



**HAL**  
open science

# Graph-based Spatio-angular Prediction for Quasi-Lossless Compression of Light Fields

Mira Rizkallah, Thomas Maugey, Christine Guillemot

► **To cite this version:**

Mira Rizkallah, Thomas Maugey, Christine Guillemot. Graph-based Spatio-angular Prediction for Quasi-Lossless Compression of Light Fields. DCC 2019 - Data Compression Conference, Mar 2019, Snowbird, United States. pp.379-388, 10.1109/DCC.2019.00046 . hal-02116369

**HAL Id: hal-02116369**

**<https://hal.science/hal-02116369v1>**

Submitted on 30 Apr 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Graph-based Spatio-angular Prediction for Quasi-Lossless Compression of Light Fields

Mira Rizkallah\*, Thomas Maugey†, and Christine Guillemot†

\*Universite de Rennes 1  
Campus de Beaulieu  
Rennes, 35042, France  
mira.rizkallah@inria.fr

†INRIA Rennes Bretagne Atlantique  
Campus de Beaulieu  
Rennes, 35042, France

## Abstract

Graph-based transforms have been shown to be powerful tools for image compression. However, the computation of the basis functions becomes rapidly untractable when the support increases, i.e. when the dimension of the data is high as in the case of light fields. Local transforms with limited supports have been investigated to cope with this difficulty. Nevertheless, the locality of the support may not allow us to fully exploit long term dependencies in the signal. In this paper, we describe a graph-based prediction solution that allows taking advantage of intra prediction mechanisms as well as of the good energy compaction properties of the graph transform. The approach relies on a separable spatio-angular transform and derives low frequency spatio-angular coefficients from one single compressed reference view and from the high angular frequency coefficients. In the tests, we used HEVC-Intra, with QP=0, to encode the reference frame with high quality. The high angular frequency coefficients containing very little energy are coded using a simple entropy coder. The approach is shown to be very efficient in a context of high quality quasi-lossless compression of light fields.

## 1 Introduction

The availability of commercial light field cameras has given momentum to the development of light field compression algorithms. Many solutions proposed so far adapt standardized image and video compression solutions (in particular HEVC) to light field data (*e.g.* [1] [2] [3]). Recent work investigated the use of homography-based low rank models for reducing the angular dimension [4], or local Gaussian mixture models in the 4D ray space [5]. The authors in [6], use a depth based segmentation of the light field into 4D spatio-angular blocks with prediction followed by JPEG-2000.

In this paper, we consider instead geometry adapted local graph transform with a support defined by super-rays for light field compression. Super-rays have been first introduced in [7] as an extension of super-pixels in the 3D domain 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. The method performs a k-means clustering of all light rays based on color and distance in the 3D space. Despite the limited support, computing the basis functions of the transform (i.e. the eigenvectors of the Laplacian of the graph) with a non separable spatio-angular graph (i.e. which

---

This work has been supported by the EU H2020 Research and Innovation Programme under grant agreement No 694122 (ERC advanced grant CLIM).

connects pixels both within super-pixels and the ray directions), remains intractable. We therefore consider separable transforms using the optimization procedure introduced in [8] to preserve angular correlation in the spatial transform domain. The local support however does not allow us to capture long term spatial dependencies, unlike efficient predictive schemes used in state of the art codecs as HEVC. To cope with this limitation, we propose a graph based spatio-angular prediction leading to a novel coding scheme for light fields.

The approach proposed in this paper is based on the observation that, when using a spatio-angular non separable transform, most of the light field energy is compacted in the low angular frequency coefficients. This motivates the design of approaches that would allow the best compression of these coefficients, e.g. by exploiting spatial correlation beyond the limits of the local graph transform support. More precisely, in this paper, we derive the equations allowing us to recover the low angular frequency coefficients of the spatio-angular transform from an encoded reference view and the high angular frequency coefficients. Thanks to these equations, the encoding of the reference view (using any classical encoder having efficient spatial predictors to exploit spatial correlation) is a way to efficiently encode the low angular frequency coefficients containing most of the light field energy. The proposed method can be seen as a graph-based prediction deriving low frequency spatio-angular coefficients from one single compressed reference view (e.g. top-left view in the experiments) and from the high angular frequency coefficients. The proposed method keeps the advantages of reduced basis function computational complexity due to the limited support, and of the energy compaction of the graph transform.

The method has been assessed in the context of quasi-lossless encoding of light fields. The compression algorithm proceeds as follows. A reference view is first compressed using a high quality encoding (we used HEVC with  $QP = 0$  in the experiments). All the coefficients (except for the low angular frequency coefficients) of the spatio-angular graph transforms are encoded using a simple arithmetic encoder. Experimental results show that, when coupled with a powerful intra-prediction tool, the graph-based spatio-angular prediction brings a substantial gain in bitrate reaching almost 30%.

The paper is organized as follows. In section 2, a brief overview of the concept of separable graph transforms based on super-rays is provided. Section 3 gives more insight on our main contribution, the so-called graph-based spatio-angular prediction before introducing in section 4 the quasi-lossless compression scheme proposed to evaluate our prediction. Experimental results are shown in Section 5.

## 2 Separable Graph Transforms

### 2.1 Local Support

The first step to construct the graph transform is to define its support. Here, we consider local supports based on super-rays. To be well suited for local separable graph transforms, the super-rays need to be consistent across the views. We hence consider the approach described in [8] to first compute super-pixels in the top-left

view, and then project, using the top left view disparity, its segmentation labels to all the other views. The occluded pixels are assigned the label of the neighboring super-ray corresponding to the foreground objects (*i.e.* having the higher disparity). As for appearing pixels, they will be clustered with the background super-rays (*i.e.* having the lower disparity).

## 2.2 Spatial Graph Transforms

We define the luminance signal in a view  $v$  of the light field as  $\mathbf{x}_v$ . A graph is then defined in each super-ray  $k$  of each view  $v$ . More precisely, pixels  $\mathbf{x}_{k,v}$  with the same label  $k$  and belonging to the same view  $v$  are gathered in a graph  $\mathcal{G}_{k,v} = \{\mathcal{V}_{k,v}, \mathcal{E}_{k,v}\}$ , whose edges are defined between horizontal and vertical adjacent pixels. For each graph  $\mathcal{G}_{k,v}$ , an adjacency matrix  $\mathbf{A}_{k,v}$  is defined, depicting the connections between pixels. The luminance of the pixels is considered as a real signal  $\mathbf{x}_{k,v}$  defined on the vertices  $\mathcal{V}_{k,v}$  of the graph. As proposed in [9], a graph transform is defined based on the Laplacian matrix  $\mathbf{L}_{k,v} = \mathbf{D}_{k,v} - \mathbf{A}_{k,v}$ , where  $\mathbf{D}_{k,v}$  is a diagonal degree matrix whose  $i^{\text{th}}$  diagonal element is equal to the sum of the weights of all edges incident to node  $i$ . The spatial graph transform basis is given by the eigenvectors  $\mathbf{U}_{k,v}$  of the Laplacian  $\mathbf{L}_{k,v}$ . The spatial graph transform coefficients for each graph are obtained by calculating:

$$\hat{\mathbf{x}}_{k,v} = \mathbf{U}_{k,v}^{\top} \mathbf{x}_{k,v}. \quad (1)$$

Inversely, the luminance values of the pixels belonging to the graph are retrieved from

$$\mathbf{x}_{k,v} = \mathbf{U}_{k,v} \hat{\mathbf{x}}_{k,v}. \quad (2)$$

The spatial graph Laplacian in the super-ray  $k$  undergoes slight changes in different views  $v$  due to occlusions. Therefore, the basis functions of each spatial graph transform  $\mathbf{U}_{k,v}$  are not the same in the different views  $v$ , resulting in decreased correlation of the spatial transform coefficients  $\hat{\mathbf{x}}_{k,v}$  across views. To circumvent this issue, we have proposed in [8], a basis optimization procedure to find the basis functions  $\mathbf{U}_{k,v}$  that approximately diagonalize their respective Laplacians  $\mathbf{L}_{k,v}$  while being coherent across the views, given the scene geometry. By coherence, we mean the similarity of the eigenvectors in different views along the epipolar lines of the light field. In the following,  $\mathbf{U}_{k,v}$  denotes the optimized basis functions. Inside a super-ray  $k$ , the spatial transform coefficients  $\hat{\mathbf{x}}_{k,v}$  are thus correlated between the views  $v$ . In the next section, we explain how an angular transform is used to capture this correlation.

## 2.3 Angular Graph Transforms

The purpose of this transform is to capture the correlation existing between the spatial transformed coefficients of each super-ray along the rays direction. For a super-ray  $k$ , we denote by spatial-band  $b$ , the vector formed by all the  $b^{\text{th}}$  spatial transform coefficients existing in the views. More precisely,  $\hat{\mathbf{x}}_k^b = [\hat{\mathbf{x}}_{k,v}(b)]_{v=1:V}$ , s.t.  $b < |\mathbf{x}_{k,v}|$ , where  $V$  is the number of views.

We then define the graphs across views, for each band  $b$ . We link the horizontal and adjacent views together with edges with weights of 1. If we have isolated nodes,

we connect them to their nearest neighbor based on the euclidean distance between view indices. Since we do not have necessarily the same number of pixels for a specific super-ray in all the views, the number of spatial coefficients after the spatial transform differs along the views. Therefore, for each spatial-band  $b$  as previously defined, we build a different graph between the views where the band  $b$  exists.

For each super-ray  $k$ , and for each band  $b$ , the Adjacency  $\mathbf{A}_k^b$  and degree  $\mathbf{D}_k^b$  matrices are used to compute the inter-view Laplacian as  $\mathbf{L}_k^b = \mathbf{D}_k^b - \mathbf{A}_k^b$ . The eigenvectors  $\mathbf{V}_k^b$  of  $\mathbf{L}_k^b$  are then used to define the angular graph transform basis. The angular transform coefficients are obtained by calculating:

$$\hat{\mathbf{x}}_k^b = \mathbf{V}_k^{b\top} \hat{\mathbf{x}}_k^b. \quad (3)$$

### 3 Graph-based Spatio-Angular Prediction

Due to the high level of correlation between the different views of a light field, the energy of the transformed coefficients  $\hat{\mathbf{x}}_k^b$  is highly compacted in the low angular frequencies. However, some correlation remains between the coefficients of different super-rays  $k$ , and is not yet captured by the graph-transforms introduced previously. We thus propose to use already existing tools that have been proven to efficiently tract these correlations: efficient image coders involving intra prediction methods. More precisely, a reference view is coded using one of these efficient coders, and is used to recover a part of the spatio-angular spectrum at the receiver side, precisely, the one with the highest energy. In other words, we propose to compact most the light field energy on a very small set of coefficients thanks to the spatio-angular transform introduced before. And then, instead of simply coding these coefficients, we recover them from the reference view and the high angular frequencies. For that aim, we introduce a prediction tool that works as follows.

Let us consider that the view 1 is coded as a reference. For a given super-ray, its spatial transform is  $\hat{\mathbf{x}}_{k,1} = \mathbf{U}_{k,1}^\top \mathbf{x}_{k,1}$  according to notations introduced in Sec. 2.2. For a given band  $b$ , the inverse angular transform is defined as

$$\hat{\mathbf{x}}_k^b = \mathbf{V}_k^b \hat{\mathbf{x}}_k^b \quad (4)$$

$$\Leftrightarrow \begin{bmatrix} \hat{\mathbf{x}}_k^b(1) \\ \hat{\mathbf{x}}_k^b(2) \\ \vdots \\ \hat{\mathbf{x}}_k^b(N_b) \end{bmatrix} = \begin{bmatrix} \mathbf{V}_k^b(1,1) & \mathbf{V}_k^b(1,2) & \cdots & \mathbf{V}_k^b(1,N_b) \\ \mathbf{V}_k^b(2,1) & \mathbf{V}_k^b(2,2) & \cdots & \mathbf{V}_k^b(2,N_b) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{V}_k^b(N_b,1) & \mathbf{V}_k^b(N_b,2) & \cdots & \mathbf{V}_k^b(N_b,N_b) \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_k^b(1) \\ \hat{\mathbf{x}}_k^b(2) \\ \vdots \\ \hat{\mathbf{x}}_k^b(N_b) \end{bmatrix} \quad (5)$$

where  $N_b$  denotes the number of views where the  $b^{th}$  band of the  $k^{th}$  super-ray is defined. Since the view 1 is transmitted separately,  $\hat{\mathbf{x}}_k^b(1)$  is available at the decoder. If we only transmit  $\hat{\mathbf{x}}_k^b(2), \dots, \hat{\mathbf{x}}_k^b(N_b)$ , then we are able to retrieve  $\hat{\mathbf{x}}_k^b(1)$  from the following equation:

$$\hat{\mathbf{x}}_k^b(1) = \frac{1}{\mathbf{V}_k^b(1,1)} \left( \hat{\mathbf{x}}_k^b(1) - [ \mathbf{V}_k^b(1,2) \cdots \mathbf{V}_k^b(1,N_b) ] \begin{bmatrix} \hat{\mathbf{x}}_k^b(2) \\ \vdots \\ \hat{\mathbf{x}}_k^b(N_b) \end{bmatrix} \right). \quad (6)$$

Equation (6) is our so-called graph-based spatio-angular prediction. The high-frequency coefficients  $\hat{\mathbf{x}}_k^b(2), \dots, \hat{\mathbf{x}}_k^b(N_b)$  can be seen as prediction coefficients, transmitted to recover the exact light field at the decoder. The basis of the linear prediction is the graph-transform basis, which makes these coefficients low-energetical and thus easy to transmit. The spatial coefficients of all the views are then retrieved from the following equation

$$\begin{bmatrix} \hat{\mathbf{x}}_k^b(2) \\ \vdots \\ \hat{\mathbf{x}}_k^b(N_b) \end{bmatrix} = \begin{bmatrix} \mathbf{V}_k^b(2, 1) \\ \vdots \\ \mathbf{V}_k^b(N_b, 1) \end{bmatrix} \hat{\mathbf{x}}_k^b(1) + \begin{bmatrix} \mathbf{V}_k^b(2, 2) & \cdots & \mathbf{V}_k^b(2, N_b) \\ \vdots & \ddots & \vdots \\ \mathbf{V}_k^b(N_b, 2) & \cdots & \mathbf{V}_k^b(N_b, N_b) \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}_k^b(2) \\ \vdots \\ \hat{\mathbf{x}}_k^b(N_b) \end{bmatrix}$$

We will show in the experimental section that the approach is very efficient when the quantization noise on the reference view is limited, hence for quasi-lossless compression. If the latter is too coarsely compressed, drift and noise amplification might appear during this prediction step. This is due to the fact that, in Equation (6), the prediction uses the spatial transform coefficients estimated on the reference view available at the decoder side. Further study will be dedicated to addressing this problem in the case of lossy compression.

## 4 The proposed coding scheme

### 4.1 Overall description of the coding scheme

The proposed quasi-lossless coding scheme is shown in Figure 1. In order to explore the inter super-rays dependencies (dependencies between different super-rays), the top left view  $\mathbf{x}_1$  is at first compressed with HEVC as an intra frame. It is then separated into uniform regions using the SLIC algorithm ([10]) to segment the image into super-pixels, and its disparity map is estimated with [11]. The disparity values are encoded using simple arithmetic coder. Using both the segmentation map and the geometrical information, we build consistent super-rays and graphs in and across all views as explained in section 2.

We apply the separable coherent spatio angular graph transforms on the coded version of the first view (quasi-lossless coding) and the original versions of all other frames to compact their energy in fewer coefficients. Since the first view is coded with very small QP, we are then almost sure that we are not adding angular incoherence between the first view and the rest of the views. Once we have the spatio-angular coefficients, instead of sending the whole spectrum with simple arithmetic coding, we propose to make use of our graph-based prediction and therefore deriving low frequency spatio-angular coefficients in the decoder side from the top-left reference view coded with HEVC-Intra and from the high angular frequency coefficients.

We thus send, for all super-rays, the AC coefficients of the  $N_{k,1}$  first bands after the second angular graph transform. ( $N_{k,1}$  is the number of pixels belonging to the super-ray  $k$  in view 1). We come back to the detailed coding of those coefficients in the sequel.

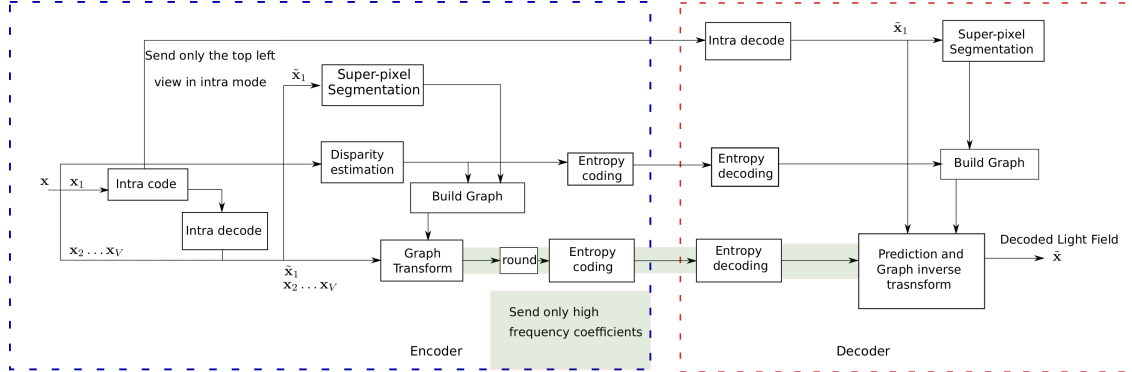


Figure 1: Overview of proposed coding scheme.

Note that to construct the spatial and angular graphs on the decoder side, we do not need extra information. Since the decoder already receives the top left image and the disparity values, with the SLIC algorithm, it can deduce the segmentation map and the super-ray segmentation used for local graph transforms design and application. Furthermore, the decoder have received the spatio-angular high frequency coefficients. Using those two kinds of information, it can predict the DC spatio-angular components obtained after the angular graph transform and then reconstruct the luminance values of all the views as explained in Section 3.

#### 4.2 The coding of the spatio-angular high frequency coefficients

After applying the spatio-angular graph transforms on the whole super-ray, all frequency coefficients are grouped into a three-dimensional array  $\mathbf{y}$  where  $\mathbf{y}(k, b, v)$  is the  $v^{th}$  transformed coefficient of the band  $b$  for the super-ray  $k$ . Using the observations on all the super-rays in a training dataset (*Rock*), we can find the best ordering for quantization. We first sort the variances of coefficients with enough observations in decreasing order. Using this scanning order, we assign a class number to each super-ray. More precisely, each super-ray with  $N$  coefficients belongs to class  $i$  if the mean energy per high frequency coefficient is less than 1, where the high frequency coefficients are the last  $\text{round}(N \times i/4)$  coefficients following the scanning order of the super-rays coefficients defined previously.

We start by finding the super-rays in the first class than remove them from the search space before finding the other classes, and idem for the following steps. We code a flag with an arithmetic coder to give the information of the class of super-rays to the decoder side. In class  $i$ , the last  $\text{round}(N \times i/4)$  coefficients of each super-ray are discarded. Then, the rest of the high frequency spatio-angular coefficients are quantized uniformly with a small step size  $Q = 1$ . Then, they are grouped in 32 uniform groups to enter a simple arithmetic coding assuming an iid model.

## 5 Experimental results



Figure 2: An example set of light fields used in our experiments. Only the top-left view is shown for illustration purpose. From left to right: *Flower2*, *Cars*, *StonePillarsInside* and *Friends*.



Figure 3: An example set of top-left view disparity maps used in our experiments

To evaluate our proposed spatio-angular prediction method in the full coding scheme, we apply it on real light fields captured by plenoptic cameras from the datasets used in [12] and [13]. To avoid the strong vignetting and distortion problems on the views at the periphery of the light field, we only consider the  $8 \times 8$  central sub-aperture images cropped to  $364 \times 524$  in [12], and  $9 \times 9$  cropped to  $432 \times 624$  from [13]. Some of the light fields considered are shown in Figure 2. The full set of light fields considered for the test is: *Flower2*, *Cars*, *Rock* and *Seahorse* from the dataset in [12] and *StonePillarInside* and *Friends* from the dataset of ICIP challenge 2017 and used in [13]. The method used to estimate the disparity of the top-left views is described in [11]. Examples of the disparity maps provided are shown in Figure 3. A sparse set of disparity values and the segmentation maps computed with SLIC [10] are used to construct local graph supports as described in Section 2. In our experiments, we fix the number of super-rays for all datasets to 4000. Indeed, an ideal segmentation for compression purposes, would be the result of a local rate-distortion optimization (as previously proposed in [14]). For the sake of simplicity in this paper, we use a simple non adaptive segmentation.

### 5.1 The importance of exploring the long term dependencies between super-rays

As explained before, we aim at compacting the energy of the light field luminance in fewer coefficients, then we will predict them based on the knowledge of a coded top-left view and the high frequency components that most likely are sparse. To illustrate



the effect of our spatio-angular transforms on the energy compaction, we compute in Table 1 the percentage of total energy that resides in the predicted DC spatio-angular bands  $\hat{\mathbf{x}}_k^b(1) \forall k, b$ . Those bands are actually not sent but predicted on the decoder side, and only the energy that is left is coded as extra information to reconstruct the original light field. We can observe that most of the energy is compacted in those coefficients, which shows the efficiency of the spatio-angular decorrelation of the Graph transforms.

Table 1: Percentage of energy residing in the DC spatio-angular bands  $\hat{\mathbf{x}}_k^b(1) \forall k, b$

Light Fields	Energy Percentage in $\hat{\mathbf{x}}_k^b(1) \forall k, b$
Flower 2	99.02 %
Cars	99.34 %
Rock	98.45 %
Seahorse	98.73 %
Stone Pillars Inside	98.26 %
Friends	99.80 %

Thanks to the equations introduced in Section 3, the encoding of the top-left view (using any classical encoder having efficient spatial predictors to exploit spatial correlations) can be seen as a way to efficiently encode those DC spatio-angular frequency coefficients which contain most of the light field energy. We choose the HEVC intra in order to exploit long term dependencies between super-rays. More precisely, we code the top left view with HEVC Intra. Then in Table 2, we compare its performance with respect to using a simple arithmetic coder for the DC spatio-angular frequency coefficients. The QP of HEVC is set to 0. As for the arithmetic coder, the observations of each DC spatio-angular frequency coefficient  $\hat{\mathbf{x}}_k^b(1) \forall k, b$  are coded separately. We can see for most of the datasets, the gain obtained with HEVC intra when compared to arithmetic coding without any spatial prediction.

Table 2: Rate comparison between HEVC intra-coding of the first view and entropy coding of all the DC spatio-angular bands  $\hat{\mathbf{x}}_k^b(1) \forall k, b$  using an arithmetic coder assuming an iid model

Light Fields	HEVC intra	Entropy coding
Flower 2	1.0 Mbits	1.48 Mbits
Cars	1.06 Mbits	1.54 Mbits
Rock	1.02 Mbits	1.38 Mbits
Seahorse	0.72 Mbits	1.22 Mbits
Stone Pillars Inside	1.81 Mbits	1.66 Mbits
Friends	1.62 Mbits	2.04 Mbits

## 5.2 Comparative assessment

To assess our Graph-based Spatio-Angular Prediction, we evaluate it in the context of the quasi-lossless coding scheme in Section 4 against a complete HEVC based

scheme with a QP set to 0 and a GOP of 4. The HEVC version used in the tests is HM-16.10. The light fields are coded following a raster scanning starting with the top-left view as a reference intra-coded frame. Results are reported in Table 3 where we compare mainly the rate needed to code a Light Field under quasi-lossless settings (based on visual quality assessment, we consider a PSNR higher than 50 dB as a quasi-lossless compression). A substantial gain in bitrate is observed while preserving a high quality of the reconstructed light fields. This can be justified by the efficiency of our spatio-angular graph transforms in terms of compaction along with the ability of HEVC-intra to effectively exploit spatial correlations in the reference view.

Table 3: Rate comparison between our proposed approach and a traditional HEVC-inter according to a raster scan order for high PSNR ( $> 50$  dB)

Light Fields	HEVC Raster Scan	Ours
Flower 2	3.3129 bpp (54.2033 dB)	2.4799 bpp (55.1969 dB)
Cars	3.6688 bpp (54.0812 dB)	2.6258 bpp (55.2009 dB)
Rock	3.2700 bpp (53.7601 dB)	2.0162 bpp (54.7765 dB)
Seahorse	2.4751 bpp (54.3804 dB)	1.9762 bpp (55.2844 dB)
Stone Pillars Inside	4.9017 bpp (52.1036 dB)	3.3094 bpp (55.0022 dB)
Friends	3.5400 bpp (52.7986 dB)	2.4436 bpp (54.8196 dB)

## 6 Conclusion

In this paper, we have explored local separable spatio-angular graph transforms for light fields compact representation. The limited support of local transforms may not allow us to exploit long term spatial dependencies. To cope with this limitation, we have proposed a novel approach to leverage the good spatial de-correlation properties of traditional codecs (e.g. HEVC intra), making use of efficient predictors, into local spatio-angular graph transforms. A reference view coded with any efficient codec is used to predict low angular frequency transform coefficients that, together with the transmitted high angular transform coefficients, allow recovering the entire graph-based representation. The scheme has been assessed for high quality (quasi-lossless) coding of light fields with narrow baselines. Note that the prediction mechanism, in case of coarser quantization, tends to amplify the quantization noise on the reconstructed data representation and the method need to be tuned for wider baselines. Those issues are left for further study.

## References

- [1] Dong Liu, Lizhi Wang, Li Li, Zhiwei Xiong, Feng Wu, and Wenjun Zeng, “Pseudo-sequence-based light field image compression,” in *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2016, pp. 1–4.
- [2] Caroline Conti, Luís Ducla Soares, and Paulo Nunes, “Hevc-based 3d holoscopic video coding using self-similarity compensated prediction,” *Signal Processing: Image Communication*, vol. 42, pp. 59–78, 2016.

- [3] Yun Li, Roger Olsson, and Mårten Sjöström, “Compression of unfocused plenoptic images using a displacement intra prediction,” in *2016 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*. IEEE, 2016, pp. 1–4.
- [4] Xiaoran Jiang, Mikaël Le Pendu, Reuben A Farrugia, and Christine 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.
- [5] Ruben Verhack, Thomas Sikora, Lieven Lange, Rolf Jongebloed, Glenn Van Wallendaël, and Peter Lambert, “Steered mixture-of-experts for light field coding, depth estimation, and processing,” in *2017 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2017, pp. 1183–1188.
- [6] Ioan Tabus, Petri Helin, and Pekka Astola, “Lossy compression of lenslet images from plenoptic cameras combining sparse predictive coding and jpeg 2000,” in *2017 IEEE International Conference on Image Processing (ICIP)*, 2017, pp. 4567–4571.
- [7] M. Hog, N. Sabater, and C. Guillemot, “Super-rays for efficient light field processing,” *IEEE J. on Selected Topics in Signal Processing, special issue on light field image processing*, Oct. 2017.
- [8] M. Rizkallah, X. Su, T. Maugey, and C. Guillemot, “Geometry-aware graph transforms for light field compact representation,” *IEEE Trans. on image processing, submitted*, 2018.
- [9] David I Shuman, Sunil K Narang, Pascal Frossard, Antonio Ortega, and Pierre Vandergheynst, “The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains,” *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83–98, 2013.
- [10] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine 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.
- [11] Xiaoran Jiang, Mikaël Le Pendu, and Christine Guillemot, “Depth estimation with occlusion handling from a sparse set of light field views,” in *2018 IEEE International Conference on Image Processing (ICIP)*, 2018.
- [12] Nima Khademi Kalantari, Ting-Chun Wang, and Ravi Ramamoorthi, “Learning-based view synthesis for light field cameras,” *ACM Transactions on Graphics (TOG)*, vol. 35, no. 6, pp. 193, 2016.
- [13] Irene Viola, Hermina Petric Maretic, Pascal Frossard, and Touradj Ebrahimi, “A graph learning approach for light field image compression,” in *Applications of Digital Image Processing XLI*. International Society for Optics and Photonics, 2018, vol. 10752, p. 107520E.
- [14] Xin Su, Mira Rizkallah, Thomas Maugey, and Christine Guillemot, “Rate-distortion optimized super-ray merging for light field compression,” in *European Signal Processing Conference (EUSIPCO)*, 2018.