

Article

# Adaptive Content Frame Skipping for Wyner–Ziv-Based Light Field Image Compression

Huy PhiCong <sup>1,2,3,4</sup>, Stuart Perry <sup>2</sup> and Xiem HoangVan <sup>1,\*</sup>

<sup>1</sup> Faculty of Electronics and Telecommunications, VNU University of Engineering and Technology, Vietnam National University, Hanoi 100000, Vietnam; cong Huy.phy@student.uts.edu.au

<sup>2</sup> School of Electrical and Data Engineering, University of Technology Sydney, Ultimo 2007, Sydney, Australia; stuart.perry@uts.edu.au

<sup>3</sup> Posts and Telecommunications Institute of Technology, Hanoi 100000, Vietnam

<sup>4</sup> JTIRC, VNU University of Engineering and Technology, Hanoi 100000, Vietnam

\* Correspondence: xiemhoang@vnu.edu.vn

Received: 26 September 2020; Accepted: 25 October 2020; Published: 29 October 2020

**Abstract:** Light field (LF) imaging introduces attractive possibilities for digital imaging, such as digital focusing, post-capture changing of the focal plane or view point, and scene depth estimation, by capturing both spatial and angular information of incident light rays. However, LF image compression is still a great challenge, not only due to light field imagery requiring a large amount of storage space and a large transmission bandwidth, but also due to the complexity requirements of various applications. In this paper, we propose a novel LF adaptive content frame skipping compression solution by following a Wyner–Ziv (WZ) coding approach. In the proposed coding approach, the LF image is firstly converted into a four-dimensional LF (4D-LF) data format. To achieve good compression performance, we select an efficient scanning mechanism to generate a 4D-LF pseudo-sequence by analyzing the content of the LF image with different scanning methods. In addition, to further explore the high frame correlation of the 4D-LF pseudo-sequence, we introduce an adaptive frame skipping algorithm followed by decision tree techniques based on the LF characteristics, e.g., the depth of field and angular information. The experimental results show that the proposed WZ-LF coding solution achieves outstanding rate distortion (RD) performance while having less computational complexity. Notably, a bit rate saving of 53% is achieved compared to the standard high-efficiency video coding (HEVC) Intra codec.

**Keywords:** light field image coding; Wyner–Ziv theorem; high-efficiency video coding

## 1. Introduction

### 1.1. Context and Motivations

Light field (LF) rendering is known as an attractive form of image-based rendering (IBR) [1,2], which collects immense amounts of image data due to the intensity of light rays traveling in every angle at every point in 3D space being captured [3]. Thus, the LF image data include information such as the location or point  $(x, y, z)$ , the angle or direction  $(\theta, \phi)$ , the wavelength  $(\gamma)$ , and the time  $(t)$  for light rays captured in a scene. This process is defined as the Plenoptic function,  $P_{LF}$ , and explains the huge amount of data stored in each LF image, as an LF image can include 7D information  $(P_{LF}(x, y, z, \theta, \phi, \gamma, t))$  [3]. A raw LF image is composed of micro-images (MIs), and a set of sub-aperture images (SAIs) is obtained by rearranging the co-located pixels from each MI. Each SAI corresponds to a captured image from a scene from a particular point of view, which can vary slightly between two different SAIs [4]. In addition, information about the parallax and depth of an

image scene can be provided by comparing SAIs. In practice, a set of constraints is introduced to the Plenoptic function to reduce the complexity of LF information, which is reduced to an extensive 4D function, as below:

$$P_{LF} = L(u, v, x, y) \quad (1)$$

Here, the light intensity  $P_{LF}$  is indexed by the sub-aperture image (viewpoint)  $(x, y)$  and the position (angle) within the sub-aperture image  $(u, v)$ .

As an example of an LF imaging technology, LF cameras have become a promising tool for various research areas, e.g., richer photography using Lytro Illum [5], material analysis using Raytrix [6], medical imaging [7], and biometric recognition [8]. As a result of the enormous size of the photo-realistic LF images (typically 1 GB [9]), data compression is, therefore, a challenge in terms of storage, processing, and transmission. Recently, the Joint Photographic Experts Group (JPEG) committee created a process for standardization called JPEG Pleno [10], which includes LF, point cloud, and holography [11]. The proposal provides an LF representation and coding with optimized viewing and resolution for a huge amount of data; thus, an efficient coding solution with high compression performance is of the utmost importance.

In the literature, various techniques and methods for LF compression have been introduced, especially for LF lenslet coding and four-dimensional LF (4D-LF) coding. The LF lenslet format is a compact version of the LF data, which represents the LF data as a massive hexagonal array of lenslets (MIs) and requires additional camera metadata in order to render images of a scene. In [12–16], exploiting the LF lenslet compression, most of the conventional image and video coding methods were applied to exploit the existing spatial redundancy of MIs within a raw LF lenslet, such as JPEG, JPEG2000, or high-efficiency video coding (HEVC) intra coding. This idea is based on the concept of self-similarity compensated prediction [12]. A block-based matching algorithm is utilized to manage the most suitable predictor block for the current block, which is compared to the previously coded and reconstructed range of the current image. With two different candidate blocks, the predictor block can be generated. Additionally, [13,14] proposed adding new coding modes to the HEVC coding tools (i.e., locally linear embedding-based prediction) and adapting the intra prediction scheme in HEVC coding tools. In addition, to exploit the data geometry for dimensionality reduction of LF, [15,16] presented coding schemes for LF based on low rank approximation. Likewise, in [17], the author used the disparity compensated prediction method to take advantage of the existing spatial redundancy. In addition, the high-order prediction (HOP) model has also been considered as a method to achieve compression, such as in [18]. Based on a geometric transformation between the current block and the reference region, this method provides a high-order intra-block prediction method by adding HOP to HEVC intra prediction modes. Moreover, in recent works, an objective performance assessment of LF lenslet representation was investigated in [19]. The LF lenslet is used with YUV 4:4:4 encoding at 10 bit/sample, which performs well in terms of coding efficiency for different colored sub-sampling formats. In regard to the repeating patterns of lenslets in this representation, screen content coding (SCC) [20] is an efficient encoder for LF image compression. The work in [21] presented an efficient lenslet image coding model, which applies SCC to encode LF lenslets. Based on the plentiful repeating patterns of the LF lenslet representation, this approach is faster and more powerful than the SCC standard, with an even faster decoding time.

On the other hand, 4D-LF represents the LF data as a stack of sub-aperture images (SAIs) generated from lenslets of an LF camera. In the 4D-LF coding approach, generating the 4D-LF pseudo-sequence is a well-known approach for LF compression. This approach involves shifting LF data from the still image coding aspect into the video coding aspect. The sub-aperture array is defined as a pseudo-sequence of different views of LF images and is compressed as a video sequence. Since the first exploration of the LF scanning order in [22,23], several approaches and a variety of scanning orders have been examined, seeking a higher redundancy among SAIs and increased compression efficiency [24–26]. For the inter-frame coding mode of a video codec, the similarity between SAIs is a significant parameter in the compression performance. In [24], a 4D-LF pseudo-sequence was created by organizing SAIs from the lenslet array structure. Nevertheless, the

coding order and reference frame management are implemented coarsely in a way that does not adapt to specific scenarios. In [25], the author presented a solution to fully exploit information among different views. A hierarchical coding order is applied to encode the 2-D coding structure with the selected number of frames used. Based on different scanning orders in [26], the greater the viewpoint distance between SAIs, the less similarity between SAIs. Additionally, [27] recently presented an efficient coding strategy to convert the model parameters into a bitstream, which is well suited for 4D-LF compression.

According to the literature, LF coding can achieve encouraging results with predictive video coding methods (i.e., H.264/AVC, H.265/HEVC). However, the conventional predictive video coding paradigm mostly focuses on one-to-many applications, which result in complex encoders but simple decoders—it is not suitable with simpler encoders for emerging applications, such as visual sensor networks, remote sensing, or visual-based Internet of Things (IoTs). In regard to the other alternative coding possibilities, three-dimensional discrete wavelet transform-based video coding (3-D DWT) [28] and compressive sensing (CS)-based video coding [29] may also be selected for emerging video applications due to their low encoding complexity requirements. However, in spite of the fast video coding provided by these techniques, 3-D DWT and CS-based video coding approaches still require a large amount of encoding memory and have inferior rate distortion performance when compared to the relevant intra-frame encoding codecs (e.g., H.265/HEVC). In this context, Wyner–Ziv (WZ) coding [30], a lossy distributed coding paradigm [31], introduces a low encoding complexity capability, whereby the motion estimation part on the encoder side is shifted to the decoder side. This coding approach has successfully been applied to many different forms of video and emerging applications, such as natural image analysis, hyperspectral images, sensor networks, and wireless video. WZ coding provides different coding techniques compared to conventional video coding, as well as notably providing a flexible distribution of the codec complexity, high compression, and inherent error robustness [32].

This type of coding manages to separately encode individual frames, which are in turn decoded conditionally to achieve similar efficiency to standard coding. The first WZ coding approach in [33,34] was applied to video signals in the real world, giving improved error resilience. Regarding WZ coding with LF images, several LF image compression approaches have been proposed [35–38]. In particular, the performance of distributed video coding for light field content was analyzed in [39]. In [40], the LF images were compressed by WZ coding for random access. Taking advantage of the WZ coding structure, the images are independently encoded by a WZ encoder while previously reconstructed images are applied as Side Information (SI) at the receiver to exploit the similarities among LF images. The results show significant compression performance compared to intra coding while maintaining the random access capabilities. Hence, this is a promising coding solution for LF images.

## 1.2. Contributions and Paper Organization

Regarding LF image coding requirements and WZ coding, the biggest challenge is the transmission of LF content to multiple end users with different display devices and applications while controlling and retaining the quality of an immense amount of data. In this sense, an efficient LF coding architecture is of utmost importance. Thus, extending and improving the work in [41], we propose a novel adaptive content frame skipping approach for LF image compression by following the distributed coding approach in order to achieve efficient compression performance for LF data with low encoding complexity. The contributions of this paper are summarized below.

- An advanced WZ-based LF image compression solution: The well-known WZ coding approach is enhanced by improving the compression performance at the key frame encoder–decoder with state-of-the-art video compression using H.265/HEVC [42], while the advantage of the low complexity of the WZ procedure is utilized on the side of the WZ frame encoder–decoder. Additionally, an advanced channel codec (i.e., LDPC codec [43]) is applied in this WZ coding approach to achieve capacity approaching the performance requirements and flexible code designs using density evolution [44];

- An efficient content-driven LF image reordering mechanism: The different scanning methods may affect the results depending on the video content and characteristics. Based on the high correlation of SAIs and different content types of LF images, 4 scanning methods (i.e., spiral scan, hybrid scan, U-shape scan, and raster scan) are evaluated thoroughly in order to select the most efficient scanning methods for LF images, and also to further improve the performance of our WZ coding solution;
- An adaptive skip mode decision algorithm: To further improve the proposed WZ-LF image coding paradigm, an adaptive skip mode decision is introduced using a decision tree rule-based method, which is based on the changes of spatial and temporal features of the LF content sequences. The associated side information is used as the final reconstructed frame when the skip mode is applied to WZ frames.

The remainder of this paper is organized as follows. Section 2 gives an overview of the proposed LF coding architecture. Section 3 presents the novel adaptive content frame skipping algorithm. Afterward, Section 4 analyzes the experimental results, while Section 5 presents the conclusions and describes directions for future work.

## 2. Overall Wyner–Ziv-Based Light Field Image Compression

This section presents the WZ-LF image compression solution in detail. In order to achieve the best performance for the solution, an efficient scanning order based on LF content is analyzed and a content skipping algorithm is introduced.

### 2.1. Proposed WZ-LF Architecture

To achieve efficient compression performance for transmission and storage of LF images, Figure 1 illustrates the proposed WZ coding-based LF image compression architecture. The proposed WZ-LF coding method is strengthened compared to the original WZ architecture proposed by Girod [31] by improving the compression performance at the key frame encoder–decoder using the state-of-the-art video compression codec H.265/HEVC Intra. As shown, the LF image can be processed in the following steps.

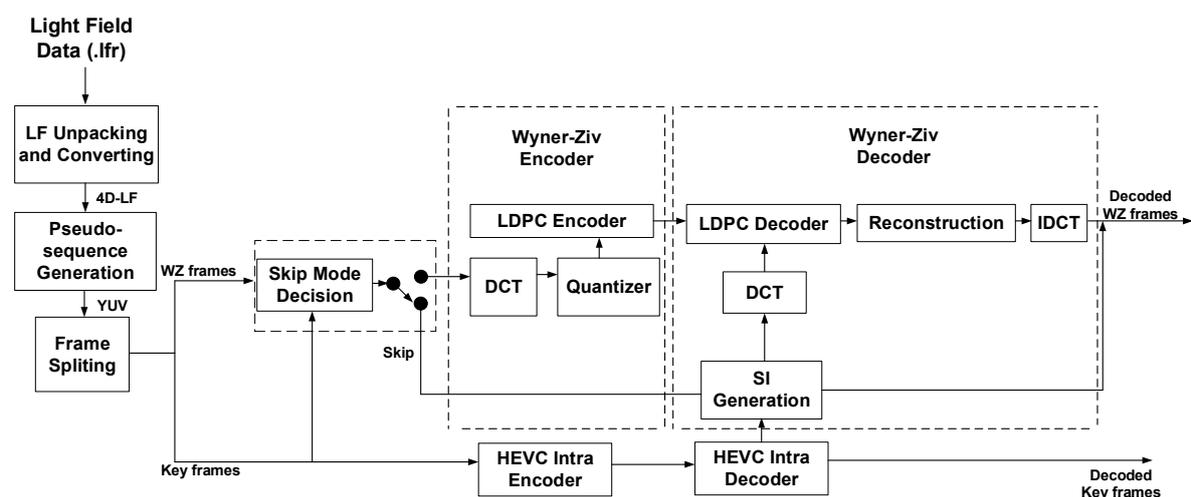


Figure 1. Proposed WZ-LF image compression architecture.

- At the encoder:

The LF data are firstly unpacked and decoded into the 4D-LF representation. The SAIs within the 4D-LF are then grouped into a pseudo-sequence using an efficient scanning order, which is described in the next sub-section. The LF image compression problem is then cast as a common video coding problem. The first frame of every group of pictures (GOP), called a key frame, is encoded using the recent H.265/HEVC intra coding approach [42], with only the spatial correlation

employed; thus, low complexity and error robustness can be achieved. For the remaining WZ frames, the following steps are performed:

1. Skip mode decision:

In this module, the skipping decision is activated based on a decision tree algorithm [45]. The key frames and WZ frames are used to determine the skip or non-skip WZ frames by identifying texture information and motion activity in the 4D-LF pseudo-sequence. The several features are computed to detect changes of spatial and temporal characteristics in the video sequence, e.g., the sum absolute difference (SAD), gradient magnitude similarity (GMS), and variance of block (VAR). These features will be explained in the next sub-section. A rule-based method with decision tree calculates the values of decision nodes based on the features in order to make the skipping decision. When the skip mode decision is activated, the WZ frames are skipped in the normal WZ encoding and decoding procedure and the associated side information is used for the final reconstruction frames. This process is explained in detail in next sub-section.

2. Discrete cosine transforms (DCT):

For WZ frames, the discrete cosine transform (DCT) is used to exploit the statistical dependencies within a frame. The DCT is applied to each  $4 \times 4$  block for WZ frames. By breaking down the image into a  $4 \times 4$  block of pixels arranged from left to right and top to bottom, the WZ frames are transformed using a  $4 \times 4$  DCT. Since the DCT operation has been started, the standard zig-zag scan order [46] within the  $4 \times 4$  DCT coefficient blocks will group the DCT coefficient bands together. The coefficients are organized into 16 bands after being processed in zig-zag scan order. The direct current (DC) band and the alternating current (AC) band are defined as low-frequency information for the first band and as high-frequency information for the remaining bands.

3. Uniform quantization:

In order to encode WZ frames, a quantizer is then applied to each DCT band individually utilizing a predefined number of levels, which depend on the target quality for the WZ frame. By utilizing a uniform scalar quantizer with a greater number of levels (i.e., with lower step sizes), the lower spatial frequencies of the DCT coefficients are processed. Meanwhile, with a lower number of levels, the higher frequency coefficients are more coarsely quantized without significant degradation of the visual quality of the decoded image. Similar to [47], 8 different types of quantization matrices are adopted in the proposed LF compression scheme to target various quality levels and data rates.

4. Low-density parity check (LDPC) encoding:

In this work, to achieve lower complexity in contrast to turbo codes [48], we employ a known low-density parity check accumulator (LDPCA) channel encoder as the WZ encoder. A LDPCA encoder comprises an LDPC syndrome former integrated with an accumulator. By using LDPC code and modulo 2, syndrome bits are established, producing the accrued syndrome for every bit plane. The accrued syndromes are saved in a buffer of the encoder, then transmission of only a few of the syndromes in chunks is started. In case of failure at the decoder, a feedback channel is utilized in the encoder buffer in order to transmit more accrued syndromes. By transmitting an 8-bit cyclic redundancy check (CRC) sum of the encoded bit plane, the decoder is provided with the ability to detect residual errors.

- At the decoder:

1. SI generation:

The SI is known as WZ frame estimation and is generated by a frame interpolation algorithm [49], with two consecutive decoded key frames at the decoder side. The SI is also considered a noisy version of the original WZ frame, with a reciprocal relationship between the number of parity bits (or bit rate) and the quality of noise estimation, i.e., the better the quality of estimation, the smaller the received bit rate. By estimating the correlation between the original WZ frame and the SI correctly, the decoding performance can be greatly improved. The better the quality of the SI that is

interpolated, the better the quality of the final reconstructed WZ frame that can be achieved. Regarding correlations between frames, the 4D-LF pseudo-sequence is a series of frames with high correlation due to the characteristics of the LF image. Thus, achieving the best quality for SI gives a huge advantage in achieving impressive decoding performance.

## 2. LDPC decoding:

In this part, we describe the decoding of a bit plane given the soft input estimations of the SI and the parity bits transmitted from the encoder. From the decoder, in the case of an increase in the number of parity bits, the decoding procedure is then looped. Additionally, for inverse accumulation activity from the encoder, the syndrome bits are removed from the received parity bits before the beginning of the procedure. On these syndrome bits a sum product decoding operation is performed. These instructions are considered as a soft decision algorithm, with the probability of each received bit treated as an input. Additionally, when the decoded bit plane matches the value received from the encoder with the CRC sum registration, this is considered to be an effective decoding process. Then, the decoded bit plane is sent to the inverse quantization and reconstruction module.

## 3. WZ frame reconstruction:

In WZ frame reconstruction, the decoded quantized symbol streams relating to each DCT band are formed through all the bit planes related to these bands. When all decoded quantized symbols are received, all DCT coefficients are reconstructed with the support of the corresponding SI coefficients and the estimated correlation information between the original WZ and SI frames. It should be noted that in the proposed scheme, a correlation noise estimation process is performed at the decoder side and used as a decoder rate control mechanism. The corresponding DCT SI bands are chosen when the DCT coefficients bands with no parity bits are transmitted. The WZ frames and the reconstructed frames are then applied to the reconstruction function to bound the error.

### 2.2. Efficient Sub-Aperture Image Arrangement

Recently, a scanning order method was developed based on optimized reference picture selection for LF image coding using a low-delay configuration with H.265/HEVC [50]. However, this method is not suitable for our proposed WZ-LF codec, which encodes and decodes *KEY* and WZ frames with an intra coding approach. Therefore, in order to select an efficient SAI arrangement, several scan paths of sub-aperture images are examined, such as spiral, raster scan, U-shape, and hybrid scanning approaches [26], as shown in Figure 2. Combining the raster and U-shape scanning order, the hybrid scanning order takes advantage of the similarity of adjacent views, both horizontally and vertically. However, due to varying angles between SAIs, the temporal correlation along SAIs may be changed by different scanning orders and with different LF content. Moreover, the compression performance of the 4D-LF pseudo-sequence can be affected by specific content. Therefore, in this section we thoroughly evaluate the scanning order to verify the most effective order for LF images.

Beginning with content-driven considerations, a set of LF data is collected from [51] containing different content types and categorized into two types: wide and narrow. The wide LF content type includes wide depth-of-field (WDOF), wide depth of field with subject layer (WDOF-L), and blurry content (BC), while the narrow type includes narrow depth of field (NDOF), narrow depth of field with a focus on one main subject (NDOF-1), and narrow depth of field with a focus on more than two subjects (NDOF-2).

Regarding scanning methods, the four types of scanning orders (i.e., spiral scan, hybrid scan, U-shape scan, and raster scan) are applied and computed to determine the most efficient scanning order for LF images. The three following LF images, i.e., spear fence 2 (NDOF), stairs (NDOF-1), and swan 1 (WDOF-L), are selected for evaluation with a temporal frequency of 15 Hz, 193 frames, and encoded by the H.265/HEVC codec.

From the RD performance results in Figure 3, the spiral scanning method may be considered the most suitable for LF images, as it achieves better results than the other scanning methods. Therefore, to achieve the best performance with our proposed WZ-LF coding solution, the spiral scanning method is chosen. Its performance is evaluated in detail in the next section.

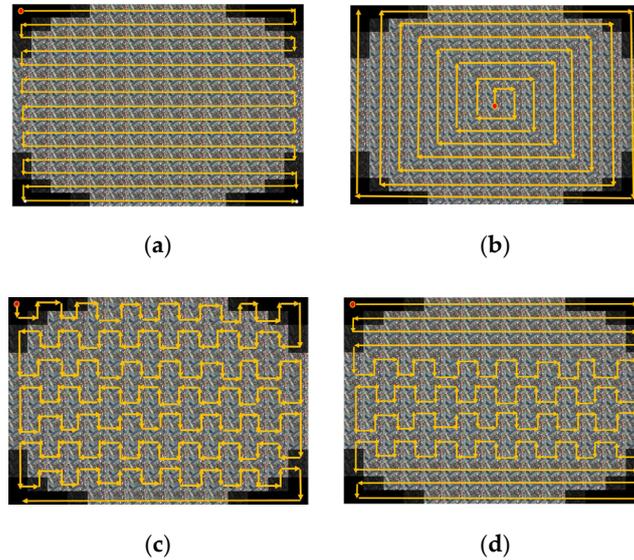


Figure 2. Four scanning order types. (a) Raster. (b) Spiral. (c) U-shape. (d) Hybrid.

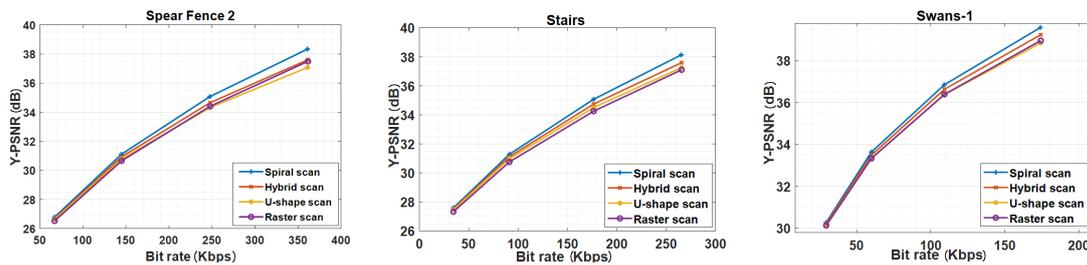


Figure 3. Rate distortion (RD) performance for different scanning methods.

### 3. Adaptive Content Frame Skipping Algorithm

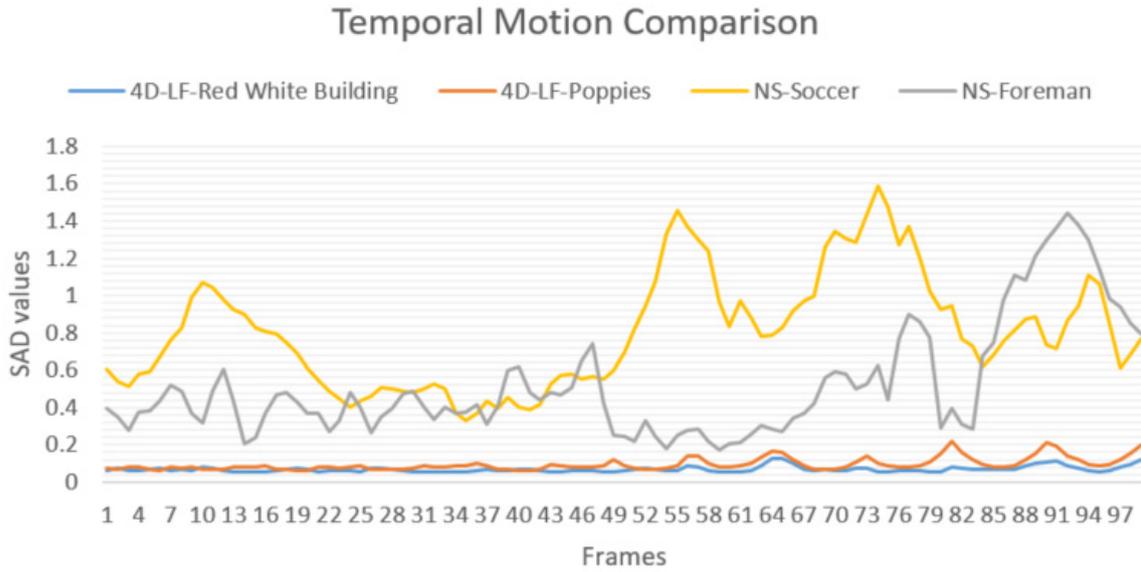
#### 3.1. Observation

Distributed video coding is well known for having low encoding complexity and for providing various advantageous coding techniques, i.e., flexible distribution of the codec complexity, high compression, and inherent error robustness [32]. This coding method is suitable for many different forms of video in emerging applications, e.g., sensor networks, wireless video, and surveillance video.

The different video sequence types (i.e., low-motion and high-motion sequences) are considered to affect to the compression performance of the codec. The low-motion and high-motion sequences refer to high correlation and low correlation between each frame, respectively. Based on the sequence motion, the distributed video coding approach in common low-motion sequences (e.g., hall monitor, Akiyo) achieves better compression performance in comparison to traditional codecs, while the compression performance declines for high-motion sequences (e.g., Soccer, Foreman) [47].

According to the LF characteristic, adjacent views in the 4D-LF pseudo-sequences both horizontally and vertically exhibit higher similarity with each other. Therefore, the 4D-LF pseudo-sequences are mostly considered low-motion sequences compared to natural videos according to the SAD values shown in Figure 4.

It is noted that the frames of 4D-LF pseudo-sequence are extremely highly correlated, as shown in Figure 4, so skipping the most similar frames may achieve efficient compression performance. Therefore, an adaptive frame skipping mechanism based on a decision tree is introduced in our WZ-LF coding solution and is described in detail in the following section.



**Figure 4.** Motion comparison between four-dimensional light field (4D-LF) pseudo-sequences and natural sequences (NS).

### 3.2. Decision Tree Based Adaptive Frame Skipping

Following the analysis of LF data types in the previous sub-section, different data types and different scanning orders can lead to different values of these features, because each SAI represents a different perspective. In this work, we apply the iterative dichotomiser 3 (ID3) algorithm [45] to the frame skipping decision based on an offline training model with spatial and temporal features of the 4D-LF pseudo-sequence.

Based on the high correlation of SAIs and the WZ-LF architecture, it is important to identify the motion activity of the key frames. Thus, the two discriminative temporal features are utilized to detect changes in the motion of the key SAI frames, i.e.,  $FT_{SAI\_SAD}$ , the sum of absolute difference of SAI key frames; and  $FT_{SAI\_GMS}$ , the similarity of the gradient magnitude employed with the Scharr operator [52]. The temporal features are computed as follows:

$$FT_{SAI\_SAD} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} |KEY_a(x, y) - KEY_b(x, y)| \tag{2}$$

where  $KEY_a$  and  $KEY_b$  are two consecutive SAI key frames,  $(x, y)$  is the pixel location in the SAI key frames with size of  $N \times M$ .

$$FT_{SAI\_GMS} = \frac{2GKEY_a(i)GKEY_b(i) + C}{GKEY_a^2(i) + GKEY_b^2(i) + C} \tag{3}$$

where  $GKEY_a(i)$  and  $GKEY_b(i)$  are the gradient magnitude of the two consecutive SAI key frames at  $i$  pixel location and  $C$  is a positive constant for equation stability.  $GKEY_a(i)$  and  $GKEY_b(i)$  employ the convolution operation  $\otimes$  in the horizontal  $D_x$  and vertical  $D_y$  directions following the Scharr filter, computed as:

$$GKEY_a(i) = \sqrt{(D_x \otimes KEY_a)^2 + (D_y \otimes KEY_a)^2} \tag{4}$$

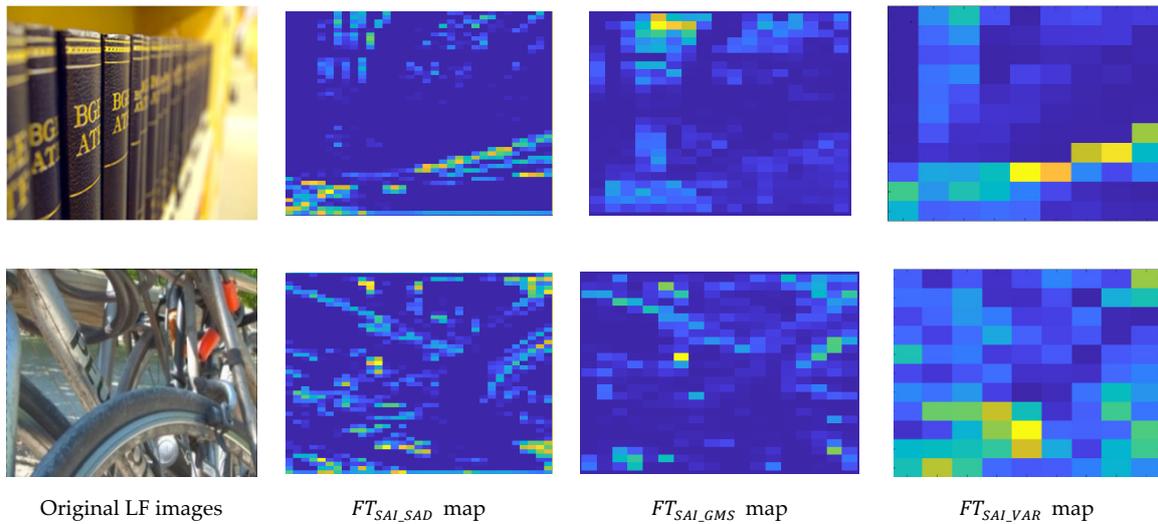
$$GKEY_b(i) = \sqrt{(D_x \otimes KEY_b)^2 + (D_y \otimes KEY_b)^2} \tag{5}$$

Regarding texture information, the spatial feature is also an essential element in order to identify flat and non-flat regions in the SAI WZ frames of the 4D-LF pseudo-sequence. By identifying the difference in texture information, the block variance is selected for content image assessment, i.e.,  $FT_{SAI\_VAR}$ , and is computed as

$$FT_{SAI\_VAR} = \sigma_{WZ}^2 \tag{6}$$

where  $\sigma_{WZ}^2$  is the variance of the SAI WZ frames in the 4D-LF pseudo-sequence.

Figure 5 shows the discriminative spatial and temporal features of the SAI key frames and WZ frames. Notably, the value of the spatial feature covers most of the flat regions (i.e., blurred regions or regions with low texture), while the values of the temporal features cover non-flat regions (i.e., regions with depth, contrast, and saturation complexity).



**Figure 5.** The visualization of the spatial and temporal features, showing books in the top frames and bikes in the bottom frames.

The frame skipping mechanism is based on a technique wherein the texture and motion activity of the two consecutive key frames and neighbor WZ frames of a 4D-LF pseudo-sequence are used for the selection of frames to be skipped through a decision tree rule [45]. In order to establish the skip and non-skip rules from the tree structure, an offline trained model is applied to the binary decision tree. The optimal weights for the offline model are determined by computing all temporal and spatial features for each LF content type as described in Algorithm 1. Based on this, the skip mode decision is considered for activation or not. It should be noted that neither the sample data nor the weights are updated for the offline model, thus, the offline model should be maintained continuously for the best accuracy.

The proposed algorithm is constructed as below

---

**Algorithm 1** The decision-tree-based adaptive frame skipping

---

Input: 4D-LF pseudo-sequence

Output: Skip mode decision (i.e., skip or non-skip)

Initialize the data partitioning with WZ frames ( $WZ_{t+1}$ ) and two consecutive KEY frames ( $Key_t; Key_{t+2}$ ).

Extract the attribute feature as following  $FT_{SAI\_VAR}; FT_{SAI\_GMS}; FT_{SAI\_SAD}$

Determine the threshold  $ThD$  based on the average value of each features.

1.  $ThD_0 = 0;$
  2. **for**  $t = 0, 1, 2, \dots, (total\_frame-2)$  **do**:
  3.  $FT_{SAI\_SAD} = \sum_{x=0}^{N-1} \sum_{y=0}^{M-1} |KEY_{t+2}(x, y) - KEY_t(x, y)|$
  4.  $ThD_{SAI\_SAD} = ThD_0 + FT_{SAI\_SAD}$
  5.  $AveThD_{SAI\_SAD} = ThD_{SAI\_SAD} / Total\_frame$
-

---

```

6.  if ( $ThD_{SAI\_SAD} < AveThD_{SAI\_SAD}$ )
7.   $FT_{SAI\_SAD} = 1$ 
8.  end if
9.   $FT_{SAI\_GMS} = \frac{2GKEY_t(i)GKEY_{t+2}(i)+C}{GKEY_t^2(i)+GKEY_{t+2}^2(i)+C}$ 
10.  $ThD_{SAI\_GMS} = ThD_0 + FT_{SAI\_GMS}$ 
11.  $AveThD_{SAI\_GMS} = ThD_{SAI\_GMS}/Total\_frame$ 
12. if ( $ThD_{SAI\_GMS} < AveThD_{SAI\_GMS}$ )
13.  $FT_{SAI\_GMS} = 1$ 
14. end if
15.  $FT_{SAI\_VAR} = \sigma_{WZ_{t+1}}^2$ 
16.  $ThD_{SAI\_VAR} = ThD_0 + FT_{SAI\_VAR}$ 
17.  $AveThD_{SAI\_VAR} = ThD_{SAI\_VAR}/Total\_frame$ 
18. if ( $ThD_{SAI\_VAR} < AveThD_{SAI\_VAR}$ )
19.  $FT_{SAI\_VAR} = 1$ 
20. end if
21. end for

```

Establish a selection method for the optimal weights  $W$

```

22.  $W = 2$ 
23. if ( $FT_{SAI\_SAD} + FT_{SAI\_GMS} + FT_{SAI\_VAR} \geq W$ )
24. Skip_mode_decision = 1
25. else
26. Skip_mode_decision = 0
27. end if
28. The skip mode decision is activated if the frame meets the optimal weight ( $W$ )
29. Generate the mode decision

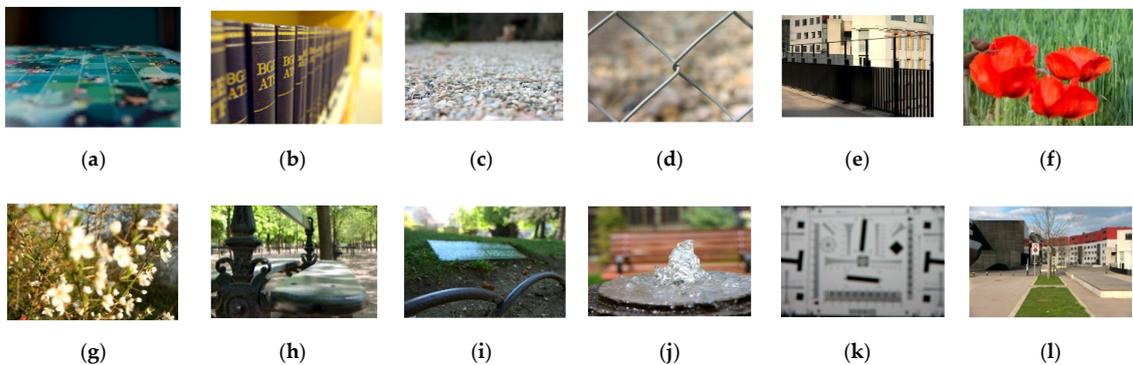
```

---

## 4. Performance Evaluation

### 4.1. Test Conditions

For emerging application scenarios such as visual sensor networks, remote sensing, or camera surveillance, low-resolution imagery is more common than high-resolution imagery; thus, we examine in this paper low-resolution versions of 12 common LF images (shown in Figure 6) by downsampling to Quarter Common Intermediate Format (QCIF) resolution with a temporal frequency of 15 Hz.

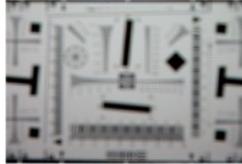


**Figure 6.** Thumbnails of light field images: (a) game board, (b) books, (c) grave garden, (d) chain link fence 1, (e) black fence, (f) poppies, (g) Mirabelle prune tree, (h) bench in Paris, (i) caution bees, (j) fountain and bench, (k) ISO chart 13, (l) red and white building.

Based on the high correlation of SAIs in the 4D-LF, the proposed WZ-LF coding solution is especially suitable for these emerging applications. Similarly, the datasets used for training are

presented in Table 1 with 16 LF training samples. This dataset was collected from [48] and covers different categories and content types. To assess the performance of the proposed LF compression solution, these LF images are examined with the relevant coding benchmark H.265/HEVC [42] and HEVC-based DVC codecs [53]. The comparison analyzes two parts, i.e., the overall rate distortion (RD) performance and the specific coding tool performance. Regarding the development environment, the proposed WZ-LF coding solution is developed using the C language through Visual Studio 2015 and integrated with the state-of-the-art H.265/HEVC Intra.

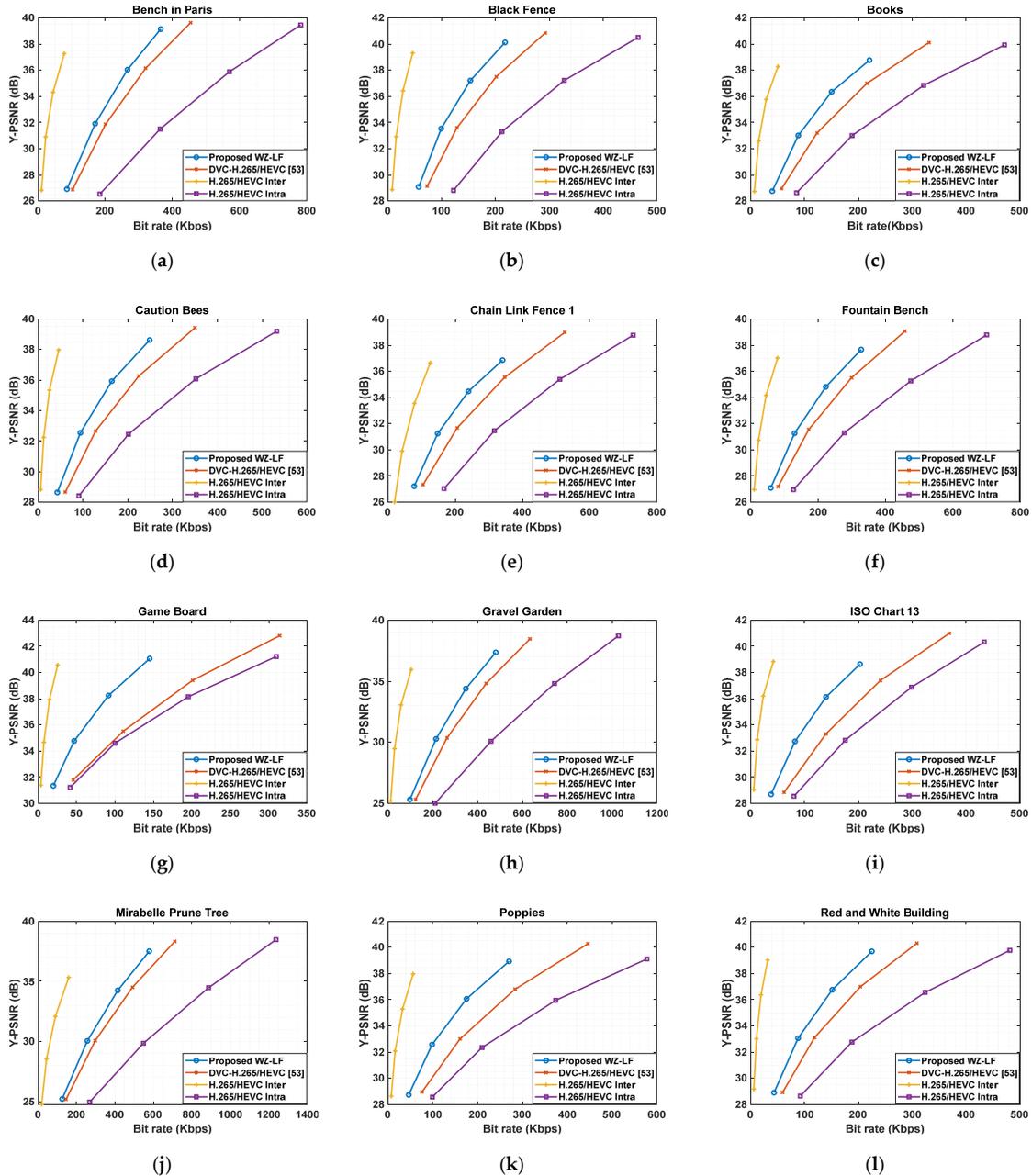
**Table 1.** List of datasets used for training.

LF Training Samples	Category	Content Types	Thumbnail Samples
Houses and lake	Landscapes	WDOF	
Backlight 2	Light		
Rolex learning center	Buildings	WDOF-L	
Reeds	Landscapes		
Backlight 1	Light		
ISO chart 15	ISO and color charts	BC	
Perforated metal 2	Grids		
Slab and lake	Landscapes	NDOF	
Bush	Nature		
Wall decoration	Urban		
Sewer drain	Urban		
Sophie and Vincent 2	People	NDOF-1	
Ankylosaurus and Diplodocus 2	Studio		
Bikes	Urban	NDOF-2	
Danger de mort	Grids		
Stone pillars outside	Urban		

#### 4.2. Overall WZ-LF Compression Performance Evaluation

Regarding the compression performance, the RD performance is widely utilized to quantify video coding schemes through use of the Bjøntegaard delta-Peak Signal to Noise Ratio (BD-PSNR) and Bjøntegaard delta rate [54]. Figure 7 presents the RD curve comparison for the

proposed WZ-LF coding solution and the other relevant benchmarks, i.e., HEVC inter and intra coding [42] and HEVC-based Distributed Video Coding labeled as DVC-H.265/HEVC [53], while the BD rate and BD-PSNR are computed in Table 2. Some conclusions can be derived from the observed results, as shown below.



**Figure 7.** Overall rate distortion (RD) performance evaluation for the proposed Wyner–Ziv (WZ)-LF coding solution: (a) Bench in Paris. (b) Black Fence. (c) Books. (d) Caution Bees. (e) Chain Link Fence 1. (f) Fountain Bench. (g) Game Board. (h) Gravel Garden. (i) ISO Chart 13. (j) Mirabelle Prune Tree. (k) Poppies. (l) Red and White Building.

- **WZ-LF versus DVC-H.265/HEVC:** The proposed codec uses a similar approach to DVC-H.265/HEVC, however with significantly improved coding tools. The proposed WZ-LF achieves impressive results for compression performance compared to DVC-H.265/HEVC by reducing the bit rate by about 25%. Additionally, taking into account the adaptive mode decision for frame skipping, the WZ-LF architecture achieves a significant gain in compression performance of almost 2 dB with different content types.

- WZ-LF versus H.265/HEVC Intra: The advantage of the proposed coding solution is applied to the intra coding solution at the encoder side and to the inter coding solution at the decoder side. Thus, the proposed WZ-LF can significantly improve the RD performance for all 4D-LF pseudo-sequences with a variety of content types. As shown, the average BD rate reductions are 53.14%, 52.53%, 53.22%, and 53.18% for the proposed WZ-LF solution with WDOF, NDOF, NDOF-1, and NDOF-2, respectively. Hence, the obtained performance improvement confirms the efficiency of the proposed skip mode decision in the WZ-LF architecture.

**Table 2.** Bjøntegaard delta (BD) rate (%) and Bjøntegaard delta-Peak Signal to Noise Ratio (BD-PSNR) (dB) compared to high-efficiency video coding (HEVC) [42] and DVC-H.265/HEVC [53].

LF Sequences	Content Types	WZ-LF vs. DVC-H.265/HEVC (Anchor) [53]		WZ-LF vs. H.265/HEVC Intra (Anchor) [42]		WZ-LF (Anchor) vs. H.265/HEVC Inter [42]		DVC-H.265/HEVC [53] (Anchor) vs. H.265/HEVC Inter [42]	
		BD Rate	BD-PS NR	BD Rate	BD-PS NR	BD Rate	BD-PS NR	BD Rate	BD-PSNR
Red and white building	WDOF	-24.40	1.89	-54.75	5.27	-87.06	11.78	-90.24	13.31
Black fence	WDOF-L	-21.41	2.04	-53.73	6.60	-81.84	11.70	-85.73	13.15
ISO chart 13	BC	-33.99	2.67	-50.95	4.60	-84.59	9.72	-89.74	11.45
Content Type Average		-26.60	2.20	-53.14	5.49	-84.49	11.06	-88.57	12.63
Grave Garden		-17.59	1.51	-53.18	6.23	-82.75	10.38	-85.82	11.48
Chain link fence 1	NDOF	-22.06	1.71	-50.44	5.01	-63.54	6.35	-71.59	8.10
Game Board		-48.29	3.58	-53.98	3.82	-82.60	8.19	-91.09	10.73
Content Type Average		-29.31	2.26	-52.53	5.02	-76.29	8.30	-82.83	10.10
Bench in Paris		-15.88	1.48	-54.69	6.95	-82.85	10.91	-85.59	11.95
Caution Bees Fountain and bench	NDOF-1	-24.53	1.67	-53.15	4.50	-84.37	9.03	-88.24	10.57
Content Type Average		-20.23	1.54	-53.22	5.40	-82.10	9.48	-85.73	10.82
Poppies		-32.92	2.45	-54.23	4.63	-80.49	8.31	-86.95	10.44
Mirabelle Prune Tree	NDOF-2	-13.52	1.18	-53.40	6.42	-75.70	8.71	-79.05	9.60
Books		-25.03	1.79	-51.91	4.63	-80.68	8.48	-85.53	9.91
Content Type Average		-23.82	1.80	-53.18	5.22	-78.95	8.50	-83.84	9.98
Total Average		-24.99	1.95	-53.01	5.28	-80.46	9.34	-85.25	10.89

- WZ-LF versus H.265/HEVC Inter: In the case of high correlation between SAIs of the 4D-LF pseudo-sequences, the compression performance of H.265/HEVC Inter is obviously better than the proposed WZ-LF with the asymmetric compression. The H.265/HEVC Inter codec can be considered as the upper bound of DVC-H.265/HEVC outperforming DVC-H.265/HEVC by 10.9 dB in compression performance, however the proposed WZ-LF improves the compression performance by narrowing the gap by about 1.5 dB compared to DVC-H.265/HEVC. Additionally, the major problem with the H.265/HEVC Inter codec is its high complexity, meaning it is not compatible with the emerging applications considered in this work, whereas the proposed WZ-LF is a suitable solution. Regarding the H.265/HEVC Inter codec with no motion, i.e., without motion compensation, the compression performance of this codec is similar to H.265/HEVC Inter because of the high correlation between SAIs, as shown in Table 3. Thus, H.265/HEVC Inter is a suitable comparison.
- Notes on performance variation with content types: Since the proposed WZ-LF solution outperforms the relevant benchmarks, i.e., DVC-H.265/HEVC and H.265/HEVC Intra, the WZ-LF compression performance changes with content type differently to other codecs. By encoding individual frames in H.265/HEVC Intra, the comparison with the WZ-LF solution maintains an approximate 53% bit rate saving for all content types. Notably, in comparison to DVC-H.265/HEVC, the WZ-LF solution achieves the best compression performance for the NDOF content type, with BD rate reductions of 29.3%, while NDOF-1 and NDOF-2 are improved respectively by approximately 20.2% and 23.8%, respectively. The WDOF and BC

content types represent the low motion and high correlation between SAIs of the 4D-LF pseudo-sequence. Thus, the bit rate of the proposed coding solution is reduced by 34% and 54% compared to DVC-H.265/HEVC and H.265/HEVC Intra for the BC and WDOF content, respectively. It is noticed that some specific LF sequences with high contrast and saturation contents, i.e., game board, poppies, chain link fence 1, and books, also vary in their bit rate savings and compression performance compared to DVC-H.265/HEVC and H.265/HEVC Intra. For instance, the proposed WZ-LF solution achieves bit rate savings of 48.2% and 53.9% in comparison to DVC-H.265/HEVC and H.265/HEVC Intra, respectively, for the game board sequence.

**Table 3.** BD rate (%) and BD-PSNR (dB) comparison between HEVC Inter and HEVC Inter with no motion.

LF Sequences	Content Types	H.265/HEVC Inter vs H.265/HEVC Inter No Motion (Anchor) [42]	
		BD-Rate	BD-PSNR
Red and white building	WDOF	-0.03	0
Black fence	WDOF-L	-0.13	0.01
ISO chart 13	BC	-6.33	0.32
Content Type Average		-2.16	0.11
Grave Garden		-1.97	0.10
Chain link fence 1	NDOF	-18.51	1.16
Game Board		-0.87	0.04
Content Type Average		-7.11	0.43
Bench in Paris		-0.72	0.04
Caution Bees	NDOF-1	-0.98	0.04
Fountain and bench		-5.42	0.28
Content Type Average		-2.37	0.12
Poppies		-0.41	0.02
Mirabelle Prune Tree	NDOF-2	-1.55	0.08
Books		-1.49	0.07
Content Type Average		-1.15	0.07
Total Average		-3.20	0.18

### 4.3. WZ-LF Codec with Various Coding Tools

#### 4.3.1. Scanning Method Assessment

Four scanning order types, i.e., spiral, hybrid, U-shape scanning, and raster, are evaluated based on BD rate [54]. Regarding the common scanning method used for video coding, raster scanning is utilized as an anchor in order to compute the BD rate. Broken down into different data types, the BD rate results for the scanning methods are shown in Table 4. The hybrid and U-shape scanning orders achieve bit rate savings of approximately 3% compared to the raster scan for most content types, however for NDOF and NDOF-1 types, the BD rate performance changes compared to the raster scan, with bit rate savings of approximately 3% and 24%, respectively. Regarding the spiral scan, this method achieves an outstanding result, with an average bit rate saving of 10% for all data types compared to the raster scan. In particular, for the NDOF and NDOF-1 data types, this method still achieves impressive performance, with bit rate savings of approximately 11% and 9%,

respectively. Thus, we could tentatively conclude that the spiral scan is the most efficient scanning method, especially for LF images.

**Table 4.** Average (%) BD rate saving comparison for different content types.

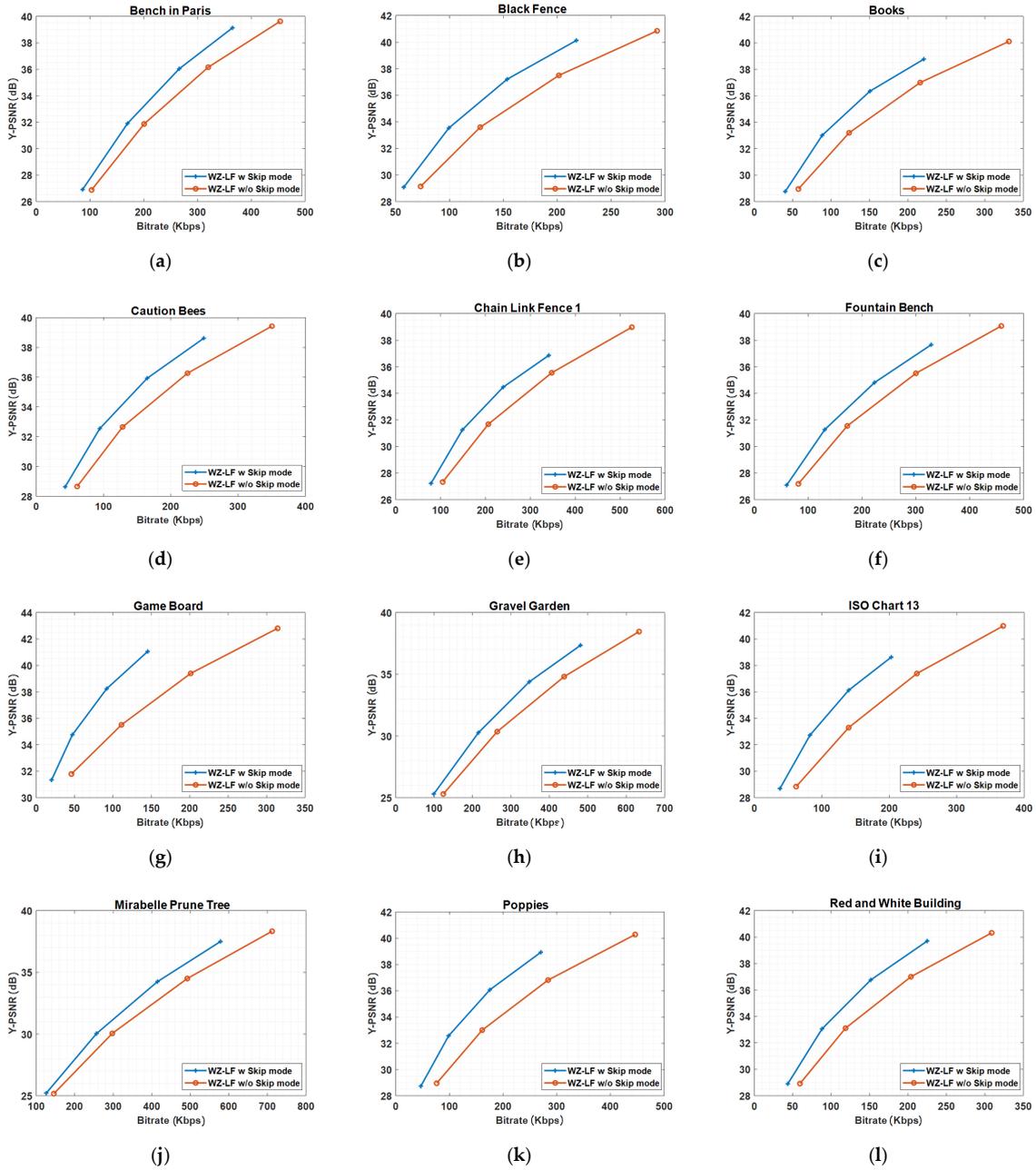
Sequences	Content Type	BD Rate (Raster Scan as Anchor)		
		Spiral	Hybrid	U-Shape
Red and white building	WDOF	-9.50	-2.81	-1.62
Black fence	WDOF-L	-11.30	-3.59	-1.47
Chain link fence 1	NDOF	-11.20	13.34	24.83
Fountain and bench	NDOF-1	-9.14	3.11	8.53
Poppies	NDOF-2	-9.58	-3.03	-1.16
ISO chart 13	BC	-13.81	-2.97	1.07
Average		-10.75	0.67	5.03

#### 4.3.2. Skip Mode Assessment

##### Compression Performance

Considering the high correlation of SAIs in the 4D-LF pseudo-sequences, the decision tree method is applied to determine the skipping process at the encoder side of the WZ-LF architecture in order to enhance the compression efficiency of the WZ-LF coding solution. The spatial-temporal features of the 4D-LF pseudo-sequences are selected based on the depth of field changes in the content. According to the rules created by the offline trained model, the skip mode decision determines whether or not to skip the WZ encode procedure and instead encode images as normal 4D-LF pseudo-sequences.

Table 5 and Figure 8 show comparisons of the WZ-LF coding solution with and without the skip mode decision. Examining the RD performance results, it is clear that WZ-LF with skip mode has lower complexity than WZ-LF without skip mode, with a bit rate saving of 25%. Notably, the NDOF content type shows a significant improvement with a bit rate saving of 29.3%, while WDOF, NDOF-1, and NDOF-2 achieve BD rate reductions of 26.6%, 20.2%, and 23.8%, respectively. Therefore, we can observe that the skip mode performs outstandingly with blurry content (34% bit rate saving) or content with a narrow depth of field (48% bit rate saving for the game board sequences).



**Figure 8.** RD performance evaluation of WZ-LF with skip and non-skip mode decisions: (a) Bench in Paris. (b) Black Fence. (c) Books. (d) Caution Bees. (e) Chain Link Fence 1. (f) Fountain Bench. (g) Game Board. (h) Gravel Garden. (i) ISO Chart 13. (j) Mirabelle Prune Tree. (k) Poppies. (l) Red and White Building.

**Table 5.** BD rate (%) and BD-PSNR (dB) of the proposed skip mode.

LF Sequences	Content Types	Skip Mode (Non-Skip Mode as Anchor)	
		BD-Rate	BD-PSNR
Red and white building	WDOF	−24.40	1.89
Black fence	WDOF-L	−21.41	2.04
ISO chart 13	BC	−33.99	2.67
Content Type Average		−26.60	2.20
Grave Garden	NDOF	−17.59	1.51
Chain link fence 1		−22.06	1.71

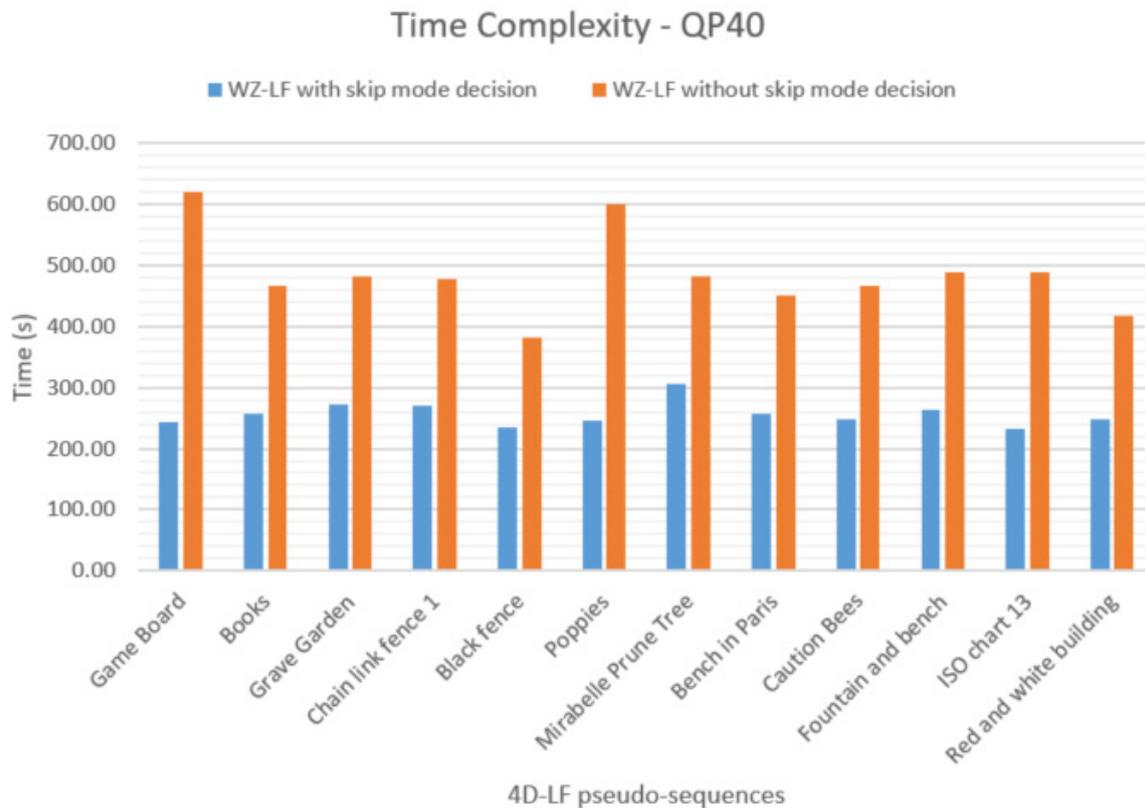
Game Board		-48.29	3.58
Content Type Average		-29.31	2.27
Bench in Paris		-15.88	1.48
Caution Bees	NDOF-1	-24.53	1.67
Fountain and bench		-20.29	1.47
Content Type Average		-20.23	1.54
Poppies		-32.92	2.45
Mirabelle Prune Tree	NDOF-2	-13.52	1.18
Books		-25.03	1.79
Content Type Average		-23.82	1.81
Total Average		-24.99	1.95

### Compression Complexity

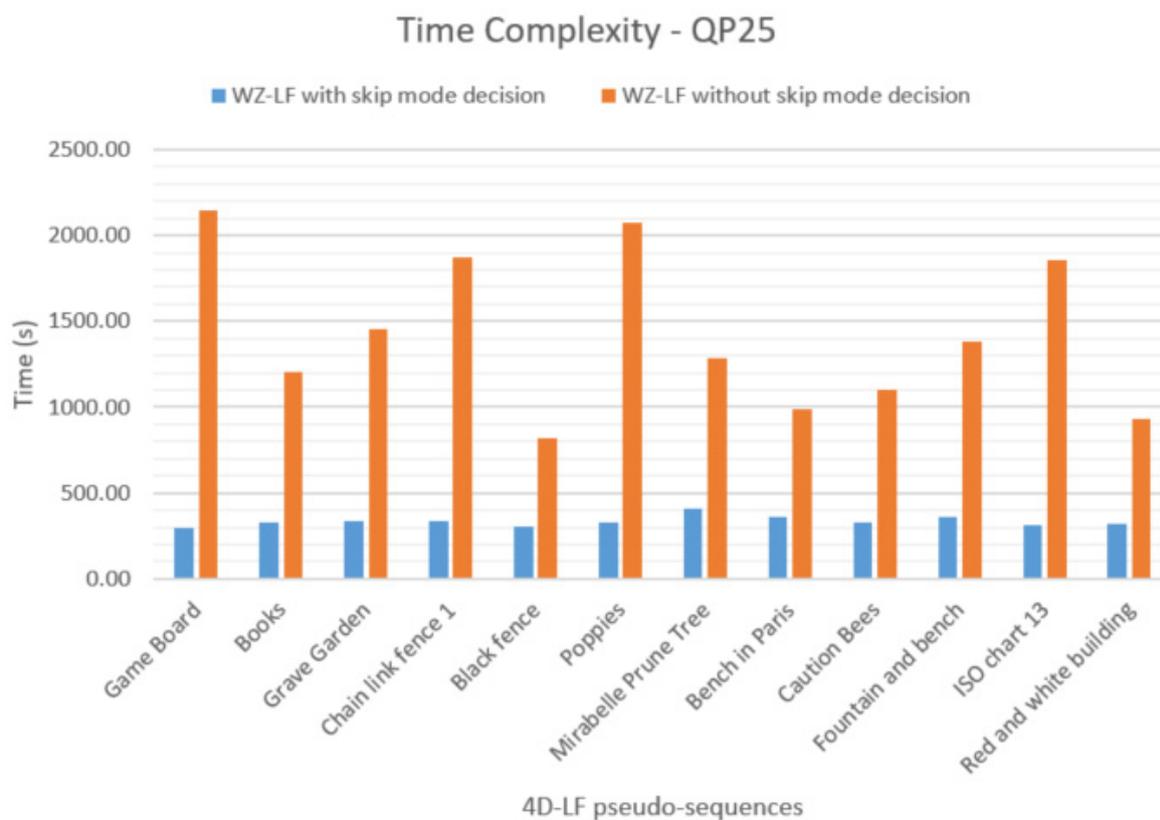
Examining the compression complexity is an essential part of the performance evaluation. For this evaluation, the coding solutions are tested on the same PC with an Intel Core i7-7700HQ (2.8 GHz) processor, 16 GB RAM, and Windows 10-Home OS. The results are shown in Figures 9 and 10, respectively, for Quantization parameter (QP) of 40 and of 25, with and without the skip mode decision. To avoid the effect of multi-thread processing during the test, the results of five repetitions of the same compression setting are averaged. Additionally, the time saving (%) is measured as:

$$Time - Saving = \frac{|T_{skip} - T_{Non\_skip}|}{T_{Non\_skip}} \times 100 \quad (7)$$

where  $T_{skip}$  and  $T_{Non\_skip}$  are the processing time of the WZ-LF codec with and without the skip mode decision, respectively.



**Figure 9.** Time complexity of the codec with and without skip mode at Quantization parameter (QP) 40.



**Figure 10.** Time complexity of the codec with and without skip mode at QP25.

From these complexity results, it can be observed that the WZ-LF codec with the skip mode decision saves a significant amount of time in encoding compared to the WZ-LF codec without the skip mode decision. The WZ-LF with the skip mode can encode approximately 46% and 74% faster on average than the WZ-LF without skip mode at QP40 and QP25, respectively.

## 5. Conclusions

This paper introduces an LF adaptive content frame skipping compression solution following the WZ coding approach by analyzing the spatial and temporal correlation between sub-aperture pictures. The proposed WZ-LF coding paradigm combines the state-of-the-art H.265/HEVC codec with an adaptive frame skipping mechanism, along with an efficient scanning order. The proposed LF compression architecture provides an efficient scanning order that adapts to LF content. This provides optimized performance for almost all LF content data types. In addition, the up-to-date WZ coding solution based on embedded adaptive frame skipping decisions significantly outperforms the relevant H.265/HEVC Intra and DVC-H.265/HEVC codecs. In particular, the proposed coding solution improves the compression performance and has lower computational complexity than both of the relevant benchmarks. Hence, the proposed WZ-LF coding solution meets the requirements for many emerging applications, e.g., visual sensor networks, video surveillance, and remote space transmission.

In future research, other LF image components, i.e., noise and depth maps, could be analyzed in order to provide better quality LF reconstruction. Thus, the proposed WZ-LF coding solution may be further improved.

**Author Contributions:** Conceptualization, X.H. and H.P.; methodology, S.P.; software, H.P.; validation, X.H., H.P., and S.P.; formal analysis, X.H.; investigation, H.P.; resources, S.P.; data curation, S.P.; writing—original draft preparation, H.P.; writing—review and editing, X.H.; visualization, H.P.; supervision, X.H. and S.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 102.01-2020.15.

**Acknowledgements:** This work has been supported in part by the Joint Technology and Innovation Research Center, a partnership between the University of Technology Sydney and Vietnam National University.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Levoy, M.; Hanrahan, P. Light Field Rendering, in SIGGRAPH'96. In Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, New Orleans, LA, USA, 4–9 August 1996; pp. 31–42.
2. Gortler, S.J.; Grzeszczuk, R.; Szeliski, R.; Cohen, M.F. The lumigraph, in SIGGRAPH'96. In Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, New Orleans, LA, USA, 4–9 August 1996; pp. 43–54.
3. Wu, G.; Masia, B.; Jarabo, A.; Zhang, Y.; Wang, L.; Dai, Q.; Chai, T.; Liu, Y. Light Field Image Processing: An Overview. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 926–954.
4. Adelson, E.H.; Wang, J.Y.A. Single Lens Stereo with a Plenoptic Camera. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 99–106.
5. Lytro Camera. Available online: <https://www.lytro.com/> (accessed on 2 March 2020).
6. Raytrix. Available online: <https://www.raytrix.de/> (accessed on 2 March 2020).
7. Xiao, X.; Javidi, B.; Martinez-Corral, M.; Stern, A. Advances in three-dimensional integral imaging: Sensing display and applications. *Appl. Optics.* **2013**, *52*, 546–560.
8. Raghavendra, R.; Raja, K.B.; Busch, C. Presentation attack detection for face recognition using light field camera. *IEEE Trans. Image Process.* **2015**, *24*, 1060–1075.
9. Levoy, M. The Digital Michelangelo Project: 3D Scanning of Large Statues, in SIGGRAPH'00. In Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, New Orleans, LA, USA, 23–28 July 2000; pp. 131–144.
10. JPEG pleno call for proposals on light field coding. Available online: [https://jpeg.org/items/20161102\\_cfp\\_pleno.html](https://jpeg.org/items/20161102_cfp_pleno.html) (accessed on 2 March 2020).
11. JPEG Pleno Holography Uses Cases and Requirements. Available online: [http://ds.jpeg.org/documents/jpegpleno/wg1n86016-REQ-JPEG\\_Pleno\\_Holography\\_Uses\\_Cases\\_and\\_Requirements.pdf](http://ds.jpeg.org/documents/jpegpleno/wg1n86016-REQ-JPEG_Pleno_Holography_Uses_Cases_and_Requirements.pdf) (accessed on 2 March 2020).
12. Conti, C.; Nunes, P.; Soares, L.D. HEVC-Based Light Field Image Coding with Bi-Predicted Self-Similarity Compensation. In Proceedings of the IEEE International Conference on Multimedia Expo Workshops: Seattle, WA, USA, 11–15 July 2016; pp. 1–4.
13. Monteiro, R.; Lucas, L.F.R.; Conti, C.; Nunes, P. Light Field HEVC-Based Image Coding Using Locally Linear Embedding and Self-Similarity Compensated Prediction. In Proceedings of the IEEE International Conference on Multimedia Expo Workshops, Seattle, WA, USA, 11–15 July 2016; pp. 1–4.
14. Li, Y.; Sjostrom, M.; Olsson, R.; Jennehag, U. Efficient Intra Prediction Scheme for Light Field Image Compression. In Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing, Florence, Italy, 4–9 May 2014; pp. 539–543.
15. Jiang, X.; Le Pendu, M.; Farrugia, R.A.; Hemami, S.S.; Guillemot, C. Homography-Based Low Rank Approximation of Light Fields for Compression. In Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing, New Orleans, LA, USA, 5–9 March 2017; pp. 1313–1317.
16. Kamal, M.H.; Vandergheynst, P. Joint Low-Rank and Sparse Light Field Modelling for Dense Multiview Data Compression. In Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing, Vancouver, BC, Canada, 26–31 May 2013; pp. 3831–3835.
17. Chang, C.-L.; Zhu, X.; Ramanathan, P.; Girod, B. Light field compression using disparity-compensated lifting and shape adaptation. *IEEE Trans. Image Process.* **2006**, *15*, 793–806.
18. Monteiro, R.J.S.; Nunes, P.J.L.; Rodrigues, N.M.M.; Faria, S.M.M. Light field image coding using high-order intra block prediction. *IEEE J. Sel. Top. Signal Process.* **2017**, *11*, 1120–1131.
19. Monterio, R.J.S.; Rodeigues, N.N.M.M.; Faria, S.M.M.; Nunes, P.J.L. Light Field Image Coding: Objective Performance Assessment of Lenslet and 4D LF Data Representations. In Proceedings of the SPIE Optical

- Engineering, Applications of Digital Image Processing XLI, San Diego, CA, USA, 20–23 September 2018; Volume 107520D.
20. Peng, W.H.; Walls, F.G.; Cohen, R.A.; Xu, J.; Ostermann, J.; MacInnis, A.; Lin, T. Overview of Screen Content Video Coding: Technologies, Standards, and Beyond. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2016**, *6*, 393–408.
  21. Tsang, S.H.; Chan, Y.L.; Kuang, W. Standard compliant light field lenslet image coding model using enhanced screen content coding framework. *J. Electron. Imaging* **2019**, *28*, 053027.
  22. Fecker, U.; Kaup, A. H.264/AVC-Compatible Coding of Dynamic Light Fields using Transposed Picture Ordering. In Proceedings of the 13th European Signal Processing Conference (EUSIPCO), Antalya, Turkey, 4–8 September 2005.
  23. Vieira, A.; Duarte, H.; Perra, C.; Tavora, L.; Assuncao, P. Data Formats for High Efficiency Coding of Lytro-Illum Light Fields. In Proceedings of the International Conference on Image Processing Theory, Tools and Applications (IPTA), Orleans, LA, USA, 10–13 November 2015.
  24. Liu, D.; Wang, L.; Li, L.; Xiong, Z.; Wu, F.; Zeng, W. Pseudo-Sequence-Based Light Field Image Compression. In Proceedings of the IEEE International Conference on Multimedia & Expo Workshops, Seattle, WA, USA, 11–15 June 2016.
  25. Li, L.; Li, Z.; Li, B.; Liu, D.; Li, H. Pseudo Sequence Based 2-D Hierarchical Coding Structure for Light-Field Image Compression. In Proceedings of the 2017 Data Compression Conference, Snowbird, UT, USA, 4–7 April 2017.
  26. Zhao, S.; Chen, Z.; Yang, K.; Huang, H. Light Field Image Coding with Hybrid Scan Order. In *SPIE Visual Communications and Image Processing*; IEEE: Chengdu, China, 2016.
  27. Verhack, R.; Sikora, T.; Wallendael, G.V.; Lambert, P. Steered Mixture-of-Experts for Light Field Images and Video: Representation and Coding. *IEEE Trans. Multimed.* **2020**, *22*, 579–593.
  28. Belyaev, E.; Egiazarian, K.; Gabbouj, M. A Low-Complexity Bit-Plane Entropy Coding and Rate Control for 3-D DWT Based Video Coding. *IEEE Trans. Multimed.* **2013**, *15*, 1786–1799.
  29. Belyaev, E. Compressive Sensed Video Coding Having Jpeg Compatibility. In Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 25–28 October 2020; pp. 1128–1132.
  30. Wyner, A.; Ziv, J. The Rate-Distortion Function for Source Coding with Side Information at the Decoder. *Ieee Trans. Inf. Theory* **1976**, *22*, 1–10.
  31. Girod, B.; Aaron, A.; Rebollo-Monedero, D. Distributed video coding. *Proc. IEEE* **2005**, *93*, 71–83.
  32. Pereira, F.; Torres, L.; Guillemot, C.; Ebrahimi, T.; Leonardi, R.; Klomp, S. Distributed video coding: Selecting the most promising application scenarios. *Signal Process. Image Commun.* **2008**, *23*, 339–352.
  33. Aaron, A.; Zhang, R.; Girod, B. Wyner-Ziv Coding of Motion Video. In Proceedings of the Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 26–29 October 2002.
  34. Aaron, A.; Rane, S.; Zhang, R.; Girod, B. Wyner-Ziv Coding for Video: Applications to Compression and Error Resilience. In Proceedings of the Data Compression Conference 2003, Snowbird, UT, USA, 25–27 March 2003.
  35. Zhu, X.; Aaron, A.; Girod, B. Distributed Compression for Large Camera Arrays. In Proceedings of the IEEE Workshop on Statistical Signal Processing (SSP '03), St. Louis, MO, USA, 28 September–1 October 2003.
  36. Jagmohan, A.; Sehgal, A.; Ahuja, N. Compression of Light Field Rendered Images Using Coset Codes. In Proceedings of the 37th Asilomar Conference on Signals, Systems, and Computers: Special Session on Distributed Coding, Pacific Grove, CA, USA, 3–6 November 2003.
  37. Toffetti, G.; Tagliasacchi, M.; Marcon, M.; Tubaro, S.; Sarti, A.; Ramchandran, K. Image Compression in a Multi-Camera System Based on a Distributed Source Coding Approach. In Proceedings of the EUSIPCO '05, Antalya, Turkey, 4–8 September 2005.
  38. Yeo, C.; Ramchandran, K. Robust Distributed Multi-View Video Compression for Wireless Camera Networks. In Proceedings of the SPIE Visual Communications and Image Processing, San Jose, CA, USA, 28 January–1 February 2007.
  39. Zhu, X.; Aaron, A.; Girod, B. *Distributed Compression of Light Fields*; Technical Report; Stanford University: Stanford, CA, USA. Available online:

- <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.77.9398&rep=rep1&type=pdf> (accessed on 16 April 2020).
40. Aaron, A.; Ramanathan, P.; Girod, B. Wyner-Ziv Coding of Light Fields for Random Access. In Proceedings of the IEEE 6th Workshop on Multimedia Signal Processing 2004, Siena, Italy, 29 September–1 October 2004.
  41. Cong, H.P.; HoangVan, X.; Perry, S. A Low Complexity Wyner-Ziv Coding Solution for Light Field Image Transmission and Storage. In Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, Jeju, Korea, 5–7 June 2019.
  42. Sullivan, G.J.; Ohm, J.R.; Han, W.J.; Wiegand, T. Overview of the High Efficiency Video Coding (HEVC) Standard. *IEEE Trans. Circuits Syst. Video Technol.* **2012**, *22*, 1649–1668.
  43. MacKay, D. Good error-correcting codes based on very sparse matrices. *IEEE Trans. Inf. Theory* **1999**, *45*, 399–431.
  44. Richardson, T.; Shokrollahi, M.; Urbanke, R. Design of capacity-approaching irregular low-density parity-check codes. *IEEE Trans. Inf. Theory* **2001**, *47*, 619–637.
  45. Quinlan, J. Induction of decision trees. *Mach. Learn.* **1986**, *1*, 81–106. Available online: <http://dx.doi.org/10.1023/A:1022643204877> (accessed on 3 April 2020).
  46. ITU-T Rec. H.264 (11/2007). Advanced Video Coding for Generic Audio Visual Services. Available online: <https://www.itu.int/rec/T-REC-H.264-201906-I/en> (accessed on 3 April 2020).
  47. Artigas, X.; Ascenso, J.; Dalai, M.; Klomp, S.; Kubasov, D.; Ouaret, M. The Discover Codec: Architecture, Techniques and Evaluation. In Proceedings of the Picture Coding Symposium, Lisboa, Portugal, 7–9 November 2007.
  48. Varodayan, D.; Aaron, A.; Girod, B. Rate-Adaptive Codes for Distributed Source Coding. *Eurasip. Signal Process. J. Spec. Sect. Distrib. Source Coding* **2006**, *86*, 11.
  49. Ascenso, J.; Brites, C.; Pereira, F. Content Adaptive Wyner-ZIV Video Coding Driven by Motion Activity. In Proceedings of the 2006 International Conference on Image Processing, Atlanta, GA, 8–11 October 2006.
  50. Ricardo Monteiro, J.S.; Nuno Rodrigues, M.M.; Sérgio Faria, M.M.; Paulo Nunes, J.L. Optimized Reference Picture Selection for Light Field Image Coding. In Proceedings of the 2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2–6 September 2019; pp. 1–5.
  51. Řeřábek, M.; Ebrahimi, T. New Light Field Image Dataset. In Proceedings of the 8th International Conference on Quality of Multimedia Experience, Lisbon, Portugal, 6–8 June 2016.
  52. Xue, W.; Zhang, L.; Mou, X.; Bovik, A.C. Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index. *IEEE Trans. Image Process.* **2014**, *23*, 684–695.
  53. Brites, C.; Pereira, F. Distributed Video Coding: Assessing the HEVC upgrade. *Signal Process. Image Commun.* **2015**, *32*, 81–105.
  54. Bjøntegaard, G. Calculation of average PSNR differences between RD-curves, Doc. ITU-T SG16 VCEG-M33. In Proceedings of the 13th VCEG Meeting, Austin, TX, USA, 2–4 April 2001.

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).