

# Light Field Image Coding with Image Prediction in Redundancy

**A. Sathesh,**

Department of EEE,  
Eritrea Institute of Technology,  
Eritrea.  
Email: sathesh4you@gmail.com

**Abstract:** The proposed work involves a hybrid data representation using efficient light field coding. The existing light field coding solution are implemented using sub-aperture or micro-images. However, the full capacity in terms of intrinsic redundancy in light field images is not completely explored. This paper represents a hybrid data representation which explores four major redundancy types. Using coding block, the most predominant redundancy is exploited to find the optimum coding solution that provides maximum flexibility. To show how efficient the hybrid representation works, we have proposed a combination of pseudo-video sequence coding approach with pixel prediction methods. The observed experimental results shows a positive bit rate saving when compared to other similar methods. Similarly, the proposed method is also said to outperform other coding algorithms such as WaSP and MuLE when compared on a HEVC-based benchmark.

**Keywords:** HEVC; spatial pixel prediction; light field representation; Light field coding;

## 1.Introduction

With the help of lenslet LF (light field) camera, light field imaging technology is able to capture both the angular information and scene radiance of an image using a wider baseline or a narrow baseline [1]. A typical lenslet LF camera comprises of sensor and lens along with microlens array (MLA). The inclusion of MLA gives the camera the ability to capture both angular as well as spatial information when light reaches the sensors. Based on the LF camera, various degrees of freedom are used and the information captured in then converted into 3D information, instead of the typical 2D information captured by other simpler cameras. Based on the angular information gathered using the LF camera, it is possible to develop several a posterior image processing manipulations like refocusing and changing the perspective of the image. LF is also used in richer content capturing technology which has found its roots in numerous applications such as 3D [2] television, medical imaging and image recognition with the help of views from 3D and 2D perspective to create multiview (MV) signal [3]. This gives way to the introduction of many interactive media application like interactive streaming, free viewpoint video streaming and MV videos that are captured using lenslet LF camera or HDCA. In the recent years, the use of LF technology has been researched by many researchers as well as standardization bodies like MPEG and JPEG [4] due to its interesting features and the need for normalization of the LF data. There is also need for coding solutions that gives the image compatibility with displays including features such as viewpoint access as well as scalability. Hence there is much involvement among these professional bodies to develop a coding standard for the new technological inventions such as holographic, point cloud, LR [5-6] etc. Due to their advancement in technologies, the standard 2D image and video coding algorithms [7] find it difficult to adapt to these features and are found to lack efficiency in coding. The inefficient [8] exploitation of inter and intra MI redundancy that is seen in the LF content is noted to be a limitation of the proposed work [9-10]. There are three alternative ways which are found to be even more efficient than the straightforward image and video codecs:

1. Use post and pre-processing tools used to transform the LF image into pseudo-video sequence and further use a video codec to encode it.
2. The image codec uses novel prediction tolls that have the capability to make use of the MI redundancies.
3. Develop other coding approaches that are built for LF images.

For these approaches, some methodologies have been proposed and it is identified that the common denominator between them deals with LF data notation. These methodologies will depend on sub-aperture images (SAI) [11] or MI. Here the pixels are extracted from the image in a fixed position [12] and the view is generated to place them in a matrix. However, redundancy of the coding is limited thereby affecting the efficiency when these two techniques are chosen. In this paper we have proposed a new methodology which combines both MI and SAI representation resulting in a hybrid LF data representation [13]. This approach will ensure the use of the following redundancies:

- Inter-MI between neighbouring MI
- Intra-MI within each MI
- Inter-view redundancy between SAIs
- Intra spatial redundancy within each SAI

Using the proposed hybrid methodology, both wide and narrow baseline LF images are analyzed and it is found that the efficiency has improved significantly.

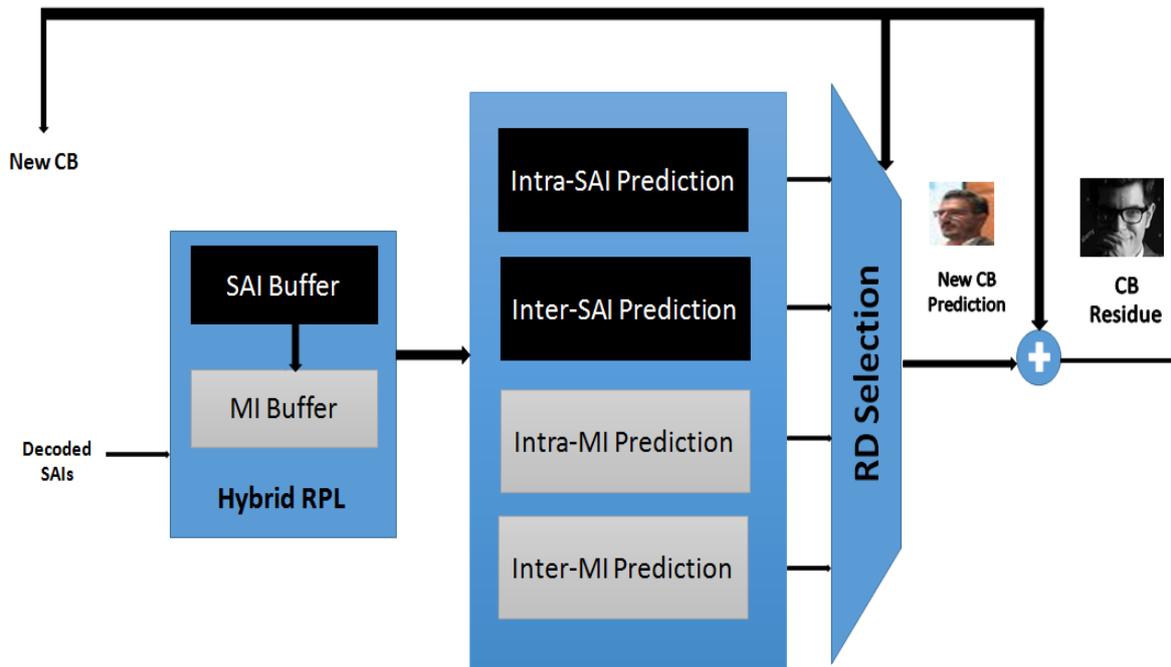
## 2. Combining LF Data Representation

SAIs and MIs are used to represent the lenslet LF data. It is possible to represent the given LF raw image as a 4D LF image [14] by converting it. During conversion, four dimension are used to represent the 4D LF image as LF (h; v; x; y) which consists of SAIs stack that is obtained from the raw image. In the dimensions, the first two indicate SAI location using vertical as well as horizontal coordinates while the other two indicate the spatial positions [15-25]. The primary use of 4D LF image is that it make SAI and MI conversions easy and reversible. On converting the LF image to its corresponding 4D LF image, it is also possible to convert it back to its raw LF image format such that it can be sampled using SAI and MI techniques. Using 4D LF, the data can be called as PVS or MI. Here, the SAIs are scanned strategically to form the PVS.

Both PVS and MI representation of LF data have their advantages:

- PVS-based representation: In each SAI, this will give rise to better exploitation of inter-view as well as spatial redundancy.
- MI-based representation: This representation ensures that both spatial and inter-view redundancy takes place in a more efficient manner.

Since coding solutions will generally depend on either PVS or MI, it is not possible to completely exploit the different redundancies which can be used in LF data. However, when hybrid LF data is used, it is possible to overcome this set-back.



**Fig.1. LF Prediction Module**

### 3. Hybrid Data Representation

In this paper we take advantage of the reversible and seamless conversion between MI and PVS representation such that it is possible to incorporate more prediction modes. Fig.1 represents the integration of the proposed hybrid LF data representation to be used as an encoder. Further, SAI is segmented into coding blocks (CB) that are sent to the prediction module where every coding block is equipped with a specific prediction block to decrease the cost expenditure. The difference between the new input CB and the generated prediction block will give the resultant CB residue. This block will initially be transformed and then entropy coded and quantized along with the signalling data. It can be observed in Fig.1.that there are 2 decoded buffers: new MI buffer (which uses MI representation to save the LF image) and the SAI buffer (HEVC decoded picture buffer). HRPL (Hybrid Reference Picture List) is a combination of both the picture buffers and uses the redundancy through the prediction types in data representation models as given below:

- Intra- and Inter- MI prediction: This mode uses MI decoder picture buffer to use the intra-MI redundancy and inter-MI redundancy.
- Intra-SAI prediction: It represents the prediction modes of the intra-picture such as Directional, Planar and DC mode in HEVC using SAIs' spatial redundancy.
- Inter-SAI prediction: This represents the prediction modes of inter-pictures like Skip, merge and motion compensation modes in HEVC using SAIs' inter-view redundancy.

### 3.1 MI Decoded Picture Buffer:

In this decoder, PVS data representation is used for the output and input data using LF coding. Hence SAI decoder is used to generate the complete LF image with the correspondence between MI and PVS data.  $(i,j)$  is the MI image's pixel position and they can be further defined as

$$\begin{pmatrix} i \\ j \end{pmatrix}_{MI} = \begin{pmatrix} h(f_n) + xM_w \\ v(f_n) + yM_h \end{pmatrix}_{MI}$$

where  $M_h$  and  $M_w$  represent the height and width of MI while  $f_n$  denotes the PVS frame number. Here the coordinates  $(h,v)$  and  $(x,y)$  index the SAI position within the LF and pixel position within each SAI. Here every MI image is build using 9 decoded SAIs. Using the above equation, the pixels from  $(x,y,f_n)$  to  $(i,j)$  in MI representation. Spiral scan is used in this proposed paper as it increases the scalability. Moreover, the proposed methodology can be implemented in any scanning order. The process of converting the image takes place in a progressive manner and the resultant image is that of a sparse LF image. Instead of blockwise prediction, both inter MI and intra MI prediction modes are done in a pixelwise manner, similar to that of PVS representation.

### 3.2 Choosing the Prediction Mode

Between the inter-SAI and inter-SAI prediction modes, the HEVC encoder will choose the right mode by generating a predictive block which has low Lagrangian RD cost such that

$$J = D + \lambda R$$

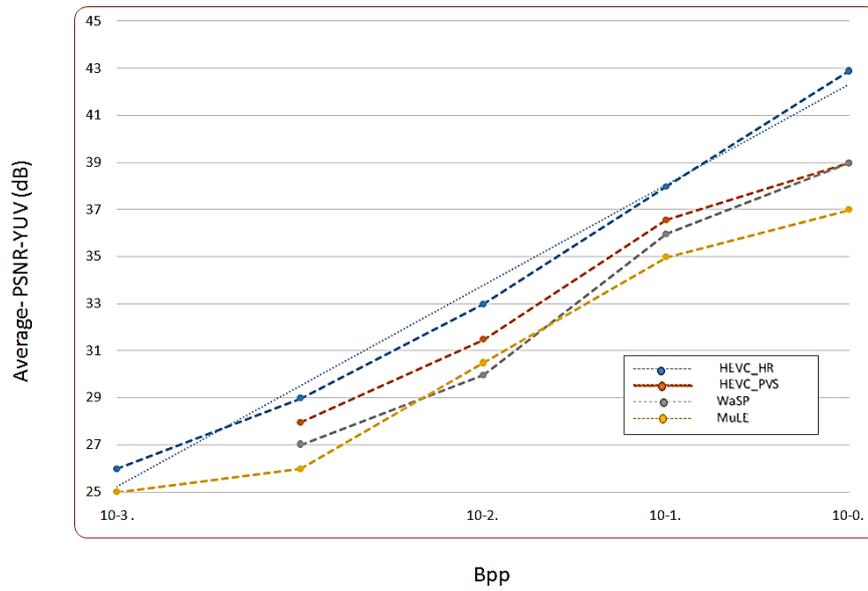
Here  $R$  is the rate representing the total bits that are needed to signal the mode towards the decoder,  $D$  is the distortion that is computed by finding the difference between the prediction block and the original block. In the proposed work, a similar process is provided to the other prediction modes that can be implemented in MI data representation.

### 3.3 Prediction Mode Signalling

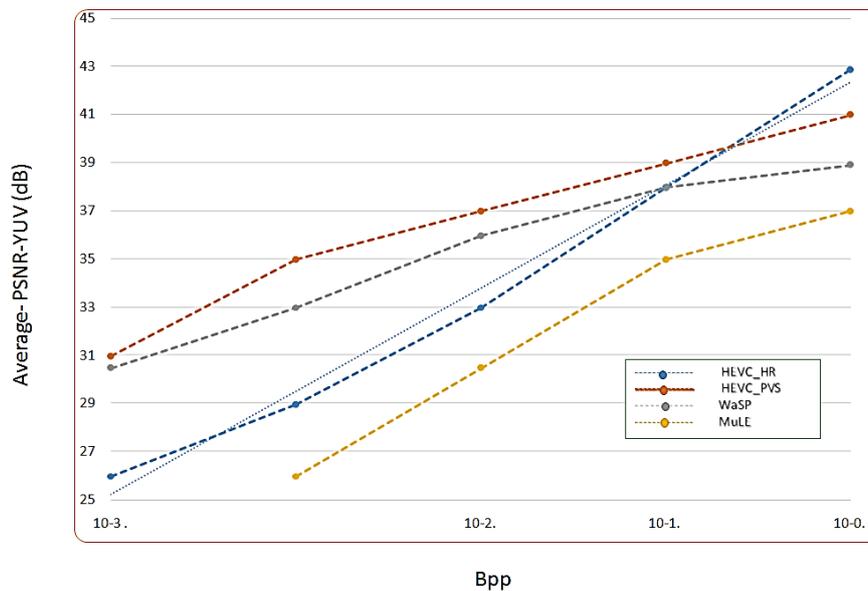
In this paper, we have used the redundancies of four different types using different prediction modes and it is important to exploit the various modes of prediction in an efficient model. The HEVC standard are set at as the standard inter and intra modes. The proposed work also makes use of a similar signalling logic.

## 4. Results and Discussion

Fig.2 (a) and Fig.2 (b) represents the LF image codecs for two images in encoding process. From the experimental results it is observed that the HEVC-HR is found to be more efficient when compared with HEVC-HR. It is also found that the inter-SAI prediction modes are used for encoding the pixels as they have higher inter-view redundancy for similar content of LF.



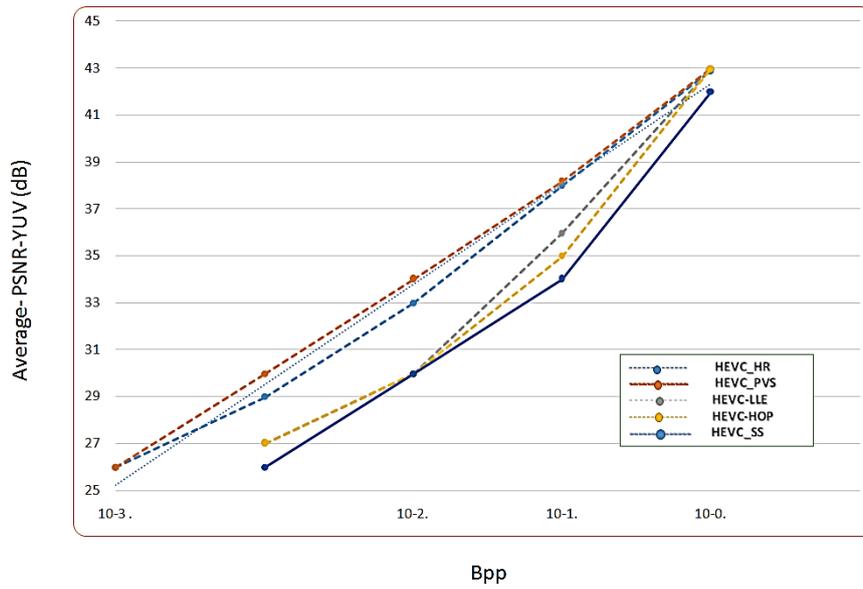
(a)



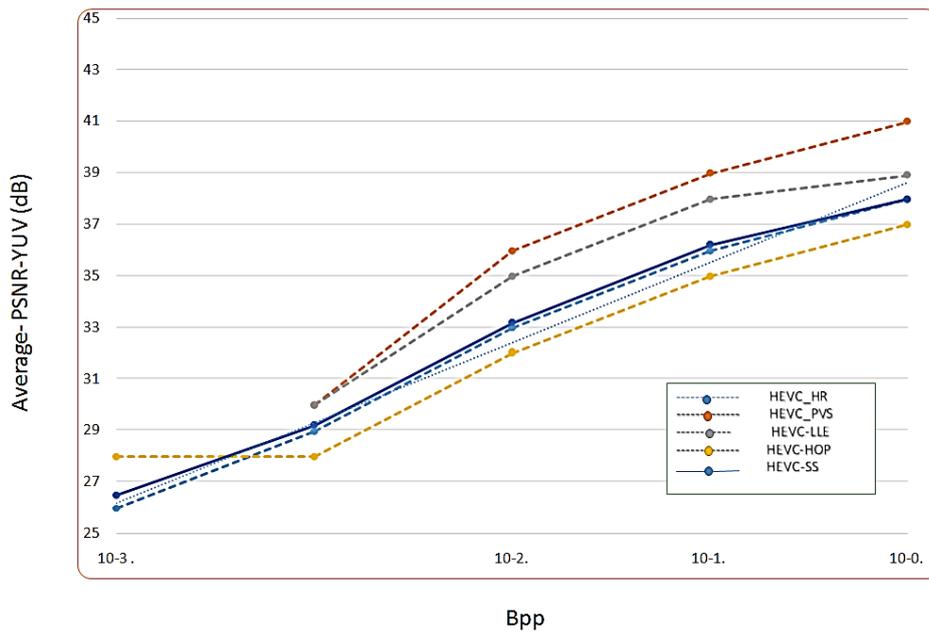
(b)

**Fig.2 (a) and (b) RD curve comparing HEVC PVS and HEVC HR for two different images**

A similar RD curve that compares the coding approaches can be observed in Fig.3 (a) and Fig.3 (b). It can be observed that the performance of PVS outperforms that of MI as the SAIs' inter-view redundancy is very large and can be utilized by the HEVC prediction tools.



(a)



(b)

Fig.3 (a) and (b) RD performance of the proposed Coding Approach for two different images

## 5. Conclusion

We have introduced a novel hybrid LF data representation using LF coding in this paper. The LF images are compressed effectively using a codec (HEVC) along with prediction modes to determine the pixel. The proposed methodology uses both SAI and MI based representation for compliance with the prediction mode. An additional performance evaluation for the proposed HEVC-HR is made with HEVC-PVS which shows that the proposed method has better bit rate saving when compared with the latter. Future work for this methodology includes addition of SAI scalability as it will result in more efficient LF navigation and improved compatibility.

## References

- [1] Abbott, L. F., Rolls, E. T., & Tovee, M. J. (1996). Representational capacity of face coding in monkeys. *Cerebral Cortex*, 6(3), 498-505.
- [2] Lupiani, B., Lee, L. F., & Reddy, S. M. (2001). Protein-coding content of the sequence of Marek's disease virus serotype 1. In *Marek's disease* (pp. 159-190). Springer, Berlin, Heidelberg.
- [3] Smys, S., João Manuel RS Tavares, Valentina Emilia Balas, and Abdullah M. Iiyasu, eds. *Computational Vision and Bio-Inspired Computing: ICCVBIC 2019*. Vol. 1108. Springer Nature, 2020.
- [4] Suma, V. (2019). Security and privacy mechanism using blockchain. *Journal of Ubiquitous Computing and Communication Technologies (UCCT)*, 1(01), 45-54.
- [5] Champel, M. L., Huguenin, K., Kermarrec, A. M., & Le Scouarnec, N. (2010, June). LT network codes. In *2010 IEEE 30th International Conference on Distributed Computing Systems* (pp. 536-546). IEEE.
- [6] Prajwalasimha, S. N., S. R. Kavya, A. S. Hema, and H. C. Anusha. "Modified Gingerbreadman Chaotic Substitution and Transformation Based Image Encryption." In *International Conference On Computational Vision and Bio Inspired Computing*, pp. 606-614. Springer, Cham, 2019.
- [7] Jeon, S. Y., Ahn, J. H., & Lee, T. J. (2016). Reliable broadcast using limited LT coding in wireless networks. *IEEE Communications Letters*, 20(6), 1187-1190.
- [8] Yuan, X., & Ping, L. (2008). On systematic LT codes. *IEEE Communications letters*, 12(9), 681-683.
- [9] Azath Mubarakali, Ninoslav Marina, and Elena Hadzieva. "Analysis of Feature Extraction Algorithm Using Two Dimensional Discrete Wavelet Transforms in Mammograms to Detect Microcalcifications." *Computational Vision and Bio-Inspired Computing: ICCVBIC 2019* 1108 (2020): 26.
- [10] Lu, H., Cai, J., & Foh, C. H. (2010, December). Joint unequal loss protection and LT coding for layer-coded media delivery. In *2010 IEEE Global Telecommunications Conference GLOBECOM 2010* (pp. 1-5). IEEE.
- [11] Bose, Subash Chandra, Murugesh Veerasamy, Azath Mubarakali, Ninoslav Marina, and Elena Hadzieva. "Analysis of Feature Extraction Algorithm Using Two Dimensional Discrete Wavelet Transforms in Mammograms to Detect Microcalcifications." In *International Conference On Computational Vision and Bio Inspired Computing*, pp. 26-39. Springer, Cham, 2019.
- [12] Kumar, T. S. (2019). A Novel Method for HDR Video Encoding, Compression and Quality Evaluation. *Journal of Innovative Image Processing (JIIP)*, 1(02), 71-80.
- [13] Hemanth, D. Jude, and S. Smys, eds. *Computational Vision and Bio Inspired Computing*. Vol. 28. Springer, 2018.
- [14] Wu, N., Nguyen, T. D., Wei, C. Y., Yang, L. L., & Hanzo, L. (2008, May). Integrated LT coding, bit interleaved differential space time coding and sphere packing modulation for the wireless internet. In *VTC Spring 2008-IEEE Vehicular Technology Conference* (pp. 344-348). IEEE.
- [15] Kumar, T. Senthil. "A Novel Method for HDR Video Encoding, Compression and Quality Evaluation." *Journal of Innovative Image Processing (JIIP)* 1, no. 02 (2019): 71-80.
- [16] Yu, J., Zhong, J., Zhao, M., Cai, Y., & Wu, J. (2012, October). Novel LT coding scheme with limited feedback in broadcast systems. In *2012 International Conference on Wireless Communications and Signal Processing (WCSP)* (pp. 1-5). IEEE.
- [17] Chang, S. K., Yang, K. C., & Wang, J. S. (2008, May). Unequal-protected LT code for layered video streaming. In *2008 IEEE International Conference on Communications* (pp. 500-504). IEEE.
- [18] D. Liu, P. An, R. Ma, C. Yang, and L. Shen, "3D holoscopic image coding scheme using HEVC with Gaussian process regression," *Signal Process., Image Commun.*, vol. 47, pp. 438-451, Sep. 2016.

- [19] D. Liu, P. An, R. Ma, W. Zhan, X. Huang, and A. A. Yahya, "Contentbased light field image compression method with Gaussian process regression," *IEEE Trans. Multimedia*, vol. 22, no. 4, pp. 846-859, Apr. 2020.
- [20] Shakya, Subarna. "Analysis of Artificial Intelligence based Image Classification Techniques." *Journal of Innovative Image Processing (JIIP)* 2, no. 01 (2020): 44-54.
- [21] Y. Li, M. Sjostrom, and R. Olsson, "Coding of plenoptic images by using a sparse set and disparities," in *Proc. IEEE Int. Conf. Multimedia Expo (ICME)*, Turin, Italy, Jun. 2015, pp. 1-6.
- [22] Y. Li, M. Sjostrom, R. Olsson, and U. Jennehag, "Scalable coding of plenoptic images by using a sparse set and disparities," *IEEE Trans. Image Process.*, vol. 25, no. 1, pp. 80-91, Jan. 2016.
- [23] Koresh, Mr H. James Deva. "Quantization with Perception for Performance Improvement in HEVC for HDR Content." *Journal of Innovative Image Processing (JIIP)* 2, no. 01 (2020).
- [24] Himabindu, Y., R. Manjusha, and Latha Parameswaran. "Detection and Removal of RainDrop from Images Using DeepLearning." In *International Conference On Computational Vision and Bio Inspired Computing*, pp. 1355-1362. Springer, Cham, 2019.
- [25] CHITRA, Ms K., and V. PRASANNA VENKATESAN. "An Antiquity to the contemporary of Secret Sharing Scheme." *Journal of Innovative Image Processing (JIIP)* 2, no. 01 (2020): 1-13.