OPTIMIZED INTER-VIEW PREDICTION BASED LIGHT FIELD IMAGE COMPRESSION WITH ADAPTIVE RECONSTRUCTION

Chuanmin Jia[*1], *Yekang Yang*[*], *Xinfeng Zhang*[†], *Xiang Zhang*[*], *Shiqi Wang*[§], *Shanshe Wang*[*], *Siwei Ma*[*]

[*] Institute of Digital Media, Peking University, Beijing, China
[†] Rapid-Rich Object Search (ROSE) Lab, Nanyang Technological University, Singapore
[§] Department of Computer Science, City University of Hong Kong, Hong Kong, China

ABSTRACT

In this paper, we explore the structure of light field (LF) and efficiently improve the performance of the pseudo sequence based lenslet image compression, by optimized sub-view rearrangement, enhanced illumination compensation and adaptive reconstruction filtering. First, the decomposed sub-view images are rearranged into a pseudo sequence according to our optimized scan order based on sub-view correlation. Second, the generated pseudo sequence is compressed with JEM codec in which we enhance the illumination compensation by adaptively selecting reference pixels used in parameter derivation. Finally, to reduce the distortions in lenslet image decomposition and reconstruction, we propose an lenslet reconstruction method by applying adaptive filters to the reconstructed lenslet images to compensate the reconstruction errors. Each filter is derived by minimizing the distortions between the original and reconstructed pixels with same geolocation in the lenslet image. Extensive experimental results show that the proposed method achieves up to 53.7% bit rate reduction over HEVC intra coding and 34.8% over JEM intra coding in terms of BDBR.

***Index Terms*—** Lenslet Image, Compression, Reorder, Enhanced Illuminance Compensation, Adaptive Filter

1. INTRODUCTION

With the standardization progress of JPEG Pleno [1], light field (LF) compression attracts tremendous attention and many coding techniques have been proposed recent years [2]. LF imaging captures the intensity of objects as well as the direction of the light ray which results in more redundancy. Hence, the high efficiency compression for LF images plays an important role in its widely application.

In this paper, we improve the pseudo sequence based lenslet compression method by optimizing the three components, reordering, enhanced illuminance compensation and adaptive lenslet reconstruction. To reduce LF redundancies during compression, we first decompose the lenslet image into sub-view images using MATLAB LF Toolbox [3], and design a novel sub-view image arrangement mechanism to reorder them into a pseudo sequence for compression. Subsequently, an enhanced illuminance compensation algorithm is integrated into JEM [4] codec by adaptively selecting pixels when training the parameters for illuminance compensation. Finally, to reconstruct lenslet, we design a set of adaptive filters for lenslet reconstruction to alleviate the distortion induced by lenslet decomposition.

The remainder of the paper is organized as follows. Section 2 briefly reviews related work. Section 3 introduces the proposed method. Extensive experimental results are reported in Section 4 and Section 5 concludes this paper.

2. RELATED WORK

Recently, many specific compression methods for LF images have been proposed in literature, which can be roughly classified into two categories: i) non-local prediction based intra compression [5] [6] [7] and ii) pseudo sequence based compression [8] [9] [10] [11] [12]. Alves *et al.* [2] assessed the compression performance by directly applying the existing widely used image/video codecs such as JPEG, JPEG2000, H.264/AVC etc.

The first kind of methods focuses on compressing the lenslet images directly. Li *et al.* [6] proposed an intra prediction algorithm by combining sparse image set and its associated disparities for the prediction. Perra *et al.* [13] proposed a preprocessing method which directly divided lenslet into multiple small images with same resolution for compression. To further reduce the redundancy, Conti *et al.* [5] formulated a self-similarity model into HEVC by searching similar blocks in a nonlocal reconstructed region for prediction. Zhong [7] proposed a super-pixel based l1 norm linear weighted intra prediction scheme for lenslet compression.

However, the above methods do not fully exploit the LF structure. For the multi-view based compression, the lenslet image is first decomposed into sub-view images according to

[1]This work was supported in part by National Natural Science Foundation of China (61632001, 61571017, 61421062), the National High-tech R&D Program of China (863 Program, 2015AA015903) and the Top-Notch Young Talents Program of China, which are gratefully acknowledged.

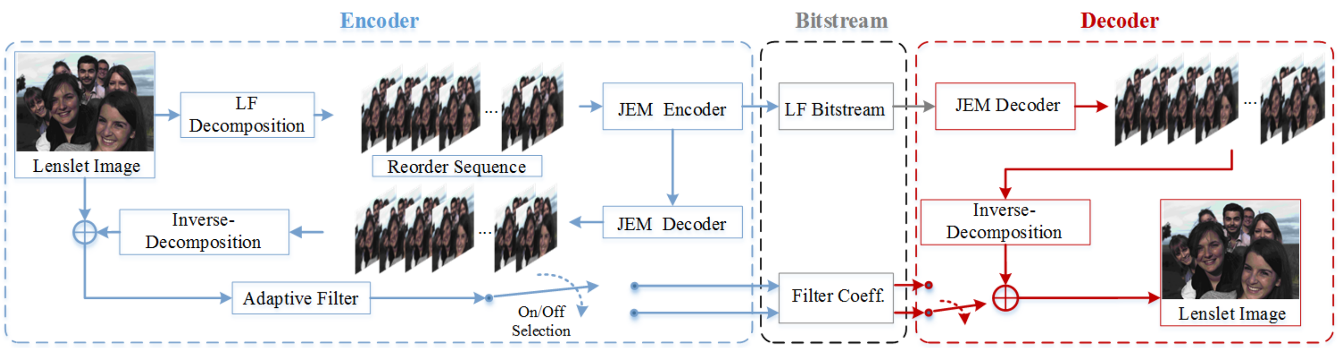
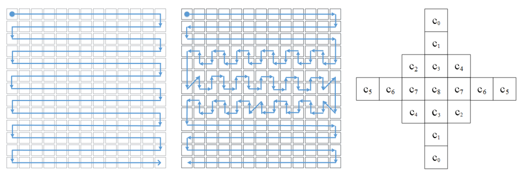


Fig. 1. Diagram of the proposed LF compression: (a) Encoder (b) Decoder.



(a) (b) (c)

Fig. 2. (a) JPEG CfP [15] sub-view arrangement method; (b) Proposed rearrangement method of sub-views; (c) Designed filter shape for lenslet reconstruction.

its calibration information, and then these sub-view images are compressed sequentially as a pseudo-sequence. For this kind of methods, there are some typical problems which limits the compression performance, e.g., the reference relationship for pseudo-sequence compression and the distortions introduced by the lenslet decomposition. Dai *et al.* [12] firstly provided multi-view based compression for lenslet images. Liu *et al.* [8] [14] proposed a reference management and rate allocation scheme for pseudo sequence compression. Chen *et al.* [9] integrated the sparse coding into LF compression and selected key sub-views images to train sparse dictionary for coding. This algorithm is efficient under low bit rate coding circumstance, which may lack its generalization ability. Zhao *et al.* [10] provided an arrangement order for sub-view images by hybridly combing z scan order and U-shape order.

Our method can be categorized to multi-view based compression. However, different from [10] [12], the proposed method improves the reconstructed quality of lenslet image by reducing the distortions during lenslet decomposition, while the methods in [10] [12] did not deal with these distortions. Furthermore, we optimize the reordering the scheme in [10] when generating pseudo sequence which will be discussed in following section.

3. PROPOSED LF COMPRESSION METHOD

In this section, we will introduce the proposed LF compression method. Fig. 1 illustrates the method, which mainly consists of LF decomposition & sub-view image reordering, enhanced illuminance compensation and lenslet reconstruction. These modules will be introduced respectively in the following subsections.

3.1. Light Field Decomposition & Sub-view Reordering

To better utilize the correlations among sub-views, we decompose the lenslet images into 5-D LF using MATLAB LF Toolbox [16]. The decomposition process includes i) affine transform for lenslet images based on the camera calibration information, ii) interpolation and iii) resamping pixels into sub-views. Sub-view images can be obtained from the 5-D LF decomposed by the lenslet.

According JPEG Pleno Call for Proposal (CfP) for LF coding [15], half of the EPFL LF dataset [17] is chosen as test images. These lenslet images with resolution 7728×5368 will be first decomposed into the 5-D LF structure with $15 \times 15 \times 434 \times 541 \times 3$. It is worth noting that the dimension of LF here is different from the dimension mentioned in chapter 2.3.1.1 of [15] because we have turned off the interpolation during lenslet decomposition. The reason for turning off interpolation is because interplotation makes resampling process irreversible when reconstructing lenslet.

Once we have obtained all sub-view images, pseudo sequence can be generated by a specific arrangement method to reorder these sub-views. Figure. 2(a) illustrates the sub-view arrangement method provided in JPEG CfP [15]. Obviously, the sub-view arrangement method will directly affect the inter-prediction efficiency. We optimize the reorder scheme in [10] and propose a well-designed reordering scheme (Figure. 2(b)) for sub-view images which aggregates similar sub-view images temporally to improve the coding efficiency of the pseudo sequence by promoting the accuracy of inter-prediction.

3.2. Enhanced Illuminance Compensation

The illuminance compensation has been widely discussed and used in HEVC and Multi-View Coding (MVC) [18].

In our proposed method for lenslet coding, an enhanced illuminance compensation technique is provided for promoting coding efficiency. The goal of illuminance compensation is adjusting predicted pixel value based on a linear model to compensate for the illumination changes. In particular, the scaling factor *a* and an offset *b* are used as the linear model for adjustment. The conventional local illuminance compensation [18] does not consider the consistency of luminance change and subsamples the left and above neighbor pixels of coding unit (CU) for parameter training.

In our implementation, the illuminance compensation is performed for each prediction unit (PU) and is enabled or disabled adaptively for each inter-mode coded coding unit (CU). The parameters *a* and *b* are derived by adaptively using neighboring samples of the CU and the corresponding pixels in the reference picture. During the sample selection, we first traverse all neighboring samples of current CU with their corresponding pixels in the reference picture to compute the average intensity of the difference between neighboring samples and their reference ones, after which the pixels whose intensity is lower than the computed average intensity are selected for parameter derivation of illuminance compensation. Then, the least square error method is employed to derive the parameters *a* and *b* based on the above mentioned selecting strategy. Finally, the parameters are derived and applied for each prediction direction separately during coding.

3.3. Adaptive Reconstruction for Lenslet

This subsection describes the lenslet reconstruction from subview images. Our method conducts inverse decomposition (see Figure. 1) to reconstruct lenslet from the coded pseudo sequence. However, the affine transform during decomposition makes the lenslet reconstruction irreversible therefore the upper bound of reconstructed lenslet quality can not be very high. To address this issue, region based adaptive reconstruction is proposed into the method.

At the encoder side, both the original lenslet image and the reconstructed lenslet image are available (see Figure.1), which enables us to improve the quality of reconstruct lenslet by reducing the distortions introduced by decomposition. Pixels at the same horizontal and vertical location within each super pixel are classified to the same regions which has been illustrated in Figure. 3. Hence, there are $15 \times 15 = 225$ regions in total and pixels in the same region share a set of filter parameters. To reduce the overhead of filters, we merge the filters for different regions based on rate-distortion cost. Our proposed reconstruction method differs from the adaptive loop filter (ALF) in HEVC [19] because we take into consideration the structure of LF where regions are determined from the geo-location of super pixel. Moreover, our filters are applied to reconstructed lenslet not to pseudo sequence.

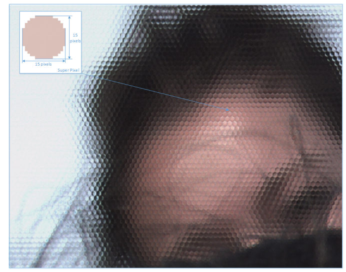


Fig. 3. Illustration of lenslet samples classification into regions.

For each pixel, a specific linear filter is applied as follows,

$$\hat{\boldsymbol{x}}[r] = \sum_{n=0}^{N-1} \boldsymbol{c}(n)\boldsymbol{y}[r+p_n], \qquad (1)$$

where $\hat{\boldsymbol{x}}[r]$ is the filtered pixel, p_n denotes the location offset of samples, and $\boldsymbol{y}[r]$ is the reconstructed lenslet sample before filtering. $\boldsymbol{c}(n)$ is the filter parameters trained by minimizing the distortions between the original lenslet and reconstruction lenslet at the encoder side by solving the pixel level optimization problem,

$$\min_{\boldsymbol{c}_k} \sum_{t \in C_k} \sum_{n=0}^{N-1} \|\boldsymbol{c}_k(n)\boldsymbol{y}(t+p_n) - \boldsymbol{x}(t)\|_2^2, \qquad (2)$$

where C_k is the k^{th} region of pixels, which share the same filter parameters while $\boldsymbol{x}(t)$ denotes the original lenslet samples. The learnt parameters $\boldsymbol{c}_k$ are normalized into closed interval [0, 1]. Meanwhile, each color channel has its own filter coefficients.

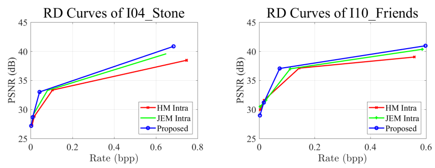
Since the filter parameters need to be signalled to the decoder side, we design a point-symmetrical filter shape with a 3×3 square superposed by a 7×7 cross [19] to reduce the overhead caused by filter parameters. As shown in Fig. 2(c), the 17 taps filter contains 9 different coefficients.

4. EXPERIMENTAL RESULTS

In this section, we will report the test conditions of the experiments and coding performance in terms of BDBR.

4.1. Test Conditions

The test images are selected according to JPEG Pleno Call for Proposal [15] and submission details [20]. Five test images are chosen from EPFL LF dataset [17] for evaluation, each of which has resolution of 7728×5368 in RGB444 color space. The bit depth of them is 10-bit.



(a) (b)

Fig. 4. (a) RD Curves for $I04\_Stone$; (b) RD Curves for $I10\_Friends$;

Anchor. HEVC (HM16.8) intra coding [21] with common test condition and main_444_10_intra profile is chosen as anchor. To further compare the coding performance, we additionally choose JEM (HM 16.8-JEM 2.0) intra and inter coding [4] with common test condition and main_444_10_intra profile as anchor.

Proposed. Each test lenslet image will be decomposed into 5-D LF with $15 \times 15 \times 434 \times 541 \times 3$ to obtain sub-view images, after which pseudo sequences can be generated by the proposed rearrangement method for sub-view images. Without color space transformation, we directly compress the pseudo sequence in RGB444 space with 10-bit. Hence, the resolution for each frame of pseudo sequences is 434×541 and there are 169 frames within this sequence since only 13×13 sub-views are considered. Subsequently, the generated pseudo sequences will be encoded by JEM-2.0 codec [4]. Adaptive reconstruction is applied to lenslet after inverse decomposition from coded pseudo sequence.

The testing conditions in proposed method are listed as follows [20].

i) Target bit rates: 0.75 bpp, 0.1 bpp, 0.02 bpp and 0.005 bpp;
ii) Objective Evaluation: Bjontegaard's method [22];
iii) Coding configuration: Random Access (RA) configuration in JEM common test condition;
iv) Profile: main_444_10 profile, bit depth: 10-bit;

Evaluation. Since the output of both anchor and proposed method is lenslet image. The lenslet images are rendered into LF with dimension $15 \times 15 \times 434 \times 625 \times 3$. All the evaluation metric follows the description of chapter 2.3.1.3 and 2.3.4.1 in JPEG call for proposal [15] which clarifies that the inner 13×13 sub-views of the whole LF is adopted for compression and evaluation. We report average PSNR-RGB (PSNR_RGB$_{mean}$) in this paper since we skip color space transform.

4.2. Objective Quality

This subsection provides the performance of the proposed method over HEVC intra coding and JEM intra coding in terms of BD rate. All experiments are evaluated according to the submission details [20]. Table. 1 shows the BD-rate performance of the proposed method compared with the HM intra coding [21] and JEM intra coding [4]. It is observed that

Test Images	Performance over Intra	
	Anchor: [21]	Anchor: [4]
I01 Bikes	-41.0%	-23.1%
I02 Danger_de_Mort	-33.8%	-32.8%
I04 Stone_Pillars_Outside	-54.8%	-32.7%
I09 Fountain_&_Vincent 2	-53.7%	-34.8%
I10 Friends_1	-29.4%	-15.2%
Overall	**-42.5%**	**-27.7%**

Table 1. The performances of the proposed method over HM intra coding and JEM intra coding.

Test Images	Performance over Inter
	Anchor: JEM Inter Coding
I01 Bikes	-2.1%
I02 Danger_de_Mort	-3.7%
I04 Stone_Pillars_Outside	-5.5%
I09 Fountain_&_Vincent 2	-6.0%
I10 Friends_1	-0.2%
Overall	**-3.5%**

Table 2. The performances of the proposed method. Anchor: JEM inter coding with original illuminance compensation, reorder scheme in [10] and no adaptive reconstruction filter.

the proposed LF coding method achieves remarkable coding performance where 42.5% bit rate reduction on average is achieved over HM intra coding. The right-most column provides the performance over JEM intra coding in which 27.7% BD-rate on average can be reduced for test lenslet images.

Table. 2 shows the BD-rate performance of proposed method over JEM inter coding and averagely 3.5% BD-rate can be reduced. To make our performance evaluation more persuasive, we also compare our performance over JEM inter coding which directly compresses the pseudo sequence with the reordering scheme proposed in [10], no enhanced illuminance compensation and no adaptive reconstruction. The JEM inter coding configurations are the same as the description of *proposed* section 4.1. The RD-curves of image $I04$ and $I10$ over intra coding are shown in Fig. 4 to better show the coding performances in target bit-rate range required in [20].

5. CONCLUSION

In this paper, a novel pseudo-sequence based lenslet images compression scheme with sub-view rearrangement, enhanced illuminance compensation and adaptive lenslet reconstruction is proposed which significantly improves the lenslet image coding efficiency by exploring the structure of LF. The proposed compression method utilizes the JEM codec to compress the pseudo sequence which is generated by decomposing the lenslet image into sub-views and rearranging them according to designed order. An enhanced illuminance compensation algorithm is proposed for pseudo sequence compression. Then region-based learnt filters are introduced to recover irreversible decomposition distortions. Experimental results demonstrate that our proposed scheme obviously improves the lenslet coding efficiency over existing codecs.

6. REFERENCES

[1] Touradj Ebrahimi, Siegfried Foessel, Fernando Pereira, and Peter Schelkens, "Jpeg pleno: Toward an efficient representation of visual reality," *IEEE MultiMedia*, vol. 23, no. 4, pp. 14–20, 2016.

[2] Gustavo Alves, Fernando Pereira, and Eduardo AB da Silva, "Light field imaging coding: Performance assessment methodology and standards benchmarking," in *Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–6.

[3] Donald G Dansereau, Oscar Pizarro, and Stefan B Williams, "Decoding, calibration and rectification for lenselet-based plenoptic cameras," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 1027–1034.

[4] "Jem reference software," `https://hevc.hhi.fraunhofer.de/trac/jem`.

[5] Caroline Conti, Paulo Nunes, and Luis Ducla Soares, "New hevc prediction modes for 3d holoscopic video coding," in *2012 19th IEEE International Conference on Image Processing*. IEEE, 2012, pp. 1325–1328.

[6] Yun Li, Märten Sjöström, Roger Olsson, and Ulf Jennehag, "Efficient intra prediction scheme for light field image compression," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 539–543.

[7] Rui Zhong, Shizheng Wang, Bruno Cornelis, Yuanjin Zheng, Junsong Yuan, and Adrian Munteanu, "L1-optimized linear prediction for light field image compression," in *Picture Coding Symposium (PCS), 2016*. IEEE, 2016, pp. 1–5.

[8] Dong Liu, Lizhi Wang, Li Li, Zhiwei Xiong, Feng Wu, and Wenjun Zeng, "Pseudo-sequence-based light field image compression," in *Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–4.

[9] Jie Chen, Junhui Hou, and Lap-Pui Chau, "Light field compression with disparity guided sparse coding based on structural key views," *arXiv preprint arXiv:1610.03684*, 2016.

[10] Shengyang Zhao, Zhibo Chen, Kun Yang, and Hongrui Huang, "Light field image coding with hybrid scan order," in *Visual Communications and Image Processing (VCIP), 2016*. IEEE, 2016, pp. 1–4.

[11] Mira Rizkallah, Thomas Maugey, Charles Yaacoub, and Christine Guillemot, "Impact of light field compression on focus stack and extended focus images," in *Signal Processing Conference (EUSIPCO), 2016 24th European*. IEEE, 2016, pp. 898–902.

[12] Feng Dai, Jun Zhang, Yike Ma, and Yongdong Zhang, "Lenselet image compression scheme based on subaperture images streaming," in *Image Processing (ICIP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 4733–4737.

[13] Cristian Perra, "Light field image compression based on preprocessing and high efficiency coding," in *Telecommunications Forum (TELFOR), 2016 24th*. IEEE, 2016, pp. 1–4.

[14] Li Li, Zhu Li, Bin Li, Dong Liu, and Houqiang Li, "Pseudo sequence based 2-d hierarchical coding structure for light-field image compression," in *Data Compression Conference (DCC), 2017*. IEEE, 2017, pp. 131–140.

[15] "Jpeg pleno call for proposals on light field coding," `https://jpeg.org/downloads/jpegpleno/wg1n74014_pleno_final_cfp.pdf`.

[16] Donald G Dansereau, "Light field toolbox for matlab," *software manual*, 2015.

[17] Martin Rerabek, Lin Yuan, Léonard Antoine Authier, and Touradj Ebrahimi, "Epfl light-field image dataset," *ISO/IEC JTC*, vol. 1, 2015.

[18] Nikola Sprljan, Stavros Paschalakis, and Ping Wu, "Local intensity compensation for inter prediction in hevc," *Doc. JCTVC-B096, JCT-VC of ITU-T VCEG and ISO/IEC MPEG, Geneva, Switzerland*, 2010.

[19] Chia-Yang Tsai, Ching-Yeh Chen, Tomoo Yamakage, In Suk Chong, Yu-Wen Huang, Chih-Ming Fu, Takayuki Itoh, Takashi Watanabe, Takeshi Chujoh, Marta Karczewicz, et al., "Adaptive loop filtering for video coding," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 6, pp. 934–945, 2013.

[20] "Jpeg pleno call for proposals submission details," `https://jpeg.org/downloads/jpegpleno/wg1n75024-REQ-JPEG_Pleno_CfP_-_Submission_Process_Details.pdf`.

[21] Frank Bossen, David Flynn, and Karsten Sühring, "Hevc reference software manual," *JCTVC-D404, Daegu, Korea*, 2011.

[22] Gisle Bjontegaard, "Calcuation of average psnr differences between rd-curves," *Doc. VCEG-M33 ITU-T Q6/16, Austin, TX, USA, 2-4 April*, 2001.