Journal of Korea Multimedia Society Vol. 21, No. 8, August 2018(pp. 813-820)
https://doi.org/10.9717/kmms.2018.21.8.813

Adaptive Frame Rate Up-Conversion Algorithms using Block Complexity Information

Kangjun Lee†

ABSTRACT

This paper proposes new frame rate up-conversion algorithms. Adaptive motion estimation based on block complexity information are used to obtain more accurate motion vectors. Because the information on block complexity is extracted from the motion estimation prediction size from the original frame, additional computational complexity is not imparted. In experimental results, the proposed algorithms provide robust frame interpolation performance for whole test sequences. Also, the computational complexity of the proposed algorithm is reduced to a benchmark algorithm.

Key words: Frame Rate Up-conversion, Motion Compensated Interpolation, Motion Estimation

1. INTRODUCTION

Frame rate up-conversion (FRUC) produces image sequences with higher frame rates by adding new frames into image sequences compared to lower frame rates. As FRUC is commonly used for video format conversion, the image sequence is displayed in smooth motion and with high quality. Recently, FRUC has attracted attention as an application to reduce the motion blur of hold-type displays such as liquid crystal displays[1]. With the appearance of high-resolution display devices, studies on the impact of frame rate on perceived video quality have become popular[2], [3].

Among the existing FRUC approaches, motion-compensated frame rate up-conversion (MC-FRUC) has been shown to be an effective scheme for reducing motion judder[4]. The basic elements of MC-FRUC are motion estimation (ME) and motion compensated interpolation (MCI). Since ME has a large impact on performance, the structure of the overall FRUC method is determined using ME approaches. The block-matching algorithm (BMA) for ME is commonly used because it is simple and easy to implement [5 - 7].

An ME method based on the BMA for FRUC is unidirectional ME (UME) [8-10]. In this method, a motion vector is estimated between previous and following frames, and then the MCI generates the new frame through an interpolation process using the motion trajectories. In the UME, when a pixel in an interpolated frame is passed through by multiple motion trajectories and no motion trajectory, overlapping and hole problems occur. Due to the overlaps and holes with irregular shapes and sizes, most UME methods are quite complex because of the post processing required to address these problematic regions.

To overcome the problem of holes and overlapping regions in the UME, bilateral ME (BME) uses a unique motion vector (MV) that is assigned for each interpolated block[11]. It is based on the assumption of temporal symmetry between the previous blocks and subsequent frames from the viewpoint of the interpolated block. However, the BME cannot ensure the accuracy of an ME when

※ Corresponding Author : Kangjun Lee, Address: (02707) Kookmin University, Jeongneung-ro 77, Seongbuk-gu, Seoul, Korea, TEL : +82-2-910-5542, FAX : +82-2-910-5263, E-mail : kangjun@kookmin.ac.kr

Receipt date : Jul. 9, 2018, Approval date : Jul. 20, 2018
† Department of Automobile and IT Convergence, Kookmin University

the MV of an object between the previous and subsequent frames has no temporal symmetry with respect to the interpolated block. The BME has a critical weakness when an object with complex texture and a background with homogeneous texture are simultaneously present[12]. To overcome the problems of BME, several methods have been introduced[13 - 18]. Although the accuracy of MV search is improved by proposed algorithms, its computational complexity is also increased. The main factors of the increased computational complexity are the excessive search range, the iterative search in the ME process, and the additional MV refinement (MVR) with variable size.

This paper proposed an adaptive ME algorithm based on image block complexity. Depending on the block complexity, various BME methods are selectively adopted. Additionally, the search range of the BME is adaptively selected. This paper is organized as follows. Section 2 presents the different characteristics of the ME process according to block complexity. Section 3 describes the proposed FRUC algorithm in detail. Experimental results are reported in Section 4. Finally, the conclusion is presented in Section 5.

2. CHALLENGES IN EXISTING APPROACHES AND A PROPOSED SOLUTION

2.1 Motion estimation in a unidirectional approach

The UME approach utilizes the BMA that matches blocks between two adjacent frame images.

$$SAD(dx,dy)=\sum_{(x,y)\in B_{ij}}\left|f_{n-1}(x+dx,y+dy)-f_{n+1}(x,y)\right|$$
$$V_{ij}=\underset{(dx,dy)\in SR}{\arg\min}\left\{SAD(dx,dy)\right\} \quad (1)$$

where $f_{n-1}(x,y)$ and $f_{n+1}(x,y)$ are the pixel values in the previous and subsequent frames, respectively. V_{ij} is the MV for block B_{ij} that minimizes the SAD in (1). SR is the predefined search range. Although the UME finds the MV close to real MV field in complex texture, this SAD criterion causes hole and overlap regions. A number of algorithms have been proposed to handle these problems[19 - 22]. However, these methods demand complicated operations and can produce undesirable artifacts.

2.2 Motion estimation in a bilateral approach

To avoid the problem of holes and overlaps, many algorithms based on the BME have been proposed. The BME estimates an MV using the temporal symmetry between blocks in the previous and subsequent frames from the viewpoint of the interpolated block (IB_{ij}) of the interpolated frame. The sum of the bilateral absolute difference (SBAD) is used to determine the best MV using (2).

$$SBAD(dx,dy,IB_{ij})=\sum_{(x,y)\in IB_{ij}}\left|f_{n-1}(x+dx,y+dy)-f_{n+1}(x-dx,y-dy)\right|$$
$$V_{ij}=\underset{(dx,dy)\in SR}{\arg\min}\left\{SBAD(dx,dy,IB_{ij})\right\} \quad (2)$$

The BME provides acceptable motion trajectory if the image sequences contain simple translational motions. However, the BME has a critical weakness when an object with complex texture and a background with homogeneous texture are simultaneously present[14]. In the system in Fig. 1, MV1 is selected as the best MV in SR1 and is matched to the true motion trajectory. As the search range is extended in SR2 including a homogenous background, MV2 is selected due to its minimal SBAD value, which is an incorrect MV search result. To overcome these problems, the cost function based on side match distortion is used to determine a more accurate MV[13]. However, massive compu-

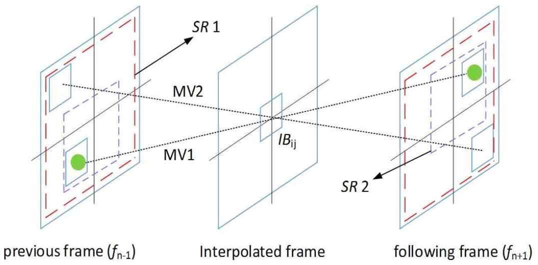


Fig. 1. Incorrect ME result in the BME process.

tational complexity is required for the iterative MV search process. In extended bilateral motion estimation (EBME), the accuracy of the ME is improved by the expanded range of the motion trajectory[14]. Although the quality of interpolated frames is increased by the EBME, the computational complexity is also increased. Kim et al.[12] proposed the BME method based on a similarity value considering texture (SVCT).

Although the computational complexity is reduced by using the new cost function and the early termination of the ME process with a threshold, the visual quality of interpolated frames fluctuates according to the image sequence. In Ref. [15-16], the UME is first used to estimate the initial MV, and then BME-based motion vector refinement algorithms are used to improve the accuracy of the MV search result. Although the accuracy of the MV search is improved by these approaches, its computational complexity is also increased by the additional UME process.

2.3 Block complexity information

In previous research, when the characteristic image sequence is not considered in the ME process, unnatural defects in interpolated images and additional computational complexity are created. To overcome these problems, an adaptive ME method based on image complexity characteristics, which indicate texture and edges of the image, is proposed. However, measuring the image complexity characteristic causes additional computational complexity. To predict the image complexity without additional computational burden, we use the ME prediction size of original frames from the transmitted bit-stream. Fig. 2 shows various block sizes for the inter-prediction of the H.264/ AVC standard.

The region coded with the small ME prediction size is the edge of the moving object or an object with a complex texture[23]. Therefore, the image complexity characteristic is estimated from the ME

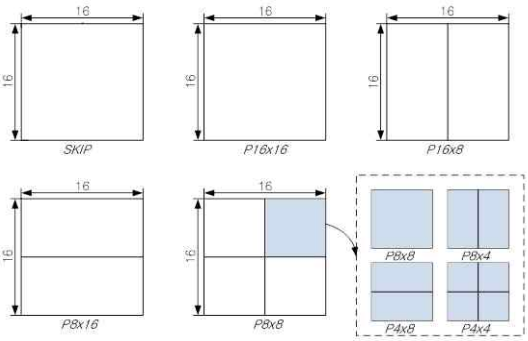


Fig. 2. Block partitions of inter-modes for inter-blocks in H.264/AVC.

prediction size in original frames. Therefore, according to the image complexity characteristic, the following two processes are adaptively used.

- The original BME or the BME with the predictive MV.
- Search range in the BME.

3. PROPOSED FRUC ALGORITHMS

To adjust adaptive ME algorithms according to image complexity characteristics, the IB is used for the basic decision unit for determining the block complexity. As shown in Fig. 3, IB_{12} and IB_{32} represent the basic decision unit in the interpolated frame. To determine the block complexity of the basic decision unit, the coded ME prediction size in SR is used. $FSR(IB_{ij})$ and $BSR(IB_{ij})$ indicate the ME prediction size in the SR of the IB in frames f_{n-1} and f_{n+1}, respectively. In Fig. 3, FSR

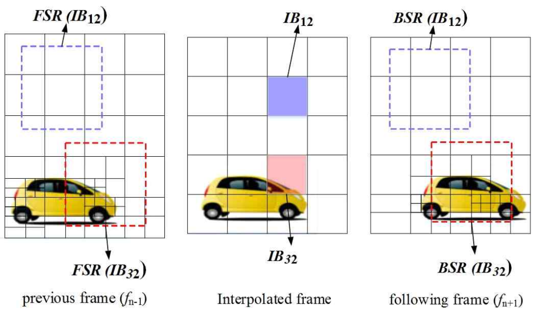


Fig. 3. Block complexity according to ME prediction size.

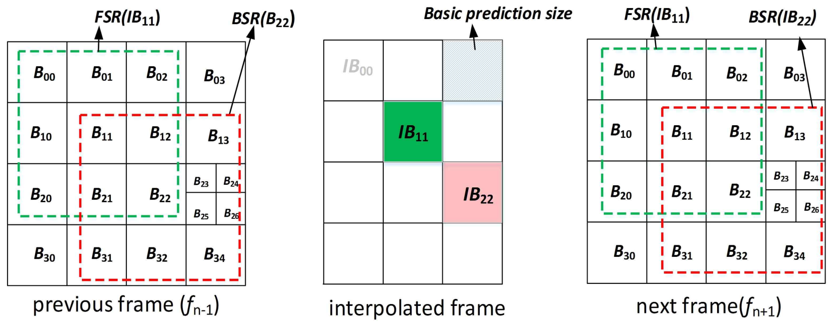


Fig. 4. Various block complexities of *IBs*.

(IB_{12}) and BSR(IB_{12}) are composed in basic prediction size in previously coded frames. In this situation, the probability that IB_{12} is a part of a background with homogeneous texture is high[23]. Therefore, the BME around the original point with the small search range is sufficient to determine the true motion trajectory.

On the other hand, FSR(IB_{32}) and BSR(IB_{32}) for IB_{32} are composed using a smaller prediction size in previously coded frames, which signifies that this region contains a moving object with a complex texture. Therefore, the BME with the predictive MV (PMV) is required in this region. The block complexity (COM_{ij}) of IB_{ij} is determined as follows.

$$COM_{ij} = \begin{cases} 1, & if\ FSR(IB_{ij})\ and\ BSR(IB_{ij}) = basic \\ 0, & otherwise \end{cases} \tag{3}$$

In Equation (3), *basic* means the basic prediction size in frames f_{n-1} and f_{n+1}. As shown in Fig. 4, FSR(IB_{11}) and BSR(IB_{11}) are composed of {B_{00}, B_{01}, B_{02}, B_{10}, B_{11}, B_{12}, B_{20}, B_{21}, and B_{22}}, respectively. The prediction size here is the same as the basic prediction size. For example, in the H.264/AVC video coding standard, 16×16 is the basic prediction size. As previously mentioned, the probability that IB_{11} is part of a background with a homogeneous texture is high. In the case of IB_{22}, a smaller prediction size is included in PSR(IB_{22}) and NSR (IB_{22}). In this situation, there is high probability that this region is part of a complex moving object.

Therefore, we propose adaptive ME selection algorithm. Fig. 5 shows the sequence diagram of these adaptive ME algorithms according to the information on block complexity. If the COM_{ij} is 1 (homogeneous region), the BME is executed with a reduced search range. If the COM_{ij} is 0 (complex region), the PMV is required for a more precise ME search result. The UME, which shows good MV search performance in the complex texture region, is first used for finding the PMV. Then, the PMV for the BME is estimated from the projection process. As shown in Fig. 6, the motion vector with the largest mapping area is assigned as the PMV for the BME[24]. Experimental results confirm that the proposed algorithm provides good performance when the search range is 3 in the homogeneous region and the search range is 2 in the complex region.

4. EXPERIMENTAL RESULTS

In this section, various experiments are con-

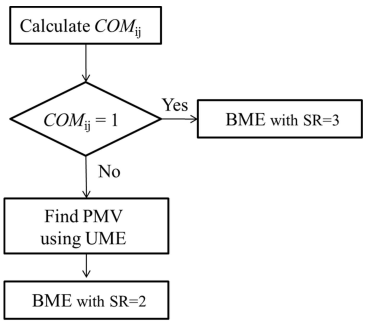
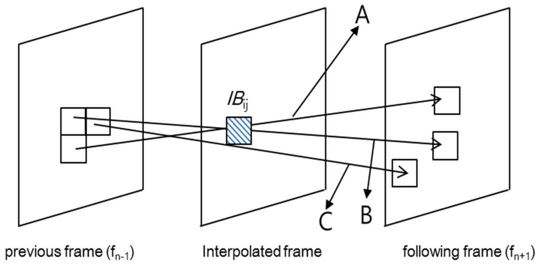
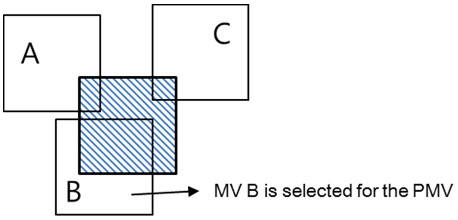


Fig. 5. The sequence diagram of the proposed ME algorithm.



(a)



(b)

Fig. 6. The PMV estimation from UME. (a) MV projection to interpolated frame, (b) PMV selection.

ducted to evaluate the performance of the proposed FRUC algorithms. The experiments use *Akiyo*, *Bus*, *Coastguard*, *Football*, *Foreman*, *Mobile*, *Paris*, *Table*, and *Stefan* image sequences with standard CIF (352×288) format. The 50 odd frames of each test sequence are removed and interpolated using each algorithm. To compare the visual performance and the computational complexity, the BME with initial MV (BMEI), which is the same process as when COM_{ij} = 0, and the SVCT[12] are used. In the proposed algorithms, since H.264/AVC is used in experiments, the basic prediction size for the ME is 16×16. The basic block is divided into 8×8 sub-blocks in variable size refinements. The initial search range for the ME is 16. For the overlapping block motion compensation, an average method with width 1 is used [11]. For an objective image quality comparison, the peak signal-to-noise ratio (PSNR) between the interpolated frame and the original frame is calculated. For the MVR, the candidate MV of neighboring 16×16 blocks is used in the SVCT and the BMEI. The results of the average PSNR comparison are presented in Table 1. The results indicate that the PSNR of the

Table 1. COMPARISON OF AVERAGE PSNR OF FRUC ALGORITHMS (dB)

Image sequence	BMEI	SVCT	Proposed
Akiyo	46.49	46.28	46.44
Bus	25.73	24.57	26.04
Coastguard	30.99	31.39	31.07
Football	22.83	21.7	23.03
Foreman	31.94	31.85	32.02
Mobile	26.43	24.34	26.35
Paris	33.13	31.69	33.7
Table	29.29	28.06	29.47
Stefan	26.96	26.74	26.71
Average	30.42	29.62	30.54

proposed algorithm is better than that of the conventional FRUC algorithms in most sequences.

The proposed algorithm achieves an increase in average PSNR of 0.12 dB compared with BMEI and 0.92 dB compared with SVCT. Compared to SVCT, remarkable performance is shown in *Mobile* and *Paris*, which have complex texture objects with homogenous backgrounds.

Fig. 7 shows the interpolated image of the *Mobile* sequence for each FRUC algorithm. The SVCT algorithm produces matching error in the number on the calendar, as shown in Fig. 7 (b). This error is caused by the false ME result with a homogenous background. As shown in Fig. 7 (c), the visual quality is dramatically improved by the proposed algorithm. This performance improvement is based on the reduced ME search range and the precise ME search using the PMV in the proposed algorithm.

To evaluate the computational complexity, the processing time is compared. Table 2 shows the processing time of each FRUC algorithm. The average processing time of the proposed algorithm is 20% less than that of BMEI. Compared to the SVCT, the average processing time of the proposed algorithms is increased by 23%.

5. CONCLUSION

In this paper, new frame rate up-conversion al-

(a) (b) (c)

Fig. 7. Interpolated image of the *Mobile sequence*. (a) BMEI, (b) SVCT, (c) Proposed.

Table 2. COMPARISON OF COMPUTATIONAL COMLEXITY (SECONDS)

Image sequence	BMEI	SVCT	Proposed
Akiyo	49	24	28
Bus	52	31	47
Coastguard	49	29	45
Football	51	42	46
Foreman	50	26	42
Mobile	53	38	50
Paris	50	47	38
Table	54	28	50
Stefan	50	30	44

gorithms are proposed. To obtain a more accurate MV without excessive computational complexity, an adaptive ME algorithm using block complexity information is used. Using this algorithm, the artifacts and computational complexity are also reduced. As the information on block complexity is extracted from the ME prediction size in original frames, no additional computational complexity is required. Experimental results show that the proposed algorithm provides a robust interpolation performance for the whole test sequence. Also, the computational complexity of the proposed algorithms is reduced by an average of 20% compared with that of BMEI. Although the computational complexity is 23% greater than that of SVCT, this increase is less important than the robust image interpolation performance.

As the size of the display in consumer electronics increases, computational complexity and visual quality are important considerations in the implementation of the FRUC application. However, these two factors have a trade-off relationship in existing algorithms. In this paper, new FRUC algorithms based on information from the previously coded original images are proposed, securing the robust visual quality of interpolated frames without additional computational complexity. Therefore, the presentation of this algorithm is expected to motivate the development of many algorithms based on the previously coded information.

REFERENCE

[1] J. Someya, N. Okuda, and H. Sugiura, "The Suppression of Noise on a Dithering Image in LCD Overdrive," *IEEE Transactions on Consumer Electronics*, Vol. 52, No. 4, pp. 1325-1332, 2006.

[2] P. Gunawan and M. Ghanbari, "Efficient Reduced-reference Video Quality Meter," *IEEE Transactions on Broadcasting*, Vol. 54, No. 3, pp. 669-679, 2008.

[3] R. Feghali, F. Speranza, D. Wang, and A. Vincent, "Video Quality Metric for Bit Rate Control Via Joint Adjustment of Quantization and Frame Rate," *IEEE Transactions on Broadcasting*, Vol. 33, No. 1, pp. 441-446, 2007.

[4] C.L. Huang and T.T. Chai, "Motion-comsated Interpolation for Scan Rate Up-conversion," *Optical Engineering*, Vol. 35, No. 1, pp. 166-

176, 1996.

[5] P. Haavisto, J. Juhola, and Y. Neuvo, "A Method for Motion Adaptive Frame Rate Up-conversion," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 6, No. 5, pp. 436-446, 1996.

[6] T. Koga, K. Iinuma, Y. Iijima, and T. Ishiguro, "Motion-compensated Interframe Coding for Video Conferencing," *Proceeding of National Telecommunication Conference*, G5.3.1-5.3.5, 1981.

[7] G.D. Haan, P.W.A.C. Biezen, H. Huijgen, and O.A. Ojo, "True Motion Estimation with 3-D Recursive Search Block Matching," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 3, No. 5, pp. 368-379, 1993.

[8] D. wang, A. Vincent, P. Blanchfield, and R. Klepko, "Motion-compendated Frame Rate up Conversion? Part II: New Algorithms for Frame Interpolation," *IEEE Transactions on Broadcasting*, Vol. 56, No. 2, pp. 142-149, 2010.

[9] U.S. Kim and M.H. Sunwoo, "New Frame Rate Up-conversion Algorithms with Low Computational Complexity," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 24, No. 3, pp. 384-393, 2014.

[10] C. Wang, L. Zhang, Y. He, and Y.P. Tan, "Frame Rate Up-conversion Using Trilateral Filtering," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 20, No. 6, pp. 886-893, 2010.

[11] B.T. Choi, S.H. Lee, and S.J. Ko, "New Frame Rate Up-conversion Using Bi-directional Motion Estimation," *IEEE Transactions on Consumer Electronics*, Vol. 46, No. 3, pp. 603-609, 2000.

[12] J.H. Kim, Y.H. Ko, H.S. Kang, S.W. Lee, and J.W. Kwon, "Frame Rate Up-conversion Method Based on Texture Adaptive Bilateral Motion Estimation," *IEEE Transactions on Consumer Electronics*, Vol. 60, No. 3, pp. 445-452, 2014.

[13] B.D. Choi, J.W. Han, C.S. Kim, and S.J. Ko, "Motion-compensated Frame Interpolation Using Bilateral Motion Estimation and Adaptive Overlapped Block Motion Compensation," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 17, No. 4, pp. 407-416, 2007.

[14] S.J. Kang, K.R. Cho, and Y.H. Kim, "Motion Compensated Frame Rate Up-conversion Using Extended Bilateral Motion Estimation," *IEEE Transactions on Consumer Electronics*, Vol. 53, No. 4, pp. 1759-1767, 2007.

[15] T.H. Tsai and H.Y. Lin, "High Visual Quality Particle Based Frame Rate up Conversion with Acceleration Assisted Motion Trajectory Calibration," *Journal of Display Technology*, Vol. 8, No. 6, pp. 341-351, 2012.

[16] D.G. Yoo, S.J. Kang, and Y.H. Kim, "Direction-select Motion Estimation for Motion-compensated Frame Rate Up-conversion," *Journal of Display Technology*, Vol. 9, No. 10, pp. 840-850, 2013.

[17] S.C. Han and J.W. Woods, "Frame-rate Up-conversion Using Transmitted Motion and Segmentation Field for very Low Bit-rate Video Coding," *Proceeding of International Conference on Image Processing*, pp. 747-750, 1997.

[18] Y.H. Jung, J.H. Kim, and Y.H. Ko, "Frame Rate up Conversion Method Using Bilateral Motion Estimation Based on Texture Activity and Neighboring Motion Information," *Journal of Korea Multimedia Society*, Vol. 17, No. 7, pp. 797-805, 2014.

[19] T.Y. Kuo and C.C.J. Kuo, "Motion-compesated Interpolation for Low-bit-rate Video Quality Enhancement," *Proceeding of Applications of Digital Image Processing*, pp. 277-288, 1998.

[20] A. Kaup and T. Aach, "Efficient Prediction of Uncovered Background in Interframe Coding

Spatial Extraprediction," *Proceeding of International Conference on Acoustics, Speech and Signal Processing*, pp. 501-504, 1994.

[21] R.J. Schutte and G.D. Haan, "Real-time 2-3 Pull-down Elimination Applying Motion Estimation/Compensation in a Programmable Device," *IEEE Transactions on Consumer Electronics*, Vol. 44, No. 3, pp. 501-504, 1998.

[22] B.W. Jeon, G.I. Lee, S.H. Lee, and R.H. Park, "Coarse-to-fine Frame Interpolation for Frame Rate Up-conversion Using Pyramid Structure," *IEEE Transactions on Consumer Electronics*, Vol. 46, No. 3, pp. 603-609, 2000.

[23] B.G. Kim, "Novel Inter-mode Decision Algorithm Based on Macroblock(MB) Tracking for the P-slice in H.264/AVC Video Coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 18, No. 2, pp. 273-279, 2008.

[24] R. Han and A. Men, "Frame Rate Up-conversion for High-definition Video Applications," *IEEE Transactions on Consumer Electronics*, Vol. 59, No. 1, pp. 229-236, 2013.

Kangjun Lee

received his B.S degree in electronic engineering from Kangnung National University, Korea, in 2003, and his M.S. and Ph.D. degree in electronic communications engineering from Hanyang University, Korea, in 2006 and 2010, respectively. From 2010 to 2015, he was with Hyundai Mobis, where he helped develop infortainment system and safe driving system for vehicle. Since 2017, he has conducted research at Kookmin University. His research interests include image/video coding, image processing, and computer vision system for vehicle.