated in the IP packet. Because the size of the image frames is larger than the maximum transmission unit (MTU) of the IP packet, many IP packets are used for transmission. First it sends all high-priority images (m packets) containing a single I-frame (m packet). In the next, mid-priority is to transmit packets containing sensitive image data for decoding, such as P-frame, the referenced B-frames, and non-referenced frame, header data, and image data for decoding, such as image control data.

Finally, packets containing low priority image data are transmitted. The image receiving the image received (Video Client) restores the original sequence of images sent for decoding. The sequential sequence to restore the original order is inserted into the video packet and used to identify the sequence. Figure 13 shows the overall flow chart of the proposed technique. The proposed technique has a data partitioning structure that takes into account the

dependencies and importance of the GOP based picture groups. The header data and image data for high-quality images that contain metadata for high-quality images are grouped according to the type of image and the encoded packet is re-assigned according to the classification. To realign the segmented data, the receiver (Video Client) rearranges the packets in order to decode the images.

Since considering the dependencies of the images within the picture group, the encoding efficiency of existing compression techniques has also been maintained and the error restoration efficiency was also factored into the decoding of video packets to prepare for the loss of the transmission. Thus, the proposed technique allows for improved transmission overhead and data utilization, improvement of objective quality, and subjective quality. In addition, the re-assignment of the re-transmission packet for the larger loss of packets can increase the effectiveness of the resource in terms of network bandwidth. Thereby the proposed scheme can improve the efficiency of the resource use and preventing transmission losses for enhancing QoS.

4 Simulation

In this section, we explain the system configuration for experiment. To evaluate the performance of the proposed selective retransmission technique, x265 is used as HEVC encoding system by FFMPEG library using [31]. Figure 14
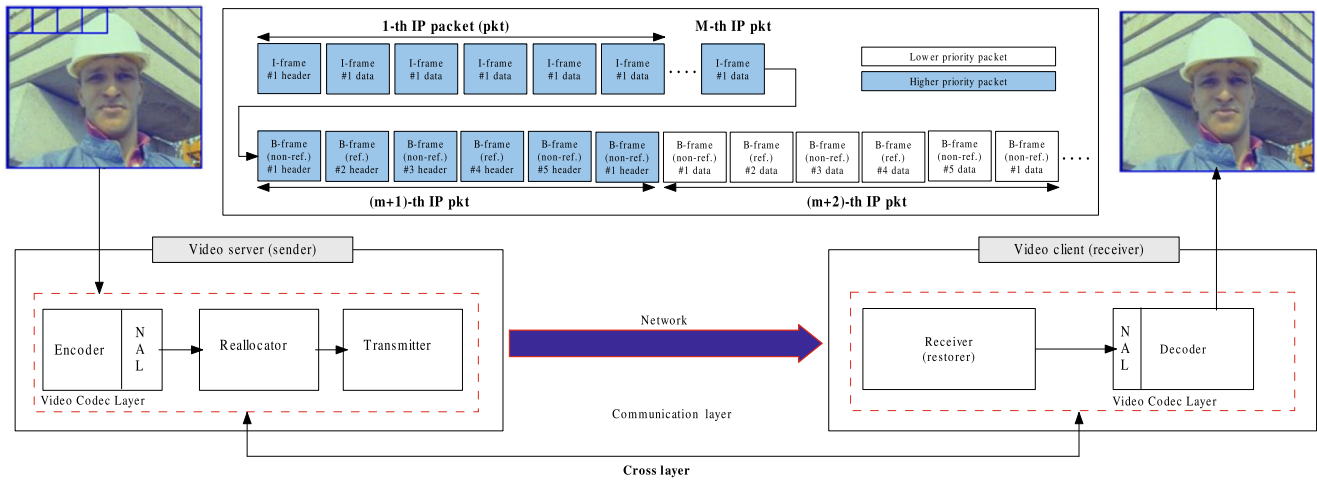
Fig. 13 Overall procedure of the proposed scheme

shows wireless streaming service for simulation. "Cross Module" refers to the proposed technique in the system. For wireless channel, adaptive white gaussian noise (AWGN) is assumed. Also, for performance evaluation in phase streaming services, we can employ test sequences with 1080i (1920 × 1080) resolution.

In terms of users, the PSNR and mean square error (MSE) are employed to measure the service quality in the decoder part. Also the delay performance evaluation may be performed to determine the impact on the entire network through the proposed techniques because of its re-assignment procedure.

5 Conclusion and discussions

In this study, we proposed a QoS enhancement scheme to prevent the loss of high-quality media streaming and video transmissions from the converged 5G network. Based on the layer of the network layer to transmit high-quality media images and video compression images, a new

cross-layer approach was introduced to improve the quality of service improvement by employing more efficient HEVC than H.264/AVC video standard.

The proposed technique has a data partitioning structure that takes into account the dependencies and importance within a cross-layer based picture group. Also it categorizes the priority of header data and image data, including metadata for high-quality images, according to image types. To realign the classified data, we re-configure and re-assign the packet by organizing the sending part and receiving part between the encoding and decoding sides.

By re-assigning the video packet after encoding, the efficiency of the video coder is maintained and the error restoration error is increased. Therefore, it could be useful for streaming services for high-quality multimedia content on a wide variety of handsets, including the next-generation multimedia content, which is available in the converged 5G network, because the proposed technique enables to consider transmission overhead and data utilization, objective quality, and subjective quality.

Fig. 14 Simulation Setup for wireless video streaming service
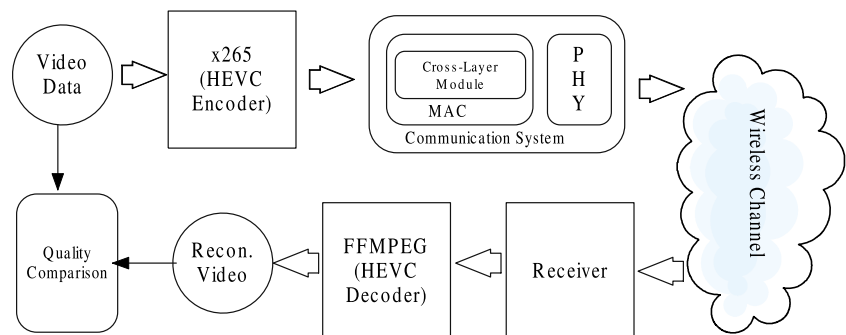
References

1. Abdel KA, Caramanis C, Heath RW (2012) A cross-layer design for perceptual optimization of H. 264/SVC with unequal error protection. IEEE J Sel Areas Commun 30:1157–1171
2. Aign S, Fazel K (1995) Temporal and spatial error concealment technique for hierarchical MPEG-2 video codec. In: IEEE international conference on communications, pp 1778–1783
3. Atzori L, Iera A, Morabito G (2010) The internet of things. A survey. Comput Netw 54:2787–2805
4. A.V.C. for Generic Audio-Visual Services, ITU-T Rec. H.264 and ISO/IEC 14496- 10(AVC). ITU-T and ISO/IEC JTC 1 (2003)
5. Azfar M, Fapojuwo AO (2009) A cross-layer framework for efficient streaming of H. 264 video over IEEE 802.11 networks. J Comput Syst Netw Commun 682813:1–13
6. Chen Y, Yu K, Li J, Li S (2004) An error concealment algorithm for entire frame loss in video transmission. In: IEEE picture coding symposium, vol 1, pp 1–4
7. Chien J-T, Li G-L, Chen M-J (2010) Effective error concealment algorithm of whole frame loss for H.264 video coding standard by recursive motion vector refinement. IEEE Trans Consum Electron 56:1689–1695
8. Cisco (2014) Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018
9. Ericsson (2013) More than 50 billion connected devices. Technical Report (Ericsson) 1
10. Fang T, Chau L (2006) GOP-based channel rate allocation using genetic algorithm for scalable video streaming over error-prone networks. IEEE Trans Image Process 15:1323–1330
11. Foukalas F, Gazis V, Alonistioti N (2008) Cross-layer design proposals for wireless mobile networks: a survey and taxonomy. IEEE Commun Surv Tutor 10:70–85
12. Go K (2016) A systematic reallocation and prioritization scheme for error-resilient trans- mission of video packets. Multimedia Tools Appl 76:6755–6783
13. Goyal VK (2001) Multiple description coding: compression meets the network. IEEE Signal Process Mag 18:74–93
14. Greengrass J, Evans J, Begen AC (2009) Not all packets are equal part II: the impact of network packet loss on video quality. IEEE Internet Comput 13:74–82
15. Guang-Tung J, Chen M-J, Chi M-C (2006) Effective error concealment algorithm by boundary information for H. 264 Video Decoder. In: IEEE international conference on multimedia and expo, vol 1, pp 2021–2024
16. Ha H, Yim C (2008) Layer-weighted unequal error protection for scalable video coding ex- tension of H.264/AVC. IEEE Trans Consumer Electron 54:736–744
17. Jassal A et al (2015) A packet prioritization scheme for 3D-HEVC content transmission over LTE networks. In: International conference on communication workshop (ICCW), pp 1788–793
18. Kim M-H, Park J-H, Na M-S, Jo S-H (2015) Trend of 5G Wireless mobile Communication. Report KICS 1:46–54
19. Kokkonis G, Psannis K, Roumeliotis M, Schonfeld D (2017) Real-time wireless multi- sensory smart surveillance with 3d-hevc streams for internet of things (iot). J Supercomput 73:1044–1062
20. Ksentini A, Naimi M, Guroui A (2006) Toward an improvement of H. 264 video transmission over IEEE 802.11 e through a cross-layer architecture. IEEE Commun Mag 44:107–114
21. Lung-Jen W, Wu C-E, Chang C-Y (2015) A cross-layer based bandwidth and queue adaptations for wireless multimedia networks. Computational intelligence, communication systems and networks (CICSyN), pp 227–232
22. MCKinsey (2014) The Internet of Things: sizing up the opportunity. MCKinsey, USA
23. M of Korea (2014) Statistics of wireless data traffic. In: Government Report
24. Psannis K, Ishibashi Y (2006) Impact of video coding on delay and jitter in 3G wireless video multicast services. EURASIP J Wirel Commun Netw 24614:1–7
25. Psannis K, Ishibashi Y (2008) Enhanced H.264/AVC stream switching over varying bandwidth networks. IEICE ELEX J 5:827–832
26. Shankar S, van der Schaar M (2007) Performance analysis of video transmission over IEEE 802. 11a/e WLANs. IEEE Trans Veh Technol 56:2346–2362
27. Stockhammer T (2011) Dynamic adaptive streaming over HTTP: standards and design principles. In: ACM multimedia systems conference, pp 133–144
28. Sullivan GJ, Ohm J-R, Han W-J, Wiegand T (2012) Overview of the High Efficiency Video Coding (HEVC) Standard. IEEE Trans Circuits Syst Video Technol 22:1649–1668
29. van der Schaar M (2005) Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms. IEEE Wirel Commun Mag 12:55–58
30. Wenger S (2003) H.264/AVC over IP. IEEE Trans Circ Syst Video Technol 13:645–656
31. x265, http://x265.org/. Multicoreware Inc. (2017)
32. Zhang XJ et al (2008) Robust video transmission over lossy network by exploiting H. 264/AVC data partitioning. In: International conference on broadband communications networks and systems, vol 1, pp 1–8

Jong-Hyeok Lee received his BS and MS degrees in the Department of Computer Engineering from SunMoon University, Korea, in 2011 and 2013, respectively. In 2013, he joined the Department of Computer Engineering for pursuing his Ph. D. degree at SunMoon University.

His research interests include image/video signal processing, deep learning algorithm, and video coding standards such as VVC.

Gwang-Soo Hong received his BS and MS degrees in the Department of Computer Engineering from SunMoon University, Korea, in 2011 and 2013, respectively. In 2013, he joined the Department of Computer Engineering for pursuing his Ph. D. degree at SunMoon University.

His research interests include image coding algorithms, video coding technique, deep learning algorithm, and depth estimation algorithms in stereo vision.

Young-Woon Lee received his BS and MS degrees in the Department of Computer Engineering from SunMoon University, Korea, in 2016 and 2018, respectively. In 2018, he joined the Department of Computer Engineering for pursuing his Ph. D. degree at SunMoon University.

His research interests include image coding algorithms, quality enhancement for encoded video, deep learning algorithm as a loop filter for the video codec.

Chang-Ki Kim received BS and MS degrees in Electronics Engineering from the Pusan National University, Pusan, Korea, in 1995 and 1997, respectively. He also received MS degree in software engineering from KAIST in 2006. In 2016, he received a PhD degree in the Department of Computer Engineering from SunMoon University. He worked as an engineer in the Samsung Electronics from 1997 to 2000. Since 2001, he has been a principal researcher at Electronics and Telecommunications Research Institute (ETRI).

His main research interests include mobile communication, radio network protocols design, video coding and multimedia transport protocol.
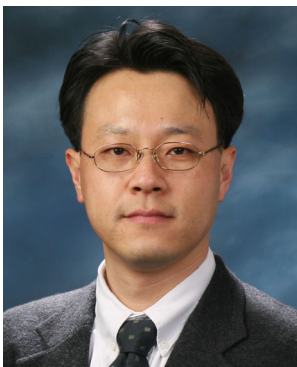
Noik Park received BS, MS and PhD degree in Industrial Engineering from the Sung Kyun Kwan University Korea, in 1993, 1995, 2000 respectively. Since 2000, he works in Electronics and Telecommunications Research Institute (ETRI) Korea.

His main research interests include mobile communication network, network optimization, etc.

Byung-Gyu Kim has received his BS degree from Pusan National University, Korea, in 1996 and an MS degree from Korea Advanced Institute of Science and Technology (KAIST) in 1998. In 2004, he received a PhD degree in the Department of Electrical Engineering and Computer Science from Korea Advanced Institute of Science and Technology (KAIST). In March 2004, he joined in the real-time multimedia research team at the Electronics and Telecommunications Research Institute (ETRI), Korea where he was a senior researcher. In ETRI, he developed so many real-time video signal processing algorithms and patents and received the Best Paper Award in 2007. From February 2009 to February 2016, he was associate professor in the Division of Computer Science and Engineering at SunMoon University, Korea. In March 2016, he joined the Department of Information Technology (IT) Engineering at Sookmyung Women's University, Korea where he is currently an associate professor. In 2007, he served as an editorial board member of the International Journal of Soft Computing, Recent Patents on Signal Processing, Research Journal of Information Technology, Journal of Convergence Information Technology, and Journal of Engineering and Applied Sciences. Also, he is serving as an associate editor of Circuits, Systems and Signal Processing (Springer), The Journal of Supercomputing (Springer), The Journal of Real-Time Image Processing (Springer), and International Journal of Image Processing and Visual Communication (IJIPVC). From March 2018, he is serving as the Editor-in-Chief of The Journal of Multimedia Information System and associate editor of IEEE Access Journal. He also served or serves as Organizing Committee of CSIP 2011, a Co-organizer of CICCAT2016/2017, and Program Committee Members of many international conferences. He has received the Special Merit Award for Outstanding Paper from the IEEE Consumer Electronics Society, at IEEE ICCE 2012, Certification Appreciation Award from the SPIE Optical Engineering in 2013, and the Best Academic Award from the CIS in 2014. He has been honored as an IEEE Senior member in 2015. He is serving as a professional reviewer in many academic journals including IEEE, ACM, Elsevier, Springer, Oxford, SPIE, IET, MDPI, and so on.

He has published over 200 international journal and conference papers, patents in his field. His research interests include image and video signal processing for the content-based image coding, video coding techniques, 3D video signal processing, deep/reinforcement learning algorithm, embedded multimedia system, and intelligent information system for image signal processing. He is a senior member of IEEE and a professional member of ACM, and IEICE.