라이트필드 영상 압축을 위한 Affine 움직임 보상 분석

Thuc Nguyen Huu, Vinh Van Duong, Motong Xu, 전병우
성균관대학교 전자전기컴퓨터공학과
thuckechsu@skku.edu, duongvinh@skku.edu, xumotong@skku.edu, bjeon@skku.edu

Analysis of Affine Motion Compensation for Light Field Image Compression

Thuc Nguyen Huu, Vinh Van Duong, Motong Xu, and Byeungwoo Jeon
Department of Electrical and Computer Engineering
Sungkyunkwan University

요 약

Light Field (LF) image can be understood as a set of images captured by a multi-view camera array at the same time. The changes among views can be modeled by a general motion model such as affine motion model. In this paper, we study the impact of affine coding tool of Versatile Video Coding (VVC) on LF image compression. Our experimental results show a small contribution by affine coding tool in overall LF image compression of roughly 0.2% - 0.4%.

1. Introduction

While a conventional 2D image carries only spatial information, Light Field (LF) image carries not only spatial but also angular information [1]. Many applications can benefit from the extended information, for example, for re-focusing, depth estimation, or 3D reconstruction [1]. However, its additional amount of LF image makes it important to more efficiently compress the LF data, which has attracted considerable attentions of engineering research. LF image data can be represented either in (i) lenslet format, or in (ii) multi-view format. The first format (lenslet) represents the LF image in terms of spatial macro pixels which contain the angular information, whereas, the second format (multi-view) represents the LF image in an array of conventional images each of which can be treated as a conventional 2D image captured at different view angles. The difference between these formats is shown in Figure 1.

Compression methods using lenslet format (lenslet-based coding) usually exploits the spatial correlation among macro pixels. Most recently, X. Jin et al. [2] have proposed the macro pixel alignment technique and three additional intra prediction modes for HEVC codec. These authors have claimed to achieve 47% bit-rate reduction comparing to normal HEVC intra coding tool.

The multi-view format is easier to exploit the temporal correlation among views for compressing LF data. For example,

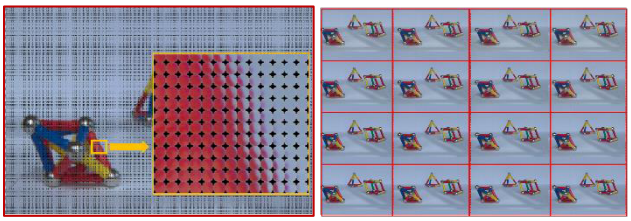

Figure 1. Two common LF representations: Lenslet (left), and Multi-view (right) formats

Friends_1 Flowers Bikes

Figure 2. Center view of LF dataset

the different views can be ordered into video sequence and then encoded by inter coding tool of H.264 or HEVC. Note that D. Liu et al. [3] have encoded the LF image in a form of video sequence and reported gain up to 6.6 dB against JPEG compression of the LF data fin the lenslet format. An extended work in [4] has further refined the architecture [3] by having multiple scan orders. It has shown improvement up to 28.4% against the previous work.

Most methods for LF compression until now are built on the existing video compression standards (H.264, HEVC) by modifying their intra or inter coding tools. It is noted that a new video coding standard known as VVC (Versatile Video Coding) is expected to be standardized in 2020 [5] with employing a lot of new coding tools. Therefore, it will be very interesting to investigate its performance in LF compression. In this context, we study the impact of the affine coding tool (which is a newly introduced motion model of VVC) on LF image compression. The paper structure is organized as follows. Section 2 briefly summarizes the key concept of the affine coding tool and the reason why we choose this for testing LF image compression. Section 3 describes the experiment setup and its experimental results. Lastly, Section 4 concludes our paper.

2. Affine Coding Tool in VVC

H.264 or HEVC considers the translation motion only while the real motion can be much more complex, for example, zoom

Table I. Summary of experiment setup

LF Dataset	Image	Friends_1, Flowers, Bikes (Figure 2)
	Resolution	Spatial: 624x420; Angular: 15x15
VVC configuration	Coding condition	Random-access with QP = 9, 12, 16, 20 (only the first frame in intra) Affine: On/Off test
	Software	VTM-5.0 [8]
PC spec.	CPU: i7 6850K; RAM 32GB; Windows 10 - 64bit	

in/out, rotation, shear. To model the several motion types, VVC employs a block-based affine motion model under which a coding block can choose either the 4-parameter or 6-parameter affine motion model [5].

We note that the view changes among different LF views can be modeled by the affine motion [6]. A recent work [6] has proposed the homography approximation (which is a more general form of affine motion) to compensate the changes among views in multi-view format. However, this recent work did not examine affine motions in a block-based context, which motivates us to investigate more on this.

3. Experiment and Results

In our experiment, we used the LF images in JPEG Pleno dataset [7] which were taken by the Lytro Illum camera. Those images are first converted to the multi-view format, and then are ordered following the raster-scan order to form a video sequence before being applied to VVC encoder as input. For experiment with VVC, we use the reference software VTM-5.0 [8] under the Random-access configuration. Only the first frame is encoded as intra, and quantization parameters are set to {9, 12, 16, 20} respectively to make the compression ratio similar to the ones in the JPEG Pleno test condition [9]. The affine coding tool is turned on/off to evaluate its compression performance on LF image. The whole configuration is summarized in Table I.

Given the anchor being the affine coding tool OFF, we compare compression performance (i.e., bitrate reduction) as well as encoding time when the affine is ON. From Table II, the case with the affine ON is shown to give bitrate reduction by 0.2% to 0.4% against the anchor condition. Also, the encoding time slightly increases by about 3-4%. Because the performance difference is very small, we only show the rate-distortion performance in case of affine on (Figure 3). It can be seen that the quality of images Friends_1 and Bikes scale very well at high bitrate, whereas images Flowers gives very poor-quality scale also at high bitrate.

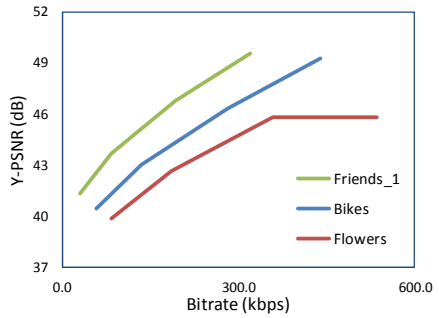

Figure 3. Rate-distortion performance when Affine is ON

Table II. Experimental results

LF images	Affine on vs. Affine off (anchor)	
	Bitrate reduction (%)	Encoding time (%)
Friend_1	-0.2%	104%
Flowers	-0.4%	103%
Bikes	-0.2%	104%

4. Conclusion

Although the affine motion model is a very powerful coding tool in VVC for normal video data, from our very simple first trial to use it for LF image compression, only very small performance gain is observed. At the same time, the encoding time complexity increase is also not too much. It certainly calls for more in-depth investigation on the effectiveness of affine model in encoding LF data.

Acknowledgments

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).

Reference

[1] G. Wu et al., "Light Field Image Processing: An Overview," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 7, pp. 926-954, Oct. 2017.

[2] X. Jin et al, "Plenoptic Image Coding Using Macropixel-Based Intra Prediction," in *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 3954-3968, Aug. 2018.

[3] D. Liu et al, "Pseudo-sequence-based light field image compression," *2016 IEEE International Conference on Multimedia & Expo Workshops* (ICMEW), Seattle, WA, 2016, pp. 1-4.

[4] L. Li et al, "Pseudo Sequence Based 2-D Hierarchical Coding Structure for Light-Field Image Compression," *2017 Data Compression Conference (DCC)*, Snowbird, UT, 2017, pp. 131-140.

[5] Versatile Video Coding (Draft 4) - Joint Video Experts Team (JVET) in *13th Meeting: Marrakech*, MA, 9–18 Jan. 2019.

[6] X. Jiang et al, "Light Field Compression With Homography-Based Low-Rank Approximation," in *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 7, pp. 1132-1145, Oct. 2017.

[7] M. Rerabek et al, "New light field image dataset," in *Proc. 8th Int. Conf. Quality Multi*. Exp., pp. 1 – 2, 2016.

[8]. VTM5 Reference software - https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM.

[9] *ICME 2016 Grand Challenge*: Light-Field Image Compression July 11th – 15th, 2016, Seattle, USA.