Rate-Distortion Optimized Super-Ray Merging for Light Field Compression

Xin Su, Mira Rizkallah, Thomas Maugey, Christine Guillemot

Rate-Distortion Optimized Super-Ray Merging for Light Field Compression

Xin Su
INRIA
Rennes, FRANCE
Xin.Su@inria.fr

Mira Rizkallah
IRISA
Rennes, FRANCE
Mira.Rizkallah@irisa.fr

Thomas Maugey
INRIA
Rennes, FRANCE
Thomas.Maugey@inria.fr

Christine Guillemot
INRIA
Rennes, FRANCE
Christine.Guillemot@inria.fr

Abstract—In this paper, we focus on the problem of compressing dense light fields which represent very large volumes of highly redundant data. In our scheme, view synthesis based on convolutional neural networks (CNN) is used as a first prediction step to exploit inter-view correlation. Super-rays are then constructed to capture the inter-view and spatial redundancy remaining in the prediction residues. To ensure that the super-ray segmentation is highly correlated with the residues to be encoded, the super-rays are computed on *synthesized residues* (the difference between the four transmitted corner views and their corresponding synthesized views), instead of the synthesized views. Neighboring super-rays are merged into a larger super-ray according to a rate-distortion cost. A 4D shape adaptive discrete cosine transform (SA-DCT) is applied per super-ray on the prediction residues in both the spatial and angular dimensions. A traditional coding scheme consisting of quantization and entropy coding is then used for encoding the transformed coefficients. Experimental results show that the proposed coding scheme outperforms HEVC-based schemes at low bitrate.

Index Terms—Super-Ray (SR) Merging, Rate-Distortion Minimization, Light Field (LF) Compression, Shape-Adaptive DCT (SA-DCT)

I. INTRODUCTION

Light fields (LF) are defined as the representation of radiance of light rays emitted along several directions by the different points in a 3D scene. Several devices have been developed for light fields capture, either based on camera arrays [1], on single moving cameras, or on arrays of microlenses [2], *etc*. Light fields have recently gained in popularity due to the variety of potential applications in computational photography and computer vision, however they represent very large volumes of data with challenges in terms of storage, transmission and processing.

In this paper, we focus on the problem of compressing dense light fields captured by plenotic cameras. First methods for compressing synthetic light fields appeared late 90's essentially based on classical coding tools as vector quantization or using JPEG coding for each sub-aperture view, yielding however limited compression performances. It is only recently that compression solutions have been proposed for dense real light fields captured by plenoptic cameras. The proposed solutions can be classified into two categories: either coding the array of sub-aperture images extracted from the lenslet image as a pseudo video sequence in [3], [4], or directly encoding the lenslet images captured by plenoptic cameras In [4]–[11], with extensions of HEVC with

dedicated prediction modes. Multiview video compression and disparity compensation techniques are considered in [7]. A homography-based low rank approximation [12] is used to exploit angular correlation of LF. Besides being represented and encoded as images or videos, the LF is represented by 4D Gaussian mixture models in [13] and by graphs containing minimum amount of color and disparity information in [14].

In this paper, we propose a compression scheme based on view synthesis. Four corner views of the LF are first encoded by HEVC-Inter and transmitted. The whole LF is then synthesized from the four corner views using the convolutional neural networks (CNN) based architecture proposed in [15]. The prediction residues are then transformed using a 4D-shape adaptive Discrete Cosine Transform (4D SA-DCT) which exploits both spatial and angular correlation remaining in the residue signals. The support of the 4D SA-DCT is defined by a segmentation of the light field into super-rays. Super-rays can be seen as a set of super-pixels that are coherent across all light field views, taking into account disparity information. Note that local transforms have also been investigated for light fields compression in [16], however, the support of the local transform was defined by co-located super-pixels in the different views, not taking into account disparity.

To ensure that the super-ray segmentation is highly correlated with the residual signals, the super-rays are computed on *synthesized residues* (the difference between the four transmitted corner views and their corresponding synthesized views). Neighboring super-rays with similar homogeneous residues are merged into a larger super-ray to have a better spatial energy compaction by optimizing a rate-distortion cost. Experimental results show that the proposed coding scheme yields rate-distortion gains at low bitrates (e.g. < 0.04 bpp corresponding to a PSNR quality up to 35 dB) compared with HEVC-based coding schemes, while being comparable or slightly worse at higher bitrates.

II. LIGHT FIELD CODING SCHEME

A. Scheme Overview

Fig.1 depicts the proposed coding scheme. Let $\mathbf{LF} = \{I_{u,v}\}$ denote a light field, where $u = 1, \ldots, U$ and $v = 1, \ldots, V$ are the view indices. Four views at the corners $\mathbf{LF}^{\text{cor}} = \{I_{1,1}, I_{1,V}, I_{U,1}, I_{U,V}\}$ are encoded using HEVC-Inter and used to synthesize the whole light field with the
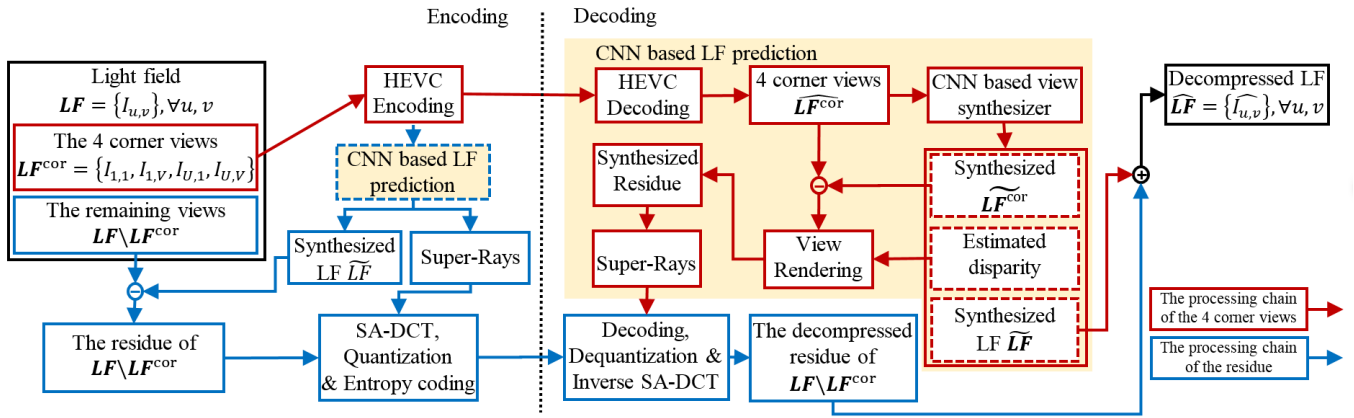
Fig. 1. Overview of proposed coding scheme.

CNN based synthesis method [15], as shown in Fig.1 (red arrows). To improve the quality of the synthesized light field, the residuals between the synthesized and original views are encoded by applying local super-ray based shape adaptive DCT (SA-DCT) (see Fig.1, blue arrows).

1) The processing chain of the four corner views: At the decoder, the decompressed four corner views are used to synthesize the whole LF using the CNN-based architecture of [15], as shown by the yellow region in Fig.1. The first CNN is trained to model the disparity from the four input views, while the second CNN is used to estimate the color of the synthesized views. The synthesis quality depends on the QP value of the HEVC-inter coder.

As shown by the yellow region in Fig.1, we compute the residual signals as the difference between the decompressed four corner views $\widehat{\mathbf{LF}^{\mathrm{cor}}}$ and their corresponding synthesized views $\widetilde{\mathbf{LF}^{\mathrm{cor}}}$. The four images of residues are then warped onto the other views using the disparity estimated by the CNN. The super-ray segmentation is then computed by applying the SLIC algorithm [17] on the set of residue images, but also taking into account disparity when performing the clustering. Computing the segmentation on the residue images rather than on the synthesized views, similar residues signals are more likely to be grouped into one segment which can benefit the following energy compaction in transform domain.

2) The processing chain of the residues: The synthesized views $\widetilde{\mathbf{LF}}$ and the super-ray segmentation are computed in the same manner at the encoder. We apply a local spatial SA-DCT on the residuals for each view using the super-ray as a support of the transform. Spatial SA-DCT coefficients of each super-ray corresponding to the same frequency form a $U \times V$ block in the angular domain, on which a second angular SA-DCT is applied to capture angular dependencies. A traditional coding scheme consisting of quantization and entropy coding is then used for encoding the transformed coefficients. At the decoder side, the decoded residuals are added to the synthesized views to obtain the final decompressed LF.
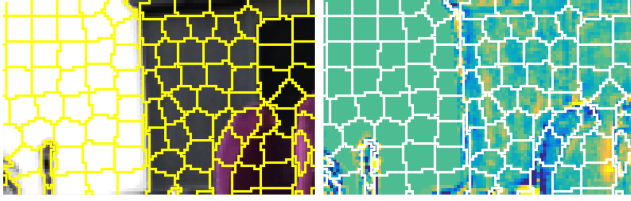
B. Super-Ray Segmentation on Residues

The concept of super-ray has been introduced in [18] as an extension of super-pixels [17] to group light rays coming from

the same 3D object, i.e. to group pixels having similar color values and being close spatially in the 3D space. While the authors in [18] estimate disparity only at the centroid of the super-rays, here we consider a scheme using dense disparity maps to synthesize the entire light field from a sparse set of views. The disparity maps used in the tests reported below, have been estimated by the first CNN of the view synthesis architecture of [15] from the four corner views which are available at both the encoder and decoder. Having a dense disparity map for each view, the pixels in all the views are clustered using a method similar to SLIC [17] with a weighted combination of a color distance, a spatial distance and in terms of depth.

Fig.2 (a) shows that the super-rays computed on the set of synthesized color images are well aligned with the edges of the objects, but are not really suited for the residue images they are supposed to represent (see the edge of the building). Indeed the residues generally lie on both sides of object edges. Since the real residue signal is not available at the decoder, we generate an estimation of it (called synthesized residue), that is used to build a more accurate segmentation. For that purpose, the residue images at the four corner views are first computed as the difference between the coded/decoded corner views and their synthesized versions (see Fig.1). The residue images at the other viewpoints are then obtained by warping the corner residue images to the other positions using the estimated disparity. We see in Fig.2 (b) that the super-ray segmentation computed on these synthesized residue images has better correlation with the real residual signals, which is proved by the energy compaction comparison between Fig.2 (a) and (b) in section IV-A. Since the synthesized residues are available at both the encoder and decoder, we can obtain the same super-ray segmentation on both sides, and do not need to transmit it.

C. 4D Shape Adaptive DCT (SA-DCT)

While 2D DCT applied on a square or rectangular support may fail to capture correlation at image edges, we consider here a separable 4D shape-adaptive DCT with a support defined by the super-ray segmentation. Fig.3 illustrates how a 4D SA-DCT is applied on the i-th super-ray $\mathbf{SR}_i$ in the LF.

(a) Super-ray segmentation computed on synthesized color is shown with corresponding color image (left) and real residual image (right).
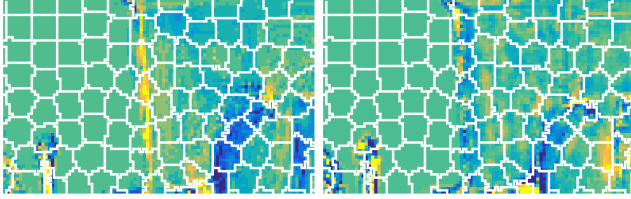
(b) Super-ray segmentation computed on synthesized residues is shown with corresponding synthesized residual image (left) and real residual image (right).

Fig. 2. Super-ray segmentation computed on (a) synthesized color and (b) synthesized residues. Only the center view ($u = 4$, $v = 4$) has been shown here. Segments in (a) are well aligned with the edges of color image, however, not aligned with the discontinuities within the residue images, see the border of the building. Segments in (b) obtained using synthesized residues better follow the discontinuities of the real residue signals.

A spatial 2D SA-DCT is first applied per view on each super-ray (i.e. on the super-pixel of the view which belongs to the considered super-ray). In each view, the obtained coefficients form a rectangular block with non-zero coefficients only in the top-left area with the DC component located at the left-top corner, as shown in the middle of Fig.3. These coefficients are sorted from low to high frequency by following a Zig-Zag order. Coefficients corresponding to the same frequency band but from different views form an $U \times V$ block in angular domain, as shown on the right of Fig.3. Another SA-DCT, *i.e.* an angular SA-DCT, is then applied on this block. Note that some values may be missing in the $U \times V$ block, since the size of $\mathbf{SR}_i$ varies in different views. Generally, the 4D DCT $\mathfrak{T}(\cdot)$ can be computed as

$$\{X_{i,b}\} = \mathfrak{T}(\mathbf{SR}_i) \ ,$$
$$\mathfrak{T}(\cdot) = \underbrace{\mathbf{DCT}_u \otimes \mathbf{DCT}_v}_{\text{Spatial SA-DCT}} \otimes \underbrace{\mathbf{DCT}_y \otimes \mathbf{DCT}_x}_{\text{Angular SA-DCT}} \quad (1)$$

where $\{X_{i,b}\} \in \mathbb{R}^{U \times V \times N}$, N is the maximum size of $\mathbf{SR}_i$ in different views, $b = 1, 2, \cdots, U \times V \times N$. For some values of b, $X_{i,b}$ may be missing due to the non-regular shape of $\mathbf{SR}_i$. The positions of missing elements in $\{X_{i,b}\}$ are available at both encoder and decoder, since the super-ray segmentation is known. $\mathbf{DCT}_* \in \mathbb{R}^{n \times n}$ corresponds to a n-point DCT (n is the number of elements of $\mathbf{SR}_i$ in corresponding coordinate), and $\otimes$ denotes a Kronecker product operator.

D. Quantization and Entropy Coding

At the end of those two transform stages, coefficients are grouped into a 2-dimensional array $\mathbf{X}$ where $\mathbf{X}(i, b)$ is the b-th band in super-ray $\mathbf{SR}_i$. Using the observations on all the super-rays in a training dataset (*Rose* [12]), we can find
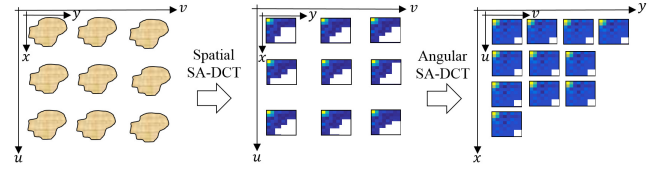

Fig. 3. An illustration of 4D SA-DCT performed on super-ray.

the best ordering for quantization. We first sort the variances of coefficients with enough observations in decreasing order. We then split them into 64 groups $\mathbf{g}$: $\{\mathbf{X}_{i,b}, \forall i, b \in \mathbf{g}\}$. All the remaining coefficients with less observations will be considered in the last group. We use the zigzag ordering of the *JPEG* quantization matrix to assign the quantization step size Q for each. A simple rounding procedure then results in the quantized coefficients $\mathbf{X}_q(i, b) = round(\mathbf{X}(i, b)/Q(\mathbf{X}(i, b))$. that are further coded using an arithmetic coder.

III. RATE-DISTORTION OPTIMIZED SUPER-RAY MERGING

In order to increase compression performances, we improve the super-ray segmentation with a rate-distortion optimized super-ray merging. Four initial segmentations are performed using different initial numbers of clusters leading to different super-ray sizes as shown in Fig.4.a. The resulting segmentations are referred to as layers. Note that we modify the super-ray segmentation in layer l respecting to the boundaries of super-rays in layer $l-1$, to make sure the boundaries are coherent at different layers. For instance, if super-ray $\mathbf{SR}_i^{l-1}$ in layer $l-1$ is across two (or maybe more) super-rays in layer l, one of these super-rays in layer l is enlarged to completely contain $\mathbf{SR}_i^{l-1}$, and the other super-rays are reduced correspondingly. We choose to enlarge the super-ray that initially contains most parts of $\mathbf{SR}_i^{l-1}$.

The merging results $\{SR_i\}$ are initialized by the super-rays $\{SR_i^{l=0}, \forall i\}$ at layer 0 and the merging starts from layer 1 to layer 3. At each time, we only consider one super-ray SR_i^l at layer l, which consists of several super-rays $\{SR_j^{l-1}\}$ at layer $l-1$, *i.e.* $SR_i^l = \{SR_j^{l-1}\} = \{SR_j^{l-1} \in SR_i^l, \forall j\}$, as shown in Fig.4 (b). $\{SR_j^{l-1}\}$ will be merged into SR_i^l, *i.e.* $\{SR_j^{l-1}\} \Rightarrow SR_i^l$, if and only the rate-distortion cost $\mathcal{J}$ reduces after merging.

The rate cost of the 4D SA-DCT coefficients (after quantization) is modeled by their entropy computed per coefficient group as

$$R = \sum_{\mathbf{g}} \text{entropy} \left(\{Q(X_{i,b}), \forall i, b \in \mathbf{g}\}\right). \quad (2)$$

The distortion is computed as the distortion of SA-DCT coefficients before and after quantization as

$$D = \sum_i \sum_b [X_{i,b} - Q(X_{i,b})]^2 \ . \quad (3)$$

Thus, the rate-distortion cost is

$$\mathcal{J} = D + \lambda R \ , \quad (4)$$

where $\lambda = 1$ represents the relation between rate and distortion. The merging procedure based on the minimization of the rate-distortion cost $\mathcal{J}$ is detailed in Algorithm 1. Fig.4 (c) gives an illustration of merging results using *Flower1* dataset in [15].

(a) Super-rays in different layers.

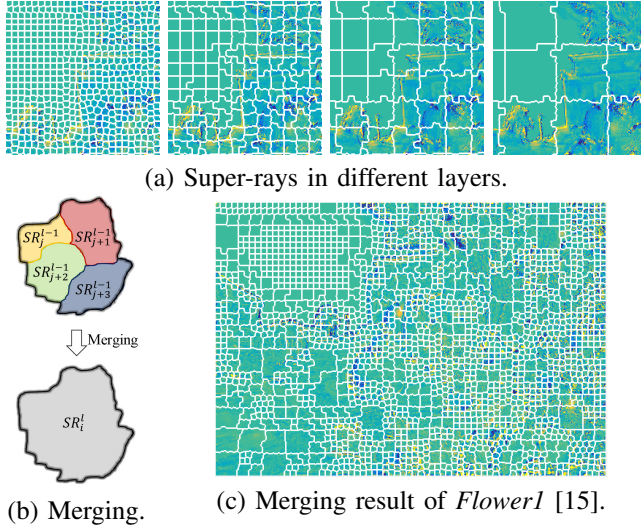(b) Merging.

(c) Merging result of *Flower1* [15].

Fig. 4. An illustration of super-ray merging. (a) Super-rays at different layer $l = 0, 1, 2, 3$ from left to right. (b) Super-rays $\{SR_j^{l-1}\} = \{SR_j^{l-1} \in SR_i^l, \forall j\}$ in layer $l - 1$ are merged into SR_i^l, i.e. $\{SR_j^{l-1}\} \Rightarrow SR_i^l$, if the rate-distortion cost $\mathcal{J}$ has been reduced. (c) A merging result of *Flower1* [15].

Algorithm 1: Merging based on rate-distortion minimization

Data: Super-rays at different layer $\{SR_i^{l=0}\}$, $\{SR_i^{l=1}\}$, $\{SR_i^{l=2}\}$, $\{SR_i^{l=3}\}$

Result: Merged super rays $\{SR_i\}$

Initialization: $\{SR_i\} = \{SR_i^{l=0}\}$;

for *Layer $l = 1$ to 3* **do**
  for *Each super-ray SR_i^l in layer l* **do**
    Compute $\mathcal{J}$ with $\{SR_i\}$ **Eq.4**;
    Find super-rays $\{SR_j^{l-1}\} = \{SR_j^{l-1} \in SR_i^l, \forall j\}$ in layer $l - 1$;
    Compute $\mathcal{J}'$ with $\{\{SR_i\} \setminus \{SR_j^{l-1}\}\} \bigcup SR_i^l$, *i.e.* $\{SR_j^{l-1}\}$ are replaced by SR_i^l **Eq.4**;
    if $\mathcal{J} > \mathcal{J}'$ **then**
      Merge, $\{SR_j^{l-1}\} \Rightarrow SR_i^l$;
      Update $\{SR_i\}$;
    end
  end
end

IV. EXPERIMENTS

We test our coding scheme on four real LF with 8×8 sub-aperture images of size $(X = 536, Y = 376)$ from the dataset used in [15], called *Flower1*, *Flower2* and *Cars*.

A. Energy Compaction

We first evaluate the effectiveness of our 4D segmentation, by analyzing the compaction of the energy after transfor-mation. Therefore, 4D SA-DCT is applied on: 1) super-pixels used in [16] (computed on synthesized color without disparity compensation or merging) 2) super-rays computed on synthesized residues without disparity compensation or merging 3) super-rays computed on synthesized residues with disparity compensation but without merging 4) super-rays computed on synthesized residues with disparity compensation and merging. Fig. 5 shows the percentage of energy carried by a given percentage of coefficients obtained by SA-DCT on these segmentations. The blue curve is the baseline method [16]. The red curves show the impact of using the synthesized residues to compute the segmentation. However, due the error in the synthesized residue, the improvement is limited. The Yellow curve shows the impact of using the disparity information, while the purple curve measures the effect of the merging. Thanks to the merging operation which compensates the errors in the super-ray segmentation, the proposed contributions bring a significant increase in terms of energy compaction ($\sim 10\%$) with respect to a direct use of a super-pixel segmentation per view [16].
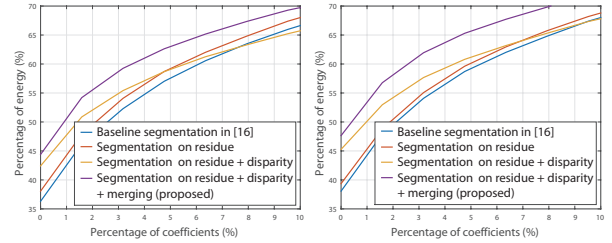

Fig. 5. Energy compaction of the transformed coefficients of *Flower1* and *Cars* using different super-ray computations. The super-rays computed on synthesized residues with merging and disparity compensation yield higher energy compaction.

B. Rate-PSNR results

The final rate-distortion performance of the proposed scheme is evaluated in comparison with three baseline meth-ods: 1) *HEVC-lozenge* [3], the whole LF is considered as a video and compressed by HEVC with a lozenge sequence, 2) *CNN+HEVC* [3], the same CNN based view synthesis is applied here, while the residues are compressed by HEVC, 3) *CNN+SA-DCT (no merging, no disparity)* [16], our previous coding scheme presented in [16] using the same CNN based view synthesis, however, there is no disparity compensation or super-merging strategy. Note that the coding methods using CNN based prediction are performed with best pairs of parameters (Q, QP) where Q is the quality parameter used to compress the residues and QP is used in the HEVC inter-coding of the four corners. The obtained rate-distortion curves are shown in Fig. 6.

The proposed CNN+SA-DCT coding scheme yields better or comparable rate-PSNR performance at low bitrate than HEVC based reference methods. The improvement of the proposed CNN+SA-DCT compared with the baseline method in [16] indicates the effectiveness of super-ray merging. It allows the proposed coding scheme to capture more informa-tion with fewer bits (at low bitrate), compared with HEVC
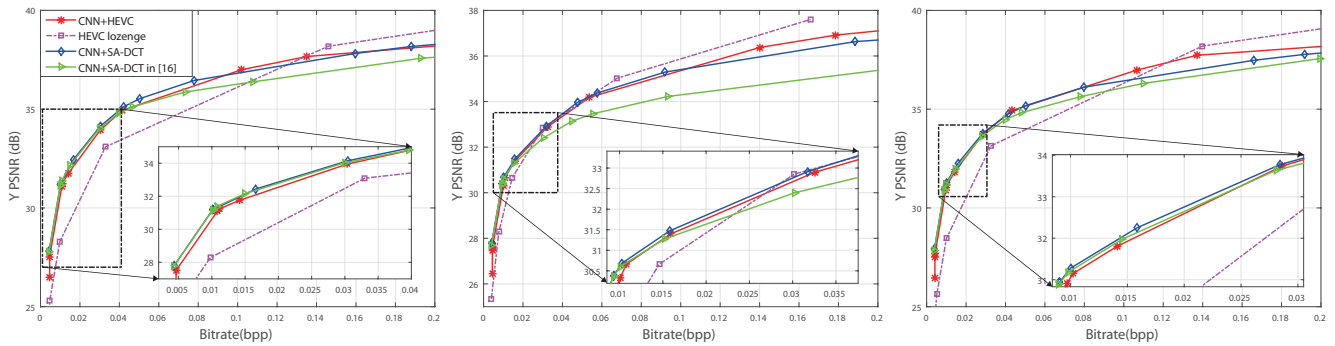
Fig. 6. Rate-distortion comparison. From left to right: *Flower 1*, *Cars* and *Flower 2*.

based encoders. Tab. I shows the improvement in terms of bjontegaard metric at low bitrate (< 0.04 bpp corresponding to a PSNR quality up to 35 dB) obtained by our coding scheme. However, as shown in Tab. I, at bitrates higher than 0.04 bpp, the HEVC based encoders (HEVC lozenge and CNN+HEVC) generally outperform the proposed coding scheme at high bitrates. This is due to the fact that the proposed scheme does not have very complex and high quality prediction strategies in residue coding which is useful at high bitrate.

TABLE I
BJONTEGAARD COMPARISON (ΔPSNR (DB)) AT *low* BITRATE (< 0.04 BPP) AND *high* BITRATE (> 0.04 BPP)

| | Our CNN+SA-DCT *vs* | | | | | |
| | CNN+HEVC | | HEVC lozenge | | CNN+SA-DCT in [16] | |
	Low	High	Low	High	Low	High
Flower 1	0.22	0.05	1.92	-0.19	0.03	0.55
Cars	0.21	-0.17	0.48	-0.50	0.17	1.08
Flower 2	0.1	-0.26	1.77	-0.85	0.09	0.38

V. CONCLUSION

In this paper, we have presented a rate-distortion optimized super-ray merging to exploit the correlation in the spatial and angular dimensions of light fields. After the CNN-based view synthesis, the residue inside each super-ray is compacted into a few coefficients using 4D shape-adaptive DCT transform. The experimental results show that the proposed light field coding scheme can yield rate-distortion gains compared with HEVC based compression, especially at low bitrate.

ACKNOWLEDGMENT

REFERENCES

[1] B. Wilburn, N. Joshi, V. Vaish, E.-V. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High performance imaging using large camera arrays," *ACM Transactions on Graphics (TOG)*, vol. 24, no. 3, pp. 765–776, 2005.

[2] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Computer Science Technical Report CSTR*, vol. 2, no. 11, pp. 1–11, 2005.

[3] M. Rizkallah, T. Maugey, C. Yaacoub, and C. 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.

[4] D. Liu, L. Wang, L. Li, Z. Xiong, F. Wu, and W. Zeng, "Pseudo-sequence-based light field image compression," in *Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–4.

[5] C. Conti, P. Nunes, and L. D. Soares, "HEVC-based light field image coding with bi-predicted self-similarity compensation," in *Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–4.

[6] C. Conti, L. D. Soares, and P. Nunes, "HEVC-based 3D holoscopic video coding using self-similarity compensated prediction," *Signal Processing: Image Communication*, vol. 42, pp. 59–78, 2016.

[7] Y. Li, M. Sjostrom, R. Olsson, and U. Jennehag, "Efficient intra prediction scheme for light field image compression," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 539–543.

[8] R. Monteiro, L. Lucas, C. Conti, P. Nunes, N. Rodrigues, S. Faria, C. Pagliari, E. da Silva, and L. Soares, "Light field HEVC-based image coding using locally linear embedding and self-similarity compensated prediction," in *Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–4.

[9] C. Perra and P. Assuncao, "High efficiency coding of light field images based on tiling and pseudo-temporal data arrangement," in *Multimedia & Expo Workshops (ICMEW), 2016 IEEE International Conference on*. IEEE, 2016, pp. 1–4.

[10] W. Ahmad, R. Olsson, and M. Sjöström, "Interpreting Plenoptic Images as Multi-View Sequences for Improved Compression," in *IEEE International Conference on Image Processing, Beijing, China 17-20 September 2017*, 2017.

[11] C. Jia, Y. Yang, X. Zhang, X. Zhang, S. Wang, S. Wang, and S. Ma, "Optimized Inter-View Prediction Based Light Field Image Compression With Adaptive Reconstruction," in *IEEE International Conference on Image Processing, Beijing, China 17-20 September 2017*, 2017.

[12] X. Jiang, M. Le Pendu, R. A. Farrugia, and C. Guillemot, "Light Field Compression With Homography-Based Low-Rank Approximation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 7, pp. 1132–1145, 2017.

[13] R. Verhack, T. Sikora, L. Lange, R. Jongebloed, G. Van Wallendael, and P. Lambert, "Steered mixture-of-experts for light field coding, depth estimation, and processing," in *Multimedia and Expo (ICME), 2017 IEEE International Conference on*. IEEE, 2017, pp. 1183–1188.

[14] X. Su, M. Rizkallah, T. Maugey, and C. Guillemot, "Graph-based light fields representation and coding using geometry information," in *IEEE International Conference on Image Processing (ICIP)*, 2017.

[15] N. K. Kalantari, T.-C. Wang, and R. Ramamoorthi, "Learning-based view synthesis for light field cameras," *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, p. 193, 2016.

[16] M. Rizkallah, X. Su, T. Maugey, and C. Guillemot, "Graph-based Transforms for Predictive Light Field Compression based on Super-Pixels," in *IEEE International Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, 2018.

[17] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[18] M. Hog, N. Sabater, and C. Guillemot, "Superrays for Efficient Light Field Processing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 11, no. 7, pp. 1187–1199, 2017.