

# Lenslet Light Field Imaging Scalable Coding

João Garrote, Catarina Brites, João Ascenso, Fernando Pereira

Instituto Superior Técnico, Universidade de Lisboa - Instituto de Telecomunicações  
Av. Rovisco Pais, 1049-001  
Lisboa, Portugal

**Abstract**— Light fields have emerged as one of the most promising 3D representation formats, enabling a richer and more immersive representation of a visual scene. The lenslet light field acquisition approach consists in placing an array of micro-lenses between the camera main lens and the photosensor to allow capturing both the intensity and the direction of the light rays. This type of representation format offers new interaction possibilities with the visual content, notably *a posteriori* refocusing and visualization of different perspectives of the visual scene. However, this representation model is associated to very large amounts of data, thus requiring efficient coding solutions in order applications involving storage and transmission may be deployed. This paper proposes a novel lenslet light field imaging scalable coding solution adopting a wavelet-based approach, able to offer view, quality and spatial scalabilities, to meet the characteristics of multiple types of displays, transmission channels and user needs. The performance results show that the proposed coding solution performs better than alternative scalable coding solutions, notably JPEG 2000.

**Keywords**— lenslet light field; sub-aperture image; disparity estimation and compensation; scalability; JPEG 2000.

## I. INTRODUCTION

Replicating the visual world in the most faithful and immersive way has always been the target of visual representation technology, notably cameras/sensors, codecs and displays. For decades, a digital picture has always been a 2D array of pixels where each pixel accumulates the light incident from all directions on a specific sensor position. However, this is a limited representation model as highlighted by the so-called *plenoptic function* [1][2], which models the light in the 3D space with a 7D function of five spatial coordinates, notably three position coordinates and two directional coordinates, thus expressing the fact that light rays have a direction.

In recent years, major developments in visual acquisition technology allowed to develop the so-called *lenslet light field cameras*, such as Lytro and Raytrix, which are able to measure the intensity of light incident on a specific position coming from multiple spatial directions using an array of micro-lenses between the camera main lens and the photosensor, creating the so-called *micro-images*. This richer representation model allows the visual data to be *a posteriori* manipulated by the users, notably controlling the focus, the scene perspective or even creating stereoscopic images [2]. However, this richer representation is associated to a larger amount of data, which critically needs efficient coding in order practical applications may be deployed. Recognizing this need, both JPEG and MPEG are studying this problem, with JPEG taking the lead by launching in January 2017 a Call for Proposals on Light Field Coding [3], which considered both lenslet and high dense camera arrays (HDCA) light fields.

Because a lenslet light field includes a lot of data that may

be progressively consumed in several dimensions, e.g. more views, more resolution, more quality, scalable coding is a rather natural requirement, which has not been much addressed in the literature. In this context, this paper proposes a novel lenslet light field imaging scalable coding solution adopting a wavelet-based approach, able to offer view, quality and spatial scalabilities, to meet the characteristics of multiple types of displays, transmission channels and user needs. This novel coding solution offers significant gains regarding the most relevant scalable solution available. The remainder of this paper is organized as follows: Section II presents a brief review of the background work, while Section III presents the architecture and walkthrough of the proposed coding solution. Section IV offers a detailed description of the most relevant tool in the codec, notably the disparity compensated inter-view discrete wavelet transform. Finally, Section V presents the performance results and their analysis, while Section VI concludes with final remarks and suggestions for future work.

## II. BACKGROUND WORK

While the lenslet light field coding domain is rather recent, there are already several solutions proposed in the literature. A major distinction between the available solutions regards the adopted data structure, notably if the lenslet image (this means the set of micro-images) is directly coded or if it is organized in the set of so-called *sub-aperture (SA) images/views*, each corresponding to a specific viewpoint. Here, the available lenslet coding solutions will be grouped into four categories, depending on their relation with available coding standards:

**1. Standard compliant coding solutions:** These solutions directly code the lenslet image with a standard coding solution. While these solutions cannot exploit all available redundancy in the lenslet light field, they benefit from the standard ecosystem as standard bitstreams and decoders are used. These solutions include both still image coding standards, namely JPEG and JPEG 2000, and video coding standards used in the Intra coding mode, namely H.264/AVC Intra and HEVC Intra [4].

**2. Standard compliant coding solutions applied after some data re-organization:** These solutions involve applying standard coding solutions after some data re-organization with the target to better exploit the redundancy in the data. The most common data re-organization involves taking the set of SA images as a sequence of video frames, creating a so-called *pseudo-video* [5][6]; other solutions code the set of SA images using some appropriate 2D spatial prediction structure [7].

**3. Extended standard coding solutions:** These solutions involve extending available coding standards with additional tools and coding modes to improve the compression performance for lenslet light field images; for example, [8][9] extend the HEVC standard with additional prediction tools to exploit the redundancy of the micro-images within the lenslet

image. Other solutions may code some SA images in a standard way, e.g. using JPEG 2000 or HEVC Intra, and the remaining SA images using depth/disparity based estimation.

**4. Non-standard based coding solutions:** These solutions adopt rather different approaches. While some solutions are based on the exploitation of depth information [10], others have as cornerstone different transforms, notably the discrete wavelet transform (DWT) [11][12][13], and the Karhunen-Loeve transform (KLT) [14]. These coding solutions may use these transforms alone or even combined [15][16].

Although there are many coding solutions, most of them do not address scalability requirements as in this paper, this means offering view, quality and spatial scalabilities.

### III. DISPARITY COMPENSATED LENSLET LIGHT FIELD SCALABLE CODING: ARCHITECTURE AND WALKTHROUGH

The architecture of the proposed Disparity Compensated Lenslet Light Field Scalable (DCLLFS) coding solution is shown in Fig. 1. As this architecture targets offering view, quality and spatial scalabilities to meet the characteristics of multiple types of displays, transmission channels and user needs, it is based on the DWT applied at both view intra-coding and inter-coding levels. Due to its scalability features, some of the modules are based on the JPEG 2000 standard. The Disparity Compensated Inter-View DWT module incorporates the main novelty of this coding solution, and aims exploiting the redundancy between the SA images while offering view scalability.

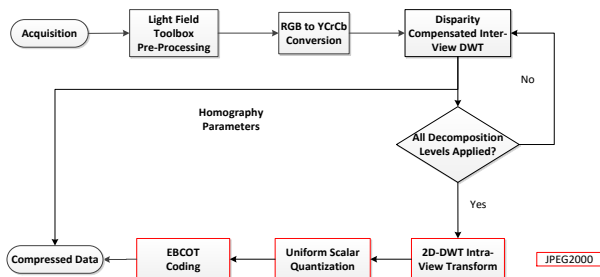


Fig. 1. Architecture of the proposed DCLLFS codec.

A brief description of each module in the DCLLFS coding solution is presented next:

**1. Light Field Toolbox Pre-Processing:** The objective of this module is to convert the raw light field image, obtained directly from the sensor, into a more suitable representation format. First, the *lenslet image* is created from the raw sensor data by applying demosaicing, devignetting, clipping, and some color processing. Then, the lenslet image, formed by thousands of micro-images, is converted into an array of SA images, each representing a different perspective view. This module uses the available Light Field Toolbox v0.4 software [17]. While the original light field image is composed by 225 SA images, it was decided to discard both the first and last row and column of the SA images array, resulting into 169 SA images, to avoid using SA images without enough quality, notably some black images in the corners due to the vignetting effect. This strategy has been also adopted by the JPEG PLENO Call for Proposals [3].

**2. RGB to YCrCb Conversion:** The objective of this module is to improve the compression efficiency by converting the RGB data into YCrCb data.

**3. Disparity Compensated Inter-View DWT:** An inter-view wavelet transform is chosen to decorrelate the various SA images and compact their energy into a small number of coefficients. This transform was designed with a lifting structure to allow including disparity estimation and compensation techniques in the prediction and updating steps [18]. The overall objective of the designed transform is to obtain low-frequency and high-frequency bands in such a way that the low-frequency band corresponds to a smoothed representation of a view and the high-frequency band to the high frequency information necessary to obtain the other view. The wavelet transform with disparity compensated lifting is applied to an array of SA images with size  $N$  and its frequency decomposition capabilities lead to  $N/2$  low-frequency bands and  $N/2$  high-frequency bands. To further exploit the correlation between the low-frequency bands, the wavelet transform can be used again in a second decomposition level, using now as input the low-frequency bands. By applying one level of transform decomposition, two scalability layers are created, the first associated to the low-frequency bands and the second associated to the high-frequency bands; for each decomposition level added, one more scalability layer becomes available. A simplified architecture of the forward transform is shown in Fig. 2. As the disparity compensated inter-view DWT is the most original tool in the proposed coding solution, it will be detailed in Section IV.

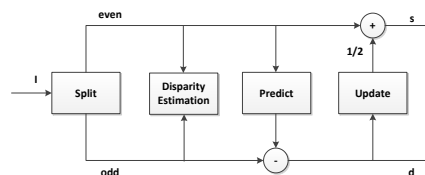


Fig. 2. Architecture of the Haar disparity compensated wavelet transform for one decomposition level.

**4. Intra-View 2D-DWT:** The objective of this module is to exploit the spatial redundancy within each SA image or high-frequency/low-frequency band. The 2D-DWT transform with six decomposition levels, as available in the OPENJPEG software [19], has been adopted for application to all the frequency bands resulting from the inter-view transform. This process consists basically in applying a 1D-DWT along the X axis (spatially horizontally) and, after, again along the Y-axis (spatially vertically) to each image/band. The result of a 1-level 2D wavelet decomposition is four filtered and subsampled images, also known as *bands*. The 2D-DWT enables resolution scalability as the SA images can be decoded at full resolution or only at a fraction of the full resolution [19].

**5. Quantization:** The objective of this module is to reduce the accuracy of the 2D-DWT coefficients to obtain higher compression. The quantization is performed using a uniform scalar quantizer with dead-zone, which is one of the available JPEG 2000 quantization methods [20]. This quantization method allows also to progressively transmit the coefficients (quality or SNR scalability) by progressively sending the most significant bitplanes (MSB) and then advancing to the least significant bitplanes (LSB). All the bands obtained after applying the Intra-view 2D-DWT are quantized using this method [20].

**6. EBCOT Coding:** The objective of this module is to exploit the statistical redundancy of the 2D-DWT coefficients by performing entropy coding. First, each band is divided into small

rectangular blocks, referred to as *codeblocks*, and each codeblock is independently encoded with Embedded Block Coding with Optimized Truncation (EBCOT). All codeblocks from low to high frequency are scanned together from top to bottom and left to right with each band independently coded from the other bands. EBCOT performs multiple-pass coding of the codeblock bitplanes obtained in the previous step. Three passes are used, notably significance propagation, magnitude refinement and cleanup; more details about each pass are available in [20]. For JPEG 2000 to be compression efficient, a context-based adaptive binary arithmetic coding method is used, which exploits the correlation among bitplanes.

#### IV. DISPARITY COMPENSATED INTER-VIEW DISCRETE WAVELET TRANSFORM

As the Haar disparity compensated DWT (illustrated in Fig. 2 for one decomposition level), which performs disparity compensation using a perspective geometric transform, is the key novelty of this paper, it is displayed with more detail in Fig. 3.

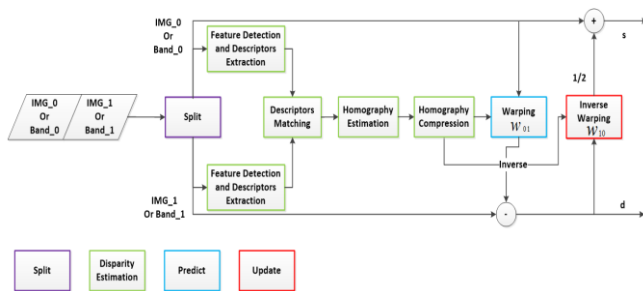


Fig. 3. Disparity compensated inter-view DWT applied to two SA images or low-frequency/high-frequency bands, highlighting the relationship with the modules in Fig. 2.

The Disparity Compensated Inter-View DWT module performs the following steps (see Fig. 3):

**1. Split:** The input set of SA images (or bands) is divided into two different, complementary sets, where the even SA images and the odd SA images (or bands) are grouped in different sets.

**2. Feature Detection and Descriptor Extraction:** The objective of this module is to detect distinctive features in the images associated to keypoints or blobs, and extract descriptors for those positions; the descriptors represent the features in some space that is invariant to common deformations such as translation, scaling, rotation, perspective changes, and partially invariant to illumination changes. In this case, the popular SIFT (Scale Invariant Feature Transform) descriptor [21] has been adopted.

**3. Descriptors Matching:** The objective of this module is to match a set of (SIFT) descriptors extracted from one SA image with the descriptors extracted from another SA image, obtaining a set of one-to-one correspondences. In this case, a simple approach has been followed, which consists in taking each descriptor in the first set and matching it with all the descriptors in the second set, using some distance metric, e.g. the Euclidean distance. Then, a ratio test is applied. This test compares the ratio of distances between the two top matches for a given keypoint. If this ratio is above the threshold of 0.7, the match is rejected. The objective of this test is to increase the reliability of the matching procedure, thus avoiding some incorrect matches between keypoints [21].

**4. Homography Estimation:** The objective of this module is to estimate the geometric transformation between one SA image (or a low-frequency band) and the other by establishing a relationship between corresponding positions in the two SA images (or low-frequency bands); these correspondences were obtained in the previous step. Several formulations for this transformation are possible, such as affine, perspective, bilinear, and polynomial transforms. Considering the lenslet light field imaging characteristics, the most adequate transform for modelling the data seems to be the perspective transform, also known as homography, since it is able to model complex geometric relationships between different perspectives of the objects in the visual scene. The perspective transform (or homography) is defined by a 9-parameter matrix which is able to describe the object displacements in the visual scene when the perspective changes. When applied to SA images, this transform should describe well the disparity between the SA images, which is mainly determined by the characteristics of the micro-lenses array. To estimate the homography parameters, RANSAC, an iterative method to estimate the parameters of a mathematical model that is robust even when there are some wrong matches (outliers), has been adopted. To avoid that outliers reduce the accuracy of the estimated transformation matrix, RANSAC attempts to identify inliers, i.e. the data fitting well a set of model parameters (typically estimated with a standard regression method) and, therefore, not considering outliers in the estimation.

**5. Homography Parameters Compression:** The perspective transform parameters are initially represented with 8 bytes (64 bits), assuming double precision floating-point format. Since this precision may require a significant rate, as these parameters have to be transmitted to the decoder, it is important to adopt a quantization technique to compress this type of data. Note that, each time the inter-view wavelet transform is applied, a different homography matrix is used and, therefore, since there are many pairs of SA images (bands) for which a transform is applied, the number of parameters to be coded may be rather high. The compression solution must be applied to each homography matrix obtained for a given decomposition level.

**6. Warping or Disparity Compensation ( $W_{01}$ ):** The objective of this module is to warp an input even SA image in such a way that it becomes similar to the odd SA image, which in this case corresponds to a slightly different perspective, in practice performing *disparity compensation*. This warping process is performed by using the decoded homography matrix. The SA image prediction is computed by multiplying each sample position in the input image (in homogeneous coordinates) by the homography matrix. By computing the difference between an odd view and the warped even view, the high-frequency band is obtained.

**7. Inverse Warping ( $W_{10}$ ):** The objective of this module is to inversely warp the high-frequency band resulting from the previous step such that it becomes similar to the even SA image, thus allowing to obtain a smoothed representation of the even SA image. To perform this process, the inverse transformation (homography) is needed. As the scene disparity involved in this kind of data mostly corresponds to translations, because it is mostly due to the spatial separation between the micro-lenses and only a little due to optical defects in the micro-lenses, the homography matrix from a reference view into another view can be inverted, and thus an inverse homography matrix can be

obtained. This is also a requirement from the disparity compensated wavelet transform, which can only be applied when the warpings  $w_{01}$  and  $w_{10}$  are symmetric as otherwise the process may end up adding a residual to the odd view that is not aligned (prediction step), thus creating ghost artifacts. By computing a weighted sum between the even view and the warped high-frequency band, a low-frequency band SA image is obtained.

The lifting scheme for the inverse disparity compensated wavelet transform to be performed at the decoder follows the scheme presented in Fig. 4. As the homography parameters are transmitted to the decoder, the inverse transform only needs to perform the *predict* and *update* steps in the reverse order flipping the signal in arithmetic operations, thus resulting in the original signal.

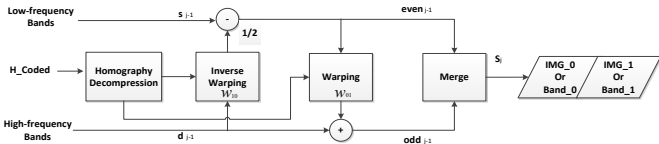


Fig. 4. Inverse disparity compensated inter-view DWT architecture.

V. PERFORMANCE ASSESSMENT

The objective of this section is to assess the rate-distortion (RD) performance of the proposed DCLLFS coding solution.

A. Test Material, Benchmarks and Metrics

To evaluate the RD performance, five lenslet light field images have been selected from the MMSPG EPFL Light Field Dataset [22]; this dataset has also been selected as the test set for the Light Field Compression Grand Challenge organized at ICME 2016 [23] and for the JPEG Pleno Call for Proposals on Light Field Coding [3]. The set of selected images is: *Bikes*, *Danger\_de\_Mort*, *Stone Pillars\_Outside*, *Friends\_1*, and *Fountain & Vincent\_2*. The images were chosen by their content, aiming to have a diversified dataset, with both high and low frequency content and objects at different depths. To simplify the text, in the following, the names will just be *Bikes*, *Danger*, *Stone*, *Friends*, and *Fountain*. Each light field is structured as a matrix of 225 SA images; however, for compression purposes, only 169 SA images, each with a spatial resolution of 625×434 pixels, will be considered. Because a scalable codec is proposed, the main benchmark will be the JPEG 2000 standard. However, due to its huge popularity, also the JPEG standard will be used as benchmark. When coding with JPEG 2000, the SA images are coded as a single “super image”, which is coded all at once. The RD points are defined by the rate spent in the “super image” and the PSNR is computed as the average PSNR of all SA images extracted from the decoded “super image”. Because this is usually enough, the performance assessment will be made only for the luminance (Y) component of the SA images. For the proposed DCLLFS coding solution, the rate is measured in bit-per-pixel (bpp) and includes both the rate for each compressed band and the homography parameters. To obtain the bpp rate, the total number of bits is divided by the number of coded SA images (169) multiplied by their resolution (625×434).

B. Performance Results and Analysis

The proposed solution and the benchmarks will be compared using the well-known Bjøntegaard Delta metrics [24]. The

DCLLFS solution will be labeled as *DCLLFS\_HX\_VY* where *X* and *Y* are the number of wavelet decompositions applied in the horizontal and vertical directions, respectively, see Fig. 5.

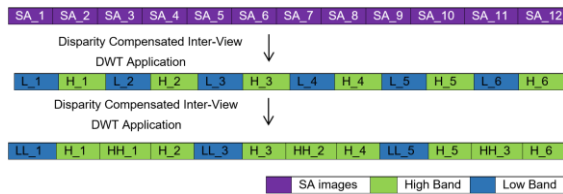


Fig. 5. Applying one (central row) and two (bottom row) decomposition levels along the horizontal direction.

The first set of results concerns the application of the proposed coding solution to horizontally neighboring SA images. Table 1 provides the BD-Rate and BD-PSNR for horizontal wavelet decompositions with two (DCLLFS\_H2) and three levels (DCLLFS\_H3) in comparison with JPEG 2000. The DCLLFS coding solution always performs much better than JPEG 2000 as this codec does not exploit the redundancy between the views. In terms of the BD-PSNR performance, it is possible to conclude that *Friends* is the light field exhibiting the highest gains, while *Danger* is the one with the lowest PSNR gains. This is understandable as *Friends* exhibits a more homogenous background while *Danger* includes letters and many more details, thus reducing the DCLLFS RD performance as there is less redundancy across the views. Regarding the number of decomposition levels impact, it is possible to conclude that increasing the number of decomposition levels of the proposed inter-view transform allows increasing the RD performance although with a reducing gain for any additional level. This implies that, at some stage, it is not worthwhile to keep increasing the number of decomposition levels. The proposed DCLLFS solution was also applied over the vertical direction, yielding similar, although slightly reduced, RD performance gains; due to length constraints, the results are not included here. Because DCLLFS\_H3 does not bring major RD performance improvements regarding DCLLFS\_H2 and the complexity increases, DCLLFS\_H2 is taken here as the best 1-direction decompositions solution.

TABLE 1 - BJØNTEGAARD DELTA RESULTS REGARDING JPEG 2000 FOR: (LEFT) DCLLFS\_H2; (RIGHT) DCLLFS\_H3.

Light Field	BD-PSNR[dB]	BD-Rate [%]	Light Field	BD-PSNR[dB]	BD-Rate [%]
Friends	4.24	-61.44	Friends	4.49	-64.24
Fountain	3.02	-46.75	Fountain	3.10	-48.75
Stone	3.48	-52.95	Stone	3.70	-56.76
Danger	3.00	-49.69	Danger	3.02	-50.75
Bikes	3.13	-47.56	Bikes	3.21	-49.36
<b>Average</b>	<b>3.38</b>	<b>-51.68</b>	<b>Average</b>	<b>3.51</b>	<b>-53.97</b>

After, the proposed DCLLFS solution was applied with decompositions in both directions. The wavelet decomposition is first applied to horizontally neighboring SA images and after to vertically neighboring SA images, as this was the order achieving better RD performance. When two decomposition levels are applied, DCLLFS\_H1\_V1 performs better than DCLFC\_H2, with an average BD-Rate saving of 5.81%. Table 2 presents results for the case where three decomposition levels are applied using as reference DCLLFS\_H1\_V1. While both solutions exploit the correlation between neighboring SA images, horizontally and vertically, the gains are rather similar for all light fields and both configurations result in rather similar RD performance improvements.

TABLE 2 - BJØNTEGAARD DELTA RESULTS USING AS REFERENCE DCLLFS\_H1\_V1 FOR: (LEFT) DCLLFS\_H2\_V1; (RIGHT) DCLLFS\_H1\_V2.

Light Field	BD-PSNR[dB]	BD-Rate [%]	Light Field	BD-PSNR[dB]	BD-Rate [%]
Friends	0.83	-18.56	Friends	0.75	-16.32
Fountain	0.47	-10.57	Fountain	0.47	-10.19
Stone	0.77	-16.73	Stone	0.66	-14.19
Danger	1.98	-37.78	Danger	2.13	-39.46
Bikes	0.49	-10.66	Bikes	0.50	-10.62
<b>Average</b>	<b>0.91</b>	<b>-18.86</b>	<b>Average</b>	<b>0.90</b>	<b>-18.16</b>

Next, Table 3 shows performance results for four decomposition levels. Both the first and second solutions were implemented in the usual way, this means first processing the horizontal decompositions and after the vertical decompositions. However, for DCLLFS\_H2\_V2, the horizontal and vertical decompositions were applied alternately. While DCLLFS\_H3\_V1 shows a reduced RD performance, DCLLFS\_H1\_V3 shows almost no performance differences and DCLLFS\_H2\_V2 is the only configuration able to increase the DCLLFS\_H2\_V1 RD performance, showing that a balanced approach between the horizontal and vertical decompositions is the best solution. At this stage, DCLLFS\_H2\_V2 is taken as the best decomposition configuration as the additional RD performance gains of additional decompositions should not compensate the additional complexity.

TABLE 3 - BJØNTEGAARD DELTA RESULTS REGARDING DCLLFS\_H2\_V1 FOR: (TOP-LEFT) DCLLFS\_H3\_V1; (TOP-RIGHT) DCLLFS\_H1\_V3; (BOTTOM) DCLLFS\_H2\_V2.

Light Field	BD-PSNR[dB]	BD-Rate [%]	Light Field	BD-PSNR[dB]	BD-Rate [%]
Friends	-0.38	10.34	Friends	-0.02	1.34
Fountain	-0.36	8.75	Fountain	0.01	-0.53
Stone	-0.37	11.06	Stone	0.02	-1.06
Danger	-0.52	14.01	Danger	0.10	-2.90
Bikes	-0.44	10.62	Bikes	-0.02	0.11
<b>Average</b>	<b>-0.41</b>	<b>10.95</b>	<b>Average</b>	<b>-0.01</b>	<b>0.18</b>

Light Field	BD-PSNR[dB]	BD-Rate [%]
Friends	0.32	-7.48
Fountain	0.22	-5.36
Stone	0.25	-6.73
Danger	0.18	-5.04
Bikes	0.23	-5.34
<b>Average</b>	<b>0.24</b>	<b>-5.99</b>

Finally, Table 4 shows Bjøntegaard delta results for the DCLLFS\_H2\_V2 solution in comparison with JPEG 2000 and JPEG. The proposed DCLFC\_H2\_V2 solution is able to outperform both the JPEG and JPEG 2000 standards, which is expectable as none of these coding solutions provides decorrelation capabilities between neighboring SA images. The overall BD-Rate savings go up to 62.85% and 78.80% for JPEG 2000 and JPEG, respectively.

TABLE 4 - BJØNTEGAARD DELTA RESULTS USING DCLLFS\_H2\_V2 AS REFERENCE FOR: (LEFT) JPEG 2000; (RIGHT) JPEG.

Light Field	BD-PSNR[dB]	BD-Rate[%]	Light Field	BD-PSNR[dB]	BD-Rate[%]
Friends	5.61	-72.67	Friends	9.30	-86.66
Fountain	3.91	-57.51	Fountain	7.62	-79.81
Stone	4.67	-65.65	Stone	5.42	-70.56
Danger	3.92	-60.97	Danger	6.64	-78.18
Bikes	3.84	-57.44	Bikes	7.43	-78.79
<b>Average</b>	<b>4.39</b>	<b>-62.85</b>	<b>Average</b>	<b>7.28</b>	<b>-78.80</b>

## VI. FINAL REMARKS

The proposed DCLLFS coding solution offers view, quality and spatial scalabilities to meet the characteristics of multiple types of displays, transmission channels and user needs, which

does not happen for most lenslet light field coding solutions in the literature. Future work should consider using different wavelet transforms and different geometric transformations for different regions of the SA images to obtain better disparity compensations.

## REFERENCES

- [1] E. H. Adelson and J. R. Bergen, "The plenoptic function and the elements of early vision", The MIT Press, Cambridge, Mass., 1991.
- [2] F. Pereira and E. A. da Silva, "Efficient plenoptic imaging representation: why do we need it?", IEEE ICME, Seattle, WA, USA, Jul. 2016
- [3] JPEG Convenor, "JPEG Pleno call for proposals on light field coding", Doc. ISO/IEC JTC1/SC29/WG1/N/74014, JPEG Geneva Meeting, Jan. 2017.
- [4] G. Alves, F. Pereira and E. A. da Silva, "Light field imaging coding: performance assessment methodology and standards benchmarking", IEEE ICME Workshops, Seattle, WA, USA, Jul. 2016.
- [5] A. Vieira, H. Duarte, C. Perra, L. Tavora and P. Assunção, "Data formats for high efficiency coding of Lytro-Illum light fields", Int. Conf. on Image Processing Theory, Tools and Applications, Orléans, France, Nov. 2015.
- [6] D. Liu *et al.*, "Pseudo-sequence-based light field image compression", IEEE ICME Workshops, Seattle, WA, USA, Jul. 2016.
- [7] L. Li, Z. Li, B. Li, D. Liu and H. Li, "Pseudo sequence based 2-D hierarchical coding structure for light-field image compression", Data Compression Conference, Snowbird, UT, USA, Apr. 2017.
- [8] C. Conti, P. Nunes and L. Soares, "HEVC-based light field image coding with bi-predicted self-similarity compensation", IEEE ICME Workshops, Seattle, WA, USA, Jul. 2016.
- [9] R. Monteiro *et al.*, "Light field HEVC-based image coding using locally linear embedding and self-similarity compensated prediction", IEEE ICME Workshops, Seattle, WA, USA, Jul. 2016.
- [10] C. Choudhury and S. Chaudhuri, "Disparity based compression technique for focused plenoptic images", Indian Conf. on Computer Vision Graphics and Image Proc., Bangalore, KA, India, Dec. 2014.
- [11] H. Zayed, S. Kishk and H. Ahmed, "3D wavelets with SPIHT coding for integral imaging compression", Int. Journal of Computer Science and Network Security, vol. 12, no. 1, pp. 126-133, Jan. 2012.
- [12] A. Aggoun and M. Mazri, "Wavelet-based compression algorithm for still omnidirectional 3D integral images", Signal, Image and Video Processing, vol. 2, no. 2, pp. 141-153, Jun. 2008.
- [13] A. Aggoun, "Compression of 3D integral images using 3D wavelet transform", Journal of Display Technology, vol. 7, no. 11, pp. 586-592, Sep. 2011.
- [14] H. Kang, D. Shin and E. Kim, "Compression scheme of sub-images using Karhunen-Loeve transform in three-dimensional integral imaging", Optics Communications, vol. 281, no. 14, pp. 3640-3647, Jul. 2008.
- [15] E. Elharar, A. Stern, O. Hadar and B. Javidi, "A hybrid compression method for integral images using discrete wavelet transform and discrete cosine transform", Journal of Display Technology, vol. 3, no. 3, pp. 321-325, Aug. 2007.
- [16] S. Kishk, H. Ahmed and H. Helmy, "Integral images compression using discrete wavelets and PCA", Int. Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 4, no. 2, pp. 65-77, Jun. 2011.
- [17] D. G. Dansereau, "Light field toolbox v0.4", Feb. 2015. [Online]. <https://www.mathworks.com/matlabcentral/fileexchange/49683-light-field-toolbox-v0-4>. [Accessed 10 01 2017].
- [18] C. Chang, X. Zhu and P. Ramanathan, "Light field compression using disparity-compensated lifting and shape adaptation", IEEE Transactions on Image Processing, vol. 15, no. 4, pp. 793-806, Mar. 2006.
- [19] M. Darbois, "DocJ2KCodec", OPENJPEG, [Online]. Available: <https://github.com/uclouvain/openjpeg/wiki/DocJ2KCodec>. [Accessed 15 08 2017].
- [20] M. Marcellin *et al.*, "An overview of quantization in JPEG 2000", Signal Processing: Image Communication, vol. 17, no. 1, pp. 73-84, Jan. 2002.
- [21] D. Lowe, "Distinctive image features from scale-invariant key", Int. Journal on Computer Vision, vol. 60, no. 2, pp. 91-110, Jan. 2004.
- [22] M. S. P. Group, "Light-field image dataset", [Online]. Available: <http://mmspg.epfl.ch/EPFL-light-field-image-dataset>. [Acc. 12 06 2017].
- [23] M. Kerabek, T. Bruylants, T. Ebrahimi, F. Pereira and P. Schelkens, "ICME 2016 grand challenge: light-field image compression", IEEE ICME, Seattle, USA, Jul. 2016.
- [24] G. Bjøntegaard, "Calculation of average PSNR differences between RD-curves", VCEG-M33, Austin, Texas, Apr. 2011.