electronics

MDPI

Article

Content Adaptive Frame Skipping for Wyner-Ziv based Light Field Image Compression

Huy PhiCong [1,2,3,4,5], Stuart Perry [3] and Xiem HoangVan [1,2,*]

[1] Faculty of Electronics and Telecommunications, VNU – University of Engineering and Technology
[2] Vietnam National University, Ha Noi, Viet Nam, 17028025@vnu.edu.vn
[3] School of Electrical and Data Engineering, University of Technology Sydney, Australia; stuart.perry@uts.edu.au; conghuy.phi@student.uts.edu.au
[4] Posts and Telecommunications Institute of Technology, Hanoi, Vietnam
[5] JTIRC, VNU University of Engineering and Technology, Hanoi, Vietnam
* Correspondence: xiemhoang@vnu.edu.vn; Address: 144 Xuan Thuy Rd, Dich Vong Hau, Cau Giay Dist, Hanoi, Vietnam, 10000

Received: date; Accepted: date; Published: date

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 transmission bandwidth but also due to the complexity requirements of various applications. In this paper, we propose a novel LF content adaptive 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 high 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, angular information, etc. Experimental results show that the proposed WZ-LF coding solution achieves an outstanding rate-distortion (RD) performance while requiring less computational complexity. Notably, bit rate savings of 53% compared to the standard High Efficiency Video Coding (HEVC) Intra are reported.

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 because the intensity of light rays traveling in every angle at every point in 3D space is captured [3]. Thus the LF image data includes information such as location/point, x, y, z, angle/direction, $\theta, \emptyset$, wavelength, γ, and the time, t, for light rays captured in the scene. This 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, $P_{LF}(x, y, z, \theta, \emptyset, \gamma, t)$ [3]. A raw LF image is composed of micro-images (MIs), and a set of sub-aperture images (SAIs) are obtained by rearranging the co-located pixels from each MI. A SAI corresponds to a captured image of a scene from a particular point of view, which varies 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 a still extensive 4D function as below:

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

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), respectively.

As an example of a LF imaging technology, LF cameras have become a promising tool for various research areas, e.g. richer photography using the Lytro Illum [5], material analysis using the Raytrix [6], medical imaging [7], and biometric recognition [8]. As a result of the enormous size of large photo-realistic LF images, typically 1GB [9], data compression is, therefore, a challenge in LF data for storing, processing, and transmission. Recently, the Joint Photographic Experts Group (JPEG) committee has started an activity for standardization called JPEG Pleno [10] which includes LF, Point Cloud, and the Holography [11]. The proposal provides a 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 in LF lenslet coding and 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 the scene. In [12 - 16], exploiting the LF lenslet compression, most of the conventional image and video coding methods are applied to exploit the existing spatial redundancy of MIs within a raw LF lenslet such as JPEG, JPEG2000 and HEVC intra coding. The idea is based on the concept of self-similarity compensated prediction [12]. To explain, a block-based matching algorithm is utilized to manage the most suitable predictor block for the current block, which compares to the previously coded and reconstructed range of the current image. With two different candidate blocks, the predictor block can be generated. Also, works in [13, 14] propose to add new coding modes to the HEVC coding tools (i.e., locally linear embedding based prediction) and adapt the intra prediction scheme in HEVC coding tools, respectively. In addition, to exploit data geometry for dimensionality reduction of LF, works in [15, 16] present coding schemes for LF based on low rank approximation. Likewise, in [17], the author uses the disparity compensated prediction method to take advantage of the existing spatial redundancy. In addition, the High Order Prediction (HOP) model is also considered as a method to achieve compression 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 is investigated in [19]. The LF lenslet is used with YUV 4:4:4 encoding with 10 bit/sample which performs well in terms of coding efficiency for different color subsampling 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; thus, the work in [21] presents an efficient lenslet image coding model which applies SCC to encode LF lenslet. 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 side, 4D-LF represents the LF data as a stack of sub-aperture images (SAIs), generated from lenslets of LF camera. In the 4D-LF coding approach, generating the 4D-LF pseudo-sequence is a well-known approach for LF compression. This approach is about 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 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 can be 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], 4D-LF pseudo-sequence is 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 presents a solution to exploit information fully 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, the recent work in [27], presents an efficient coding strategy to convert the model parameters into a bitstream, which is well suited for 4D-LF compression.

According to the literature, the LF coding has achieved encouraging results with predictive video coding, i.e., H.264/AVC, H.265/HEVC. However, the conventional predictive video coding paradigm mostly focuses on one-to-many applications which results in complex encoders but simple decoders, it is not suitable with simpler encoders of the emerging applications recently 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 that the motion estimation part at 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, etc. WZ coding provides different coding techniques compared to conventional video coding, notably a flexible distribution of the codec complexity, high compression and inherent error robustness [32].

This type of coding manages to encode separately 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 the video signals in real-world with improved error resilience. In regards to WZ coding with LF images, the several LF image compression approaches have been proposed [35-38]. In particular, the performance of distributed video coding for light field content is analyzed in [39]. In [40], the LF images are 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 SI at the receiver to exploit similarity among LF images. The results present a significant compression performance compared to intra coding while maintaining the random access capabilities. Hence, this is a promising recent coding solution for LF images.

1.2. Contributions and paper organization

In regards to 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, meanwhile 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 content adaptive frame skipping for LF images compression solution by following the distributed coding approach, to achieve efficient compression performance for LF data while providing a low encoding complexity solution. The contributions of this paper can be summarized as:

- An advanced WZ based LF image compression solution: the well-known WZ coding approach is enhanced by improving compression performance at the key frame encoder/decoder with state-of-the-art video compression H.265/HEVC [42] while the advantage of the low-complexity of the WZ procedure is utilized on the side of WZ frame encoder/decoder. Also, an advanced channel codec, i.e. LDPC codec [43] is applied in this WZ coding approach to get capacity approaching performance and flexible code designs using density evolution [44].
- An efficient content-driven LF image reordering mechanism: the different scanning methods may affect different results depending on video contents 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 improve further 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 with 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 will be 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 describes the overview of proposed LF coding architecture. Section 3 presents the novel content adaptive frame skipping algorithm. Afterward, Section 4 analyzes the experimental results, while Section 5 sums up all the conclusions in the paper and describes the future work.

2. Overall Wyner-Ziv based Light Field image compression

This section presents in detail the WZ-LF image compression solution. In order to get the best performance of the solution, an efficient scanning order based on LF content is analyzed, together with the introduction of a content skip algorithm.

2.1. Proposed WZ-LF architecture

To achieve the 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 is strengthened compared to the original WZ architecture from B. Girod [31], by improving compression performance at the key frame encoder/decoder with state-of-the-art video compression codec H.265/HEVC Intra. As shown, the LF image can be processed in the following steps.

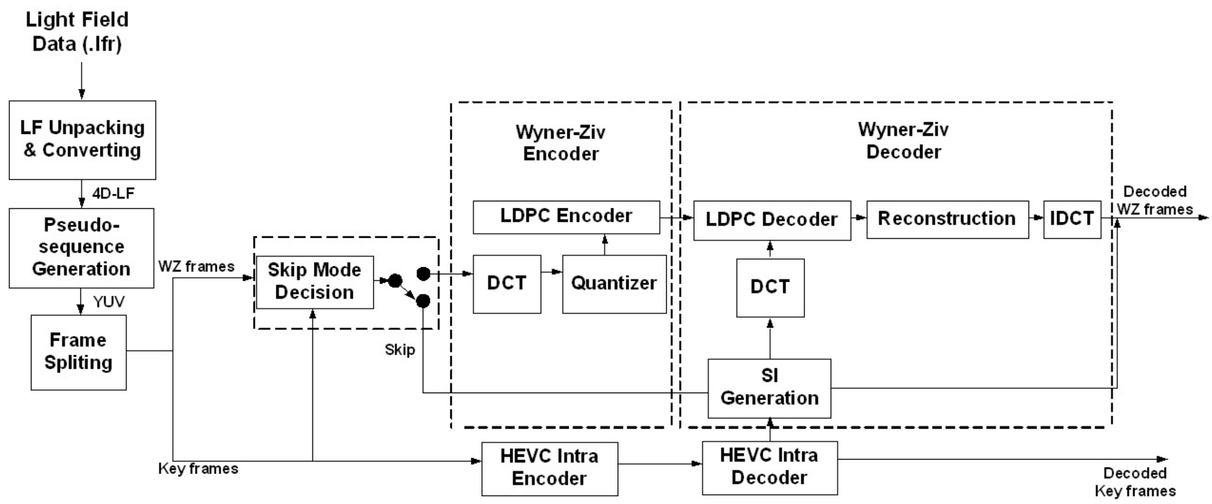


Figure. 1. Proposed WZ-LF image compression architecture

- At the encoder:

The LF data is 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 now 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., 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 will calculate the value of decision nodes based on the features to make the skipping decision. When the skip mode decision is activated, the WZ frames will be skipped in the normal WZ encode and decode procedure and the associated side information will be used as 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 4×4 block for WZ frames. By breaking down the image into 4×4 block of pixels, arranged from left to right and top to bottom, the WZ frames are transformed using a 4×4 DCT. Since the DCT operation has been started, the standard Zig-Zag scan order [46] within the 4×4 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 high frequency information for the remaining bands, respectively.

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 depends on the target quality for the WZ frame. By utilizing a uniform scalar quantizer with higher 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 in the visual quality of decoded image. Similar to [47], 8 different types of quantization matrices were adopted in the proposed LF compression scheme to target various qualities and data rates.

4. Low density parity check (LDPC) encoding:

In this work, to take advantage of 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 and produce 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 frames with a reciprocal relationship between the number of parity bits (or bit-rate) and its quality of noise estimation, i.e., the better quality of estimation, the smaller the bit-rate received. By estimating the correlation between original WZ frame and the SI correctly, the decoding performance can be greatly improved. The better the quality of SI that is interpolated, the better the quality of the final reconstructed WZ frame that can be achieved. Regarding correlation between frames, the 4D-LF pseudo-sequence are a series of frames with high correlation due to the characteristics of the LF image. Thus, achieving the best quality of SI gives a huge advantage in achieving impressive decoding performance.

2. LDPC decoding

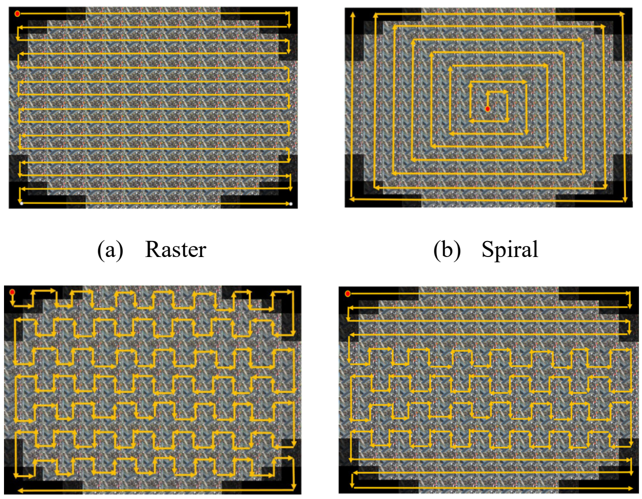
In this part, we describe the decoding of a bit plane given the soft input estimations of the SI and the parity bits transmitted from the encoder. From the decoder, in the case of an increase in the number of parity bits, the decoding procedure is then looped. Additionally, to do an inverse accumulation activity from the encoder, the syndrome bits are removed from the received parity bits before the beginning of the procedure. On these syndrome bits, the operation of Sum-product decoding is performed. These instructions are considered as a soft decision algorithm with the probability of each received bit as input. Additionally, when the decoded bit plane matches the value received from the encoder with the CRC sum registration, it is considered as effective decoding. 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 stream related to each DCT band is formed through all the bit planes related to that band. When all decoded quantized symbols are received, all DCT coefficients are reconstructed with the support of the corresponding SI coefficients and 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 images arrangment

Recently, a scanning order method has been developed based on the optimized reference picture selection for LF image coding with H.265/HEVC low-delay configuration [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 SAIs arrangement, several scan



(a) Raster (b) Spiral

(c) U-shape (d) Hybrid

Fig. 2. Fourth types of scanning order

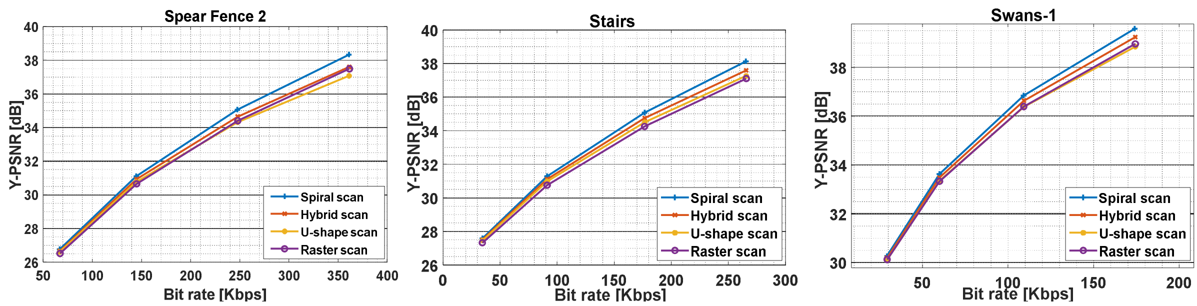


Fig. 3. RD performance between different scanning methods

paths of sub-aperture images as shown in Figure. 2 are examined such as spiral, raster scan, U-shape, and hybrid scanning approaches [26]. Combing the raster and U-shape scanning order, the hybrid scanning order takes the most advantage from the similarity of adjacent views both horizontally and vertically. However, due to varying angles between SAIs, it is observed that the temporal correlation along SAIs may be changed by different scanning orders, also with different LF content. Moreover, the compression performance of the 4D-LF pseudo-sequence can be affected by some specific content. Therefore, in this section we thoroughly evaluate scanning order to verify the most effective scanning order with LF images.

To begin with the content-driven considerations, a set of LF data is collected from [51] with different contents 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).

In regard to scanning methods, the four types of scanning order, 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 the better results than the other scanning methods. Therefore, to exploit the best performance of our proposed WZ-LF coding solution, the spiral scanning method is chosen and the performance is evaluated in detail in the next section.

3. Content adaptive frame skipping algorithm

3.1. Observation

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

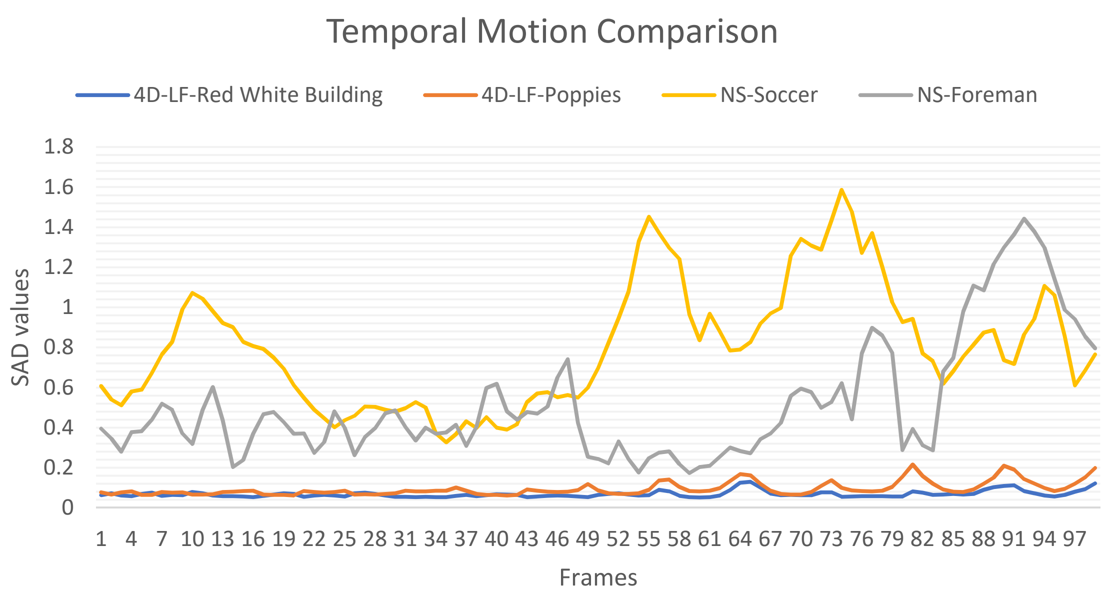


Fig. 4. Motion comparison between 4D-LF pseudo-sequences (4D-LF) and natural sequences (NS)

The different motion of video sequences (i.e., low-motion and high-motion) are considered to affect to the compression performance of the codec. The low-motion and high-motion sequences refer respectively to high correlation and low correlation between each frames. Based on the sequence motion, it is observed that 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 degrades for high-motion sequences (e.g., Soccer, Foreman) [47].

According to the LF characteristic, it is observed that 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 SAD values as shown in Figure. 4

It is noted that the frames of 4D-LF pseudo-sequence are extremely high-correlation as shown in Figure. 4, so, skipping the most similar frames may achieve the 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.

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, the motion activity of the key frames is important to identify. Thus, the two discriminative temporal features are utilized for detecting changes in the motion of the SAI key frames, i.e., FT_{SAI_SAD} - the sum of absolute difference of SAI key frames, and FT_{SAI_GMS} - the similarity of 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:

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

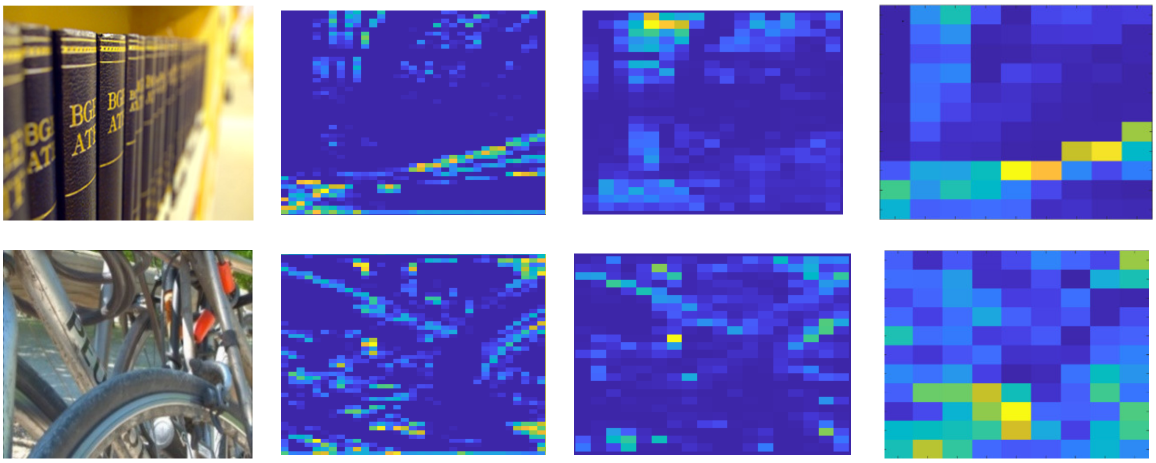
$$GKEY_b(i) = \sqrt{(D_x \otimes KEY_b)^2 + \left(D_y \otimes KEY_b\right)^2} \tag{5}$$

In regard to 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 σ_{WZ}^2 is 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 value of the temporal features covers non-flat regions (i.e., regions with depth, contrast, and saturation complexity).



Original LF images — FT_{SAI_SAD} map — FT_{SAI_GMS} map — FT_{SAI_VAR} map

Figure. 5. The visualization of the spatial and temporal features, from top to down: Books and Bikes

The frame skipping mechanism is based on the technique wherein the texture and motion activity of the two consecutive key frames and neighbor WZ frames of a 4D-LF pseudo-sequence is 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. Based on that, 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

The decision tree based adaptive frame skipping algorithm

Input: 4D-LF pseudo-sequence

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

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

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

Determine the threshold $\boldsymbol{ThD}$ based on the average value of each features.

1. $\boldsymbol{ThD_0} = 0$;
2. **for** t = 0,1,2,…,(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. $FT_{SAI\_VAR} = \sigma^2_{WZ_{t+1}}$
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 $\boldsymbol{W}$

22. $\boldsymbol{W}$**=2**
23. **if** $(FT_{SAI_{SAD}} + FT_{SAI_{GMS}} + FT_{SAI_{VAR}} >= \boldsymbol{W})$
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 weights ($\boldsymbol{W}$)
29. Generate the mode decision

4. Performance evaluation

4.1. Test conditions

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

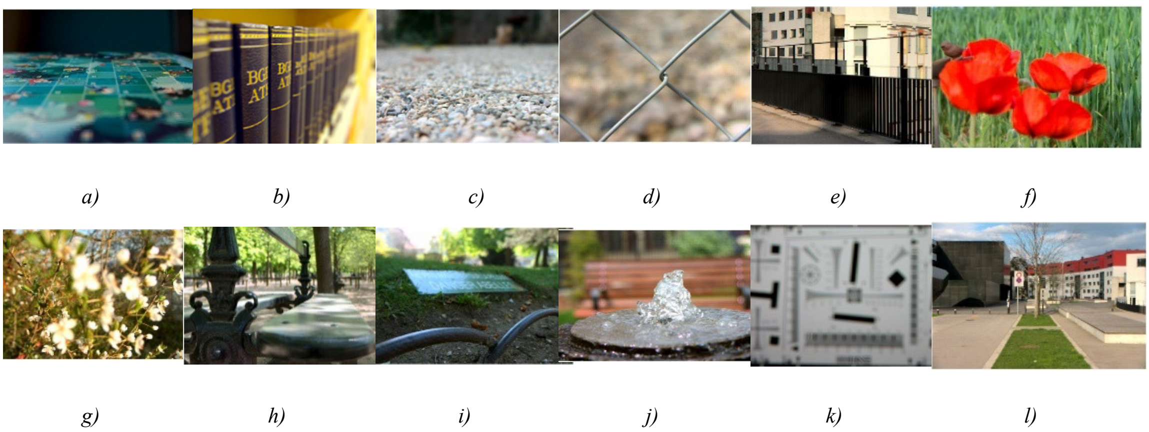
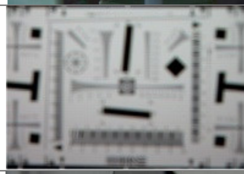


Figure. 6. Thumbnails of Light field images: (a) Game Board, (b) Books, (c) Grave Garden, (d) Chain link fence 1, (e) Black fence, (f) Poppies, (g) Mirabelle Prune Tree, (h) Bench in Paris, (i) Caution Bees, (j) Fountain and bench, (k) ISO chart 13, (l) Red and white building

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	NDOF-1	
Sophie and Vincent 2	People		
Ankylosaurus and Diplodocus 2	Studio		
Bikes	Urban	NDOF-2	
Danger de mort	Grids		
Stone pillars outside	Urban		

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 is 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 tools 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.

4.2. Overall WZ-LF compression performance evaluation

Regarding the compression performance, RD performance is widely utilized for quantifying video coding schemes through use of the Bjøntegaard Delta (BD)-PSNR and Bjøntegaard Delta (BD)-Rate [54]. Figure. 7 presents the RD curve comparison between the proposed WZ-LF coding solution and the other relevant benchmarks, i.e., HEVC inter and intra [42] and HEVC based DVC labelled as DVC-H.265/HEVC [53]; while BD-Rate and BD-PSNR are computed in Table 2. From the observed results, some conclusions can be derived as following:

- WZ-LF versus DVC-H.265/HEVC: Since the proposed codec has a similar approach to the DVC-H.265/HEVC, but with an significant improvement of coding tools. The proposed WZ-LF achieves impressive results for compression performance compared to DVC-H.265/HEVC by reducing the bit-rate in total by about 25%. Also, taking into account the adaptive mode decision

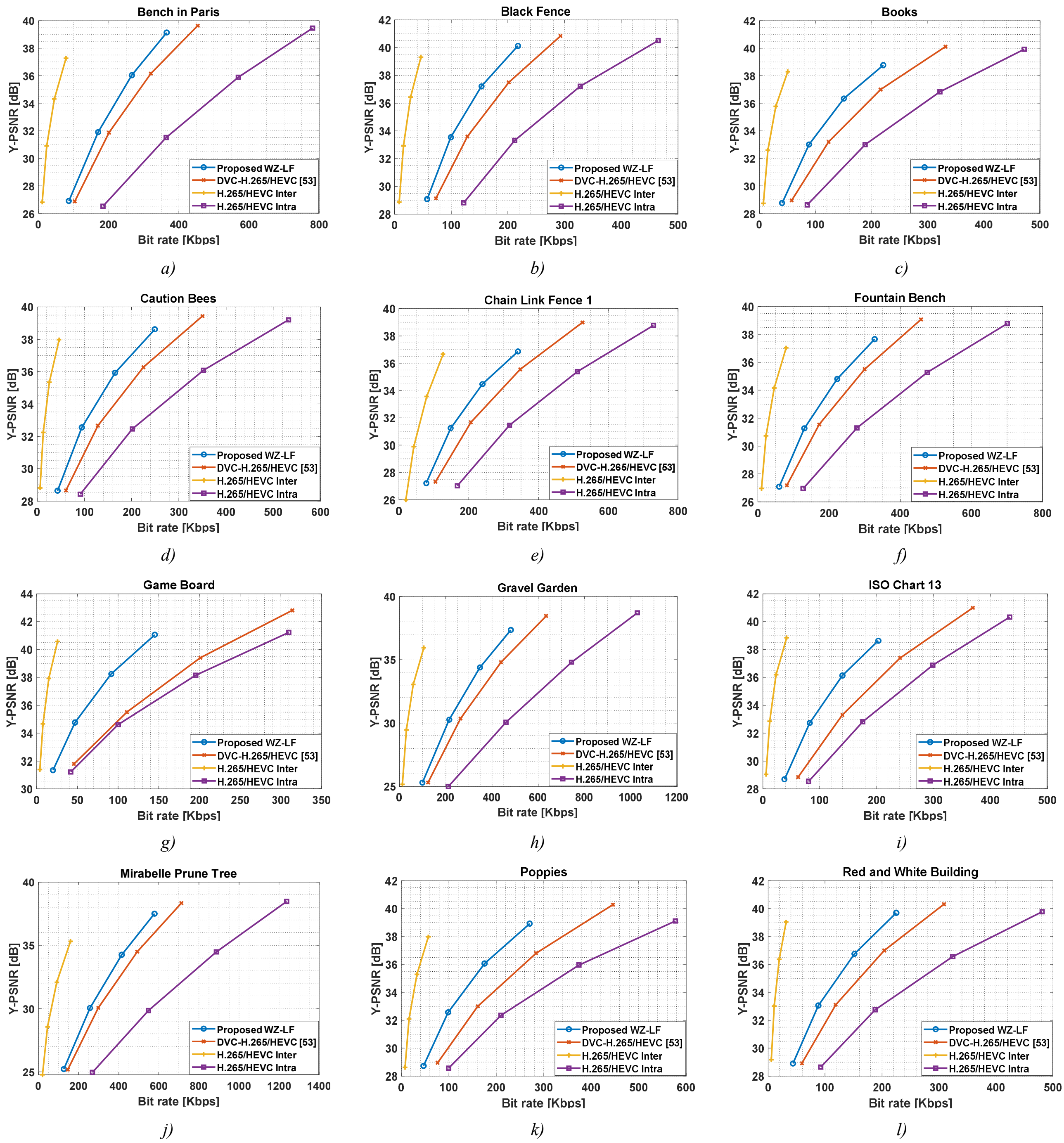


Figure. 7. Overall RD performance evaluation of the proposed WZ-LF coding solution

in frame skipping, the WZ-LF architecture achieves a significant gain of almost 2 dB in compression performance with different content types.

- WZ-LF versus H.265/HEVC Intra: The advantage of proposed coding solution is applied to the intra coding solution at the encoder side and 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 variety of content types. As shown, the content type average BD-Rate reductions of 53.14%, 52.53%, 53.22%, and 53.18% are obtained for the proposed WZ-LF solution with content types of 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. BD-RATE [%] & BD-PSNR [dB] COMPARED TO 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-PSNR	BD-Rate	BD-PSNR	BD-Rate	BD-PSNR	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 Types 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 Types 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	NDOF-1	-24.53	1.67	-53.15	4.50	-84.37	9.03	-88.24	10.57
Fountain and bench		-20.29	1.47	-51.81	4.76	-79.10	8.52	-83.37	9.96
Content Types 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 Types 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 the 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 upper bound of DVC-H.265/HEVC outperforming DVC-H.265/HEVC by 10.9 dB in compression performance, but the proposed WZ-LF improves the compression performance by narrowing the gap about 1.5 dB compared to DVC-H.265/HEVC. Additionally, the major problem of the H.265/HEVC inter codec is high complexity and not it is hence not compatible with the recent emerging applications considered in this work, whereas the proposed WZ-LF can be a suitable solution. Regarding the H.265/HEVC inter 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 representative for 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 due with content types differently compared to the other codecs. By encoding individual frames in H.265/HEVC Intra, the comparison with the WZ-LF solution keeps roughly 53% bit-rate saving for all content types. Notably, in comparison to the DVC-H.265/HEVC, the WZ-LF solution achieves the best compression performance on the NDOF content type with BD-Rate reductions of 29.3% while the other content types such as NDOF-1 and NDOF-2 are improved respectively by only about 20.2% and 23.8%. The WDOF and BC contents represent the less motion and high correlation between SAIs of 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

TABLE 3. BD-RATE [%] & BD-PSNR [dB] COMPARED BETWEEN HEVC INTER AND HEVC INTER 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 Types 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 Types 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 Types 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 Types Average		**-1.15**	**0.07**
Total Average		***-3.20***	***0.18***

contents, i.e., Game Board, Poppies, Chain link fence 1, and Books, also vary in bit-rate saving and compression performance compared to DVC-H.265/HEVC and H.265/HEVC Intra. For instance, the proposed WZ-LF solution can save respectively 48.2% and 53.9% bit-rate in comparison to DVC-H.265/HEVC and H.265/HEVC Intra for the Game Board sequence.

4.3. WZ-LF codec with various coding tools

4.3.1. Scanning method assessment

Four types of scanning order, i.e., spiral, hybrid, U-shape scanning, and raster, are evaluated based on BD-Rate [54]. Regarding the common scanning method of 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 of the scanning methods are shown in Table 4. Hybrid and U-shape scanning order saves about 3% bit-rate compared to raster scan for most content types, however, notably in NDOF and NDOF-1 types, the BD-rate performance changes compared to the raster scan with respectively about 3% to 24% bit-rate saving. Regarding the spiral scan, this method achieves an outstanding result with about 10% bitrate saving on the average for all data types compared to the raster scan. In particular for the NDOF and NDOF-1 data types, this method still achieves an impressive performance by saving about 11% and 9% bitrate, 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 ON DIFFERENT CONTENTS

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

4.3.2.1. 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 compression efficiency of the WZ-LF coding solution. The spatial-temporal features of the 4D-LF pseudo-sequences are selected based on the changes of depth of field 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 encode as a normal 4D-LF pseudo-sequences.

TABLE 5. BD-RATE [%] & 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 Types 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 Types Average		**-29.31**	**2.27**
Bench in Paris	NDOF-1	-15.88	1.48
Caution Bees		-24.53	1.67
Fountain and bench		-20.29	1.47
Content Types Average		**-20.23**	**1.54**
Poppies	NDOF-2	-32.92	2.45
Mirabelle Prune Tree		-13.52	1.18
Books		-25.03	1.79
Content Types Average		**-23.82**	**1.81**
Total Average		***-24.99***	***1.95***

Table 5 and Figure. 8 show the comparison of the WZ-LF coding solution with and without the skip mode decision. Examining the RD performance results as shown, it is clear that WZ-LF with skip mode achieves lower complexity than WZ-LF without skip mode with a bitrate saving of 25%. Notably, the NDOF content type shows a significant improvement for a bit-rate saving of 29.3%, while WDOF, NDOF-1, and NDOF-2 come with a BD-Rate reduction of 26.6%, 20.2%, and 23.8%, respectively. Therefore, we can observe that the skip mode performs outstandingly with burry content (34% bit-rate saving) or content having a narrow depth of field (48% bit-rate saving for the Game Board sequences).

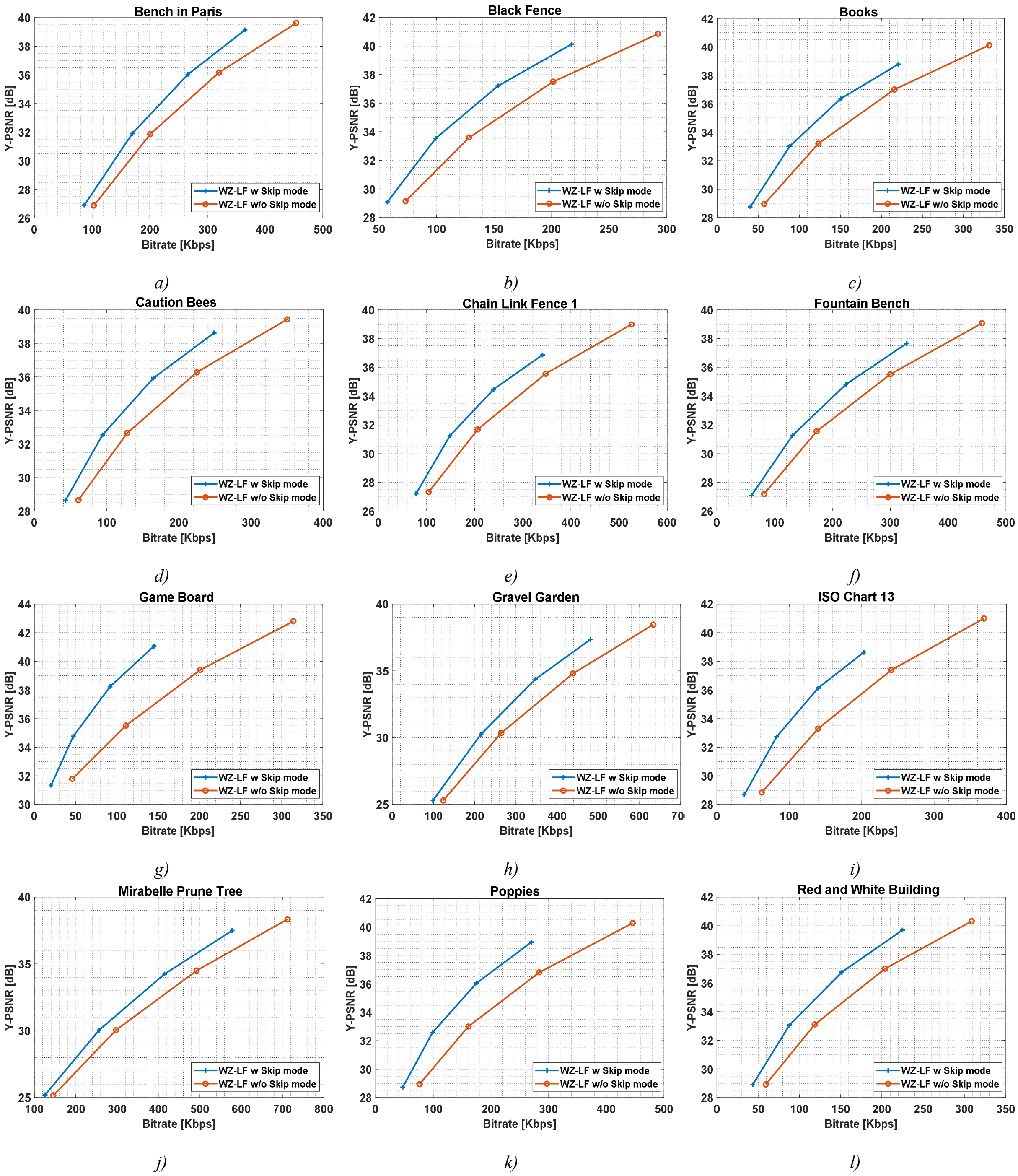


Figure. 8. RD performance evaluation of WZ-LF with skip and non-skip mode decision

4.3.2.2. Compression complexity

Examining compression complexity is an essential part of performance evaluation. For this evaluation, the coding solutions are tested on the same PC with an Intel Core i7-7700HQ (2.8 GHz) processor, 16GB RAM, and Windows 10-Home OS. The results are shown in Fig. 9 and Fig. 10, respectively for QP 40 and QP 25, with and without 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 respectively processing time of the WZ-LF codec with and without skip mode decision.

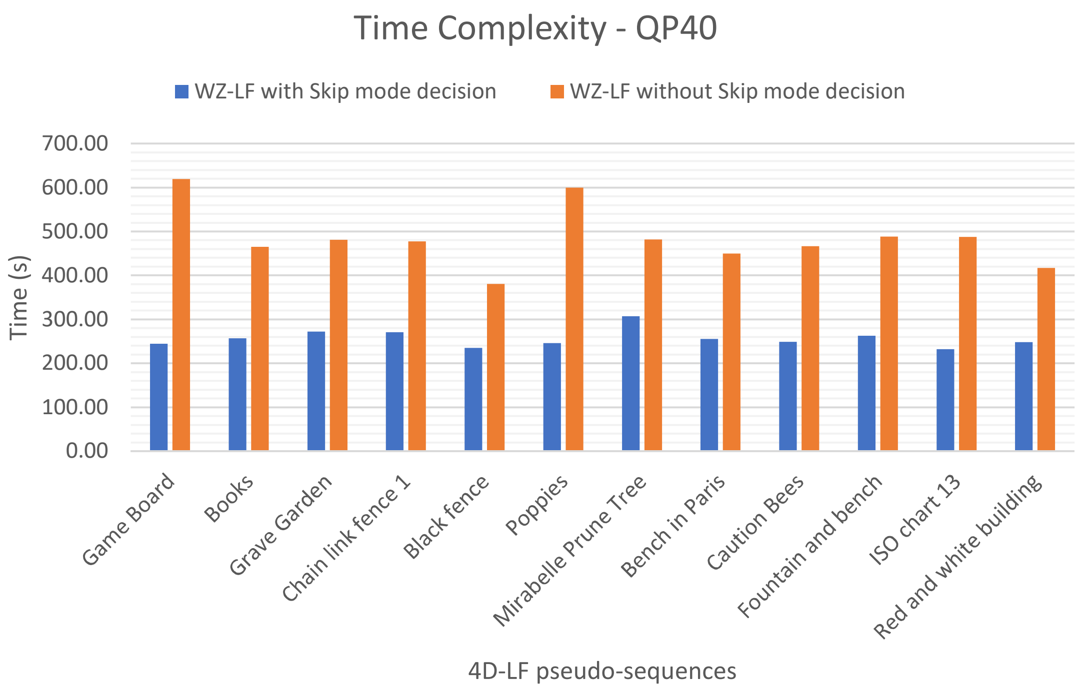


Figure. 9. Time complexity of the codec with skip and without skip mode at QP40

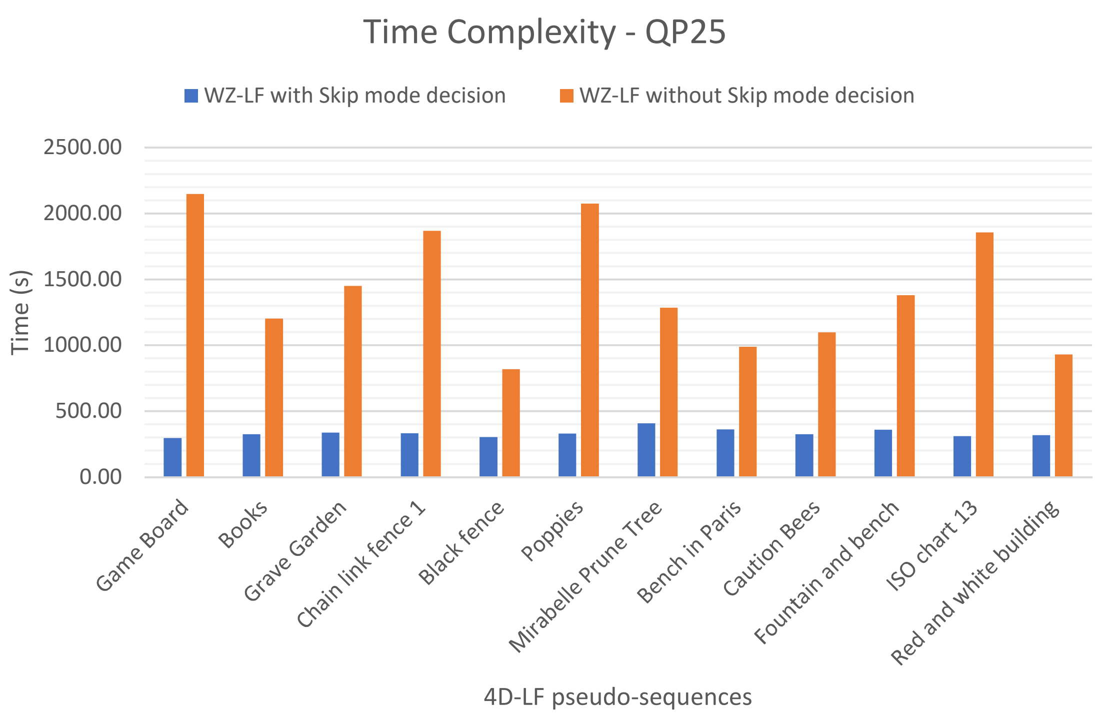


Figure. 10. Time complexity of the codec with skip and without skip mode at QP25

From these complexity results, it can be observed that the WZ-LF codec with skip mode decision is significantly time-saving for encoding compared to the WZ-LF codec without skip mode decision.

The WZ-LF with skip mode can encode faster on average about 46% and 74% than the WZ-LF without skip mode at QP 40 and QP 25, respectively.

5. Conclusion

This paper introduces a LF content adaptive 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, which adapts to LF content. This provides an optimized performance on 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 codec. In particular, the proposed coding solution improves compression performance, and retains a 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, or depth map, can be analyzed in order to provide a better quality of LF reconstruction. Thus, the proposed WZ-LF coding solution may further be improved.

Author Contributions: Conceptualization, Xiem HoangVan and Huy Phi Cong; methodology, Stuart Perry; software, Huy Phi Cong.; validation, Xiem HoangVan, Huy Phi Cong and Stuart Perry; formal analysis, Xiem HoangVan.; investigation, Huy Phi Cong; resources, Stuart Perry; data curation, Stuart Perry; writing—original draft preparation, Huy Phi Cong; writing—review and editing, Xiem Hoang Van.; visualization, Huy Phi Cong; supervision, Xiem HoangVan and Stuart Perry; 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 and by the Joint Technology and Innovation Research Centre - a partnership between University of Technology Sydney and Vietnam National University.

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

References

1. Levoy. M and Hanrahan. P, Light field rendering, in SIGGRAPH'96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, pp. 31–42, Aug. 1996.
2. Gortler. S. J, Grzeszczuk. R, Szeliski. R, and Cohen. M. F, The lumigraph, in SIGGRAPH'96: Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, pp. 43-54, Aug. 1996.
3. Wu. G, et al., Light Field Image Processing: An Overview, in *IEEE Journal of Selected Topics in Signal Processing*, **2017**, vol. 11, no. 7, pp. 926-954.
4. Adelson. E. H and Wang. J. Y. A, Single Lens Stereo with a Plenoptic Camera, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **1992**, vol. 14, no. 2, pp. 99–106.
5. Lytro camera, https://www.lytro.com/
6. Raytrix, https://www.raytrix.de/
7. Xiao. X, Javidi. B, Martinez-Corral. M, and Stern. A, Advances in three-dimensional integral imaging: Sensing display and applications, *Applied Optics*., **2013**, vol. 52, no. 4, pp. 546-560.
8. Raghavendra. R, Raja. K . B, and Busch. C, Presentation attack detection for face recognition using light field camera, *IEEE Transactions on Image Processing*., **2015,** vol. 24, no. 3, pp. 1060-75.
9. Levoy. M, The digital michelangelo project: 3D scanning of large statues, in SIGGRAPH'00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques, pp. 131-144, Aug. 2000.
10. JPEG pleno call for proposals on light field coding, *Proc. ISO/IEC JTC 1/SC29/WG1 N74014*, **2017** Switzerland, pp. 1-37.
11. JPEG Pleno Holography Uses Cases and Requirements, *Proc. ISO/IEC JTC 1/SC29/WG1 N86016*, **2020** Australia, pp. 1-9.

12. Conti. C, Nunes. P, and Soares. L. D, HEVC-based light field image coding with bi-predicted self-similarity compensation, in IEEE International Conference on Multimedia Expo Workshops, Seattle, WA, USA, pp. 1–4, Jul. 2016.
13. Monteiro. R, Lucas. L. F. R, Conti. C, and Nunes. P, Light field HEVC-based image coding using locally linear embedding and self-similarity compensated prediction, in IEEE International Conference on Multimedia Expo Workshops, pp. 1–4, Jul. 2016.
14. Li. Y, Sjostrom. M, Olsson. R, and Jennehag. U, Efficient intra prediction scheme for light field image compression, IEEE International Conference on Acoustics Speech and Signal Processing, Florence, Italy, pp. 539-543, May. 2014.
15. Jiang. X, Le Pendu. M, Farrugia. R. A, Hemami. S. S, and Guillemot. C, Homography-based low rank approximation of light fields for compression, IEEE International Conference on Acoustics Speech and Signal Processing, New Orleans, LA, USA, pp. 1313-1317, Mar. 2017
16. Kamal. M. H and Vandergheynst. P, Joint low-rank and sparse light field modelling for dense multiview data compression, IEEE International Conference on Acoustics Speech and Signal Processing, Vancouver, BC, Canada, pp. 3831-3835, May. 2013.
17. Chang. C.-L, Zhu. X, Ramanathan. P, and Girod. B, Light field compression using disparity-compensated lifting and shape adaptation, *IEEE Transactions on Image Processing*, **2006,** vol. 15, pp. 793–806.
18. Monteiro. R. J. S, Nunes. P. J. L, Rodrigues. N. M. M, and Faria. S. M. M, Light field image coding using high-order intra block prediction, *IEEE Journal of Selected Topics in Signal Processing*, **2017**, vol. 11, no. 7, pp. 1120–1131.
19. Monterio. R. J. S, Rodeigues. N. N. M. M, Faria. S. M. M, and Nunes. P. J. L, Light field image coding: objective performance assessment of Lenslet and 4D LF data representations, Proceedings of SPIE Optical Engineering, Applications of Digital Image Processing XLI, vol. 107520D, San Diego, California, USA, Sept. 2018
20. Peng. W. H, et al, Overview of Screen Content Video Coding: Technologies, Standards, and Beyond, in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, **20116**, vol. 6, no. 4, pp. 393-408.
21. Tsang. S. H, Chan. Y. L, and Kuang. W, Standard compliant light field lenslet image coding model using enhanced screen content coding framework, *Journal of Electronic Imaging*, **2019** vol. 28, no. 5, 053027 (23 October 2019)
22. Fecker. U and Kaup. A, H.264/AVC-compatible coding of dynamic light fields using transposed picture ordering,, 13th European Signal Processing Conference (EUSIPCO), Antalya, Turkey, Sep. 2005.
23. Vieira. A, Duarte. H, Perra. C, Tavora. L and Assuncao. P, Data formats for high efficiency coding of Lytro-Illum light fields, International Conference on Image Processing Theory, Tools and Applications (IPTA), Orleans, USA, 2015
24. Liu. D, et al, Pseudo-sequence-based light field image compression, IEEE International Conference on Multimedia & Expo Workshops, Seattle, WA, 2016.
25. Li. L, et al, Pseudo Sequence Based 2-D Hierarchical Coding Structure for Light-Field Image Compression, 2017 Data Compression Conference, Snowbird, UT, 2017.
26. Zhao. S, et al, Light field image coding with hybrid scan order, in SPIE Visual Communications and Image Processing, Chengdu, 2016.
27. Verhack. R, Sikora. T, Wallendael. G. V, and Lambert. P, Steered Mixture-of-Experts for Light Field Images and Video: Representation and Coding, in *IEEE Transactions on Multimedia*, **2020,** vol. 22, no. 3.
28. Belyaev. E, Egiazarian. K and Gabbouj. M, A Low-Complexity Bit-Plane Entropy Coding and Rate Control for 3-D DWT Based Video Coding, in *IEEE Transactions on Multimedia*, **2013,** vol. 15, no. 8, pp. 1786-1799.
29. Belyaev. E, Compressive Sensed Video Coding Having Jpeg Compatibility, 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 2020, pp. 1128-1132.
30. Wyner. A and Ziv. J, The Rate-Distortion Function for Source Coding with Side Information at the Decoder, *IEEE Transactions on Information Theory*, **1976,** vol. 22, no. 1, pp. 1-10.
31. Girod. B, Aaron. A, and Rebollo-Monedero. D, Distributed video coding, in *Proceedings of the IEEE (Special Issue on Advances in Video Coding and Delivery)*, **2005,** vol. 93, no. 1, pp. 71-83.
32. Pereira. F, et al., Distributed video coding: selecting the most promising application scenarios, *Signal Processing Image Communication*, **2008**, vol. 23, no. 5, pp. 339–352.
33. Aaron. A, Zhang. R, and Girod. B, Wyner-Ziv coding of motion video, Conference Record of the Thirty-Sixth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 2002.

34. Aaron. A, Rane. S, Zhang. R, and Girod. B, Wyner-Ziv coding for video: applications to compression and error resilience, Data Compression Conference, 2003. Proceedings. DCC 2003, Snowbird, UT, USA, 2003.
35. Zhu. X, Aaron,. A, and Girod. B, Distributed compression for large camera arrays, in IEEE Workshop on Statistical Signal Processing (SSP '03), St. Louis, Mo, USA, Sept. 2003.
36. Jagmohan. A, Sehgal. A, and Ahuja. N, Compression of light field rendered images using coset codes, in 37th Asilomar Conference on Signals, Systems, and Computers: Special Session on Distributed Coding, Pacific Grove, CA, USA, Nov. 2003.
37. Toffetti. G, et al, Image compression in a multi-camera system based on a distributed source coding approach, in EUSIPCO '05, Antalya, Turkey, Sept. 2005.
38. Yeo. C and Ramchandran. K, Robust distributed multi-view video compression for wireless camera networks, in SPIE Visual Communications and Image Processing, San Jose, CA, USA, Jan. 2007.
39. Zhu. X, Aaron. A, and Girod. B, Distributed compression of light fields, Stanford University, Stanford, CA, USA, Tech. Rep.
40. Aaron. A, Ramanathan. P, and Girod. B, Wyner-Ziv coding of light fields for random access, IEEE 6th Workshop on Multimedia Signal Processing, 2004, Siena, Italy, 2004
41. Cong. H. P, HoangVan. X, and Perry. S, A low complexity Wyner-Ziv coding solution for Light Field image transmission and storage, IEEE International Symposium on Broadband Multimedia Systems and Broadcasting, Jeju, Korea, Jun. 2019.
42. Sullivan. G. J, Ohm. J. R, Han. W. J, and Wiegand. T, Overview of the High Efficiency Video Coding (HEVC) Standard, in *IEEE Transactions on Circuits and Systems for Video Technology*, **2012,** vol. 22, no. 12, pp. 1649-1668.
43. MacKay. D, Good error-correcting codes based on very sparse matrices, *IEEE Transactions on Information Theory*, **1999,** vol. 45, pp. 399-431.
44. Richardson. T, Shokrollahi. M, Urbanke. R, Design of capacity-approaching irregular low-density parity-check codes, *IEEE Transactions on Information Theory*, **2001,** vol. 47, pp. 619-637.
45. Quinlan. J, Induction of decision trees, *Machine Learning*, **1986,** vol. 1, no. 1, pp. 81–106. [Online]. Available: http://dx.doi.org/10.1023/A:1022643204877
46. ITU-T Rec. H.264 (11/2007) "Advanced video coding for generic audio visual services"
47. Artigas. X, et al, The discover codec: architecture, techniques and evaluation, in Proceedings of Picture Coding Symposium, Lisboa, Portugal, Nov. 2007.
48. Varodayan. D, Aaron. A and Girod. B, Rate-Adaptive Codes for Distributed Source Coding, *EURASIP Signal Processing Journal, Special Section on Distributed Source Coding*, **2006**, vol. 86, no. 11.
49. Ascenso. J, Brites. C, and Pereira. F, Content Adaptive Wyner-ZIV Video Coding Driven by Motion Activity, 2006 International Conference on Image Processing, Atlanta, GA, 2006.
50. Ricardo Monteiro. J. S, Nuno Rodrigues. M. M, Sérgio Faria. M. M and Paulo Nunes. J. L, Optimized Reference Picture Selection for Light Field Image Coding, 2019 27th European Signal Processing Conference (EUSIPCO), A Coruna, Spain, 2019, pp. 1-5
51. Řeřábek. M and Ebrahimi. T, New Light Field Image Dataset, 8th International Conference on Quality of Multimedia Experience, Lisbon, Portugal, 2016.
52. Xue. W, Zhang. L, Mou. X, and Bovik. A. C, Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index, *IEEE Transactions on Image Processing*, **2014**, vol. 23, no. 2, pp. 684-695.
53. Brites. C and Pereira. F, Distributed Video Coding: Assessing the HEVC upgrade, *Signal Processing: Image Communication*, **2015**, vol. 32, no. 3, pp. 81-105.
54. Bjøntegaard. G, Calculation of average PSNR differences between RD-curves, Doc. ITU-T SG16 VCEG-M33, Austin, TX, USA, Apr. 2001