VVC 인트라 부호화기술을 이용한 라이트필드 영상 부호화

*Vinh Van Duong, Thuc Huu Nguyen, 이재린, 전병우
성균관대학교 전자전기컴퓨터공학과
e-mail: *duongvinh@skku.edu, thuckechsu@skku.edu, jaelin@skku.edu, bjeon@skku.edu*

Light Field Image Compression using Versatile Video Coding Intra Prediction

* Vinh Van Duong, Thuc Huu Nguyen, Jaelin Lee, and Byeungwoo Jeon
Department of Electrical and Computer Engineering
Sungkyunkwan University

Abstract

Light Field (LF) camera captures not only the light intensity but also the light direction coming to camera. While the rich information captured by LF camera enables many interesting applications such as digital refocusing, viewpoint changing, and 3D reconstruction, but it also r equires powerful coding tools to reduce its large volume of data. In this paper, we investigate using the intra prediction scheme of the versatile video coding (VVC), which is the most recent video coding technology currently under developing, to compress the LF image. The Intra Block Copy (IBC) technique in VVC is exploited considering special LF image structure. The expe rimental result shows that VVC i ntra predict ion outperforms the H.265/HEVC i ntra coding technique in encoding LF data irrespective of using the IBC mode or not.

1. Introduction

The hand-held light field (LF) camera has the micro -lens array in addition to the main lens, and the sensor placed behind each micro-lens records not only the light intensity but also the light direction coming to the sensor [1 -2]. The rich information which LF camera acquires makes it possible to realize many interesting applications such as digital refocusing, viewpoint changing, and 3D reconstruction. Fig. 1 shows an example of LF image captured by Lytro Illum camera in which one can note that the micro -lens has the circle shape. The image behind each mirco -lens is a set of marco-pixels where a macro-pixel represents 15x15 pixels or 225 light directions coming to a point. The dimension of a raw LF image is 7728x5368 pixels, corresponding to 8K resolution (7680x4320), and the data size is about 50 Mbytes. It means that LF compression is utmost necessary for pract ical transmission or storage of the data.

2. Related works

LF compression has attracted much attention among researchers in recent years. Many LF coding solutions which have been already proposed [2] can be divided into two groups based on the format of input data: a) solutions taking input in lenslet image format which contains the macro-pixels (as seen in Fig. 1). It exploits the spatial redundancy between the macro-pixels; b) solutions for inputs in 4D LF format which contains multiple sub-aperture images (SAIs). It exploits the temporal redundancy between these sub-aperture images.

The lenslet -based approach compresses the raw lenslet LF image directly via transform -based scheme or spatial predictive scheme. The transform -based schemes use the 3D Discrete Cosine Transform (3D -DCT) [3], Discrete Wavelet Transfor m (3D-DWT) [4], Karhunen -Loeve Transform (KLT) [5], or techniques combining these. It was reported that the LF coding using DWT has the best rate -distortion performance irrespective of whether used alone or combined with DCT or KLT. It was also shown that the compression efficiency is higher than JPEG but lower than High Efficient Video Coding (HEVC) intra coding. Regarding spatial predictive technique, very recently, the authors in [6] proposed to reshape the lenslet image to make the macro -pixels aligned first and then to apply conventional intra prediction modes plus three more intra prediction modes designed to better exploit the correlation between the macro -pixels. They reported bitrate saving up to 47% compared to HEVC Intra. The Intra Block Copy (IBC) mode reported effective in encoding screen contents is also shown effective in improving coding performance of LF data [6].

In the 4D LF format approach, Dong Liu [7] is seen to first propose decomposition of the macro-pixels of lens let image to SAIs, th en encode the SAIs like a pseudo-video sequence using the Joint Exploration Model (JEM) software. They report ed coding gain up to 6.6 dB compared with LF image

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017R1A2B2006518).

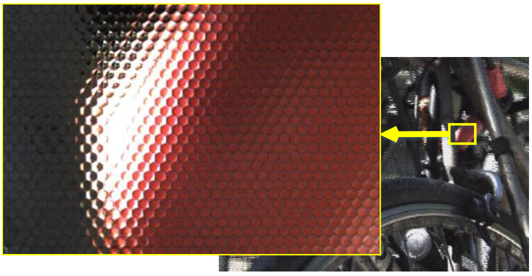

Fig. 1. Bikes LF Image

encoded by JPEG. In later works in [8], they tried to explore the best way of arranging the SAIs following the hierarchical coding structure and showed maximum 28.4% of bitrate savings compared with [7]. As another approach, the authors in [9] proposed a new scan order to re-arrange the SAIs in order to improve the coding efficiency.

The results of LF coding challenge in [10] suggested that coding is more efficient in 4D LF format than in the lenslet format. However, coding in the 4D LF format based on the inter-prediction tools has much complexity compared to the lenslet format, which might not be suitable for those applications requiring lower latency. Therefore, in this paper we investigate a coding scheme which has balance between coding efficiency and complexity.

3. VVC-based Light Field Image Coding

Although, HEVC has achieved significant improvement in compression efficiency when compared to its prior standard of H.264/AVC, it is required to have even better compression for many emerging applications serving screen content, consumer generated content, virtual reality, omnidirectional content, high dynamic range content, etc. The basic architecture of VVC is similar to that of the HEVC, however, to achieve even higher compression performance, several novel coding tools have been proposed for VVC architecture, as following [11]:

1) Block structure: larger Coding Tree Unit (up to 256x256) and transforms (up to 64x64); quadtree plus binary tree block structure (QTBT) replacing the quadtre e structure of HEVC.
2) Intra prediction: 67 prediction modes with 4-tap interpolation filter for intra prediction; cross -component linear model (CCLM)-based prediction; position dependent intra prediction combination (PDPC).
3) Transform: multiple core transfor ms; mode dependent non-separable secondary transforms in intra coding.
4) Inter prediction: sub-block level motion vector prediction; locally adaptive motion vector resolution (AMVR); 1/16 pel. motion vector storage accuracy; overlapped block motion compensation (OBMC); local illumination compensation (LIC); affine motion prediction, etc.
5) In-loop filters: adaptive loop filter (ALF); bilateral filter.
6) CABAC design: context model selection for transform coefficient level s; multi-hypothesis probability estimation, improved initialization for context models .

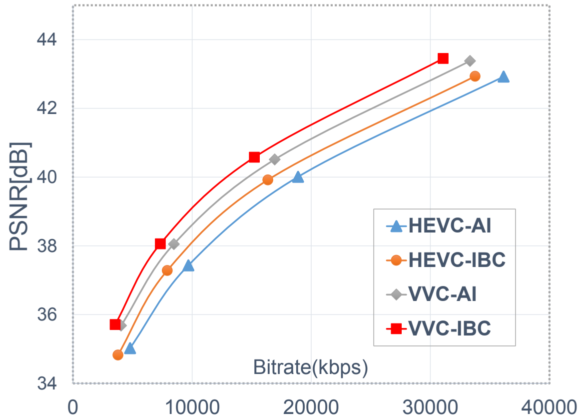

Fig. 2. Compression performance of Bikes LF Image

The resolution of lenslet image is up to 8K as mentioned earlier. Therefore, it will make more sense if compressing LF data with VVC instead of HEVC. Moreover, with many coding tools newly added in VVC, VVC may be more effective in encoding the LF data than HEVC.

4. Experimental Results

In this paper, we use the JPEG Pleno dataset [12] captured by Lytro Illlum camera. Raw LF image has been decoded using Light Field Toolbox in Ma tlab [13] and converted to YUV 420 format before being encoded. For LF image coding with HEVC, we use the reference software for the HEVC Format Range Extension (RExt), HM-16.17+ SCM8.7 [14] with configurations in the file "encoder_intra_main.cfg" (it corresponds to normal HEVC All Intra (HEVC -AI)) and "encoder_intra_main_scc.cfg"(it corresponds to HEVC Intra Block Copy (HEVC-IBC)). As for VVC, VTM 4.0 [15] is used with the configuration file "encoder_intra_vtm.cfg" with or without turning off the IBC mode (namely, VVC-IBC when IBC is turned on, and VVC-AI when IBC is turned off). The Quantization Parameter (QP) is set up at 22, 27, 32, and 37. To measure the coding gain of the proposed method, Bjontegaard Delta Bitrate (BDBR) savings [16] are measured agains t HEVC-AI as an anchor. A negative BDBR represents better coding efficiency of the proposed method compared with the anchor.

As one can see in Fig. 2 and Table I, the VVC -IBC has the best coding performance with the BDBR of -28.85% on average compared with HEVC-AI. This results demonstrate effectiveness of the new design in block partitioning, intra prediction added in VVC. The VVC -AI and HEVC-IBC also showed bitrate savings of -22.65% and -8% on average, respectively. To evaluate the computational complexi ty, execution time is calculated using a PC with Intel ® Core™ i5-8600K CPU @ 3.60GHz with 16 GB RAM and 64 -bits Windows 10 operating system. We computed relative execution time ratios of different methods with respect to that of HEVC -AI as an anchor as shown in Fig. 3. The VVC -IBC consumes the longest execution time with more than 38 times compared with HEVC-AI because the IBC works similar to inter -frame coding

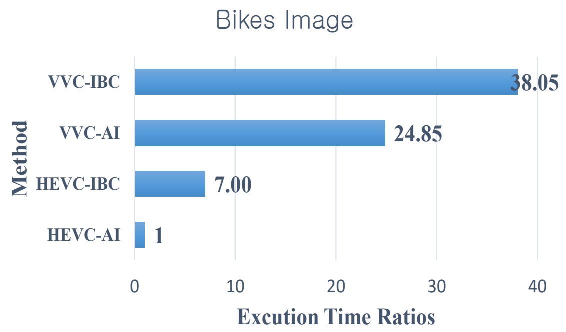

Fig. 3. Execution time ratios of different coding methods relatives to HEVC Intra at QP22

technique and much more prediction modes are added in VVC. Since the encoding complexity of VVC is significantly higher than that of HEVC, it is suggested that many fast algorithms for VVC should be investigated in future to better address the consumer markets.

5. Conclusions

In this paper, we investigated using the VVC intra coding tools to compress lenslet format of LF images. The results showed that VVC with IBC mode can achieve maximum 28.85% bitrate savings on average compared to HEVC Intra; but in terms of encoding time complexity, it also has much higher complexity than HEVC -AI. Further work may focus on deep analysis of compression efficiency and complexity trade off using VVC, and the effect of compression on different applications such as depth estimation and refocusing capabilities.

References

[1] R. Ng, et al., "Light field photography with a hand -held plenoptic camera," in Computer Science Technical Report, vol. 2, no. 11, pp. 111, 2005.

[2] G. Wu, et al., "Light field Image Processing: An Overview, " in IEEE Journal Select. Topics Signal Process., vol. 11, no. 7, pp. 926-954, Oct. 2017.

[3] A. Aggoun, "A 3D DCT compression algorithm for omnidirectional integral images, " in IEEE Inter. Confer. Acoustics Speech Signal Process., vol. 2, pp. 517-520, May. 2006.

[4] A. Aggoun, "Compression of 3D Integral Image using 3D Wavelet Transform, "Journal of Display Technology, vol. 7, no. 11, pp. 586-592, Nov. 20011.

[5] K. Ho Huyn, et al., " Compression scheme of sub-images using Karhunen-Loeve Transform in three -dimension integral image," in Optics Commutations, pp. 3640-3647, July. 2008.

[6] X. Jin, et al., "Plenoptic Image Coding Using Macropixel - based intra prediction, " IEEE Trans. Image Process. , vol. 27, no. 08, pp. 3954-3968, Aug. 2018.

Table I: BDBR (%) performance compared with HEVC -AI anchor (Y -PSNR)

Images	HEVC-IBC	VVC-AI	VVC-IBC
Bikes	-12.6%	-22.3%	-31.7%
Flowers	-3.4%	-23.0%	-26.0%
Average	-8.0%	-22.65%	-28.85%

[7] D. Liu, et al., "Pseudo-sequence based light field image compression," IEEE Int. Conf. Multi. Expo Works, pp. 1 - 4, 2016.

[8] L. Li, et al., "Pseud sequence based 2-D hierarchical coding structure for light field image compression, " IEEE Journal Select. Topics Signal Process., vol. 11, no. 7, pp. 1107-1119, Oct. 2017.

[9] T. N. Canh, V. V. Duong, and B. Jeon, "Boundary handling for video based light field coding with a new hybrid scan order, " in Inter. Workshop on Advanced Image Tec h. (IWAIT), pp. 1-4, Jan. 2019.

[10] I. Viola, et al., "Comparison and evaluation of light field image coding approaches," IEEE J. Select. Topic Sig. Process., vol. 11, no. 7, pp.1092 – 1106, 2017.

[11] J. R. Ohm and G. J. Sullivan, "Versatile Video Coding – towards the next generation of video compression, " Invited talk at Picture Coding Symposium 2018, available at http://www.pcs2018.com/uploads/1/1/0/0/110091239/pcs2018_vvc_overview_06_26.pdf

[12] M. Rerabek and T. Ebrahimi, "New light field image dataset," In Proc. 8th Int. Conf. Quality Multi. Exp., pp. 1 - 2, 2016.

[13] D. Dansereau, "Light Field Toolbox v0.4 ", available https://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4.

[14] Reference software HEVC Format R ange Extension, HM16.17+SCM8.7, available at https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/

[15] Reference software for Versatile Vid eo Coding, VVCSoftware_VTM4.0, available at https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM

[16] G. Bjontegaard, "Calculation of Average PSNR Differences Between RD Curves, " ITU -T SG16 Q.6, VCEG - M33, Geneva, Switzerland, Apr. 2001.