A LOSSLESS CODING SCHEME FOR BAYER COLOR FILTER ARRAY IMAGES USING BLOCK-ADAPTIVE PREDICTION

Toshiyuki Abe , Ichiro Matsuda , Susumu Itoh

†Department of Electrical Engineering, Science University of Tokyo
2641 Yamazaki Noda-shi, Chiba 278-8510, JAPAN
E-mail:abe@itohws01.ee.noda.tus.ac.jp

ABSTRACT

This paper proposes a novel lossless coding scheme for Bayer color filter array (CFA) images which are generally used as internal data of color digital cameras having a single image sensor. The scheme employs a block-adaptive prediction method to exploit spatial and spectral correlations in local areas containing different color signals. In order to allow adaptive prediction suitable for the respective color signals, four kinds of linear predictors which correspond to 2×2 samples of Bayer CFA are simultaneously switched block-by-block. Experimental results show that the proposed scheme outperforms other state-of-the-art lossless coding schemes in terms of coding efficiency for Bayer CFA images.

Keywords: Bayer color filter array, lossless coding, block-adaptive prediction,

1. INTRODUCTION

Most digital cameras have a single image sensor and an optical filter array to capture color image signals at once. Fig. 1 illustrates the most popular optical filter array known as the Bayer color filter array (CFA) [1]. Since each element of the Bayer CFA corresponds to a cell of the image sensor, the captured image has only one of three color signals (i.e. R, G and B signals) at each pels. Such an incomplete image is usually interpolated by a process called demosaicing and then compressed by JPEG algorithm. Through these processes, a few quality loss is unavoidable because each of them is irreversible. Therefore, some kinds of digital cameras designed for high-end applications have a function to provide unprocessed Bayer CFA images as RAW-data. Since the RAW-data generally requires a large amount of storage space, an efficient lossless compression technique is desired for practical use. From this point of view, several lossless coding schemes specially designed for the Bayer CFA images have been recently proposed. For example, [2] employs an integer wavelet transform based on the lifting scheme, both [3] and [4] use a backward prediction technique to remove signal redundancy peculiar to the Bayer CFA images, respectively.

In this paper, a novel lossless coding scheme using block-adaptive prediction is presented. The scheme is based on our previous study dedicated to monochrome images [5], but formation of linear predictors is extended in conformity with structure of the Bayer CFA images. Furthermore a quadtree-based variable block-size partitioning method is introduced to improve accuracy of the adaptive prediction. The effectiveness of the proposed coding scheme is confirmed through performance comparison with several state-of-the-art lossless coding schemes.

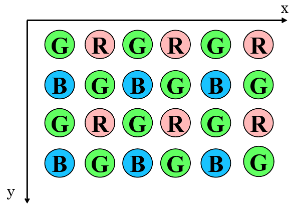
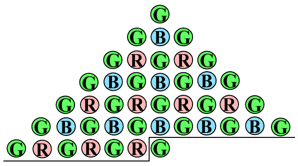


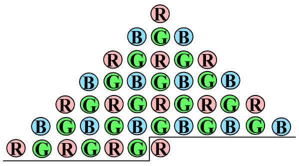
Fig. 1: Bayer color filter array.

2. BLOCK-ADAPTIVE PREDICTION

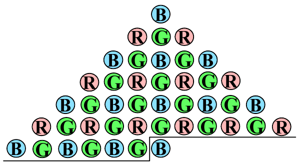
Our lossless coding scheme employs a block-adaptive prediction technique which partitions an image into square blocks and classifies them into several classes. In the case of monochrome images [5], it is enough to prepare a single predictor for every class if the classification is carried out properly. However, with respect to Bayer CFA images, we must consider that positional relationship of color signals used in the prediction is absolutely different pel-by-pel. For example, when a Bayer CFA image is encoded in raster scan order, there are four different patterns of R, G and B signals in casual areas depending on whether coordinates of the current pel $\boldsymbol{p} = (x, y)$ are even or odd as shown in Fig. 2. In general, it is reasonable to use different color signals simultaneously for prediction in the sense of exploiting both spatial and spectral correlations in a causal area. However, applying the same predictor to all of the pels within a class ignoring difference of the above mentioned patterns is undesirable. Therefore, in this paper, we prepare a set of four linear predictors corresponding to the respective patterns shown in Fig.2 for every class. When the current pel $\boldsymbol{p}$ is in a block belonging to the m-th class ($m \in \{1, 2, 3, \ldots, M\}$), a predicted value $\hat{s}(\boldsymbol{p})$ is expressed



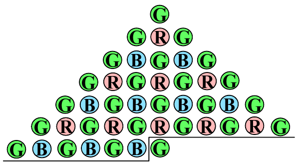
(a) (x, y) = (even,even)



(b) (x, y) = (odd,even)



(c) (x, y) = (even,odd)



(d) (x, y) = (odd,odd)

Fig. 2: Positional relationship of color signals at causal neighbors.

as:

$$\hat{s}(\boldsymbol{p}) = \sum_{k=1}^{K} a_{m,c}(k) \cdot s(\boldsymbol{p} + \boldsymbol{r}_k),$$
$$c = 2 \cdot (y \bmod 2) + x \bmod 2 \qquad (1)$$

where K is the number of pels used in prediction (i.e. prediction order), $\{a_{m,c}(k)\}$ are prediction coefficients given for the four kinds of predictors $(c \in \{0, 1, 2, 3\})$ and $s(\boldsymbol{p})$ corresponds to one of R, G and B signals available at the pel $\boldsymbol{p}$. In addition, $\boldsymbol{r}_k$ $(k = 1, 2, \dots, K)$ is a vector which indicates relative coordinates of the k-th reference pel used in the prediction. In this paper, we assign a suffix k to the reference pels in a causal area in increasing order of the city-block distance $\|\boldsymbol{r}_k\|_1$.

3. CODING OF PREDICTION ERRORS

After the prediction, context modeling for adaptive arithmetic coding of the prediction error $e = s(\boldsymbol{p}) - \hat{s}(\boldsymbol{p})$ is conducted. This context modeling is based on non-linear quantization of a context function which is defined as the wighted sum of absolute prediction errors at already encoded causal neighbors:

$$= \sum_{k=1}^{20} \frac{1}{\|\boldsymbol{r}_k\|_2} \cdot |s(\boldsymbol{p} + \boldsymbol{r}_k) - \hat{s}(\boldsymbol{p} + \boldsymbol{r}_k)| \qquad (2)$$

Since relationship between a value of this function and a probably density function (PDF) of the prediction errors seems to vary according to which color signal is being encoded, thresholds used for the quantization are switched not only by the class m but also the index c in the same way as the prediction. By using the resulting thresholds $\{Th_{m,c}(1), Th_{m,c}(2), \dots, Th_{m,c}(15)\}$, we obtain sixteen contexts $(n = 1, 2, 3, \dots, 16)$ as the quantization levels of . Here, we assume that a conditional PDF of the e observed in each context can be modeled by the generalized Gaussian function $P(e\,|\,n)$ with a variance ^2_n [5]. In addition, if each of the color signals has 8-bit precision, possible values of the prediction error e for a given value of $\hat{s}(\boldsymbol{p})$ are limited to the following 256 values:

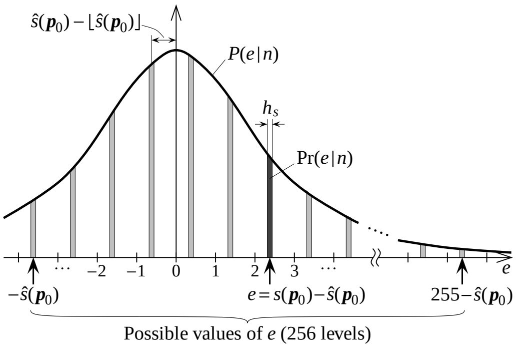


Fig. 3: Conditional probability of the prediction error e.

$$e \in \{s - \hat{s}(\boldsymbol{p}) \mid s = 0, 1, \dots, 255\} \qquad (3)$$

Therefore, a conditional probability of occurrence for each possible value of e, when both the context n and the predicted value $\hat{s}(\boldsymbol{p})$ are known, is derived from the above PDF model.

$$\Pr(e\,|\,\hat{s}(\boldsymbol{p}), n) = \frac{\Pr(e\,|\,n)}{\sum_{s=0}^{255} \Pr(s - \hat{s}(\boldsymbol{p})\,|\,n)}, \qquad (4)$$

$$\Pr(e\,|\,n) = \int_{-h_s\ 2}^{h_s\ 2} P(e + \varepsilon\,|\,n)\, \mathrm{d}\varepsilon \qquad (5)$$

As a matter of fact, the predicted value $\hat{s}(\boldsymbol{p})$ is explicitly rounded to nearest multiple of $h_s = 1/8$ to avoid accumulation of unexpected rounding errors. Hence the value of h_s is used an interval for integration in equation (5). Adaptive arithmetic coding of the actual value of e is carried out according to the conditional probabilities calculated by using equations (4) and (5). Note that the numerator of equation (4) corresponds to the area shown in dark gray and the denominator is the sum of shaded area in Fig. 3.

4. OPTIMIZATION OF CODING PARAMETERS

In the proposed coding scheme, parameters listed below are needed as side information.

- Class label m for each block.
- Prediction coefficients $\{a_{m,c}(k)\}$.
- Thresholds $\{Th_{m,c}(n)\}$.
- Shape parameter of the PDF model $P(e\,|\,n)$.

Values of these parameters are iteratively optimized for each image so that the following cost function can have a minimum.

$$J = -\sum_{\boldsymbol{p}} \log_2 \Pr(e\,|\,\hat{s}(\boldsymbol{p}), n) + B_{side} \qquad (6)$$

The first term of the cost function represents the number of bits required for the adaptive arithmetic coding of the prediction errors. The second term (B_{side}) is the amount of side information on the above coding parameters. Concrete procedures for minimizing the cost function J are as follows.

(1) Blocks composed of 8×8 pels are classified into M classes according to variance of a G signal within the respective blocks. Then $4M$ kinds of initial predictors are designed based on this classification.

(2) Two prediction coefficients $a_{m,c}(i)$ and $a_{m,c}(j)$ are chosen randomly, and partial optimization of them are carried out by varying their values gradually. Repeat this operation a certain number of times for each predictor.

(3) The thresholds $\{Th_{m,c}(n)\}$ are optimized by using the dynamic programming technique.

(4) A shape parameter of the generalized Gaussian function is modified for each PDF model.

(5) All the blocks are re-classified by selecting the optimum set of predictors.

(6) Procedures (2)–(5) are repeated until all of the coding parameters converge.

5. VARIABLE BLOCK-SIZE ADAPTIVE PREDICTION

In the block-adaptive prediction technique, accuracy of the adaptive prediction is obviously improved by using smaller block-size, while the amount of side information on class labels assigned to the blocks increases. To realize better control of such a trade-off, we introduce a quadtree-based variable block-size partitioning method into the above mentioned optimization procedures. A quadtree is built by recursive partitioning method of a square block into four sub-blocks. Each node of the tree, except in the lowest level, has a flag which indicates whether the corresponding block is further partitioned '1' or not '0'. Accordingly, an arbitrary partitioning pattern can be represented by a series of such flags as shown in Fig. 4. In the proposed scheme, five-level partitioning with a maximum block-size of 32×32 pels is performed and the best combination of both block-sizes and class labels which minimizes the cost function J is determined. The number of bits required for the above quadtree flags is also added to B_{side} when the cost function J is calculated. This quadtree decision process is performed in the optimization procedure (5) described in the previous section for every block of 32×32 pels.

6. EXPERIMENTAL RESULTS

To evaluate coding performance of the proposed scheme under the same condition with [3], Bayer CFA images used for our experiments are generated by resampling the original 24-bit RGB images known as Kodak Photo CD PCD0992 dataset (768×512 pels). Table 1 lists coding rates of our coding schemes based on the block-adaptive prediction. In this table, "FBS (Fixed Block-Size)" is a variant of the proposed scheme where block-size is always fixed to 8×8 pels for adaptive prediction. "RGB mixed" and "RGB split" means application of our previous work [5] which was designed for monochrome images. In the case of "RGB mixed", a whole Bayer CFA image is encoded at once ignoring difference of color signals, while "RGB split" separates the Bayer CFA image into three components of R, G and B signals and encodes them individually. In these schemes, coding conditions listed in Table 2 are used. These values were determined though preliminary experiments so that they gave the best performance in average for the respective schemes. In "RGB mixed" scheme, adoption of larger prediction order ($K = 72$) enables use of enough number of pels which have the same color with the current pel $\boldsymbol{p}$ in prediction. This is the reason why "RGB mixed" attains better coding performance than "RGB split" which cannot use any other color signals. On the other hand, an average coding rate of "FBS" slightly lower than that of "RGB mixed" even though the former has a restriction of fixed block-size. Furthermore, the proposed scheme ("VBS") provides the best coding performance by allowing

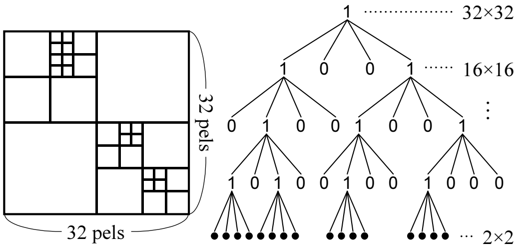


Fig. 4: Example of quadtree-based variable block-size partitioning.

Table 1: Comparison of coding rates (bits/pel).

Image	VBS (Proposed)	FBS	RGB mixed	RGB split
KD01	**5.106**	5.114	5.113	5.535
KD06	**4.401**	4.415	4.438	4.834
KD13	**5.774**	5.795	5.811	6.197
KD19	**4.338**	4.353	4.352	4.622
KD21	**4.423**	4.437	4.454	4.718
Average	***4.808***	*4.823*	*4.834*	*5.181*

Table 2: Coding conditions.

	K	M	Block-size
VBS (Proposed)	42	20	Variable
FBS	42	20	Fixed (8×8)
RGB mixed	72	30	Variable
RGB split	40/30*	30/20*	Variable

*in the case of R and B signals

Fig. 5: Test image (KD06) and the result of variable block-size partitioning

Table 3: Performance comparison with the state-of-the-art lossless coding schemes (bits/pel).

Image	Proposed	OVP[3]	JPEG -LS	JPEG 2000
KD01	**5.106**	5.51	6.398	5.810
KD06	**4.401**	4.67	5.862	5.207
KD13	**5.774**	6.07	6.735	6.369
KD19	**4.338**	4.61	5.686	4.923
KD21	**4.423**	4.71	5.469	5.035
Average	***4.808***	*5.11*	*6.030*	*5.469*

variable block-size adaptive prediction. From the result of variable block-size partitioning, an enlarged view of which is shown in Fig. 5, we can see that appropriate block-sizes according to complexity of local textures are obtained.

Table 3 also lists coding rates of the proposed scheme together with those of other lossless coding schemes. “OVP” (Optimal Vector Prediction) relies on a backward adaptive prediction technique which was specially designed for Bayer CFA images in [3]. “JPEG-LS” and “JPEG 2000” are standard coding schemes for still images and their algorithms are specified in [6] and [7] respectively. For both of the standard schemes, a whole Bayer CFA image is encoded at once in a similar way to “RGB mixed”. This causes serious degradation in coding efficiency of “JPEG-LS” because the MED (Median Edge Detector) predictor adopted in JPEG-LS uses only three pels having different color signals. This is a crucial difference with the case of “RGB mixed”. On the other hand, “JPEG 2000” indicates relatively better coding performance because sampling points in the lifting-based wavelet transform have a conformity with the Bayer pattern as noted in [2]. However, we can say that the coding schemes designed for the Bayer CFA images have much advantage in coding efficiency. Especially, the proposed scheme achieves the best results for all of the tested images. Moreover, it is worth noting that the proposed scheme has another advantage of decoding speed over the “OVP” scheme, because the optimization procedures are carried out only at the encoder side.

7. CONCLUSIONS

In this paper, we have proposed an efficient lossless coding scheme for Bayer CFA images. The scheme is based on a block-adaptive prediction method which was developed for monochrome images, but formation of linear predictors is extended to exploit both spatial and spectral correlations as far as possible. Experimental results show that the proposed scheme attains approximately 6 % better coding performance than “OVP” scheme which was specially designed for Bayer CFA images. Complexity reduction in encoding procedures as well as extension to the Bayer CFA images having a higher bit-depth will be our future tasks.

8. REFERENCES

[1] B.E. Bayer: “Color Imaging Array”, U.S. patent, no. 3,971,065 (1976).

[2] N. Zhang and X. L. Wu: “Lossless Compression of Color Mosaic Images”, IEEE Trans. Image Processing, **15**, 6, pp.1379–1388, (Jan. 2006).

[3] S. Andiriani, G. Calvagno and D. Menon: “Lossless Compression of Bayer Mask Images Using an Optimal Vector Prediction Technique”, Proc. of 14th European Signal Processing Conf. (EUSIPCO 2006), Wed.5-3, (Sep. 2006).

[4] K.H. Chung and Y.H. Chan: “A Lossless Compression Scheme for Bayer Color Filter Array Images”, IEEE Trans. Image Processing, **17**, 2, pp.134–144, (Jan. 2008).

[5] I. Matsuda, N. Ozaki, Y. Umezu and S. Itoh: ”Lossless Coding Using Variable Block-Size Adaptive Prediction Optimized for Each Image”, Proceedings of 13th European Signal Processing Conf. (EUSIPCO 2005), WedAmPo3, (Sep. 2005).

[6] ISO/IEC, ISO/IEC 14495-1:1999: “Information Technology – Lossless and Near-Lossless Compression of Continuous-Tone Still Images: baseline”, (Dec. 1999).

[7] ISO/IEC JTC 1/SC 29/WG 1, ISO/IEC FCD 15444-1: “Information Technology – JPEG 2000 Image Coding System”, (Mar. 2000).