

# JPEG PLENO LIGHT FIELD ENCODER WITH MESH BASED VIEW WARPING

*Yue Li, Reji Mathew, David Taubman*

School of Electrical Engineering and Telecommunications  
University of New South Wales, Sydney

## ABSTRACT

We introduce mesh-based view warping to the JPEG Pleno light field coding framework and replace the standardized sample-based forward warping and splatting of reference texture with mesh-based backward warping, which allows for a more disciplined interpolation of the reference texture for predicting the target view. Instead of coding depth maps with JPEG 2000, which is the default option of the JPEG Pleno framework, we employ a recent extension referred to as JPEG 2000 Part 17. This extension utilises breakpoints to describe discontinuity boundary geometry for the purpose of modifying the predict and update lifting steps in the vicinity of detected discontinuities. We directly decode breakpoints and corresponding DWT coefficients onto a mesh and describe a scheme to construct a single, consolidated mesh for a large group of views, borrowing information from multiple coded depth maps. Results show that the cumulative impact of all these modifications enable improved rate-distortion performance in comparison with the default operation of the JPEG Pleno encoder.

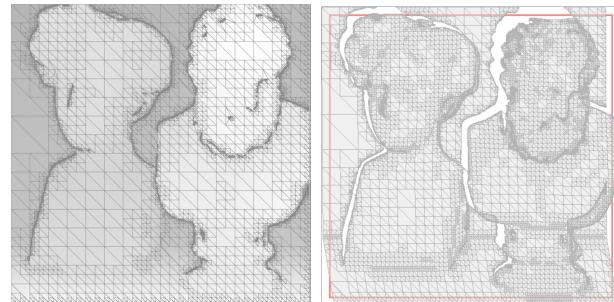
**Index Terms**— Light Fields, Triangular Mesh, DWT

## 1. INTRODUCTION

The recent JPEG Pleno light field coding standard [1] enables a 4D Prediction Mode (4DPM), where a given target view can be predictively coded using previously decoded reference views. To facilitate inter-view prediction, a small set of depth maps at select view locations are also coded, such that depth information along with corresponding camera parameters can be used to calculate the disparity between views. A prediction is formed by subjecting a set of available reference views to warping and merging in accordance with the disparity information determined at the encoder.

In this work we explore the cumulative impact of a number of modifications to the default behaviour of the JPEG Pleno light field encoder operating in 4DPM [2] [3]. Instead of defaulting to JPEG 2000 for coding depth maps, we employ a recent extension to the standard, referred to as JPEG 2000 Part 17 [4], which is specifically suited to coding piecewise smooth imagery. The Part 17 extension employs breakpoints to model discontinuities in the depth flow and these breakpoints in turn modify the DWT in its vicinity. We refer to this new transform as breakpoint dependent DWT (BD-DWT).

Another important modification relates to representing depth information with a mesh model to perform mesh-based view warping. This is in contrast to JPEG Pleno light field coding which assumes decoded depth is sample-based, such that depth samples at each pixel location are converted to a corresponding disparity vector with the knowledge of camera parameters. Each disparity vector specifies the displacement of a pixel from the reference view, where depth is available, to a target view. The JPEG Pleno encoder uses these derived disparity vectors to forward warp reference texture pixels to a



**Fig. 1.** (Left) Example of a mesh representation for a depth map of the Greek dataset. (Right) When a mesh is projected to another view, tears occur at break-induced discontinuities, leaving holes.

target view. Since warped locations may not necessarily fall at integer pixel locations at the target view, a splatting process is employed which involves splatting to the nearest neighbouring pixel location. The splatting process can cause excessive smoothing or blurring and does not allow for a disciplined filtering and up-sampling approach. In this work, breakpoints and wavelet coefficients of a depth map subject to BD-DWT, are decoded directly onto a mesh. The BD-DWT is defined on a hierarchical triangular grid allowing for triangular mesh representation which facilitates mesh-based view warping using triangular cells. The mesh describes a piece-wise continuous representation of depth with breakpoints denoting the boundary geometry along which tears may occur due to occlusion (double mappings) or dis-occlusion (holes) when performing view warping. Importantly, mesh-based warping is invertible, thereby enabling backward warping which allows for disciplined view interpolation.

A further modification which we introduce is the ability to fuse together depth information from multiple communicated depth maps to form a single comprehensive base-mesh. For a large group of views, such as views captured by a high density camera array, multiple depth maps are often provided. Additional depth maps provide depth details of the scene that may be hidden or not visible from a single view point. In our scheme, an initial base-mesh is first created by directly decoding onto a triangular mesh the BD-DWT data of a single depth map. If further depth maps are available then the base-mesh is augmented with additional information sourced from these new depth maps for regions that are not visible from the initial base view location. A single augmented base-mesh, that can describe the geometric relationship between views of a large view array, provides for a consolidated and consistent mechanism for deriving disparity for view warping.

Rate-distortion (R-D) results show that the cumulative impact of all three modifications result in significant bit-rate savings. Subjective results also showcase the advantages of our proposed augmented base-mesh, allowing for improved prediction at object boundaries and in dis-occluded regions.

## 2. RELATION TO PRIOR WORK

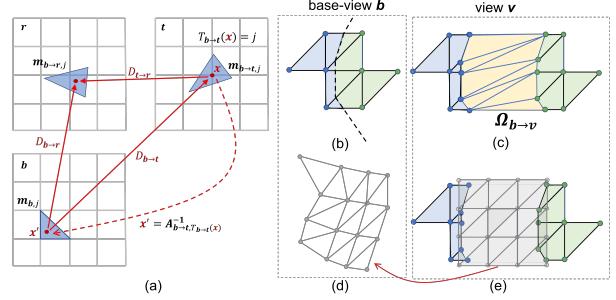
Wavelets defined over triangles which are dependent on breakpoints (i.e. BD-DWT) and the estimation of these breakpoints have been described in prior work [5] [6]. Essentially, depth samples are considered to be located on a triangular grid with lifting steps performed across the horizontal, vertical and diagonal grid line segments. Breakpoints which describe boundary geometry are placed on the arcs of the triangular grid and adapt the predict and update steps to ensure that the wavelet basis kernels do not cross over detected discontinuity boundaries. The benefits of adopting BD-DWT to the JPEG Pleno light field encoder have been reported in [6]. This prior work is limited to replacing the default JPEG 2000 coding of depth maps with the JPEG 2000 Part-17 extension which defines the BD-DWT option on a triangular grid; the view warping procedures remain unchanged. In this paper we build a mesh utilising breakpoints and BD-DWT coefficients and employ this mesh to perform backward warping of texture, replacing the sample based forward warping and splatting defined by the JPEG Pleno light field coding standard.

Mesh-based reconstruction of wavelet-transformed data has been considered earlier [7] [8]. More recently, we explored building a mesh from BD-DWT data [5] with results showing that the piecewise affine mesh representation, comprised of triangular cells, can accurately recreate depth samples at integer pixel locations. With the number of triangular cells being typically far smaller than the number of pixels, mesh-based representations also form a more compact description of the information conveyed by a depth map. In our prior efforts [9], mesh-based view warping was studied from a purely view synthesis perspective, omitting the steps of coding. In this work, we incorporate mesh-based view warping to the JPEG Pleno light field coding framework and report on R-D improvements.

We introduced the concept of a base-mesh in our earlier publication [10] in which the mesh is anchored or co-located with a communicated depth map. The base-mesh investigated in our prior study [10] is unable to incorporate depth information from additional depth maps. When performing view warping, regions of dis-occlusions or holes are filled in using a backfilling strategy which is a self-inferencing scheme as it relies upon only the base-view depth information; depth is inferred by extrapolation from the background side at mesh discontinuities. In this work we are able to augment a base-mesh with depth information from other depth maps, allowing for a single consolidated base-mesh to be used for view warping for the complete view array. Our earlier work [10] was conducted prior to the standardisation of the JPEG Pleno framework and therefore the R-D impact of incorporating mesh-based view warping to the standardized encoder was not previously investigated. Compared to previous efforts [9] [10], we provide improved dis-occlusion handling with a disciplined mesh augmentation strategy and report on benefits to JPEG Pleno encoding.

## 3. CONSTRUCTING A MESH

In this section we provide a brief overview of mesh construction from BD-DWT data. Further details can be found in [5]. We generate mesh models through progressive mesh subdivision in accordance with the appearance of non-zero wavelet coefficients or any novel breakpoints during the BD-DWT synthesis process. We consider subband samples to be located on a coarse to fine triangular grid with synthesis lifting steps performed across the horizontal, vertical and diagonal grid line segments at each scale. We start at the coarsest level of the transform and associate the LL subband data with nodes



**Fig. 2.** (a)  $D_{t-r}$  derived from base-mesh  $M_b$ . (b) Initial base-mesh  $M_b^0 = M_b$ . (c) Dis-occluded region  $\Omega_{b-v}$  when  $M_b^0$  is warped to view  $v$ . (d) Warped mesh elements comprising a new layer  $M_v^l$ . (e) mesh elements of  $M_v$  overlapping with dis-occluded region  $\Omega_{b-v}$ .

$\{n_i\}$  of a mesh, such that these nodes correspond to vertices of triangular mesh elements  $m_j = [n_g, n_k, n_q]$  whose edges form the coarsest level horizontal, vertical and diagonal arcs. As coefficients from other subbands and finer scales are progressively incorporated, if any non-zero high-pass subband coefficient appears within a triangle  $m_j$  or on any of its bounding arcs then the mesh element  $m_j$  is classified as being non-affine and subdivided into smaller triangles. A non-zero update step, from a subband sample located outside a triangle, can also render the mesh element  $m_j$  non-affine if the update is applied to a point that is not one of the three nodes (i.e. vertices)  $[n_g, n_k, n_q]$  of the triangle  $m_j$ . Non-zero updates therefore can also cause mesh elements to be sub-divided.

Breakpoints derived from the highly scalable JPEG 2000 Part 17 compressed representation provide a piecewise linear description of discontinuity boundaries. Only a small subset of breakpoints are explicitly communicated, these are novel breakpoints and termed *vertices* while the remaining breakpoints are induced along the piecewise linear geometry. Mesh elements are partitioned along break-inducing lines and any subsequent breakpoints induced along this line at finer levels are not considered novel and therefore do not result in any subdivision. In contrast, an appearance of a novel breakpoint will cause a redefinition of the boundary geometry and will result in further subdivisions of the triangular mesh elements.

A mesh constructed for a depth map of the Greek dataset is shown in Fig 1. Regions of smooth depth flow are modelled by large triangles while regions with more complex variations and non-linear boundary geometry require sub-division to finer mesh elements. Fig 1 also shows a warped mesh with dis-occluded regions.

## 4. MESH-BASED WARPING

In the JPEG Pleno coding framework, depth maps are coded only for a sparse set of views and therefore disparity needs to be inferred at other remaining views. We create a mesh at a view  $b$  for which depth is communicated and then augment this mesh to form a comprehensive description of disparity for all views. We term view  $b$  as a base view with the corresponding mesh  $M_b$  being the base-mesh. We first discuss transporting disparity information from the base view to reference and target views to allow for backward warping. Specific details of dealing with dis-occlusion and mesh augmentation are presented later.

### 4.1. Deriving Disparity from a Base-Mesh Model

Using notation  $Z_b$  to refer to the depth map at view  $b$ , the disparity field between  $b$  and another view  $v$  can be written as

$$D_{b \rightarrow v}(\mathbf{x}) = \bar{D}(C_b, C_v, \mathbf{x}, Z_b(\mathbf{x}))$$

where  $C_b$  and  $C_v$  represent the camera parameters describing the projective geometry at views  $b$  and  $v$  respectively, while the transformation  $\bar{D}$  is derived from perspective geometry.

To determine the disparity between any pair of views  $v$  and  $w$ , we first project each mesh element  $\mathbf{m}_{b,j}$  of  $\mathcal{M}_b$  to corresponding elements  $\mathbf{m}_{b \rightarrow v,j}$  by displacing the respective cell nodes according to  $\mathbf{n}_{b \rightarrow v,i} = \mathbf{n}_{b,i} + D_{b \rightarrow v}(\mathbf{n}_{b,i})$ . In this notation for mesh elements and nodes, the first subscript identifies the particular base-view at which the mesh is anchored (i.e.,  $b$ ) or a non-base-view it projects to (i.e.,  $b \rightarrow v$ ), while the second subscript corresponds to a unique triangle-ID  $j$  or node ID  $i$ . Since mesh elements are triangular, the mapping of  $\mathbf{m}_{b,j}$  to  $\mathbf{m}_{b \rightarrow v,j}$  is a 2D affine transformation  $\mathbf{A}_{b \rightarrow v,j}$ .

A triangle-ID map  $T_{b \rightarrow v}$  at view  $v$  is constructed by assigning  $T_{b \rightarrow v} = j$  for each location  $\mathbf{x} \in \mathbf{m}_{b \rightarrow v,j}$ . When multiple triangles  $\mathbf{m}_{b \rightarrow v,j}$  overlap in view  $v$ , the ambiguity due to occlusion (i.e., double mapping) is resolved by assigning  $T_{b \rightarrow v}(\mathbf{x})$  to the triangle-ID  $j$  whose depth at location  $A_{b \rightarrow v}^{-1}(\mathbf{x})$  in view  $b$  is smallest. Following this, the disparity field  $D_{v \rightarrow w}$  is obtained as

$$D_{v \rightarrow w}(\mathbf{x}) = D_{b \rightarrow w}(A_{b \rightarrow v, T_{b \rightarrow v}(\mathbf{x})}^{-1}(\mathbf{x})) - D_{b \rightarrow v}(A_{b \rightarrow v, T_{b \rightarrow v}(\mathbf{x})}^{-1}(\mathbf{x})) \quad (1)$$

Fig. 2 (a) illustrates the implementation of equation (1) for determining  $D_{t \rightarrow r}$ , the disparity field from a target view  $t$  to a reference view  $r$  using mesh elements anchored at the base view  $b$ . From  $D_{t \rightarrow r}$  a disparity vector for each pixel in  $t$  pointing to a location in  $r$  can be determined, thereby enabling backward warping.

## 4.2. Mesh augmentation

At the base view, when a breakpoint induced line intersects with an arc of the triangular grid, new nodes are added to either side of the discontinuity. Typically, breakpoint induced line segments describe the boundary between foreground and background content. New nodes placed on the background side of the discontinuity take on background depth, extrapolated from values from endpoint arc nodes on the same side of the discontinuity. Similarly, new nodes placed on the foreground side take on foreground depth extrapolated from foreground arc nodes. These nodes on either side of a boundary are connected together to form *infinitesimal mesh elements*, which have no area within the base-view, but expand to identify depth dis-occluded regions in other views. An example is shown in Fig. 2 (b) and (c). At the base view, in Fig. 2 (b), the breakpoint defined boundary is shown by the dashed line. Infinitesimal mesh elements are created at the base view  $b$  such that when warped to another view  $v$ , these elements expand to create stretched triangles highlighting dis-occluded regions  $\Omega_{b \rightarrow v}$ .

In our prior work, we performed a *backfilling* strategy where at a non-base view  $v$ , expanded infinitesimal mesh elements are filled with background depth. This is performed by replacing foreground node values of the expanded triangles at  $v$  with extrapolated background depth values. These background filled mesh elements are then warped from view  $v$  to the base view  $b$  to form a new layer of mesh elements. The initial base-mesh model  $\mathcal{M}_b$  is augmented with multiple layers, with each layer associated with a unique layer-ID  $l$ .

$$\hat{\mathcal{M}}_b = \bigcup_{l=0,1,2,\dots,L} \mathcal{M}_b^l \quad (2)$$

The augmented base-mesh  $\hat{\mathcal{M}}_b$  is a union of mesh elements from all layers, as shown by Equation (2). Mesh elements associated with

layer-ID  $l = 0$  belong to the initial base-mesh derived from  $Z_b$ ; that is  $\mathcal{M}_b^0 = \mathcal{M}_b$ . Subsequent layer-IDs  $l$  correspond to new elements derived from backfilling when warping to a particular view in the view array. When the augmented model  $\hat{\mathcal{M}}_b$  is warped to a given view, many mesh elements may map to any given location within that view. We resolve such multiple mappings by assigning absolute priority to mesh elements with lower layer-ID  $l$ .

Backfilling is a self-inferencing scheme, relying on only the base-view depth information  $Z_b$ . In this work we pursue a layered representation which is able to incorporate depth information from other communicated depth maps. The process is illustrated in Fig. 2 (d) and (e). If in addition to the base-view  $b$ , a depth map is also communicated at another view  $v$  then a new independent mesh  $\mathcal{M}_v$  is created at  $v$  in accordance with  $Z_v$ . The original base-mesh  $\mathcal{M}_b$  is warped from  $b$  to  $v$  with expanded infinitesimal mesh elements identifying dis-occluded regions  $\Omega_{b \rightarrow v}$  in  $v$ . Mesh elements  $\mathbf{m}_{v,j}$  of the new mesh  $\mathcal{M}_v$  which intersect with  $\Omega_{b \rightarrow v}$  are identified and denoted as  $\{\mathbf{m}_{v,j}\}^l$ , such that

$$\{\mathbf{m}_{v,j}\}^l = \{\mathbf{m}_{v,j} \cap \Omega_{b \rightarrow v} \neq \emptyset\}$$

This set of elements  $\{\mathbf{m}_{v,j}\}^l$  are then warped to the base view to create a new layer for the augmented representation with layer-ID  $l > 0$  as shown below.

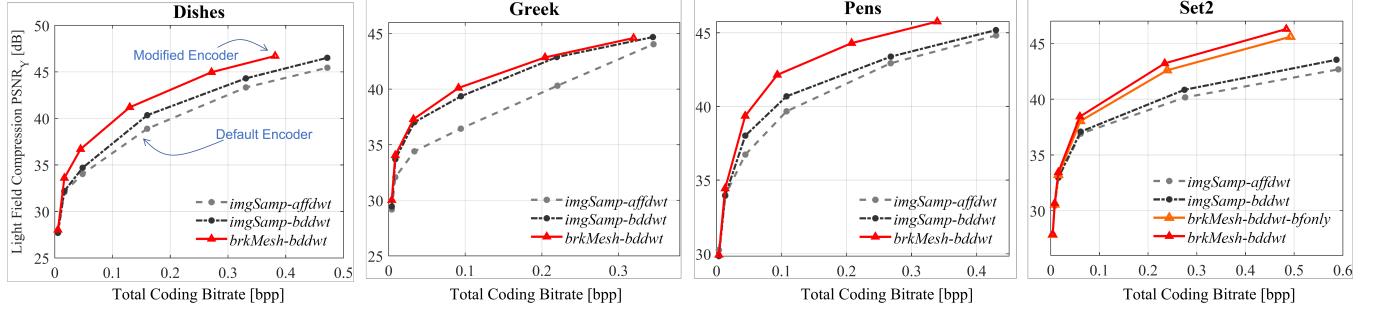
$$\mathcal{M}_b^l = \{\mathbf{m}_{v \rightarrow b,j}\}^l$$

As before, the augmented mesh is created in accordance with equation (2), with the original mesh at  $b$  forming the set for layer-ID  $l = 0$  ( $\mathcal{M}_b^0 = \mathcal{M}_b$ ). If only the base-view depth map is available, or if the communicated depth information is insufficient to fill dis-occluded regions at certain views, then our scheme defaults back to employing backfilling. In Fig. 2 (e), the grey shaded grid represents mesh elements  $\{\mathbf{m}_{v,j}\}^l$  anchored at the non-base view  $v$  that intersects with  $\Omega_{b \rightarrow v}$ . These triangular cells are warped to the base-view  $b$  as shown in Fig. 2 (d) to create a new layer in addition to the initial mesh in Fig. 2 (b). In this layered representation, dis-occluded regions at layer  $l$  are potentially filled with triangles emerging at subsequent layers with layer-ID  $l + 1$  or greater.

## 5. EXPERIMENTAL RESULTS

We modify the JPEG Pleno Light Field Verification Model 2.1 to incorporate (i) BD-DWT coding of depth maps as defined by JPEG 2000 Part 17, (ii) mesh-based backward warping and view prediction and (iii) base-mesh augmentation using backfilling and when possible augmentation using multiple depth maps. R-D results comparing the cumulative impact of these modifications to the default operation of the state-of-the-art JPEG Pleno light field encoder is presented in Fig. 3. We follow a hierarchical coding structure, as typically employed by the JPEG Pleno coding standard when operating in 4DPM. At the coarsest level a small number of texture views and depth maps are intra-coded. Views at subsequent levels are coded with reference to previously decoded views from coarser levels.

We report on target bitrates (bpp) inclusive of the operating ranges specified by the Common Test Conditions (CTC) defined by JPEG Pleno [11]. We present results for synthetic HCI light field datasets Dishes, Greek and Pens [12], along with the Set2 dataset [13] which is captured by a high density camera array. The HCI scenes are comprised of a  $9 \times 9$  array of views with a resolution of  $512 \times 512$ . These datasets are coded, at the coarsest level, with one ground-truth depth map and one intra-coded texture view at the center viewpoint [4, 4]. Set2 is composed of views of resolution



**Fig. 3.** Rate-distortion results. Average  $PSNR_Y$  for a range of rates expressed in bits per pixel (bpp)

$1920 \times 1080$  with the original size of the array being  $99 \times 21$  ( $cols \times rows$ ). We subsample this large array by 6 in the horizontal direction and 2 in the vertical direction and consider only the central  $9 \times 7$  sub-sampled array for coding. We note that there are large disparities between views in this dataset and therefore the quality of view warping and prediction is key to performance. For Set 2, at the coarsest level, we code 4 depth maps along with 5 intra-coded texture views - one at each corner along with the centre view location. All datasets are coded with 5 levels of hierarchy.

In Fig 3, (red) curves labelled *brkMesh-bddwt* correspond to our proposed augmented mesh model for backward warping of texture along with BD-DWT coding of depth maps. This case represents incorporating all our proposed modifications to the standard JPEG Pleno encoder. Curves (gray-dashed) labelled *imgSamp-affdwt* refer to the default operation of using sample-based depth to perform forward warping with splatting and employing the conventional 5/3 DWT for coding depth maps. We note that the 5/3 DWT is also well suited to represent smooth affine flows and hence we have deployed the name *affdwt* in our labelling convention. We highlight that *imgSamp-affdwt* cases correspond to the default operation of the JPEG Pleno encoder. R-D performance improvements achieved by our proposed modifications are clearly evident for all 4 datasets when comparing curves *brkMesh-bddwt* and *imgSamp-affdwt*.

Curves (black) labelled *imgSamp-bddwt* refer to cases where the default sample-based forward warping and splatting is employed but the coding of depth maps is changed to use the BD-DWT. The modification to the JPEG Pleno encoder is therefore limited to switching from the 5/3 DWT to the BD-DWT. Performance gains in relation to default operation (*imgSamp-affdwt*) is solely due to improved compression of depth maps and this has been previously reported. By comparing curves *brkMesh-bddwt* with *imgSamp-bddwt*, we can gauge the performance enhancements achieved due to our proposed mesh representation and backward warping of texture.

For Set 2, more than 1 depth is communicated and therefore our new mesh augmentation strategy that can incorporate depth information from multiple depth maps is employed. The (orange) curve labelled *brkMesh-bddwt-bfonly* shows performance for Set 2 when mesh augmentation is limited to backfilling while *brkMesh-bddwt* includes our new mesh augmentation strategy that supplements the base-mesh with information from all 4 depth maps. While the backfilling strategy works well, our new approach provides further gain.

To provide further insight to our proposed modifications, we consider coding only the coarsest level which includes intra-coded texture views and depth maps and then synthesising the remaining views at the finer levels. This coding scenario is chosen to highlight view synthesis capability using a common set of high quality decoded reference texture views and depth maps. In Table 1  $SSIM_Y$

**Table 1.** Average  $SSIM_Y$  of synthesized light field views. Shown in brackets are differences to the benchmark *imgSamp-affdwt*

Dataset	Bitrate (bpp)	<i>imgSamp-affdwt</i>	<i>brkMesh-bddwt</i>
Dishes	0.069	0.9528	0.9695(+0.0167)
Greek	0.052	0.9419	0.9593(+0.0174)
Pens	0.040	0.9314	0.9609(+0.0295)
Set2	0.166	0.9681	0.9771(+0.0090)



**Fig. 4.** Crop of target view [4, 8] synthesized from decoded coarsest level of Greek scene: (left) *imgSamp-affdwt*, (right) *brkMesh-bddwt*

results are shown, computed by averaging the  $SSIM_Y$  values of individual views, as recommended by the JPEG Pleno CTC [11]. Higher  $SSIM_Y$  values for *brkMesh-bddwt* suggest that scene object boundary structure is better synthesized by our proposed modifications. A crop of a synthesized view of the Greek dataset is shown in Fig. 4; cleaner and sharper object boundaries are recreated by our proposed set of modifications.

## 6. CONCLUSION

We modify the JPEG Pleno light field encoder to incorporate breakpoint dependent DWT (BD-DWT) coding of depth maps as defined by the recent JPEG 2000 Part 17 extension. The components of the BD-DWT are decoded directly onto a triangular mesh allowing for mesh-based backward view warping and prediction. We further introduce a strategy to augment the mesh representation to include depth information from multiple communicated depth maps. Results show that the modifications which we apply to the JPEG Pleno light field encoder provide significant gains in rate-distortion performance compared to the default operation of the standardized encoder.

## 7. REFERENCES

- [1] “Information technology - jpeg pleno plenoptic image coding system - part 2: Light field coding,” 2019, ISO/IEC JTC 1/SC29/WG1N84065.
- [2] P. Schelkens, P. Astola, E. A. B. da Silva, C. Pagliari, C. Perra, I. Tabus, and O. Watanabe, “JPEG Pleno light field coding technologies,” in *Applications of Digital Image Processing XLII*. International Society for Optics and Photonics, 2019, vol. 11137, pp. 391 – 401, SPIE.
- [3] C. Perra, P. Astola, E. A. B. da Silva, H. Khanmohammad, C. Pagliari, P. Schelkens, and I. Tabus, “Performance analysis of jpeg pleno light field coding,” in *2019 SPIE Optics and Photonics, Applications of Digital Image Processing XLII*, 2019.
- [4] “Information technology — jpeg 2000 image coding system — part 17: Extensions for coding of discontinuous media,” 2023, ISO/IEC JTC 1/SC29/FDIS 15444-17.
- [5] Y. Li, R. Mathew, and D. Taubman, “Scalable mesh representation for depth from breakpoint-adaptive wavelet coding,” in *2020 IEEE International Workshop on Multimedia Signal Processing (MMSP)*, 2020.
- [6] R. Mathew and D. Taubman, “Jpeg pleno light field encoder with breakpoint dependent affine wavelet transform for disparity maps,” in *2022 IEEE International Conference on Image Processing (ICIP)*, 2022.
- [7] M. Bertram, M. A. Duchaineau, B. Hamann, and K. I. Joy, “Generalized b-spline subdivision-surface wavelets for geometry compression,” *IEEE transactions on visualization and computer graphics*, vol. 10, no. 3, pp. 326—338, 2004.
- [8] L. Linsen, B. Hamann, K. I. Joy, V. Pascucci, and M. A. Duchaineau, ““wavelet-based multiresolution with pn 2 subdivision,” *Computing*, vol. 72, no. 1-2, pp. 129—142, 2004.
- [9] Yue Li, Reji Mathew, Dominic Ruefenacht, Aous Naman, and David Taubman, “Consistent disparity synthesis for inter-view prediction in lightfield compression,” in *2019 Picture Coding Symposium (PCS)*, 2019, pp. 1–5.
- [10] Dominic Rüfenacht, Aous Thabit Naman, Reji Mathew, and David Taubman, “Base-anchored model for highly scalable and accessible compression of multiview imagery,” *IEEE Transactions on Image Processing*, 2019.
- [11] “Jpeg pleno light field coding common test conditions,” 2018, ISO/IEC JTC1/SC29/WG1, WG1N80027.
- [12] K. Honauer, O. Johannsen, D. Kondermann, and B. Goldluecke, “A dataset and evaluation methodology for depth estimation on 4d light fields,” in *2016 Asian Conference on Computer Vision*. 2016, pp. 19–34, Springer.
- [13] Fraunhofer IIS, “Static planar light field test dataset,” in <https://www.iis.fraunhofer.de/en/fj/amm/dl/lightfielddataset.html> [online].