LOSSLESS PLENOPTIC IMAGE COMPRESSION USING ADAPTIVE BLOCK DIFFERENTIAL PREDICTION

Cristian Perra

University of Cagliari

ABSTRACT

Plenoptic images are obtained from the projection of the light crossing a matrix of microlens arrays which replicates the scene from different direction into a camera device sensor. Plenoptic images have a different structure with respect to regular digital images, and novel algorithms for data compression are currently under research. This paper proposes an algorithm for the compression of plenoptic images. The micro images composing a plenoptic image are processed by an adaptive prediction tool, aiming at reducing data correlation before entropy coding takes place. The algorithm is compared with state-of-the-art image compression algorithms, namely, JPEG 2000 and JPEG XR. Obtained results demonstrate that the proposed algorithm improves the coding efficiency.

Index Terms**—** Image compression, Plenoptic signals, JPEG2000, JPEGXR, Intra-prediction

1. INTRODUCTION

Plenoptic digital camera systems are becoming available as consumer and professional devices, or as experimental devices in research laboratories. The availability of such novel devices is leveraging considerable interest in plenoptic imaging as a next generation camera system.

A plenoptic digital camera exploits an array of small circular lenses for capturing positional and directional information related to the light ray incident on the camera.

Plenoptic raw images can be very data-intensive and it is important to define coding algorithms for optimizing the storage and transmission of plenoptic data. Some recent work related to the compression of the plenoptic data is briefly reviewed in the following.

In [1], the problem of sampling and compression of the data acquired by a camera sensor network is addressed. It is shown that the correlation in the visual data modeled by the plenoptic function can be estimated using some limited geometrical information about the scene and the position of the cameras. A distributed coding strategy for simple synthetic scenes is also proposed. It takes advantage of the estimated correlation to reduce the amount of data to be transmitted from the sensors to the receiver.

In [2], a compression method for multiview images is proposed. The data set is partitioned into planar layers characterized by a constant depth value. The separable three-dimensional wavelet transform is exploited for decorrelating the partitioned data.

Occlusions and disparity variations for different depths are also kept into account during the process. The transform coefficients are then entropy coded. The results show that the coding method is capable of outperforming the H.264/AVC video encoder for different data sets.

In [3], a system for capturing, rendering and compression of the plenoptic data is proposed. It is shown that the samples of plenoptic data are highly correlated and also that direct application of traditional compression schemes usually results in sub-optimal performance.

Zernike polynomials are exploited in [4], for the analysis and compression of plenoptic camera images. In the proposed technique images are represented with a natural set of basis functions for reconstruction of slices through the light field data and for image compression.

In [5], two multiview image compression methods are proposed. These methods outperform H.264/MVC and JPEG 2000 coding standard are proposed. The first method exploits a separable multi-dimensional wavelet transform applied across the viewpoint and spatial dimensions. The second method exploits distributed source coding (DSC) principles which reduces the inter-view redundancy and facilitates random access at the image level.

The problem of lossy plenoptic image compression using image coding standards has been addressed in [6]. Three image coding algorithms have been compared (JPEG, JPEG 2000, and SPIHT) in terms of subjective quality of the rendered view.

A method for light field compression that incorporates disparity compensation into 4D wavelet coding using disparity-compensated lifting has been presented in [7]. The wavelet coding scheme has shown higher compression efficiency and scalability in all dimensions when compared to a texture-map coder.

Two approaches to light field compression are presented in [8]. The first one applies video compression schemes to the 4D data structure of light fields by diving the light field image into blocks. The second one exploits disparity compensation of light field images before coding the data.

978-1-4673-6997-8/15/$31.00 ©2015 IEEE

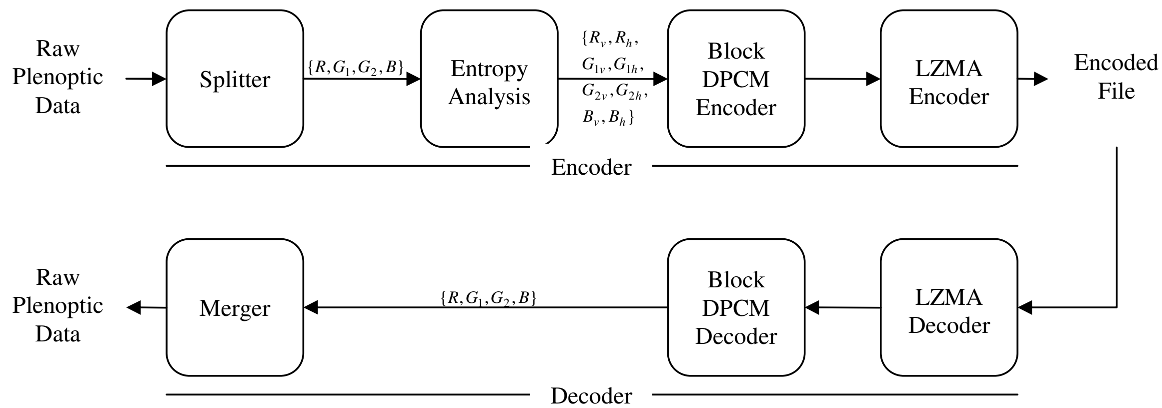

Fig. 1. Proposed coding architecture.

The problem of encoding a set of calibrated images representing a static scene from different viewpoints is addressed in [9]. At first the 3D geometry of the scene is reconstructed. The view-dependent texture maps are generated from all images. The texture maps are encoded by set partitioning in hierarchical trees (SPIHT).

Several image coding standards have been developed in the years. Nevertheless, some of these algorithms (e.g. JPEG) have not been designed for coding digital images sampled at more than 8-bit per pixel (bpp), and are not suited for the coding of plenoptic information captured with modern digital photosensors.

Two image coding standards, namely ISO/IEC JPEG 2000 and ISO/IEC JPEG XR, have been selected, among the various possibilities, for the interest in analysing and comparing the compression performances of two main approaches to image compression, that are block based (as JPEG XR) and wavelet based (as JPEG 2000). Moreover these coding standards both satisfy the constraint of providing support for reading and processing images sampled at more than 8bpp.

One of the main contributions of this paper is the proposal of a novel raw plenoptic data coding architecture designed for exploiting the structure and spatial redundancy of the plenoptic data. Another contribution is the evaluation of the performances of two image coding standards (JPEG 2000 and JPEG XR) as raw plenoptic data encoders tools.

The general objective of this research is to put into evidence what type of technologies might address the compression problem of plenoptic images acquired from modern plenoptic cameras. The particular objective, addressed in this paper, is to propose a novel algorithm for the compression of raw plenoptic data.

The paper is organized as follows. Section II presents the proposed coding architecture. Section III presents the experimental analysis and comparisons against the JPEG 2000 and JPEG XR coding of the plenoptic signal. Conclusions are drawn in section IV.

2. CODING ARCHITECTURE

A lossless coding architecture is presented in this section. The proposed architecture is shown in Fig. 1. The input signal is a raw (data from the image sensor) plenoptic image which is compressed by the encoder modules for producing the output encoded file. The output encoded file is processed by the decoder modules for the reconstruction of the original plenoptic image. The description of the modules in Fig. 1 is presented in the following.

A. Raw plenoptic data

The source image captured from the plenoptic digital camera is acquired by a color filter array (CFA) which arranges RGB color on a square grid of photosensors. A common filter pattern is named BGGR Bayer filter array which captures for every four pixels: 2 green, 1 red, and 1 blue color components.

B. Splitter and Merger

The splitter module decomposes the raw data into the different color components.

The two adjacent green pixels (G_1 and G_2) are treated as different channels since they are corresponding to two different spatial position.

A input raw plenoptic image having a resolution $W \times H$ pixels is decomposed into four matrix, as shows in Fig. 1:

$$\{raw\} \rightarrow \{R, G_1, G_2, B\} . \tag{1}$$

The merger module performs the inverse operation reconstructing the raw plenoptic data in the original data format.

C. Entropy analysis

Each color plane, R, G_1, G_2, or B, produced by the splitter module, is processed by the entropy analysis module. Since a plenoptic raw image is originated by an array of small circular lenses, there will be several replicas of small adjacent circular images.

The purpose of this module is to compute the horizontal and vertical gaps between pixels that maximize the matching between adjacent blocks of pixels for optimizing a subsequent differential prediction process.

Let $s(o_v, o_{h,} v_d, h_d)$ be a decimation of the raw image, realized keeping only every v_d^{th} sample in the vertical direction and every h_d^{th} sample in the horizontal direction, starting from the pixel at coordinates ($o_v, o_{h,}$):

$$\begin{aligned} & s(o_v, o_{h,} v_d, h_d, i, j) = raw(o_v + iv_d, o_h + jh_d), \\ & o_v \in [0, v_d - 1], \quad o_h \in [0, h_d - 1], \\ & i \in \left[0, \frac{H}{v_d}\right], \quad j \in \left[0, \frac{W}{h_d}\right], \quad \forall v_d, \quad \forall h_d. \end{aligned} \tag{2}$$

For each downsampled image $S(o_v, o_{h,} v_d, v_h)$, the 1D DPCM is computed in raster scan order using as predictor for the sample $s(o_v, o_{h,} v_d, v_h, i, j)$ the value

$$p(i,j) = \frac{s(o_v, o_{h,} v_d, v_h, i-1, j) + s(o_v, o_{h,} v_d, v_h, i, j-1)}{2} \tag{3}$$

The entropy analysis module computes the displacements that minimize the entropy of the 1D DPCM ($dpcm$) computed for all the downsampled images:

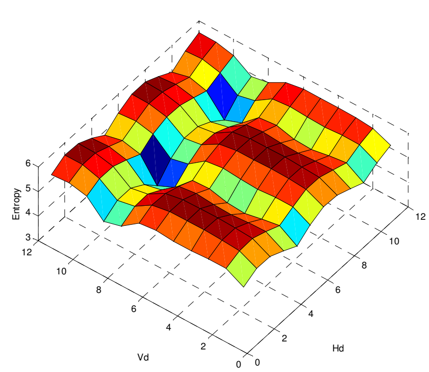

Fig. 2. Example of the entropy of the predicted signal in function of the vertical (Vd $\equiv v_d$) and horizonal (Hd $\equiv h_d$) displacement. The graph has the minimum entropy equal to H(x)=3.9 for $v_d = 9$ and $h_d = 5$.

$$D(v_d, h_d) = \bigcup_{\forall (o_v, o_h)} dpcm(S(o_v, o_{h,} v_d, h_d), \tag{4}$$

$$\underset{(v_d, h_d)}{\arg\min}\, H(D(v_d, h_d)), \tag{5}$$

where $D(v_d, h_d)$ is the 1D DPCM for all the sub-images downsampled with horizontal and vertical displacement corresponding to (v_d, h_d) pixels. Fig. 2 shows an example of the entropy $H(D(v_d, h_d))$ surface. The minimum of the graph is at coordinates $(v_d = 9, h_d = 5)$ corresponding to an entropy $H(S) = 3.9$.

The vertical (v_d) and horizontal (h_d) displacements for each color plane are denoted as (Fig. 1)

$$\begin{aligned} & (R_v, R_h) \equiv (v_d, h_d)_{RED} \\ & (G_{1v}, G_{1h}) \equiv (v_d, h_d)_{GREEN\,1} \\ & (G_{2v}, G_{2h}) \equiv (v_d, h_d)_{GREEN\,2} \\ & (B_v, B_h) \equiv (v_d, h_d)_{BLUE} \end{aligned}. \tag{6}$$

Block DPCM encoder and decoder

The displacement information computed in the previous step is used for partitioning each color plane into blocks of v_d raws and h_d columns. A horizontal block DPCM is computed by predicting each block from the previous block.

Let B denote a block of a color plane, the raw image S is obtained by

$$S = \bigcup_{\forall (i,j)} B_{(v_d, v_h)}(i, j) \tag{7}$$

The DPCM decoder performs the inverse DPCM for reconstructing the original color planes.

LZMA encoder and decoder

Lossless data compression of the block DPCM data is performed with the Lempel–Ziv–Markov chain algorithm (LZMA).

LZMA is an optimized version of LZ77 [10] providing high compression ratio with low decompression complexity and low memory requirements for decompression.

3. EXPERIMENTAL ANALYSIS

A dataset composed with eight raw plenoptic signals (16 bits per sample, 16bpp), has been selected for the experimental analysis presented in this section. The dataset and the proposed algorithm are publicly available at the authors' institution website. Considering the constraint of coding data at 16 bpp, the state-of-the-art image coding standard that have been chosen for the analysis are, as already mentioned, the JPEG 2000 and the JPEG XR, both configured for lossless coding.

JPEG2000 coding has been performed with the Kakadu software [11] enabling reversible compression (lossless) and 16 bit sample precision parameters.

JPEG XR (ITU-T T.832 | ISO/IEC 29199-2) is an image coding system primarily targeting the representation of continuous-tone still images such as photographic images [12]. JPEG XR coding has been performed with the ISO/IEC reference software [13] setting to zero the quantization parameter (lossless coding) and to 16 bit the sample precision parameter.

The compression ratio of the proposed algorithm is compared against the compression ratio of the JPEG 2000 and JPEG XR encoders.

Table I shows the results for each image and the average results. The bpp column shows the bit per pixel for the source image (16 bpp) and for the images encoded with JPEG2000, JPEG XR and with the proposed method. The '%' column shows the size of the compressed file in percentage with respect to the corresponding source file. The last raw of Table I reports the average results: the file size of the images compressed with JPEG 2000 and JPEG XR is the 74% of the size of the source images, while the file size of the images compressed with the proposed method is the 33% of the size of the source images.

These results are very promising and confirm the initial assumption that the compression of plenoptic signals deserves further studies in order to define novel coding architecture designed in function of the plenoptic signal structure.

Table 1. Experimental Results: the bpp column contains the bit per pixel of each image for the corresponding encoding; the % column contains the size of the compressed file in percentage with respect to the corresponding source file.

Image#	Source		JPEG2000		JPEGXR		Proposed	
	bpp	%	bpp	%	bpp	%	bpp	%
1	16	100	11.1	69	10.9	68	4.8	30
2	16	100	11.9	75	11.8	74	5.3	33
3	16	100	12.3	77	11.5	72	4.9	30
4	16	100	11.7	73	12.0	75	5.3	33
5	16	100	12.1	76	12.3	77	5.1	32
6	16	100	11.4	71	11.6	72	5.4	34
7	16	100	11.7	73	11.9	74	5.6	35
8	16	100	12.5	78	12.7	80	5.5	34
Average	16	100	11.8	74	11.9	74	5.2	33

4. CONCLUSIONS

Plenoptic cameras produce raw images that have a different structure with respect to raw images produced by conventional cameras. A novel coding algorithms for lossless compression of plenoptic data has been proposed in this paper. The inherent redundancy of the plenoptic signal has been evaluated by an entropy analysis process aiming at determining a block size for subsequent processing. DPCM has been then applied to color plane blocks. The proposed algorithm has been compared with state-of-the-art standard image coding algorithm, namely JPEG 2000 and JPEG XR. Experimental results have shown that the proposed method outperforms the standard algorithms. The results confirm the initial assumption that a better understanding of the plenoptic signal structure can leverage the development of novel compression algorithms providing higher compression ratio than former algorithms.

5. REFERENCES

[1] N. Gehrig and P. L. Dragotti, “Geometry-driven distributed compression of the plenoptic function: performance bounds and constructive algorithms,” IEEE Transactions on Image Processing, vol. 18, no. 3, pp. 457–70, Mar. 2009.

[2] A. Gelman, P. L. Dragotti, and V. Velisavljevic, “Multiview image compression using a layer-based representation,” Image Processing (ICIP), 2010 17th IEEE International Conference on, 2010.

[3] S.-C. Chan, K.-T. Ng, Z.-F. Gan, K.-L. Chan, and H.-Y. Shum, “The plenoptic videos: capturing, rendering and compression,” 2004 IEEE International Symposium on Circuits and Systems, vol. 3, 2004.

[4] J. Schwiegerling, G. C. Birch, and J. S. Tyo, “Analysis and compression of plenoptic camera images with Zernike polynomials,” Proc. SPIE 8487, Novel Optical Systems Design and Optimization XV, vol. 8487, no. 520, p. 84870G–84870G–7, Oct. 2012.

[5] A. Gelman, P. L. Dragotti, and V. Velisavljević, “Centralized and interactive compression of multiview images,” Proc. SPIE 8135, Applications of Digital Image Processing XXXIV, vol. 44, no. 0, p. 81350J–81350J–15, Sep. 2011.

[6] R. S. Higa, R. Fredy, L. Chavez, R. B. Leite, R. Arthur, and Y. Iano, “Plenoptic image compression comparison between JPEG, JPEG2000 and SPITH,” Cyber Journals: Multidisciplinary Journals in Science and Technology, Journal of Selected Areas in Telecommunications (JSAT), vol. 3, no. 6, pp. 1–6, 2013.

[7] C.-L. Chang, X. Zhu, P. Ramanathan, and B. Girod, “Light field compression using disparity-compensated lifting and shape adaptation,” IEEE Transactions on Image Processing, vol. 15, pp. 793–806, 2006.

[8] M. Magnor and B. Girod, “Data compression for light-field rendering,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 10, no. 3, 2000.

[9] M. Magnor and B. Girod, “Model-based coding of multiviewpoint imagery,” in Visual Communications and Image Processing 2000, 2000, pp. 14–22.

[10] J. Ziv and A. Lempel, “A universal algorithm for sequential data compression,” IEEE Transactions on Information Theory, vol. 23, no. 3, pp. 337–343, 1977.

[11] D.S. Taubman, 2014, Kakadu software, [Online]. Available: http://www.kakadusoftware.com/

[12] F. Dufaux, G. J. Sullivan and T. Ebrahimi, “The JPEG-XR image coding standard,” IEEE Signal Processing Magazine, Volume 26, Issue 6, November 2009, pp. 195-199, 204-204.

[13] ISO/IEC FCD 29199-5: Information technology—JPEG XR image coding system—Part 5: Reference software [ISO/IEC JTC 1/SC 29/WG 1 N 5020] [Online]. Available: http://www.jpeg.org.