

Performance Analysis of Future Video Coding (FVC) Standard Technology

Young-Ju Choi¹, Ji-Hae Kim¹, Jong-Hyeok Lee², Byung-Gyu Kim^{1,*}

Abstract

The Future Video Coding (FVC) is a new state of the art video compression standard that is going to standardize, as the next generation of High Efficiency Video Coding (HEVC) standard. The FVC standard applies newly designed block structure, which is called quadtree plus binary tree (QTBT) to improve the coding efficiency. Also, intra and inter prediction parts were changed to improve the coding performance when comparing to the previous coding standard such as HEVC and H.264/AVC. Experimental results shows that we are able to achieve the average BD-rate reduction of 25.46%, 38.00% and 35.78% for Y, U and V, respectively. In terms of complexity, the FVC takes about 14 times longer than the consumed time of HEVC encoder.

Key Words: Future Video Coding (FVC), High Efficiency Video Coding (HEVC), Video compression, Quadtree plus binary tree (QTBT), Inter prediction, Intra prediction

I. INTRODUCTION

The Future Video Coding (FVC) is a state of the art video compression standard that has been standardized since High Efficiency Video Coding (HEVC), has been studied by the Joint Video Exploration Team (JVET) of ITU-T VCEG and ISO/IEC MPEG. FVC provides a software model called the Joint Exploration Model (JEM) [1], which was released up to version 5.0.

With the development of display resolution and types, video compression technology is required for higher-resolution and the newest services such as VR, AR. For that reason, the FVC standard has begun to work newly.

The Joint Exploration Test Model (JEM) of FVC was developed based on the HEVC Test Model (HM) [2],[3],[4],[5]. Consequently, the basic framework of encoding and decoding is the same with the HEVC, however, the internal coding tools of modules of block structure, intra and inter prediction and transform, loop filter and entropy coding are added and modified.

In this paper, we introduce new coding tools to improve the coding efficiency and analyze them in detail.

Especially, we have three coding parts: Block structure, intra prediction, and inter prediction mechanisms. Block structure defines the coding unit in encoding procedure and intra prediction method defines how to make the best residual image signal in the picture. Finally, inter prediction usually is about how to get the minimized residual signal from the reference pictures (multiple reference pictures).

This paper is organized as follows. In section 2, an overview of FVC is presented by detailing its major aspects based on Joint Exploration Test Model (JEM) 5.0. Section 3 presents the video coding performance and the average time distribution. Section 4 concludes this paper and presents some ideas for future work.

II. OVERVIEW OF CODING FEATURES OF JEM 5.0

In this section, we would like to introduce major coding tools which have been proposed in JVET.

Manuscript received June 26, 2017; Revised June 27, 2017; Accepted June 27, 2017. (ID No. JMIS-2017-0015)

Corresponding Author (*): Byung-Gyu Kim, Dept. of IT Engineering, Sookmyung Women's University, Seoul, Korea, Tel: 82-2-2077-7293, E-mail: bg.kim@sm.ac.kr.

¹Dept. of IT Engineering, Sookmyung Women's University, Seoul, Korea, E-mail: {yj.Chio, jh.Kim}@vicl.sookmyung.ac.kr, bg.kim@sm.ac.kr

²Dept. of Computer Engineering, SunMoon University, Asan, Korea, E-mail: {jh.Lee, yw.Lee}@vicl.sookmyung.ac.kr

2.1. Block Structure

Figure 1 illustrates important block structure in FVC standard. In HEVC, CTU is split into CUs through the quadtree structure. Each CU is divided into PU or TU according to the module to be performed, and each unit size applied to each process is different. Unlike HEVC standard, FVC adopts quadtree plus binary tree (QTBT) [6] structure which eliminate this multiple partition type concept and support more flexible CU shapes.

In the QTBT structure, CTU is first divided by a quadtree types, and quadtree leaf nodes are divided by binary tree types. There are two types of binary trees: horizontal and vertical. The leaf nodes of this binary tree are called CUs.

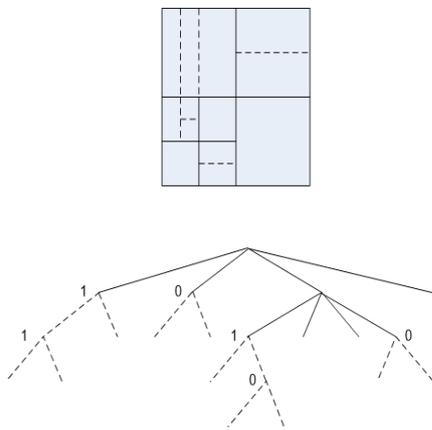


Fig. 1 Illustration of quadtree plus binary tree (QTBT) structure.

The prediction and transform are performed with the corresponding unit size type without any further partitioning. For the QTBT structure, some parameters are defined, which are CTU size (the quadtree root node size), MinQTSsize (the minimum quadtree leaf node size), MaxBTSsize (the maximum binary tree root node size), MaxBTDepth (the maximum binary tree depth), MinBTSsize (the minimum binary tree leaf node size).

In Figure 1, the first figure illustrates a QTBT structure block partitioning example, and the second illustrates the tree structure corresponding with that of above side. The solid lines mean quadtree splitting and dotted lines mean binary tree splitting. In binary tree splitting, flag 0 means horizontal splitting and flag 1 means vertical splitting.

2.2. Intra Prediction

In HEVC, there are 35 modes such as planar, DC, and 33 directional modes in the intra prediction mode. In FVC, the directional intra modes extended to 65, as a result,

FVC has 67 modes and adopts methods to reduce candidate modes for each of luma and chroma [7],[8],[9],[10]. In Figure 2, the black line means the existing HEVC directional mode, and the red line means the directional mode newly added in FVC.

In order to reduce the candidate mode for the luma intra mode coding of the FVC which nearly doubled the number of directional intra modes, 6 Most Probable Modes (MPMs) are employed as shown in Figure 3. The intra prediction modes are classified into three groups, and three steps are taken to create an MPM list with 6 candidates. The three groups are neighbouring intra modes, derived intra modes, and default intra modes. First, in neighbouring intra modes group, add modes of 5 neighbouring blocks, i.e., left(L), above(A), below left(BL), above right(AR), and above left(AL) as shown in Figure 3, planar mode, and dc mode to make a MPM list. If MPM list is not full, go to the group of derived intra modes. In derived intra modes group, -1 or +1 angular modes of the already included angular modes are added. If MPM list still is not full, add the default modes: vertical, horizontal, mode 2, and diagonal mode. As a result, a list of 6 MPM modes is created.

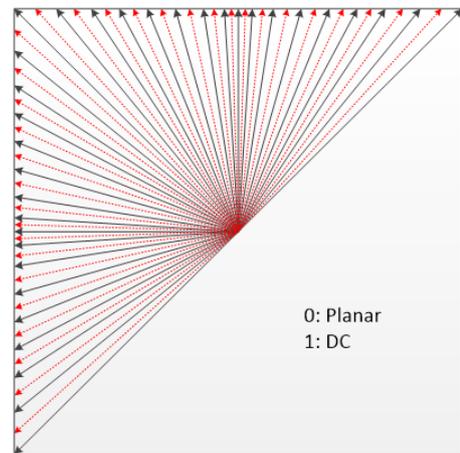


Fig. 2 Illustration of intra prediction modes.

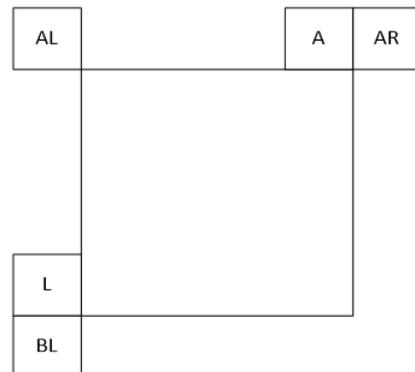


Fig. 3 Neighbouring blocks for MPM derivation.

In chroma intra mode coding, a total of 11 modes are selected as the candidate mode list. It consists of 5 traditional intra modes: non-CCLM modes and 6 CCLM [7],[12] modes. Non-CCLM mode refers to the spatial neighbouring block mode as in luma, and the CCLM mode use the luma mode of reconstructed luma sample after encoding for predicting a chroma mode.

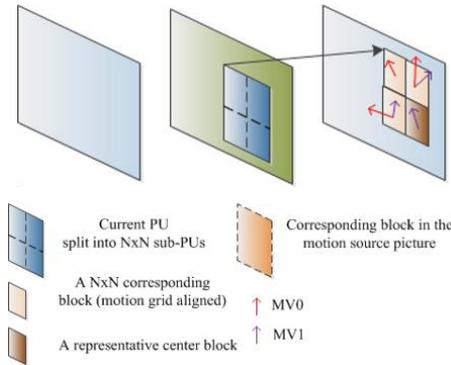


Fig. 4 ATMVP motion vector prediction.

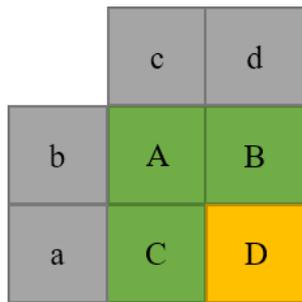


Fig. 5 Example of one CU with four sub-blocks (A-D) and its neighbouring blocks (a-d).

2.3. Inter Prediction

In comparison to the most important algorithms with HEVC, in HEVC, there is the Advance Motion Vector Prediction (AMVP) [11] method which uses motion vectors of 5 spatial neighbors, 2 temporal motion vectors, and zero motion vectors as candidate list and employs them for prediction. In FVC, the Advanced temporal motion vector prediction (ATMVP) and Spatial-temporal motion vector prediction (STMVP) methods are used which improve on AMVP [7],[12],[13]. Both of the aforementioned methods are accessed using the sub-CU. ATMVP is a method for predicting the current CU of the current frame using the motion vector multiple set information of the sub-CU at the same position of the reference frame as shown in Figure 4.

STMVP is a method for motion vector candidate set list is generated by using temporal sub-CUs as the same procedure of TMVP as specified in HEVC and spatial neighboring sub-CU as shown in Figure 5, and then used

for predicting.

III. PERFORMANCE ANALYSIS

The reference software model JEM 5.0 [14] and HM 16.10 [15] are used to experiment. Evaluation occurred under 3 profile of all intra, low delay-P and random access. Used standard sequences are 'ParkScene' in class B, 'BasketballDrill' in class C and 'FourPeople' in class E as shown in Table 1. Four base layer QP values of 22, 27, 32, and 37 were used with 32 total frames in each test sequence.

Table 1. Video sequences considered in the experiments

Class	Sequence	Resolution	Frame rate
B	ParkScene	1920x1080	24
C	BasketballDrill	832X480	50
E	FourPeople	1280X720	60

For performance comparison, we employed average BD-rate [16] and the difference of the consumed encoding time (EncT). Bjontegaard's metric (BD-rate) allows to compute the average gain in PSNR or the average percent saving in bitrate between two rate-distortion curves [16].

The difference of the consumed encoding time (EncT) can be defined as:

$$\text{EncT} = \frac{T_{JEM5} - T_{HM16.10}}{T_{HM16.10}} \times 100\%. \quad (1)$$

Here, T means the checked time. If EncT becomes large positively, then JEM5.0 encoder get more time to process with the same video content and setting. It means that JEM5.0 is slower than HM16.10 software encoder. When EncT is a negative value, it means that JEM5.0 encoder is faster than HM16.10.

Table 2. Coding efficiency comparison of FVC and HEVC in BD-rate and time complexity.

	Y	U	V	EncT
All Intra	-22.41%	-31.41%	-27.44%	6315%
Low Delay_P	-25.72%	-42.29%	-40.69%	515%
Random Access	-28.24%	-40.32%	-39.21%	854%
Overall	-25.46%	-38.00%	-35.78%	1405%

3.1. Coding Efficiency

Table 2 shows results of variation ratio of BD-rate and time complexity of JEM comparing with HM. For Y, U and V, average BD-rate reduction ratio of 25.46%, 38.00% and 35.78% are accomplished, respectively. We could see that the reduction value of luma was greater than that of chroma. The time complexity variation of all intra has resulted in surprising results. JEM takes about 6.3 times longer than HM encoder. The results of low delay P and random access are about 5.1 and 8.5 times, respectively. Consequentially, in overall, JEM takes about 14 times longer than HM.

Figure 6 presents rate distortion curves of JEM 5.0 and HM 16.10 for the BasketballDrill video sequence. These results indicate the improvement of the coding efficiency of the new video compression standard. For all intra configuration, we can observe the PSNR improvement of 2.0 dB ~ 2.7dB at same bitrate. In case of low delay P, about 2.0 dB of the quality improvement has been achieved. For random access configuration, over 2.5 dB of PSNR improvement was observed in 2000 Kbps. At this time, the PSNR is over 40dB. This is much improvement when comparing to the HEVC standard.

3.2 The Consumed Time Profile

In order to determine which of the coding components consume most time, we set a timer in encoder. Figure 7 shows the time division of the JEM 5.0 encoder for 3 profile: all intra, low delay P, random access. The Transform and Quantization represent the highest percentage of the coding time in all intra. In low delay p, the inter prediction represent nearly 50% of the encoding time. In random access, consumes more than 60% of the inter prediction encoding time and 27% of the intra prediction encoding time.

For this reason, we can assume that the I-frame coding consumes the most time in the transform and quantization and the P-frame or B-frame coding consumes the most time in the inter prediction coding.

IV. CONCLUSIONS

In this paper, a comprehensive coding features for the FVC standard comparing with HEVC have introduced and analyzed. The most important feature of the FVC is the block structure with a QTBT structure which simplify the coding units and improve the coding efficiency. The improved intra and inter prediction changes also contributed to performance improvements. Through experiments, the FVC based on JEM 5.0 software provides an average BD-rate reduction of 25.46%, 38.00%

and 35.78% for Y, U and V, respectively. Also the FVC technology took about 14 times longer than HEVC encoding system. For future work, we need to more analysis to improve the coding efficiency of the FVC.

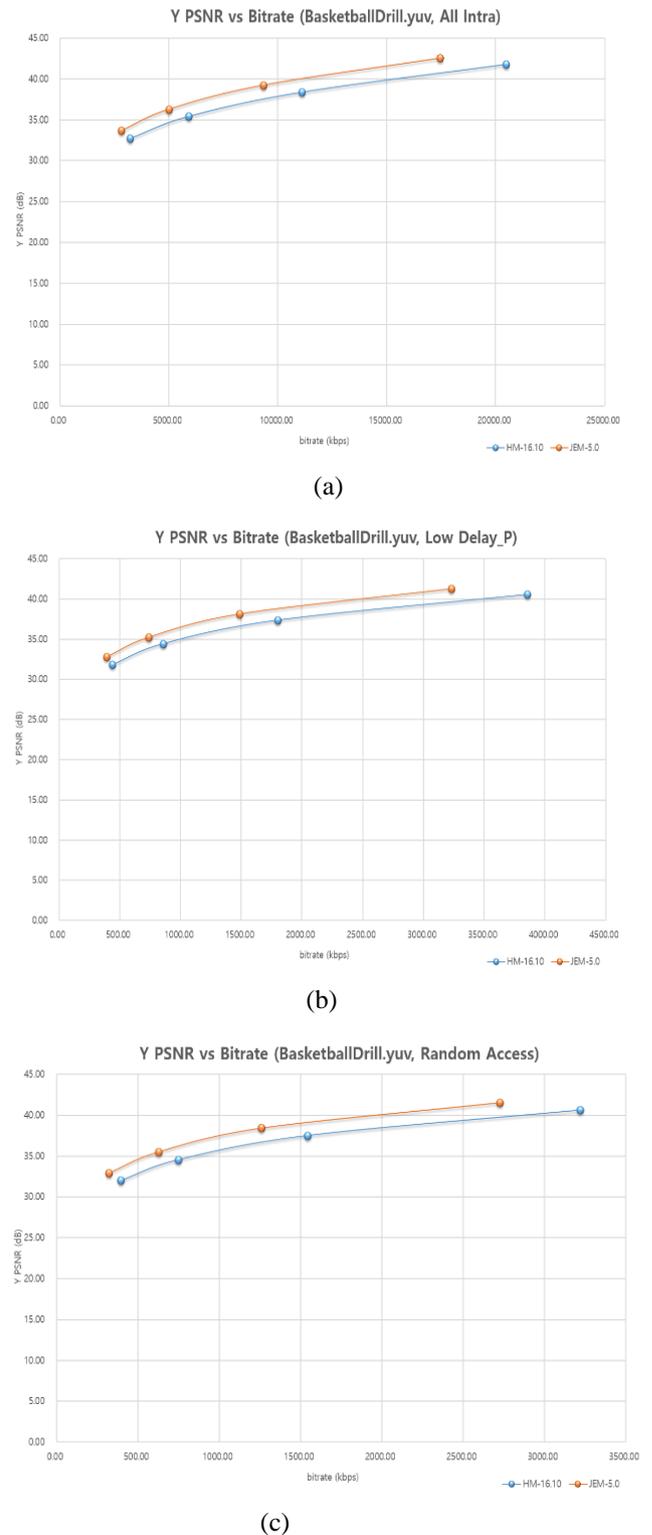
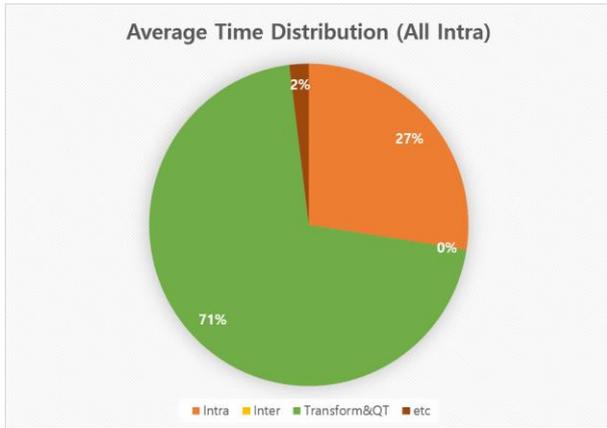
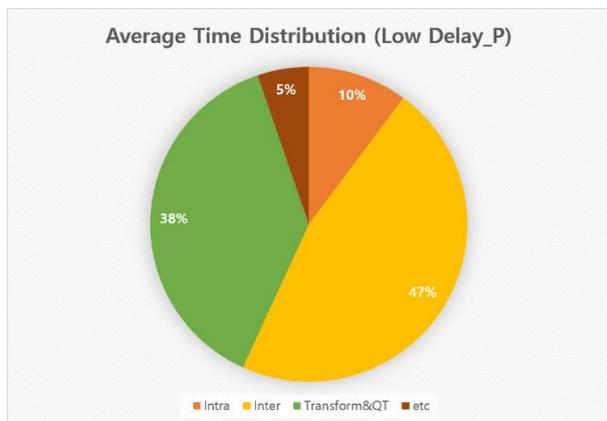


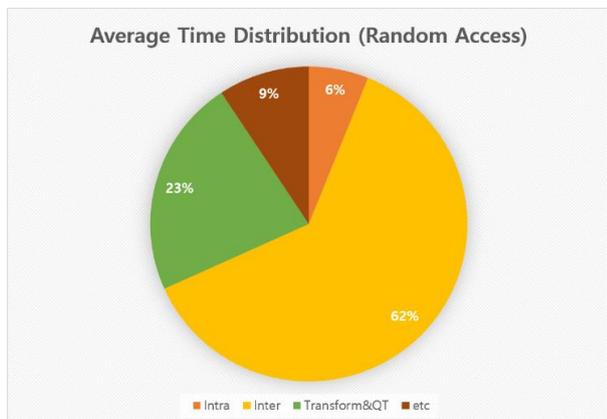
Fig. 6 Rate distortion curves for: JEM 5.0, HM 16.10 for BasketballDrill sequence: (a) all intra, (b)low delay p, and (c)random access.



(a)



(b)



(c)

Fig. 7 Average consume time profiles for: (a) all intra, (b) low delay p, and (c) random access

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2016R1D1A1B04934750).

REFERENCES

- [1] J. Chen, Qualcomm Inc., E. Alshina, Samsung Electronics, Gary J. Sullivan, Microsoft Corp., Jens-Rainer Ohm, RWTH Aachen University, Jill Boyce and Intel, "Algorithm Description of Joint Exploration Test Model 5 (JEM 5)," *Joint Video Exploration Team (JVET) document JVET-E1001, 5th Meeting*, Geneva, CH, Jan. 12-20, 2017.
- [2] G. J. Sullivan, J. R. Ohm, W.-J. Han, T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Transaction on Circuits and Systems for Video Technology*, Vol. 22, No. 12, pp. 1649-1668, DEC. 2012.
- [3] High Efficiency Video Coding (HEVC), Rec. ITU-T H.265 and ISO/IEC 23008-2, Jan. 2013.
- [4] C. Rosewarne, B. Bross, M. Naccari, K. Sharman, and G. J. Sullivan, "High Efficiency Video Coding (HEVC) Test Model 16 (HM 16) Update 4 of Encoder Description", *Joint Collaborative Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11 on Video Coding, JCTVC-V1002, 22nd Meeting*, Oct. 2015.
- [5] X. Chen, J. An, and J. Zheng, "EE3: Decoder-Side Motion Vector Refinement Based on Bilateral Template Matching", *Joint Video Exploration Team of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, JVET-E0052, 5th Meeting*, Jan. 2017
- [5] H. Schwarz, D. Marpe, T. Wiegand, "Further Results on Constrained Inter-layer prediction," *Joint Video Team, Doc. JVT-O074*, Apr. 2005.
- [6] J. An, Y.-W. Chen, K. Zhang, H. Huang, Y.-W. Huang, and S. Lei, "Block partitioning structure for next generation video coding", *MPEG doc. m37524 and ITU-T SG16 Doc. COM16-C966*, Oct. 2015.
- [7] M. Karczewicz, J. Chen, W.-J. Chien, X. Li, A. Said, L. Zhang, and X. Zhao, "Study of coding efficiency improvements beyond HEVC", *MPEG doc. m37102*, Oct. 2015.
- [8] J. Chen, W.-J. Chien, M. Karczewicz, X. Li, H. Liu, A. Said, L. Zhang, and X. Zhao, "Further improvements to HMKTA-1.0", *ITU-T SG16/Q6 Doc. VCEG-AZ07*, Jun. 2015.
- [9] E. Alshina, A. Alshin, J.-H. Min, K. Choi, A. Saxena, and M. Budagavi, "Known tools performance investigation for next generation video coding", *ITU-T SG16/Q6 Doc. VCEG-AZ05*, Jun. 2015.
- [10] K. Choi, E. Alshina, A. Alshin, and C. Kim, "Information on coding efficiency improvements over HEVC for 4K content", *MPEG doc. m37043*, Oct. 2015.
- [11] B. Bross, "High efficiency video coding (HEVC) text specification draft 6", *JCTVC-H1003, 8th JCT-VC*

Meeting, San Jose, CA, USA, 1-10 February, 2012.

- [12] J. Chen, Y. Chen, M. Karczewicz, X. Li, H. Liu, L. Zhang, and X. Zhao, "Coding tools investigation for next generation video coding", ITU-T SG16 Doc. COM16-C806, Feb. 2015.
- [13] W.-J. Chien and M. Karczewicz, "Extension of Advanced Temporal Motion Vector Predictor", ITU-T SG16/Q6 Doc. VCEG-AZ10, Jun. 2015.
- [14] "Future Video Coding FVC test model 5.0 (JEM 5.0)" https://jvet.hhi.fraunhofer.de/svn/svn_HMJEMSoftware/tags/HM-16.6-JEM-5.0/
- [15] "High efficiency video coding HEVC test model 16.10 (HM 16.10)" https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/HM-16.10/
- [16] G. Bjontegaard, Calculation of average PSNR differences between RD curves, ITU-T SG16/Q6, Texas, US, Doc. VCEG-M33, April, 2001.

Authors



Young-Ju Choi received her BS degrees in the Department of IT Engineering from Sookmyung Women's University, Korea, in 2017. In 2017, he joined the Department of IT Engineering for pursuing her MS degree at Sookmyung Women's University. Her research interests include image and video coding techniques, deep learning-

based video recognition, and intelligent/perceptual video coding algorithms.



Ji-Hae Kim received her BS degrees in the Department of IT Engineering from Sookmyung Women's University, Korea, in 2017. In 2017, he joined the Department of IT Engineering for pursuing her MS degree at Sookmyung Women's University. Her research interests include video signal processing,

facial expression recognition, emotional expression, and video coding techniques.



Jong-Hyeok Lee received his BS and MS degrees in the Department of Computer Engineering from SunMoon University, Korea, in 2011 and 2013, respectively. He is a PhD candidate in Department of Computer Engineering at SunMoon University, Korea.

His research interests include video signal processing, high efficiency video coding, and coding technique of high dynamic range/wide color gamut imaging.



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), Helyion Journal (Elsevier), and International Journal of Image Processing and Visual Communication (IJIPVC). He also served as Organizing Committee of CSIP 2011 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 has published over 150 international journal and conference papers, patents in his field. His research interests include software-based image and video object segmentation for the content-based image coding, video coding techniques, 3D video signal processing, wireless multimedia sensor network, embedded multimedia communication, and intelligent information system for image signal processing. He is a senior member of IEEE and a professional member of ACM, and IEICE.