Color Interpolation using Markov Random Field Modeling

Jayanta Mukherjee R.Parthasarathi S.Goyal

Department of Computer Science & Engineering,
Indian Institute of Technology
Kharagpur, INDIA 721302

Abstract

In digital color imaging, color filter arrays (CFAs) are obtained from single chip cameras in the form of sampled spectral components (red, green and blue) in an interleaved fashion. Color demosaicing is the process of interpolating these color filter arrays (CFAs) into the dense pixel maps for each spectral component. There are different interpolation techniques for color demosaicing operations. These approaches have their limitations regarding the improvement of the quality of the images. They perform poorly in recovering the edges of the images. In this work we have applied Markov Random Field (MRF) processing over these roughly interpolated images to improve the quality of the reconstruction. We have observed that the processing improves the image quality in many cases. The edges of the reconstructed images are significantly enhanced using MRF processing.

1 Introduction

Single-chip cameras [1] [2] for color imaging use color filter arrays (CFAs) to obtain sampled red, green and blue pixel data (or *luminance* and *chrominance* signals) in an interleaved fashion. There are different checker-board patterns [1] [2] for obtaining color filter arrays (CFAs). In our work we have used the Bayer pattern [3] in the CFA. The Bayer pattern is shown in Figure 1. In the figure the sampled color components are denoted by R (for red), G (for green) and B (for blue). Once sampled color pixel data in the CFAs are available, the sparse pixel values are required to be interpolated for obtaining dense pixel maps in all the three spectral components.

The process of interpolating these sparse data into the dense pixel maps is referred to the literature as *color interpolation* or *color demosaicing*. In the early years (late eighties and early nineties), one of the major concerns of this interpolation task was to keep the hardware cost as well as the computation time as small as possible. This is to make the digital color imaging (for both still and video images) cost-effective and technologically viable. Hence the early methods were based on low storage requirement and simple computations. In the present work, efforts are made to improve the quality of the interpolated images obtained by the existing techniques. The process may be considered as the post-operation on the roughly interpolated images obtained from the single-chip cameras.

R	G	R	G
G	B	G	B
R	G	R	G
G	B	G	B

Figure 1: The Bayer pattern.

2 Comparisons among different techniques

Different color demosaicing algorithms are available in the literature. The simplest technique for color interpolation is to replicate the known spectral component [4] [1] from the neighbors of a pixel. The performance greatly improves, if one computes with more than one neighbors having same spectral components in the CFA. In bilinear interpolation techniques the average of such neighboring values is computed. There are also techniques which compute weighted averages of the neighboring spectral components, specified in the CFA [5]. For example, B-Spline bi-cubic polynomials [6] are used for determining these weights.

In other approaches, the color demosaicing algorithms are mainly designed on the following two principles:

1. interpolation with the pixels lying with the low gradient directions [7] [8].

2. use of the homogeneity of the cross-ratios of different spectral components around a small neighborhood [9], [10].

Kimmel [11] also proposed an algorithm by weighting the cross-ratios with the gradient information around the neighborhood of a pixel.

In another variation, instead of considering only horizontal and vertical gradients, gradients in different directions are used in the interpolation process. The directions where the gradient values are less are used in interpolation.

In our work, we have compared these techniques by finding out the PSNR's (Peak Signal To Noise Ratio) with respect to the original images. We have used the Bayer pattern (refer Figure 1) for obtaining the Color Filtered Array (CFA). As the PSNR's do not always reflect the quality of the images in terms of edge reconstructions, we have used another measure PESNR (Peak Edge Signal to Noise Ratio) for reflecting how edges are recovered in the interpolated images. Let $I_s(x,y), s = 1,2,3$ be the spectral components of the benchmark images of size $M \times N$ and $I_s^{'}(x,y), s = 1,2,3$ be the respective reconstructed spectral components. Then PSNR and PESNR are defined below:

$$PSNR = 20 * log(\frac{255}{\sqrt{\frac{\Sigma_{\forall s}\Sigma_{\forall x}\Sigma_{\forall y}(I_s(x,y)-I_s^{'}(x,y))^2}{3\times M\times N}}}) \quad (1)$$

For defining PESNR we have used the binary edge map of an image, which is computed from the gradient image of the CFA. In an edge map $e(x,y)$ of an image, if the value at a pixel location is 1 it shows the presence of edge pixels and otherwise the value is 0. Then, PESNR is defined as:

$$PESNR = 20*log(\frac{255}{\sqrt{\frac{\Sigma_{\forall s}\Sigma_{\forall x}\Sigma_{\forall y}e(x,y)\times(I_s(x,y)-I_s^{'}(x,y))^2}{3\times\Sigma_{\forall x}\Sigma_{\forall y}e(x,y)}}}) \quad (2)$$

In our experimentation we have used following interpolation techniques.

1. Bilinear Interpolation Techniques (BI) [2]
2. B-Spline Interpolation (BSPI) [6]
3. Interpolation by Averaging Red and Blue Hues (ARBH)[9]
4. Laplacian Corrected Edge Correlated Interpolation Techniques (LCEC) [12]

Table 1: **PSNR of Different Interpolation Techniques**

Images	BI	BSPI	ARBH	EDCRAC	LCEC	VNGD
Lena	28.24	25.48	25.27	25.57	**29.45**	29.02
Peppers	**29.45**	21.52	24.47	25.31	27.24	26.94
Mandrill	21.00	20.09	21.48	21.71	21.91	**23.79**
T'mahal	23.38	20.02	24.76	25.67	25.36	**26.47**
L'house	25.09	22.47	27.27	30.61	30.48	**31.46**
Sails	27.14	23.93	28.74	31.36	31.00	**31.81**
Statue	27.85	23.86	29.66	31.90	32.21	**34.08**
Window	26.84	22.81	29.39	32.93	33.09	**33.62**
Flower	22.59	18.55	22.85	24.24	**24.96**	24.58
Bmw	21.68	18.87	21.73	22.27	**22.51**	21.76

Table 2: **PESNR of Different Interpolation Techniques**

Images	BI	BSPI	ARBH	EDCRAC	LCEC	VNGD
Lena	32.93	30.16	**33.36**	20.48	33.12	33.10
Peppers	11.51	10.66	10.33	11.75	12.10	**12.13**
Mandrill	22.09	21.71	21.98	20.25	**23.05**	22.55
T'mahal	19.55	17.32	22.33	20.02	21.34	**23.08**
L'house	19.61	18.91	18.57	17.39	**20.53**	20.04
Sails	18.33	16.04	18.35	23.92	**24.60**	22.95
Statue	12.91	12.44	11.88	13.48	13.70	**13.79**
Window	15.17	11.78	18.52	**24.42**	24.09	21.13
Flower	18.61	14.12	21.71	22.63	25.11	**26.60**
Bmw	22.01	13.46	25.12	20.30	28.74	**31.09**

5. Edge Directed Cross-Ratio Averaging and Corrections (EDCRAC)[11]
6. Interpolation with the values lying in Variable Number of Gradient Directions (VNGD)

We have used acronyms against each technique, which will be used in our subsequent discussions.

The PSNR and the PESNR values obtained by different techniques are presented in the Tables 1 and 2. We found that the quality of reconstruction from the VNGD method is very good and often the best among all the techniques considered here (refer Table 1 and 2). We have also noted the computation times for different interpolation techniques. We found that bilinear (BI), ARBH and LCEC are computationally more efficient than the other methods such as BSPI, EDCRAC and VNGD. Out of these last three the VNGD is the slowest one. Interestingly the LCEC method is both computationally efficient and perform very close to the VNGD in terms of quality of reconstruction. In some cases the performance of the LCEC is the best among all the techniques used in our experimentation.

3 Markov Random Field (MRF): Preliminaries

Let $X = X_s, s \in S$ denote a family of random variables indexed by site s. Let Ω denote the space for all possible configurations of X. In the present context, X_s assumes the pixel values for any spectral component from a finite set of B (=256, in our case) levels at location $s \in S$ over an $M \times N$ lattice such that $S = \{(i,j)|1 \leq i \leq M, 1 \leq j \leq N\}$. Let us define $G = \{G_s|s \in S\}$ as a *General Neighborhood System* of S as follows ,

Definition 1: G is a *General Neighborhood System* if it satisfies the following properties:

1. $G_s \subset S$
2. $s \notin G_s$
3. $\forall s, r \in S, s \in G_r \Leftrightarrow r \in G_s$

The sites $r \in G_s$ are called the *neighbors* of s. There are different neighborhood structures over S. In the present context, over a lattice of $M \times N$ the nth order neighborhood of $s(= (i,j))$ is defined below:

Definition 2: An nth order neighborhood $\eta^{(n)}(s)$ of s is defined as: $\eta^{(n)}(s) = \{r|d(s,r) \leq n, r \neq s\}$, where $d(.)$ is a distance function.

In our work we have considered the *Chess-board Distance* for the neighborhood definition and we have restricted ourselves to the *first order* neighborhood system. The pair (S,G) is called a *graph*.

Let us now define the *Markov Random Field (MRF)* as follows:

Definition 3. X is a *MRF* with respect to the general neighborhood system G if

1. $P(X = \omega) \geq 0, \forall \omega \in \Omega$
2. $P(X_s = x_s|X_r = x_r, r \neq s, \forall r \in S) = P(X_s = x_s|X_r = x_r, \forall r \in G_s)$

where $P(.)$ and $P(.|.)$ denote the probabilities and conditional probabilities respectively.

The interesting part of MRF processing is that both the probability distribution of X and conditional probability distributions $P(X_s|X_r, \forall r \in S, r \neq s)$ follow *Gibbs Distribution* [13] which is of the following form:

Definition 4. A *Gibbs Distribution* with respect to a graph (S,G) is a probability distribution π on configuration Ω such that: $\pi(\omega) = \frac{e^{-U(\omega)/T}}{Z}$, where Z is a *normalizing constant* and T is the *temperature*. U is called the *energy function*.

As Z is the normalizing constant it follows that:

$$Z = \Sigma_{\forall \omega \in \Omega} e^{-U(\omega)/T} \tag{3}$$

The *energy function* U takes the following form:

$$U(\omega) = \Sigma_{c \in C} V_c(\omega) \tag{4}$$

where $V_c(.)$ is the *potential function* defined over *cliques* C, as defined below:

Definition 5: A *Clique* $c \in C$ is defined as a subset of S such that every pair of distinct sites in c are *neighbors*. A single site s is also defined as a *clique*.

In restoring or interpolating the images, one may consider the *Bayesian approach* [13] by maximizing the *a posteriori probability*. The *MRF* modeling is suited to this approach as the functional forms of the probability distributions are known to us.

4 MRF processing for Color Demosaicing

Before going into the description of MRF processing, first we will define a few notations and terminologies which are used in our description. We denote the red, green and spectral components of a color image as $I_s(i,j), 1 \leq i \leq M, 1 \leq j \leq N$ for $s = 1, 2$ and 3 respectively. We obtain a CFA using color masking functions which are defined below.

Let

$$\begin{aligned} f_1(i,j) &= 1 \text{ if mask at } (i,j) \text{ is red} \\ &= 0 \text{ Otherwise} \end{aligned}$$

Similarly, $f_2(i,j)$ and $f_3(i,j)$ are defined for green and blue masks respectively. It may be noted that the mask functions are mutually exclusive which implies:

$$\Sigma_{s=1}^{3} f_s(i,j) = 1 \tag{5}$$

Using these color masks we can obtain Color Filter Array (CFA) as follows:

$$c(i,j) = \Sigma_{s=1}^{3}(f_s(i,j) I_s(i,j)) \tag{6}$$

The CFA's are subjected to color interpolation technique to reconstruct back the images as $I_s(i,j)$, for $s = 1, 2$ and 3. Initially we will have an interpolated image $I_s^{(0)}(i,j), s = 1, 2$ and 3 using any one of the algorithms mentioned previously. These initial estimates are subjected to the MRF processing for further enhancement of the images.

4.1 The maximum a posteriori (MAP) estimation approach

In our case we will consider the application of MRF modeling in maximizing the a posteriori probability of the interpolated image. A detailed discussion of this approach has been made in [14]. We will consider the observed image as the $I_s^{(0)}(i,j), 1 \leq i \leq M, 1 \leq j \leq$

$N, 1 \leq s \leq 3$, which have been obtained by any one of the existing interpolation technique (say by *bilinear interpolation method*). We assume here that an additive *Gaussian noise* corrupts the true spectral components $I_s(i,j)$ to produce the resulting observed image $I_s^{(0)}(i,j)$. Hence they are related as follows:

$$I_s^{(0)}(i,j) = I_s(i,j) + n(i,j) \tag{7}$$

In the above equation $n(i,j)$ follows a Gaussian distribution $N(0, \sigma(i,j))$. We also assume that $n(i,j)$'s are *spatially* and *spectrally uncorrelated*. Hence the form of the energy function to be minimized here for maximizing the a posteriori probability using the MRF model is given below (see [14] for a detailed discussion):

$$U(\omega) = \Sigma_{c \in C} V_c(\omega) + \Sigma_{\forall s} \Sigma_{\forall i} \Sigma_{\forall j} \frac{(I_s(i,j) - I_s^{(0)}(i,j))^2}{2\sigma^2(i,j)} \tag{8}$$

4.2 The Energy Model

MRF processing is heavily dependent on *the energy model*. In *the energy model* the desirable features of the images should be enhanced. We have considered mainly two factors while minimizing the energy. They are stated below.

1. The interpolated image should be smooth. This implies that the deviation of pixel values around its neighborhood should be small.
2. The cross-ratios remain more or less same around its neighborhood.

The energy model used in our work is the linear combination of the above two factors. In addition to the above factors, another constraint is imposed in the interpolated images, by which the values of the CFA's for a spectral component remain unchanged. The other aspect of the energy model is its dependence on gradient or edge information in the CFA's. We propose a Gaussian weight-age model for enhancing the factors along low gradient directions as follows:

$$\Phi(i+m, j+n) = \frac{1}{\sqrt{2\pi\sigma(i,j)}} e^{-\frac{(c(i,j) - \Sigma_{s=1}^{3} (f_s(i,j).I_s(i+m,j+n))^2}{2\sigma(i,j)}} \tag{9}$$

The $\sigma(i,j)$ of the Gaussian mask is a function of the minimum gradient value (absolute difference between two neighboring pixel values). In our work we have considered the following function for the $\sigma(i,j)$.

$$\begin{aligned} \sigma(i,j) &= min\{\frac{|c(i+m,j+n) - c(i-m,j-n)|}{2\sqrt{m^2+n^2}}\} + \sigma_{thresh} \\ &\quad \forall (m,n) \in \eta^{(1)}(0,0) \end{aligned} \tag{10}$$

The value of σ_{thresh} is taken as 10. It may be noted here that for estimating the $\sigma(i,j)$ we have assumed that the gradients of the spectral components are correlated. We have used the same estimate for estimating the strength of the noise contamination $(n(i,j))$ at (i,j) for each spectral component. Now we present *the local energy model* at any pixel by the following expressions:

For the smoothness criterion:

$$\begin{aligned} U^{(sm)}(i,j) &= \Sigma_{m=-p}^{p} \Sigma_{n=-p}^{p} \Sigma_{s=1}^{3} \{(1 - f_s(i,j)). \\ &\quad \Phi(i+m, j+n). \\ &\quad (I_s(i,j) - I_s(i+m, j+n))^2\} \end{aligned} \tag{11}$$

For the constancy of cross-ratios:

$$\begin{aligned} U^{(\rho)}(i,j) &= \Sigma_{m=-p}^{p} \Sigma_{n=-p}^{p} \Sigma_{s=1}^{3} \{(1 - f_s(i,j)). \\ &\quad \Phi(i+m, j+n). \\ &\quad (\frac{I_s(i,j)}{c(i,j)} - \frac{I_s(i+m,j+n)}{\Sigma_{t=1}^{3} f_t(i,j) I_t(i+m,j+n)})^2\} \end{aligned} \tag{12}$$

Hence *the local energy model* is *the linear combination* of the above two terms and an additional term for restoring the image from the additive Gaussian noise (refer equation 8):

$$\begin{aligned} U(i,j) &= \mu U^{(sm)}(i,j) + \lambda U^{(\rho)}(i,j) \\ &\quad + \gamma \Sigma_{s=1}^{3} \{(1 - f_s(i,j)) \frac{(I_s(i,j) - I_s^{(0)}(i,j))^2}{2\sigma^2(i,j)}\} \end{aligned} \tag{13}$$

It is interesting to note that the term $(1 - f_s(i,j))$ puts the constraint of ignoring any change in the true value of the respective spectral component at (i,j) directly obtained from the CFA. In our implementation, the constrained has been forced by not updating the respective spectral component. It may be noted that *the global energy function* is the sum of *the local energy functions* i.e.,

$$U = \Sigma_{\forall i} \Sigma_{\forall j} U(i,j) \tag{14}$$

As we have mentioned earlier, we have restricted ourselves to the first order MRF modeling only, the value of p in our experimentations is taken as 1.

4.3 The Energy Minimization using Stochastic Relaxation

We have used *simulated annealing* technique for minimizing the energy to maximize the *a posteriori probability*. The simulated annealing algorithm is well discussed and described in text books [15] and technical papers [13]. In our work we have started with an initial temperature as 1. It has been observed that the higher values do not improve the result. We iterate the process till we reach at a small temperature taken as 10^{-30}. As *Gaussian* and *Gibbsian sampling*

Table 3: **Gain/Loss in PSNR measures from IMRF processing (MAP approaches) on initial images obtained by different Interpolation Techniques**

Images	BI	BSPI	ARBH	EDCRAC	LCEC	VNGD
Lena	0.70	2.10	2.64	2.54	0.43	0.59
Peppers	1.41	2.59	2.22	2.16	0.61	0.74
Mandrill	0.33	0.77	0.98	0.89	0.65	.099
T'mahal	2.00	1.90	1.91	1.75	1.66	1.05
L'house	1.34	1.69	1.30	0.62	0.19	-0.32
Sails	1.76	2.20	2.06	0.41	0.40	0.28
Statue	1.76	2.66	1.95	0.79	0.47	-0.52
Window	2.39	2.67	2.77	0.83	0.37	-0.04
Flower	1.97	2.15	1.92	1.13	0.75	0.61
Bmw	0.97	1.50	1.05	0.65	0.48	0.49

Table 4: **Gain/Loss in PESNR measures from IMRF processing (MAP approaches) on initial images obtained by different Interpolation Techniques**

Images	BI	BSPI	ARBH	EDCRAC	LCEC	VNGD
Lena	0.77	1.99	0.59	3.31	0.45	0.40
Peppers	1.16	1.09	1.69	1.57	1.00	1.06
Mandrill	0.39	0.55	0.75	1.74	-0.19	-0.02
T'mahal	4.80	2.24	4.60	4.14	4.79	3.72
L'house	0.76	-0.21	0.81	2.08	-0.01	0.30
Sails	3.51	2.85	3.25	1.54	0.23	0.24
Statue	0.91	1.94	0.90	0.94	0.51	0.49
Window	4.31	5.49	3.35	0.96	0.99	0.76
Flower	6.22	1.04	7.91	3.78	6.09	5.99
Bmw	3.64	0.26	5.19	3.61	2.14	1.36

[13] take longer time, we have used *uniform sampling* in our experimentations. We have also adopted two different approaches for the *energy minimization process*. In one approach, all the red, green and blue values are simultaneously updated if they satisfy the *Metropolis Acceptance* [16] criterion. We denote this approach as *simultaneous MRF processing (SMRF)*. On the other hand, we have also considered the energy minimization of green, red and blue spectral components independently one after another. This approach is referred to as *individual MRF processing (IMRF)*. As we have noticed that the performance with SMRF processing is always poorer than that with the IMRF processing, here we present only the results with the IMRF processing.

5 Experimentation

We present here the respective gains or losses in the PSNR and the PESNR values in the Tables 3 and 4. One can observe that in most of the cases there are gains in the PSNR as well as in the PESNR values. One may also note that for the interpolation techniques yielding relatively lower values of PSNR and PESNR (such as BI, BSPI and ARBH), the gains are substantial. On the other hand for other techniques the gains are relatively less and in some cases they are marginal. However for some images such as Tajmahal, Peppers, Flower and Bmw, though the increases in the PSNR values are small, the PESNR values have been improved significantly.

To closely observe the improvements in reconstructed images we have given a typical examples near the edges for the image Tajmahal (Figure 2(a)). In Figure 2(e) one may observe that the edges are improved by applying the IMRF processing over the initial estimates obtained by the VNGD technique.

It may be noted that MRF processing is computationally intensive. It imposes roughly an additional overhead of 3 milli-seconds per pixel in the computation time under the same computational environment.

6 Conclusion

In this work we have applied MRF processing for color demosaicing to improve the quality of the reconstruction. We have first studied the relative performances of different exisiting demosaicing algorithms. We have observed that out of the chosen techniques the VNGD and the LCEC perform better than others. Interestingly, the LCEC takes much less time than the VNGD. We have observed that the edges of the reconstructed images are significantly enhanced using MRF processing. A quantitative measure denoted as PESNR is used here to present the relative performances in the edge reconstructions by various techniques. In some cases there are substantial gains (by 5 to 7 dB) in the PESNR values, though the PSNR values are marginally improved.

References

[1] Kenneth A Parulski. Color filters and processing alternatives for one-chip cameras. *IEEE Transaction on Electron Devices*, ED - 32(8):1381 – 1389, Aug. 1985.

[2] Y. Tim Tsai. Color image compression for single-chip cameras. *IEEE Transaction on Electron Devices*, 38(5):1226 – 1232, 1991.

[3] B.E. Bayer. Color imaging array. (U.S. Patent 3971065), 1976.

[4] Tadashi Sakamoto, Chikako Nakanishi, and Tomohiro Hase. Software pixel interpolation for digital still cameras suitable for a 32-bit mcu. *IEEE*

Transaction on Consumer Electronics, 44(4):1342 – 1352, Nov. 1998.

[5] Don P. Mitchell and A. N. Netravali. Reconstruction filters in computer graphics. *Computer Graphics, (SIGGRAPH'88 Proceedings)*, 22(4):221–228, 1988.

[6] Hsieh S. Hou et. al. Cubic splines for images interpolation and digital filtering. *IEEE Transactions on Acoustic,Speech and Signal Processing*, ASSP-26:508–517, 1987.

[7] R.H. Hibbard. Apparatus and method for adaptively interpolating a full color image utilizing luminance gradients. (US Patent 5382976).

[8] D.R. Cok. Signal processing method and appratus for sampled image signals. (U.S. Patent 4630307).

[9] James E. Adams. Interactions between color plane interpolation and other image processing functions in electronic photography. *Proceedings of SPIE*, 2416:144–151, 1995.

[10] John A. Weldy. Optimized design for a single-sensor color electronic camera system. *Proceedings of SPIE*, 1701:300–307, 1988.

[11] Ron Kimmel. Demosaicing : Image reconstruction from color ccd samples. *IEEE Transactions on Image Processing*, 8(9):1221 – 1228, Sep. 1999.

[12] Hamilton(Jr.) et. al. Adaptive color plane interpolation in single color electronic camera. (US Patent 5629734).

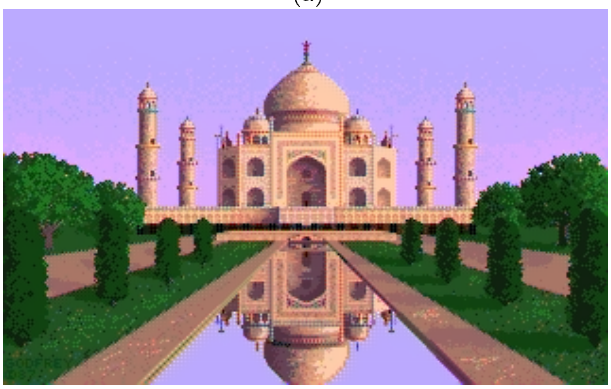
[13] S.Geman and D.Geman. Stochastic relaxation ,gibbs distributions and the bayesian distribution of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6:721–741, 1984.

[14] U.B.Desai. Markov random field models for early vision problems. (Technical Report: SPANN-LAB-93-3):1–27, 1993.

[15] E. Aarts and J. Korst. *Simulated Annealing and Boltzman Machines*. Wiley, Chichester, 1989.

[16] N.Metropolis, A.Rosenbluth, M.Rosenbluth, and E.Teller. Equation of state calculations by fast computing machines. *Journal of Chem. Physics*, 21:1087–1092, 1953.

(a)



(b)

(c)

(d)

(e)

Figure 2: (a) Original (b) VNGD (c) IMRF-VNGD (d) VNGD Reconstructed (e)IMRF-VNGD Reconstructed