Fuzzy Neural Network Based Prediction Coding for Bayer Pattern Image

Yongqiang Cheng Jiang Zhao Keming Xie Gang Zhang
College of Information Engineering, Taiyuan University of Technology
Taiyuan, P.R. China
cyqtyster@gmail.com

Abstract—**In this paper, a sequential lossless compression of raw data from image sensor with Bayer pattern is proposed. Inspired by model of JPEG-LS, the proposed encoder consists of fuzzy neural network predictor, adaptive correction part based on context and adaptive arithmetic coder. As in JPEG-LS, it is a empirically observed that the global statistics of residuals from the ANN fixed predictor in raw data , which effectively exploits structural redundancies between mosaic-like color components, are well-modeled by a TSGD centered at zero. In the meantime, we propose a context determination approach based on causal interpolation that achieves high coding efficiency. Consequently, we can encode mosaic images on the fly at low complexity level. Compared with existing methods of lossless compression for Bayer raw data, the performance of proposed method is apparently the best.**

Keywords— **lossless compression; raw data; image sensor**

I. INTRODUCTION

Most of commercially available digital cameras capture color information by single light sensitive sensor with CFA (color filter array). In CFA, only one basic color is captured per pixel position. The remaining two basic colors need to be reconstructed later on by digital image processing algorithms for CFA interpolation. Among various CFA masks the most popular one is the so-called Bayer pattern [1] Fig. 1 presents a mosaic image with Bayer CFA.

The CFA captures only one-third of the necessary color intensities and the full color image is generated from the captured raw data by interpolation. In most conventional applications, the generated color image is then compressed before storage or transmission.

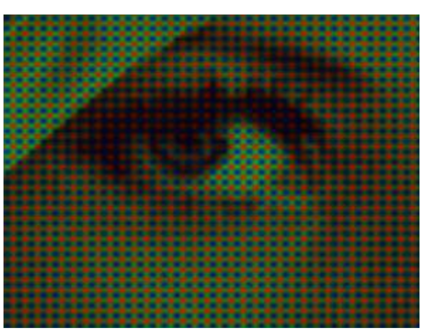


Figure 1. A mosaic image with Bayer color filter array

However, CFA image compression is of paramount importance to image-enabled consumer electronic devices, such as wireless PDAs, mobile phones, and surveillance systems. Since bandwidth reduction is crucial for transmission of captured images in wireless networks, compression of the CFA data rather than the demosaiced data basically allows a three-fold reduction of the data redundancy by moving compression operation before color pixel interpolation [2]. It was demonstrated that using the compression-first scheme retains more pertinent information, thus lower compression ratio is achieved and higher image quality are to be transmitted. Several new methods based on the compression-first scheme have been reported in [2]. On the other hand, operating on the CFA pixels arranged in the original mosaic layout may limit the compression efficiency due to the artificial high frequencies in the CFA image. So the raw data is not suitable for direct compression using the existing compression standards such as JPEG, JPEG-LS, and JPEG-2000. To overcome the mosaic-like structure of the acquired CFA data and further increase the CFA image coding efficiency, a CFA data structure conversion is used prior to image compression in the existing compression-first schemes [2]. The structure conversion step transforms the CFA pixels corresponding to the same color filters into a structure more appropriate for image coding [3]. Once the conversion is done, the existing compression standards are applied to it. Therefore, the structure conversion can help to achieve higher compression ratios compared to the direct coding of the acquired CFA image. Ning Zhang and Xiaolin Wu also developed a simpler and faster lossless mosaic image codec [4], propose a particular wavelet decomposition scheme, called Mallat wavelet packet transform, which is ideally suited to the task of de-correlating color mosaic data, and apply a low-complexity adaptive context based Golomb–Rice coding technique to compress the coefficients of Mallat wavelet packet transform.

In this paper, our work proposes an effective sequential prediction based coder for lossless compressing raw mosaic data, which performs sequentially pixel by pixel in raster-scan order, also so called one-pass scheme as opposed to multiple-pass scheme. The proposed method has four major components: fuzzy neural network predictor, context model, context based adaptive entropy coder and run coder. It has achieved best comprise between compression performance and computational complexity. This coder outperforms previous proposed lossless coder in both bit rate and low-complexity. The presentation is

978-1-4244-1674-5/08 /$25.00 ©2008 IEEE

organized as follows. Section Ⅱ introduces proposed coder structure and its neural predictor is described in section Ⅲ. In section Ⅳ, we propose a context determination approach based on causal interpolation that achieves high coding efficiency. The principle of the arithmetic coder and run length coder is interpreted in section Ⅴ Experimental results are reported in Section Ⅵ and conclusions are drawn in Section Ⅶ.

II. THE STRUCTURE OF THE CODER

Proposed coder is a sequential coding scheme that encodes and decodes in raster scan order with a single pass through the CFA image. The coding process uses prediction and context templates that involve only the two previous scan lines of coded pixels. The schematic description of proposed encoder is given in Fig. 2. As in JPEG-LS [5], proposed coder operates in two modes: normal mode and run mode. The detail of the run mode is analogous to the run mode in [5]. In normal mode, the system has four major integrated components: ANN prediction, context based error correction, interpolation for context, context modeling of prediction errors, and entropy coding of prediction errors. First, ANN fixed predictor effectively exploits structural redundancies between mosaic-like color components, and adaptive correction part is used to cancel bias which typically presents in context-conditioned prediction error signals. In context modeler, The context that conditions the encoding of the current prediction residual, as in JPEG-LS [5], is built out of the differences $g_1=d-b$, $g_2=b-c$, $g_3=c-a$, where a, b, c, and d surrounding the current encoded pixel is acquired by causal interpolation based on color correlation, see section Ⅳ. Finally, entropy coder of the residue is implemented with arithmetic coding and run-length coding.

III. ANN PREDICTOR

Benefited from artificial neural network with learning ability and generality, we design an ANN predictor, which effectively exploits structural redundancies between mosaic-like color components, and uses past five pixels (see Fig. 3) to guess the value of current encoded pixel. The average of each color component in this area including current encoded pixel could be computed and denoted as $\overline{R}$, $\overline{G}$ and $\overline{B}$, respectively.

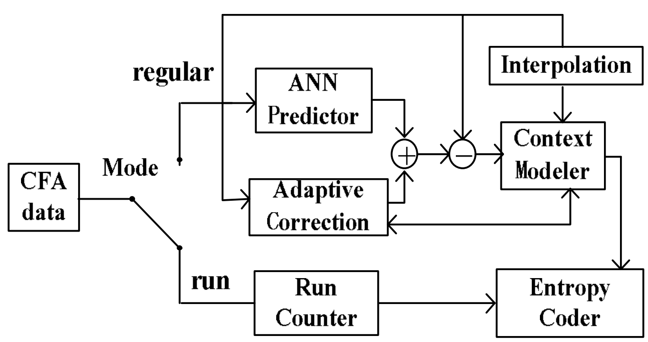


Figure 2. The proposed coder

The network structure presented for this application is presented in Fig. 4.

$y-2$	R	G	R	G	R	G	R	G
$y-1$	G	B	G	B	G	B	G	B
y	R	G	R	G	X			
			$x-2$	$x-1$	x			

Figure 3. Causal pattern of the predictor

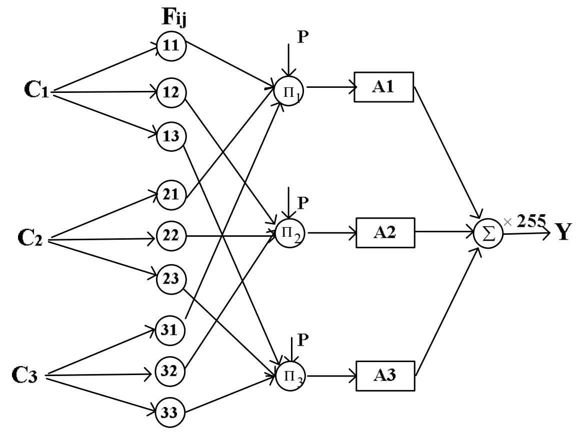


Figure 4. The structure of neural network

The network used here is the modified TFNN [5] and contains three layers, noticing that the net structure is simplified for low complexity demand. The net inputs are causal pixels surrounding the current pixel X. we limit its image buffering requirement to one scan line. The shadowed R, G and B are used to estimate the value of the pixel X on location (x, y) and represented as the vector

$$V = (f_{x+1,y-1}, f_{x,y-1}, f_{x-1,y-1}, f_{x-2,y}, f_{x-1,y}) \tag{1}$$

The basic function in each layer is described as follow

On Layer 1, Membership layer, the Gaussian functions are used as the membership function. Each membership node is responsible for mapping an input linguistic variable into a possibility distribution for that variable. Each node performs a membership function

$$O_{ij}^{\ 1} = F_{ij}(C_i) = EXP - \left(\frac{C_i - m_{ij}}{\sigma_{ij}}\right)^2 \tag{2}$$
$$C_1 = \overline{R}/255, C_2 = \overline{B}/255, C_3 = \overline{G}/255$$

where $i,j = 1,2,3$, and m_{ij} and σ_{ij} denote trained expectation and variance of each color variable , respectively.

On Layer 2, the I/O representation of each node is

$$O_j^{\ 2} = \prod_{i=1}^{3} F_{ij}(C_i)P$$
$$P = \frac{V}{255\sum_{j=1}^{3}\mu_j}, \mu_j = \prod_{i=1}^{3} F_{ij}(C_i) \tag{3}$$

On Layer 3, The output of TFNN is a linear combination producing prediction value Y

$$O^3 = Y = 255\left\lfloor \sum_{j=1}^{3} \mu_j P A_j^{\,T} \right\rfloor \tag{4}$$

Here the symbol $\lfloor\ \rfloor$ denotes downward integer conversion, μ_j is predictor weight, and each component in A_j represents prediction coefficient.

We employ BP learning algorithm to train the TFNN predictor system. The cost function is defined as

$$E = \frac{1}{2}\left\| f_{x,y} - Y \right\|^2 \tag{5}$$

Detail description of BP learning algorithm can be found in [6]. Finally, update law is as follow

$$A_j(k+1) = A_j(k) + \eta_A\left(-\frac{\partial E}{\partial A_j}\right) \tag{6}$$

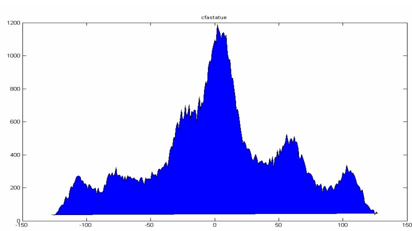
$$m_{ij}(k+1) = m_{ij}(k) + \eta_m\left(-\frac{\partial E}{\partial m_{ij}}\right) \tag{7}$$

$$\sigma_{ij}(k+1) = \sigma_{ij}(k) + \eta_\sigma\left(-\frac{\partial E}{\partial \sigma_{ij}}\right) \tag{8}$$

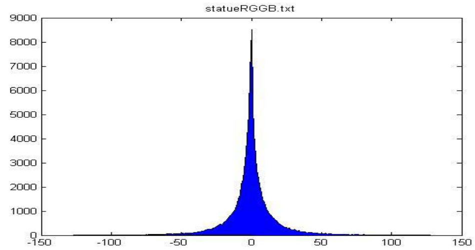
The parameters trained above are transferred to decoder as side information. Fig. 5 (a) and (b) separately show the global statistics of residuals from directly applying JPEG-LS and the proposed predictor to the CFA image of "statue". Clearly, the global statistics of residuals from the proposed predictor are well-modeled by a two-side geometric distribution (TSGD) centered at zero.

IV. CONTEXT DETERMINATION

According to JPEG-LS, the adaptive part of the predictor is context-based and it is used to "cancel" the integer part of the offset due to the TFNN predictor. This adaptive correction (or bias cancellation) could be well performed at low complexity. Context quantization is very important for improving compression performance. Prediction errors are partitioned into a predefined number of statistically homogeneous classes based on context, and each class is entropy coded by means of Golomb coding. If the classes found out are discriminated, the entropy of a context-conditioned model of prediction errors will be lower than that derived from a stationary memory-less model of the uncorrelated source.



(a). The residuals histogram from JPEG-LS



(b). The residuals histogram from proposed coder

Figure 5. The residual histogram of "statue" CFA image

c	b	d
a	Ix	

(a). Context pattern

B_{11}	G_{12}	B_{13}	G_{14}	B_{15}	G_{16}
G_{21}	R_{22}	G_{23}	R_{24}	G_{25}	R_{26}
B_{31}	G_{32}	B_{33}	G_{34}	B_{35}	G_{36}
G_{41}	R_{42}	G_{43}	R_{44}		

(b). Encoding red pixel

R_{00}	G_{01}	R_{02}	G_{03}	R_{04}	G_{05}
G_{10}	B_{11}	G_{12}	B_{13}	G_{14}	B_{15}
R_{20}	G_{21}	R_{22}	G_{23}	R_{24}	G_{25}
G_{30}	B_{31}	G_{32}	B_{33}		

(c). Encoding blue pixel

R_{00}	G_{01}	R_{02}	G_{03}	R_{04}	G_{05}
G_{10}	B_{11}	G_{12}	B_{13}	G_{14}	B_{15}
R_{20}	G_{21}	R_{22}	G_{23}	R_{24}	G_{25}
G_{30}	B_{31}	G_{32}	B_{33}	G_{34}	

(d). Encoding green pixel

Figure 6. An example of context pattern

Now we use interpolation method to obtain the context of current encoded pixel. As in JPEG-LS, both the current encoded pixel Ix and its corresponding context pattern including four elements *a*, *b*, *c* and *d* are shown in Fig. 6 (a), with the same color. In Fig 6 (b), (c) and (d), assumed that current encoded pixel is R_{44}, B_{33} or G_{34} of which contexts would be red, blue or green component being culled out through Bayer color filter array, each missing color component on location *a*, *b*, *c* and *d* could be reconstructed by interpolation as follows:

When Ix=R_{44}, its context pattern is red component and each element is computed by

$$\begin{aligned} a &= R_{43} = R_{42} + (G_{43} - G_{41}) \div 2 \\ b &= R_{34} = R_{24} + (G_{34} - G_{14}) \div 2 \\ c &= R_{33} = (R_{42} + R_{24} + R_{22}) \div 3 \\ d &= R_{35} = (R_{24} + R_{26}) \div 2 + (B_{35} - B_{15}) \div 2 \end{aligned} \tag{9}$$

When Ix=B_{33}, its context pattern is blue component and each element is computed by

$$\begin{aligned} a &= B_{32} = B_{31} + (G_{32} - G_{30}) \div 2 \\ b &= B_{23} = B_{13} + (G_{23} - G_{03}) \div 2 \\ c &= B_{22} = (B_{11} + B_{13} + B_{31}) \div 3 \\ d &= B_{24} = (B_{13} + B_{15}) \div 2 + (R_{24} - R_{04}) \div 2 \end{aligned} \tag{10}$$

When Ix=G_{34}, its context pattern is green component and each element is computed by

$$\begin{aligned} a &= G_{33} = (G_{32} + B_{33} - B_{31}) \div 2 + (G_{23} + B_{33} - B_{13}) \div 2 \\ b &= G_{24} = [(G_{23} + G_{25}) \div 2 + G_{14} + (R_{24} - R_{04}) \div 2] \div 2 \\ c &= G_{23} \\ d &= G_{25} \end{aligned} \tag{11}$$

Here, we use causal pattern to reconstruct each color component forming the context of current encoded pixel. At decoder side, the same operate is causally executed. No side-information needs be transferred.

V. Entropy Coder

Generally, CFA images are very far from being continuous-tone and have sparse histograms, containing only a subset of the possible sample values in each component, and the ANN predictor would tend to concentrate the value of the prediction residuals into a reduced set. However, prediction correction tends to spread these values over the entire range, and even if that were not the case, the probability assignment of a TSGD model which fits for Golomb coding would not take advantage of the reduced alphabet. Therefore, the arithmetic coder in [6] is borrowed to complete encoding prediction residual. Notice that the shift parameter *s* of TSGD is tuned in the range [-1/2, 1/2).

The encoder enters a run-length mode when a “flat region” context with *a*=*b*=*c*=*d* is detected. If current sample Ix is equal to *a*, then run-length increases by 1. For the next sample of Ix, its context elements *a*, *b*, *c* and *d* are not computed if it is equal to the previous sample of Ix and encoded in run-length mode. For other following up pixels, we use the same method as above to decide whether in run state.

When the run is broken by a non-matching sample Ix, the encoder goes into a “run interruption” state, where the difference (with the sample above) is encoded. Runs can also be broken by ends of lines, in which case the encoder returns to normal context-based coding. Finally, prediction errors are arithmetic-coded conditioned on one of the 12 encoding states is performed, following binarization strategy of [7]. For a state with index, we choose the corresponding binarization tree as the Golomb tree for the parameter (the run state also uses k=0).

TABLE I. Lossless Bit Rates of CFA Images by JPEG-LS, JPEG 2K, Method [4] and OURS

Images	JPEG-LS	JPEG2K	Method[4]	OURS
Lena	4.93	5.02	4.975	4.878
Sailboat	5.12	5.24	5.731	5.446
Peppers	6.44	6.52	4.535	4.430
Plane	5.72	5.76	4.000	3.982
Flower	4.82	4.97	4.294	4.510
Statue	5.33	5.39	4.717	5.022
Lighthouse	4.86	4.96	4.975	5.022
Zebra	5.28	5.38	4.242	4.342
Town	4.91	4.93	4.751	4.950
Window	5.99	6.09	4.969	5.171
Average	5.34	5.43	4.872	**4.878**

VI. Experimental Results

As the structure of proposed encoder is analogous to the one in JPEG-LS, we use similar program to implement our encoder. In our experiments, we simulated color mosaic images by sub-sampling some typical color images drawn from the ISO JPEG test set and the Kodak set, and interleaving the samples according to the Bayer pattern [4]. TABLE Ⅰ presents the lossless bit rates of the proposed method in comparison with JPEG-LS, the lossless mode of JPEG 2000 [8] based on CFA data structure conversion and method in [4], and the corresponding results are listed in TABLE Ⅰ. The average bit rate of the proposed method is lower than that of both JPEG-LS and JPEG2000. The method in [4] gives an average lossless bit rate of 4.872 bit/pixel with negligible compression gain, compared with an average bit rate of 4.878 bit/pixel for our method. Note that the complexity of proposed method is also the lowest among exiting lossless compression methods for CFA images. Obviously, the structure conversion or wavelet transform based methods require more resource both of time and memory than proposed method in which TFNN may be implemented using parallel structure while costs few hardware resource [5]. Actually, in many applications, a drastic complexity reduction can have more practical impact than a modest increase in compression ratio.

On a 2.8-GHz Pentium-4 personal computer with 1-GB RAM, we simulate the proposed compression scheme in software with C++ programming language, the average execution time to compress a 768×512 Bayer pattern image is around 91ms.

VII. CONCLUSION

In this paper, we propose a low-complexity and high-efficient lossless CFA image compression method with low memory cost suitable for a hardware design based on the Bayer format image. It turns out that raster scan based DPCM coder can be directly applied to CFA images compression and provides a higher compression ratio while keeping computational complexity reasonably low. Furthermore, our coder could be employed to near lossless compress full-color image being sub-sampled through Bayer color filter array, and the decompressed mosaic image may be demosaiced by various existing methods [9] [10] to reconstruction the full-color image.

In the future, this work still needs to be improved in some aspects. Firstly, as current contexts cannot adequately characterize some of more complex relationships between predicted pixel and its surroundings, an additional context which exploits higher order structures such as texture pattern and local activity may be combined with current contexts to form compound contexts [11] for further compression gains. Secondly, for further reducing encoding complexity, we try to use adaptive combination/switching of adaptive predictors (ACAP/ASAP) [12], instead of ANN predictor.

ACKNOWLEDGMENT

This work is supported by Natural Science Foundation of Shanxi Province [20051019].

REFERENCES

[1] B.E.Bayer, “Color Imaging Array”, U.S.Patent 3,971,065,1976.

[2] S.Y. Lee and A. Ortega, “A novel approach of image compression in digital cameras with a Bayer color filter array,” IEEE Int. Conf. Image Processing 2001, vol. 3, Oct 2001, pp. 482-485.

[3] Rastislav Lukac, and Konstantinos N Plataniotis, “Single-Sensor Camera Image Compression,” IEEE Transactions on Consumer Electronics, Vol. 52, No. 2, MAY 2006, pp. 299-307.

[4] N. Zhang and X.Wu, “Lossless compression of color mosaic images,” IEEE Int. Conf. Image Processing, vol.1, 2004, pp. 517–520.

[5] C. H. Lee, W. Y. Lai and Y. C. Lin, “A TSK-Type Fuzzy Neural Network (TFNN) Systems for Dynamic Systems Identification,” 42nd IEEE Conf. on Decision and Control, 2004, pp.4002-4007.

[6] M. J. Weinberger, G. Seroussi, and G. Sapiro, “The LOCO-I Lossless Image Compression Algorithm:Principles and Standardization into JPEG-LS,” IEEE Trans. Image Processing, vol. 9, no. 8, Aug. 2000, pp.1309-1324

[7] Activity Level Classification Model (ALCM), 1995. Proposal submitted in response to the Call for Contributions for ISO/IEC JTC 1.29.12.

[8] B. Parrein, M. Tarin, and P. Horain, “Demosaicking and JPEG2000 compression of microscopy images,” IEEE Int. Conf. Image Processing, vol. 1, 2004, pp. 521–524.

[9] Gunturk BK, Glotzbach J, Altunbasak Y, Schaffer RW, and Murserau RM, “Demosaicking: color filter array interpolation”, IEEE Signal Process Mag 22(1), 2005, pp. 44–54.

[10] R. Lukac, K. N. Plataniotis, D. Hatzinakos, and M. Aleksic, “A novel cost effective demosaicing approach,” IEEE Transactions on Consumer Electronics, vol. 50, no. 1, February 2004, pp. 256-261.

[11] X. Wu and N. D. Memon, “Context-based, adaptive, lossless image coding,” IEEE Trans. Commun., vol. 45, pp. 437–444, Apr. 1997.

[12] Matsuda, H. Mori, and S. Itoh, “Lossless coding of still images using minimum-rate predictors,” Proc. IEEE Int. Conf. Image Processing,vol. I/III, 2000, pp. 132–135.