15th European Signal Processing Conference (EUSIPCO 2007), Poznan, Poland, September 3-7, 2007, copyright by EURASIP

A LOSSLESS COMPRESSION SCHEME FOR BAYER CFA IMAGES

King-Hong Chung and Yuk-Hee Chan

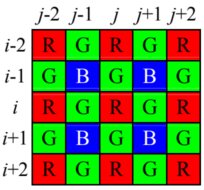
Centre for Signal Processing
Department of Electronic and Information Engineering
The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

In most digital cameras, demosaicing and compression are generally performed sequentially. Recently, it was found that compression-first schemes outperform the conventional demosaicing-first schemes in terms of image quality. An efficient lossless compression scheme for Bayer CFA images is presented in this paper. It exploits a context matching technique to rank the neighboring pixels when predicting a pixel. The prediction residue is then encoded with an adaptive coding scheme using Rice code. Simulation results show that the proposed algorithm can achieve a better compression performance as compared with conventional lossless CFA image coding methods.

1. INTRODUCTION

Most digital cameras use a single image sensor to capture scene images. In these cameras, Bayer color filter array (CFA) [1], as shown in Fig. 1, is usually coated over a sensor to record only one of the three chromatic components at each pixel location. In general, a CFA image is first interpolated via a demosaicing process [2-5] to form a full color image which is then compressed for storage. Fig. 2a shows the workflow of this imaging chain.



	j-2	*j*-1	*j*	*j*+1	*j*+2
i-2	R	G	R	G	R
i-1	G	B	G	B	G
i	R	G	R	G	R
i+1	G	B	G	B	G
i+2	R	G	R	G	R

Fig. 1 – Bayer pattern having a red sample as its center

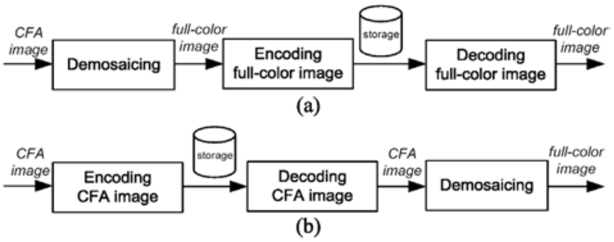


Fig. 2 Single-sensor camera imaging chain: (a) the demosaicing-first scheme, (b) the compression-first scheme

Recently, some reports [6,7] indicated that such a demosaicing-first scheme was inefficient as the demosaicing process introduced redundancy which should eventually be removed in the compression step. As a result, an alternative approach [6,7] which carries out compression before demosaicing as shown in Fig. 2b has been proposed lately. Under this new strategy, digital camera can have a simpler design and lower power consumption as the computationally heavy processes like demosaicing can be carried out in an offline powerful personal computer. This motivates the demand of CFA image compression techniques. [6-8] are some good examples in which lossy compression techniques are used to compress a CFA image.

In some high-end photography applications, original CFA images are required for producing high quality full color images directly. In such cases, lossless compression of CFA images is necessary. Some lossless image compression methods like JPEG-LS [9] and JPEG2000 [10] can be used to encode a CFA image but only a fair performance can be attained. Recently, an advanced lossless CFA image compression algorithm (LCMI) [11] was proposed. In this algorithm, the mosaic data is de-correlated by the Mallat wavelet packet transform and the coefficients are then compressed by adaptive Rice code.

In this paper, a simple prediction-based lossless CFA compression scheme is presented. It employs context matching technique to rank the neighboring pixels for predicting the current pixel. In addition, an adaptive color difference estimation technique is also used to remove the color spectral redundancy. Experimental results show that the proposed compression method can effectively and efficiently reduce the redundancy in both spatial and color spectral domains. As compared with the existing lossless CFA image coding algorithms, the proposed scheme provides the best compression performance.

This paper is structured as follows. In the next section, the proposed context matching based prediction scheme is presented. In Sections 3 and 4, the description of an adaptive color difference estimation technique and the structure of the proposed compression scheme are, respectively, provided. Section 5 demonstrates some simulation results and, finally, a conclusion is given in Section 6.

2. CONTEXT MATCHING BASED PREDICTION

The proposed prediction scheme handles the green plane and the non-green planes separately in a raster scan manner. It reorders the neighboring samples such that the one has higher context similarity to that of the current sample will contribute more to the current prediction.

Let us consider the prediction on the green plane first. Assume that we are now processing a particular green sample $g(i,j)$ as shown in Fig 3a. The four nearest neighboring green samples of $g(i,j)$ form a candidate set $\Phi_{g(i,j)}=\{g(i,j-2), g(i-1,j-1), g(i-2,j), g(i-1,j+1)\}$. The candidates are ranked by comparing their support regions (i.e. context) with that of $g(i,j)$. The support region of a green sample at position (i,j), $S_{g(i,j)}$, is defined as shown in Fig. 4a. In formulation, we have $S_{g(i,j)}=\{(i,j\text{-}2), (i\text{-}1,j\text{-}1), (i\text{-}2,j), (i\text{-}1,j\text{+}1)\}$. The matching extent of the support region of $g(i,j)$ and the support region of $g(m,n)$ for $g(m,n) \in \Phi_{g(i,j)}$ is measured by

$$D(S_{g(i,j)},S_{g(m,n)}) = \left|g_{(i,j-2)} - g_{(m,n-2)}\right| + \left|g_{(i-1,j-1)} - g_{(m-1,n-1)}\right| + \left|g_{(i-2,j)} - g_{(m-2,n)}\right| + \left|g_{(i-1,j+1)} - g_{(m-1,n+1)}\right| \quad (1)$$

Let $g(m_k,n_k) \in \Phi_{g(i,j)}$ for k=1,2,3,4 be the 4 ranked candidates of sample $g(i,j)$ such that $D(S_{g(i,j)},S_{g(m_u,n_u)}) \le D(S_{g(i,j)},S_{g(m_v,n_v)})$ for $1 \le u < v \le 4$. The value of $g(i,j)$ can then be predicted with a prediction filter as

$$\hat{g}(i,j) = round\left(\sum_{k=1}^{4} w_k g(m_k,n_k)\right) \quad (2)$$

where w_k for k=1,2,3,4 are weighting coefficients. In our study, w_k are obtained by quantizing the training result derived by linear regression with a set of training images covering half of the test images shown in Fig. 6. They are quantized to reduce the realization effort of eqn. (2). After all, the coefficients of the prediction filter used to obtain the result presented in this paper are $\{w_1, w_2, w_3, w_4\}$={5/8, 2/8, 1/8, 0} and the predicted green sample is given by

$$\hat{g}(i,j) = round\left(\frac{4g(m_1,n_1) + 2g(m_2,n_2) + g(m_1,n_1) + g(m_3,n_3)}{8}\right) \quad (3)$$

As for the case when the sample being processed is a red or blue sample, the prediction is carried out in the color difference domain instead of the green color plane as before.

Let $d(p,q)$ be the green-red (or green-blue) color difference value of a non-green sample $c(p,q)$. Its determination will be discussed in detail in Section 3. For any non-green sample $c(i,j)$, its candidate set becomes $\Phi_{c(i,j)}=\{d(i,j\text{-}2), d(i\text{-}2,j\text{-}2), d(i\text{-}2,j), d(i\text{-}2,j\text{+}2)\}$ and its support region (context) is defined as $S_{c(i,j)}=\{(i,j\text{-}1), (i\text{-}1,j), (i,j\text{+}1), (i\text{+}1,j)\}$ as shown in Fig. 3b and Fig. 4b, respectively.

The prediction for a non-green sample is carried out in color difference domain. Specifically, the predicted color difference value of sample $c(i,j)$ is given by

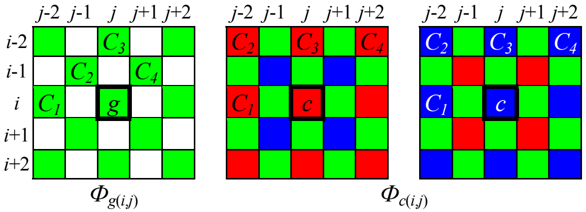
$$\hat{d}(i,j) = round\left(\sum_{k=1}^{4} w_k d(m_k,n_k)\right) \quad (4)$$

where w_k and $d(m_k,n_k)$ are, respectively, the k^{th} predictor coefficient and the k^{th} ranked candidate in $\Phi_{c(i,j)}$ such that $D(S_{c(i,j)},S_{c(m_u,n_u)}) \le D(S_{c(i,j)},S_{c(m_v,n_v)})$ for $1 \le u < v \le 4$, where

$$D(S_{c(i,j)},S_{c(m,n)}) = \left|g_{(i,j-1)} - g_{(m,n-1)}\right| + \left|g_{(i,j+1)} - g_{(m,n+1)}\right| + \left|g_{(i-1,j)} - g_{(m-1,n)}\right| + \left|g_{(i+1,j)} - g_{(m+1,n)}\right| \quad (5)$$

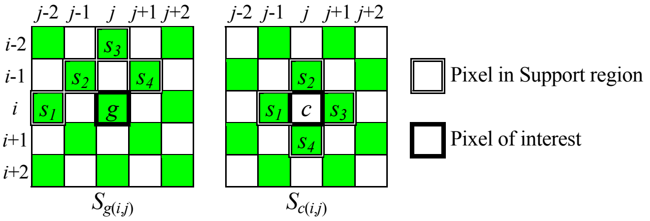
Again, w_k are trained with the same set of training images used to train the predictor coefficients in eqn.(2). For the compression results reported in this paper, the predictor used for the color difference prediction is

$$\hat{d}(i,j) = round\left(\frac{4d(m_1,n_1) + 2d(m_2,n_2) + d(m_3,n_3) + d(m_4,n_4)}{8}\right). \quad (6)$$



(a) (b)

Fig. 3 The candidate set of (a) a green sample and (b) a red/blue sample



(a) (b)

Fig. 4 The support region of (a) a green sample and (b) a red/blue sample

In the proposed compression scheme, all green, red and blue pixels are encoded respectively in a raster scan manner. The four samples used for predicting sample $g(i,j)$ in eqn.(2) are $g(i,j)$'s closest processed neighboring samples of the same color. They have the highest correlation to $g(i,j)$ in different directions and hence can provide a good prediction result even in an edge region. A similar argument applies to explain why $\Phi_{c(i,j)}$ is used to handle blue/red samples in eqn.(4).

As for the support regions, no matter the concerned sample is green or not, as shown in Fig. 4, their supports are defined based on their four closest known green samples. This is because the green channel has a double sampling rate as compared with the other channels in a CFA image and hence provides a more reliable context for matching.

In the proposed scheme, as green samples are encoded first in raster sequence, all green samples are known in the decoder and hence the support of a non-green sample can be non-causal while the support of a green sample has to be causal. This non-causal support tightly and completely encloses its sample of interest. It models image features such as intensity gradient, edge orientation and textures better such that more accurate support matching can be achieved.

3. ADAPTIVE COLOR DIFFERENCE ESTIMATION

When compressing the non-green color plane, color differences are exploited to remove the color spectral dependency. Let $c(m,n)$ be the intensity value of the available color (either red or blue) at a non-green sampling position (m,n). The green-red (green-blue) color difference of pixel (m,n), $d(m,n)$, is determined by

$$d(m,n) = \hat{g}(m,n)\text{-}c(m,n) \quad (7)$$

where $\hat{g}(m,n)$ represents an estimate of the missing green component at position (m,n). In the proposed estimation, $\hat{g}(m,n)$ is adaptively determined according to the horizontal gradient δH and the vertical gradient δV at (m,n) as follows.

$$\hat{g}(m,n) = round\left(\frac{\delta H \times G_V + \delta V \times G_H}{\delta H + \delta V}\right) \qquad (8)$$

where $G_H = \frac{g(m,n-1)+g(m,n+1)}{2}$ and $G_V = \frac{g(m-1,n)+g(m+1,n)}{2}$ donate, respectively, the preliminary green estimates obtained by linearly interpolating the adjacent green samples horizontally and vertically. Note that, in eqn.(8), the missing green value is determined in such a way that a preliminary estimate contributes less if the gradient in the corresponding direction is larger. The weighing mechanism automatically directs the estimation process along an edge if there is.

Gradients δV and δH are determined by averaging all local green gradients in the same direction within a 5×5 window as

$$\delta V = \frac{1}{5} \sum_{\substack{(p,q)\in\{(m-2,n-1),\\(m-2,n+1),(m-1,n),\\(m,n-1),(m,n+1)\}}} \left|g(p,q) - g(p+2,q)\right|$$

and

$$\delta H = \frac{1}{5} \sum_{\substack{(p,q)\in\{(m-1,n-2),\\(m+1,n-2),(m,n-1),\\(m-1,n),(m+1,n)\}}} \left|g(p,q) - g(p,q+2)\right| \cdot \qquad (9)$$

4. PROPOSED COMPRESION SCHEME

Fig. 5 shows the proposed compression scheme. In the encoding phase, a CFA image is first divided into 2 sub-images: a green sub-image which contains all green samples of a CFA image and a non-green sub-image which holds the rest samples. The green sub-image is coded first and the non-green sub-image follows based on the green sub-image as a reference.

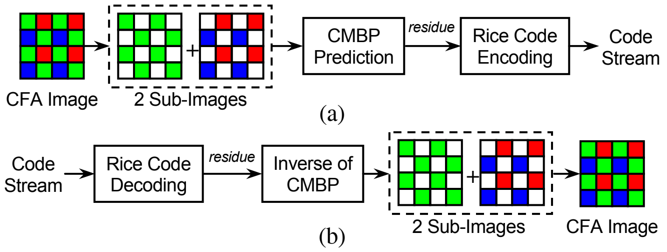


Fig. 5 Structure of the proposed compression scheme: (a) encoder and (b) decoder

To code a sub-image, the sub-image is raster-scanned and each pixel is predicted with its 4 neighboring pixels by using *context matching based prediction* (CMBP), the prediction scheme proposed in Section 2. Its prediction error, say $e(i,j)$, is given by

$$e(i,j) = \begin{cases} \hat{g}(i,j) - g(i,j) & \text{when } (i,j) \text{ is in green sub-image} \\ \hat{d}(i,j) - d(i,j) & \text{when } (i,j) \text{ is in non-green sub-image} \end{cases} \qquad (10)$$

where $g(i,j)$ and $d(i,j)$ are, respectively, the real green sample value and the color difference of pixel (i,j). $d(i,j)$ is estimated by the method described in Section 3. $\hat{g}(i,j)$ and $\hat{d}(i,j)$, respectively, represent the predicted value of $g(i,j)$ and the predicted value of $d(i,j)$. The error residue $e(i,j)$ is then mapped to a non-negative integer as follows to reshape its value distribution from a Laplacian one to an exponential one for Rice code.

$$E(i,j) = \begin{cases} 2e(i,j) & \text{if } e(i,j) \ge 0 \\ -2e(i,j) - 1 & \text{otherwise} \end{cases} \qquad (11)$$

Rice code is exploited to code $E(i,j)$ because of its simplicity and high efficiency in handling exponentially distributed sources. When Rice code is used, each mapped residue $E(i,j)$ is split into a quotient Q=floor($E(i,j)/2^k$) and a remainder $R=E(i,j)$mod(2^k), where parameter k is a non-negative integer. The quotient and the remainder are then saved for storage or transmission. The length of the codeword used for representing $E(i,j)$ is k–dependent and is given by

$$L(E(i,j)\,|\,k) = floor\left(\frac{E(i,j)}{2^k}\right) + 1 + k \,. \qquad (12)$$

Parameter k is critical to the compression performance as it determines the code length of $E(i,j)$. For a geometric source $\boldsymbol{S}$ with distribution parameter $\rho\in(0,1)$ (i.e. Prob($\boldsymbol{S}=s$) $=(1-\rho)\rho^s$ for $s=0,1,2,\ldots$), the optimal coding parameter k is given as

$$k = \max\left\{0, ceil\left(\log_2\left(\frac{\log\phi}{\log\rho^{-1}}\right)\right)\right\} \qquad (13)$$

where $\phi = (\sqrt{5}+1)/2$ is the golden ratio [12]. Since the expectation value of the source is given by $\mu=\rho(1-\rho)^{-1}$, as long as μ is known, parameter ρ and hence the optimal coding parameter k for the whole source can be determined easily.

In the proposed scheme, μ is estimated adaptively in the course of encoding $\{E(i,j)|$ for all (i,j) in a particular sub-image$\}$. In particular, it is estimated by

$$\tilde{\mu} = round\left(\frac{\alpha\tilde{\mu}_p + M_{i,j}}{1+\alpha}\right) \text{ and } M_{i,j} = \left(\frac{1}{4}\sum_{(a,b)\in\zeta_{i,j}} E(a,b)\right) \quad (14)$$

where $\tilde{\mu}$ is the current estimate of μ for selecting the k to determine the codeword format of the current $E(i,j)$, $\tilde{\mu}_p$ is the previous estimate of $\tilde{\mu}$, $M_{i,j}$ is the local mean of $E(i,j)$ in a local region defined by Set $\zeta_{i,j}$, and α is a weighting factor which specifies the significance of $\tilde{\mu}_p$ and $M_{i,j}$ when updating $\tilde{\mu}$. Set $\zeta_{i,j}$ is defined to be $\{(i,j-2), (i-1,j-1), (i-2,j), (i-1,j+1)\}$ when coding the green sub-image. For the non-green sub-image, Set $\zeta_{i,j}$ is defined to be $\{(i,j-2),(i-2,j-2),(i-2,j),(i-2,j+2)\}$. Experimental results showed that $\alpha=1$ can provide a good compression performance. $\tilde{\mu}$ is updated for each $E(i,j)$. The initial value of $\tilde{\mu}_p$ is 0 for both sub-images.

The decoding process is just the reserve process of encoding. The green sub-image is decoded first and then the non-green sub-image is decoded with the decoded green sub-image as a reference. The original CFA image is then reconstructed by combining the two sub-images.

5. COMPRESSION PERFORMANCE

Simulations were carried out to evaluate the compression performance of the proposed scheme. Twenty-four 24-bit color images of size 512×768 each as shown in Fig. 6 were Bayer sub-sampled to form a set of 8-bit testing CFA images. They were then directly coded by the proposed compression scheme. Some representative lossless compression algorithms such as JPEG-LS [9], JPEG 2000 (lossless mode) [10] and LCMI [11] were also evaluated for comparisons.

Table 1 lists the output bit-rates of the CFA images achieved by various algorithms. It clearly shows that the proposed scheme outperforms all other evaluated methods in all testing images. Especially for the images which contain many fine textures such as images 1, 5, 8, 13, 20 and 24, the bit-rates achieved by the proposed scheme are at least 0.32bpp lower than the corresponding bit-rates achieved by LCMI, the method offers the second best compression performance. On average, the proposed scheme yields a bit-rate as low as 4.63bpp. It is, respectively, around 1.27, 0.38 and 0.26 bpp lower than those achieved by JPEG-LS, JPEG2000 and LCCMI.

As a green pixel estimation method is proposed and used when compressing the non-green sub-image in color difference domain, a simulation was also carried out to evaluate its performance. For comparison, some other estimation methods such as bilinear interpolation [4] (BI), edge sensing interpolation [5] (ESI) and adaptive directional interpolation [2] (ADI) were also evaluated. To provide a clear demonstration, only the non-green sub-image was coded in this simulation.

Table 2 reveals the average bit rates of the outputs achieved by various estimation algorithms. It shows that the proposed adaptive estimation method yields the best compression performance among the evaluated estimation methods. On average, the proposed estimation method achieves a bit-rate of 4.49bpp which is around 0.09 bpp lower than that achieved by BI.

6. CONCLUSIONS

In this paper, a simple lossless compression scheme for Bayer images is proposed. A context matching technique is used to rank the neighboring pixels of a pixel for predicting the sample value of the pixel. The prediction residue is then encoded with Rice code adaptively. Experimental results show that the proposed compression scheme can efficiently and effectively decorrelate the data dependency in both spatial and color spectral domains. Consequently, it provides the best average compression ratio as compared with the latest lossless Bayer image compression algorithms.

7. ACKNOWLEDGEMENT

This work was supported by the RGC of Hong Kong Special Administrative Region (PolyU5205/04E) and The Hong Kong Polytechnic University (PolyU Grant G-YE90).

Fig. 6 Twenty-four digital color images (Refers as image 1 to image 24, from top-to-bottom and left-to-right)

Image	**JPEG-LS**	**JPEG 2000**	**LCMI**	**Ours**	**Image**	**JPEG-LS**	**JPEG 2000**	**LCMI**	**Ours**
1	6.403	5.816	5.824	5.497	**13**	6.747	6.372	6.503	6.130
2	6.787	5.134	4.629	4.329	**14**	6.289	5.555	5.487	5.168
3	5.881	4.216	3.965	3.745	**15**	6.317	4.656	4.396	4.098
4	6.682	4.931	4.606	4.367	**16**	5.289	4.552	4.521	4.387
5	6.470	5.947	5.859	5.427	**17**	4.965	4.547	4.499	4.286
6	5.871	5.210	5.139	4.894	**18**	6.184	5.570	5.538	5.274
7	5.974	4.500	4.299	3.989	**19**	5.470	4.909	4.898	4.747
8	6.295	5.899	5.966	5.635	**20**	4.317	4.026	4.054	3.544
9	5.074	4.391	4.319	4.192	**21**	5.467	5.039	4.983	4.804
10	5.395	4.556	4.415	4.226	**22**	6.188	5.218	5.060	4.842
11	5.370	4.986	4.952	4.693	**23**	6.828	4.525	3.960	3.839
12	5.628	4.485	4.307	4.097	**24**	5.719	5.223	5.257	4.895
					Avg.	**5.900**	**5.011**	**4.893**	**4.629**

Table 1 Achieved bit-rates of various lossless compression algorithms in terms of bits per pixel (bpp)

Image	**BI**	**ESI**	**ADI**	**Ours**	**Image**	**BI**	**ESI**	**ADI**	**Ours**
1	5.414	5.411	5.349	5.242	**13**	5.969	5.967	5.981	5.919
2	4.366	4.365	4.367	4.320	**14**	5.030	5.028	5.006	4.939
3	3.724	3.721	3.709	3.669	**15**	4.192	4.186	4.161	4.145
4	4.314	4.312	4.304	4.276	**16**	4.295	4.291	4.275	4.176
5	5.312	5.301	5.267	5.199	**17**	4.202	4.201	4.163	4.134
6	4.851	4.849	4.802	4.708	**18**	5.117	5.116	5.126	5.089
7	3.890	3.893	3.881	3.853	**19**	4.705	4.703	4.627	4.563
8	5.574	5.570	5.447	5.317	**20**	3.701	3.698	3.676	3.664
9	4.154	4.156	4.113	4.077	**21**	4.737	4.735	4.723	4.676
10	4.173	4.171	4.142	4.088	**22**	4.732	4.732	4.730	4.694
11	4.625	4.618	4.586	4.494	**23**	3.844	3.841	3.838	3.828
12	4.085	4.083	4.063	3.989	**24**	4.787	4.785	4.771	4.675
					Avg.	**4.575**	**4.572**	**4.546**	**4.489**

Table 2 Achieved bit-rates (bpp) for coding non-green planes with the proposed coding scheme using various estimation methods to estimate a green sample for reference

REFERENCES

[1] E. Bayer, "Color imaging array", U.S. Patent 3 971 065, July 1976.

[2] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Schafer and R.M. Mersereau, "Demosaicking: color filter array interpolation", IEEE Signal Processing Magazine, vol.2, no. 1, pp.44-54, Sept. 2005.

[3] K.H. Chung and Y.H. Chan, "Color Demosaicing Using Variance of Color Differences", IEEE Trans. on Image Processing, Vol. 15, No.10, pp.2944-2955, Oct 2006.

[4] T. Sakamoto, C. Nakanishi and T. Hase, "Software pixel interpolation for digital still camera suitable for a 32-bit MCU", IEEE Trans. Consumer Electronics, vol. 44, no. 4, pp.1342-1352, Nov. 1998.

[5] R. Lukac and K. N. Plataniotis, "Data-adaptive filters for demosaicking: A framework", IEEE Trans. Consumer Electronics, vol. 51, no. 2, pp.560-570, May 2005.

[6] X.Xie, G.L.Li, Z.H.Wang, C.Zhang, D.M.Li and X.W.Li, "A novel method of lossy image compression for digital image sensors with Bayer color filter arrays", Proc. ISCAS'05, vol.5, pp.4995-4998, May 2005.

[7] C.C. Koh, J. Mukherjee and S.K. Mitra, "New efficient methods of image compression in digital cameras with color filter array", IEEE Trans. Consumer Electronics, vol. 49, no. 4, pp.1448-1456, Nov. 2003.

[8] A. Bruna, F. Vella, A. Buemi and S. Curti, "Predictive differential modulation for CFA compression", Signal Processing Symp. Proc. 2004, pp.101-104, Jun. 2004.

[9] ISO/IEC JTC1/SC29/WG1, "Information technology – lossless and near-lossless compression of continuous-tone still image", ISO FDIS 14495-1 (JPEG-LS), 1998.

[10] ISO/IEC JTC1/SC29/WG1 N1890, "JPEG2000 Part I final draft international standard", 2000.

[11] N. Zhang and X. Wu, "Lossless compression of color mosaic images", IEEE Trans. Image Processing, vol.15, no. 6, pp.1379-1388, Jun. 2005.

[12] A. Said, "On the determination of optimal parameterized prefix codes for adaptive entropy coding", Technical Reports HPL-2006-74, Hewlett Packard Laboratories, Apr. 2006