A CLASSIFICATION APPROACH TO COLOR DEMOSAICKING

Cindy Kwan and Xiaolin Wu
Department of Electrical and Computer Engineering, McMaster University
Hamilton, Ontario, Canada L8S 4K1
{c2kwan@grads/xwu@mail}.ece.mcmaster.ca

Abstract

Color demosaicking for CCD cameras is a task of estimating missing data. Based on different hypotheses on the image structures multiple estimates can be made. Choosing the best estimate becomes a statistical decision problem. We propose an optimal classification technique based on Fisher's discriminant to solve this decision problem. The new technique is more robust and obtains superior visual quality of demosaicked images than existing methods.

I. INTRODUCTION

Digital cameras which use single charged couple device (CCD) sub-sample the primary color channels: red, green and blue (RGB). The Bayer color filter array (CFA), shown in Figure 1, is commonly used in digital cameras as the color-sampling scheme.

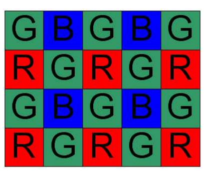

Fig. 1. Bayer color filter array pattern.

At each pixel, only one of the three primary colors is captured. The other two color samples need to be interpolated and this process is called *demosaicking*.

Most papers on color demosaicking attack the problem from the perspective of filter design [1], [2], [3]. Recently, Wu and Zhang proposed a so-called primary-consistent soft-decision (PCSD) approach to color demosaicking [4]. In the PCSD approach, statistical decision making process is introduced to augment interpolation filtering operations. A key element common to most color demosaicking methods is sample interpolation in an estimated edge direction. However, when the image signal has high frequency components beyond the Nyquist limit, it is error prone to estimate the edge direction. If the directional interpolation is carried out cross instead of along the edge, it causes objectionable color artifacts. The PCSD approach aims to minimize the risk of interpolating in a wrong direction. Instead of forcing a decision on the interpolation direction with insufficient information, the PCSD method makes multiple estimates of the missing color sample under different hypotheses (letting conflicting hypotheses compete with one another) on edge directions, and selects the estimate that agrees with the underlying hypothesis. The decision on interpolation direction is delayed until the results of demosaicking under different hypotheses can be evaluated and verified against the hypotheses in a larger context.

In this paper we seek to optimize the statistical decision in the PCSD framework for color demosaicking. This decision problem is converted to one of minimum risk classification based on a set of observable features related to the sample to be estimated and to the particular interpolation filter. To improve the robustness of the existing color demosaicking algorithms we focus on the most difficult cases for directional interpolation methods, and design a classifier of Fisher's linear discriminant aiming for best possible performance in the worst case.

II. PROBLEM FORMULATION AND CLASSIFICATION APPROACH

Among the directional interpolation filters for color demosaicking, the second order Laplacian filter is probably the best known and highly effective in most cases [1], [2], [3]. For concreteness, consider the situation as depicted by Fig. 2, where a missing green sample G_c is to be interpolated from the available green samples: G_n, G_w, G_s, and G_e. Throughout this paper we adopt geographic subscript notations such as n, w, s, e to denote the samples to the north, west, and so on. The current position for the interpolation is labelled by c. In Fig. 2 the original CCD sample at the current position is red, R_c. Depending on

0-7803-8554-3/04/$20.00 ©2004 IEEE.

whether the horizontal or vertical direction is chosen for the interpolation, the second order Laplacian filter is given by

$$\begin{aligned} G^h &= (G_w + G_e)/2 + (2R_c - R_{ww} - R_{ee})/4 \\ G^v &= (G_n + G_s)/2 + (2R_c - R_{nn} + R_{ss})/4 \end{aligned} \quad (1)$$

The missing green sample G_c will be estimated as G^h if $\Delta H < \Delta V$, or as G^v otherwise, where

$$\begin{aligned} \Delta H &= |G_w - G_e| + |2R_c - R_{ww} - R_{ee}| \\ \Delta V &= |G_n - G_s| + |2R_c - R_{nn} - R_{ss}| \end{aligned} \quad (2)$$

Wu and Zhang found that horizontal and vertical gradients ΔH and ΔV in the mosaic image are not sufficient to make the right decision between G^h and G^v in some cases [4]. In other words, it is possible that $|G_c - G_h| < |G_c - G_v|$ even if $\Delta H > \Delta V$. In order to rectify the problem, they based the decision between G^h and G^v on a richer feature vector

$$\mathbf{z} = (\Delta_h^h, \Delta_h^v, \Delta_v^v, \Delta_v^h)$$

The features Δ_h^h, Δ_h^v, Δ_v^v, Δ_v^h are extracted from the neighborhood of R_c (where G_c is missing and will be interpolated) and from the results of the directional Laplacian filter as follows. First, define for every pixel at location i the so-called primary differences γ between R and G, and β between B and G:

$$\begin{aligned} \gamma_i^h = R_i^h - G_i^h, \quad \beta_i^h = B_i^h - G_i^h \\ \gamma_i^v = R_i^v - G_i^v, \quad \beta_i^v = B_i^v - G_i^v \end{aligned} \quad (3)$$

where the superscript denotes whether a missing sample is interpolated horizontally or vertically. Note that at any given location i, one of G, R and B is the original CCD sample. In that case we simply ignore the superscripts and use the corresponding original sample values in (3). For instance, in the configuration of Fig. 2, R_c is the original sample, hence $R_c^h = R_c^v = R_c$.

Based on the quantities (features) of (3), the following four compound features in terms of $\mathbf{z}$ are extracted in a neighborhood of the concerned pixel to capture the high order statistical dependency between the samples, both original and estimated.

$$\Delta_h^h = \sum_{(p,q) \in \left\{ \begin{array}{l} (nw,n),(ne,n),(nw,ne), \\ (w,c),(e,c),(w,e), \\ (sw,s),(se,s),(sw,se) \end{array} \right\}} \|(\gamma_p^h, \beta_p^h) - (\gamma_q^h, \beta_q^h)\|_1 \quad (4)$$

$$\begin{aligned} \Delta_h^v = min\{ & \sum_{(p,q) \in \{(nw,w),(n,c),(ne,e)\}} \|(\gamma_p^h, \beta_p^h) - (\gamma_q^h, \beta_q^h)\|_1, \\ & \sum_{(p,q) \in \{(sw,w),(s,c),(se,e)\}} \|(\gamma_p^h, \beta_p^h) - (\gamma_q^h, \beta_q^h)\|_1\} \end{aligned} \quad (5)$$

$$\Delta_v^v = \sum_{(p,q) \in \left\{ \begin{array}{l} (nw,w),(ne,c),(nw,e) \\ (sw,w),(s,c),(se,e), \\ (nw,sw),(n,s),(ne,se) \end{array} \right\}} \|(\gamma_p^v, \beta_p^v) - (\gamma_q^v, \beta_q^v)\|_1 \quad (6)$$

$$\begin{aligned} \Delta_v^h = min\{ & \sum_{(p,q) \in \{(nw,n),(w,c),(sw,s)\}} \|(\gamma_p^v, \beta_p^v) - (\gamma_q^v, \beta_q^v)\|_1, \\ & \sum_{(p,q) \in \{(ne,n),(e,c),(se,s)\}} \|(\gamma_p^v, \beta_p^v) - (\gamma_q^v, \beta_q^v)\|_1\} \end{aligned} \quad (7)$$

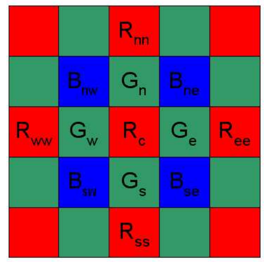

Fig. 2. Bayer pattern with location labels.

An important distinction between the gradient estimates ΔH and ΔV and the feature vector $\mathbf{z}$ is that $\mathbf{z}$ reflects the structures of the interpolated samples under different hypotheses on gradient. Specifically, $\mathbf{z}$ incorporates a look-ahead mechanism, which can optimize the decision on interpolation direction in a greater context than ΔH and ΔV. We include in $\mathbf{z}$ the features Δ_h^v and Δ_v^h in addition to Δ_h^h and Δ_v^v, because if a strong edge is present and if the interpolation direction is correct, Δ_h^h (Δ_v^v) will differ significantly from Δ_h^v (Δ_v^h). On the other hand, if Δ_h^h (Δ_v^v) and Δ_h^v (Δ_v^h) are both large, then there is high risk in interpolating horizontally (vertically). By their relative magnitudes, Δ_h^h, Δ_h^v, Δ_v^v, and Δ_v^h reveal the merits or risks in choosing different interpolation directions, and hence they form the feature vector for the demosaicking-induced classification problem to be defined below.

By associating the decision whether G^h or G^v should be the estimate of G_c with the feature vector $\mathbf{z}$, color demosaicking via directional interpolation can be viewed as an optimal statistical decision problem. Let the decision whether G^h or G^v is a better estimate of G_c be a binary random variable X. If $P(X|\mathbf{z})$ is known, then the optimal decision corresponds to the value of X that maximizes the posterior probability $P(X|\mathbf{z})$. Since $P(X|\mathbf{z})$ is generally unknown, we alternatively convert the optimal decision problem to one of binary classification. Namely, we consider $\mathbf{z} = (\Delta_h^h, \Delta_h^v, \Delta_v^v, \Delta_v^h)$ to be a vector of features that the binary random variable X exhibits. Given a suitable training set Z representing the statistics of the CCD camera data, the value of X partitions the set Z into two subsets Z_0

and Z_1. Set Z_i consists of all the training feature vectors associated with $X = i$, $i = 0, 1$. Our task is to design a classifier that separates Z_0 from Z_1 with the minimum number of misclassifications. We choose Fisher's linear discriminant to solve the above classification problem for its low computational complexity. Fisher's linear discriminant is based on the direction $\mathbf{a}$ of maximum separation between the two clusters Z_0 and Z_1, as given by

$$\mathbf{a} = \arg\max_{\mathbf{x}} \frac{[\mu_0(\mathbf{x}) - \mu_1(\mathbf{x})]^2}{\sigma_0^2(\mathbf{x}) + \sigma_1^2(\mathbf{x})} \tag{8}$$

where $\mu_i(\mathbf{x})$ and $\sigma_i^2(\mathbf{x})$ are the mean and variance of Z_i, $i = 0, 1$, respectively.

III. SELECTION OF TRAINING SETS

A key issue in designing optimal classifier for color demosaicking is the selection of a suitable training set. This is particularly important if our goal is to achieve best possible visual quality rather than just high PSNR values. If the training set is drawn at random and uniformly from representative images independent of image features, then the population of the training set will be dominated by samples drawn from relatively smooth areas, because natural images mostly consist of large smoothly shaded regions (i.e., of exponentially decaying power spectrum). These samples will overwhelmingly influence the design of the classifier because of their sheer weight. On the other hand, the human visual system is very sensitive to edges and fine textures, even though they present only a minority of the sample population. Therefore, one should design the classifier for samples drawn from the activity regions of the images. This design will not affect the visual quality of the smooth regions, for which the two estimates G^h and G^v will be close to each other anyways and either is a good estimate of G_c.

The first selection criterion for the training set is $|G^h - G^v| > \delta$, where δ is a threshold. We call this training set $\mathcal{A}$. Because the two estimates G^h and G^v differ significantly from each other, it becomes critical to choose the one that is closer to the missing sample G_c. The penalty of choosing a wrong interpolation direction is potentially high.

The worst case for color demosaicking by directional interpolation, if a decision on the edge direction is made locally, is when $|G^h - G_c| > |G^v - G_c|$ whereas $\Delta H < \Delta V$. In other words, the local gradient estimates ΔH and ΔV are contrary indicators for the correct interpolation direction. In order to improve the robustness of existing color demosaicking methods, we want to minimize the risk of misclassifications in this worst case. Therefore, we form a more selective, worst-case subset Æ of the training set $\mathcal{A}$, which contains the samples where $|G^h - G_c| > |G^v - G_c|$ but $\Delta H < \Delta V$. After an extensive investigation we have not been able to find a classifier that can separate the two clusters corresponding to $X = 0$ and $X = 1$ in Æ using the feature vector $\mathbf{z}$ with better than 50% probability consistently for all test images. This is because many samples in the set Æ are the inherently unsolvable cases, where the image signal exceeds the Nyquist frequency in both horizontal and vertical directions and there is lack of correlation between color channels.

However, by adding an additional constraint $|\Delta H - \Delta V| < \epsilon$, where ϵ is a threshold, to the selection criteria of Æ we can still correct some interpolation errors of the current demosaicking methods. Denote by $\check{\mathcal{A}}$ the resulting subset of Æ. By the constraint $|\Delta H - \Delta V| < \epsilon$ we select from Æ those milder classification errors committed by directional interpolation based on ΔH and ΔV, while screening out the irreparable errors. As we will report in the next section, the Fisher's linear discriminant classifier designed for set $\check{\mathcal{A}}$ in the feature space of $\mathbf{z}$ is more robust than the one designed for set $\mathcal{A}$, and it can correct some errors of the latter in certain difficult cases. This also demonstrates that $\mathbf{z} = (\Delta_h^h, \Delta_h^v, \Delta_v^v, \Delta_v^h)$ are more powerful features than the local gradient estimates ΔH and ΔV for the classification problem induced by color demosaicking.

IV. EXPERIMENTAL RESULTS

Training sets $\mathcal{A}$ and $\check{\mathcal{A}}$ as defined above are generated from 13 images. The corresponding minimum-risk classifiers designed for these training sets are denoted by $\mathcal{C}$ and $\check{\mathcal{C}}$. Each of the two classifiers is designed for five different selection parameters $\delta = 5, 10, 15, 20, 25$. The larger is δ, the more restrictive the training set gets with more error-prone samples being drawn. The resulting classifiers are used by the PCSD algorithm [4] to choose one of the two directional estimates.

Five test images are used to evaluate the performance of the proposed classifiers $\mathcal{C}$ and $\check{\mathcal{C}}$. These five test images are different from those included in the training sets. This ensures that the designed classifiers do not exhibit a favorable bias toward any of the test images. The performance of each proposed classifier is also compared with the empirical classifier $\mathbf{a} = [1, 0, -1, 0]$ and $\xi = 0$, used in [4]. Recall that $\mathbf{a}$ is the projection direction of the linear classifier and ξ is a threshold that determines which of the two directional estimates should be used in color demosaicking.

Table I presents the parameters of the two linear classifiers $\mathcal{C}$ and $\check{\mathcal{C}}$ designed for training sets $\mathcal{A}$ and $\check{\mathcal{A}}$ respectively for $\delta = 5$.

Table II presents the PSNR results of $\mathcal{C}$, $\check{\mathcal{C}}$ and the empirical classifier when they are applied to all the pixels in the test images. As expected both the optimized context-

Table I. The parameters of Fisher's linear discriminant of $\mathcal{C}$ and $\check{\mathcal{C}}$ for training sets $\mathcal{A}$ and $\check{\mathcal{A}}$, $\delta = 5$.

Training Set	a	ξ
$\mathcal{A}$	[0.6427, 0.3079, -0.6720, -0.2012]	1.0523
$\check{\mathcal{A}}$	[0.4124, 0.5801, -0.4528, -0.5370]	-3.4439

Table II. Average PSNRs of different classifiers.

Classifier	Red (dB)	Green (dB)	Blue (dB)
$\mathcal{C}$	35.96	39.49	36.37
$\check{\mathcal{C}}$	36.05	39.61	36.49
empirical	35.88	39.38	36.28

based classifiers outperform the empirical classifier. Somewhat surprisingly, $\check{\mathcal{C}}$ has superior average performance over $\mathcal{C}$, although $\check{\mathcal{C}}$ is designed for a very selective subset of $\mathcal{A}$.

We also experimented with the idea of switching between $\mathcal{C}$ and $\check{\mathcal{C}}$ depending on the context of the current pixel to be interpolated. The decision on the interpolation direction at pixels that satisfy $|G^h - G^v| > \delta$ and $|\Delta H - \Delta V| < \delta$ is made by the classifier $\check{\mathcal{C}}$; otherwise, by the classifier $\check{\mathcal{C}}$. We refer to this strategy as the multiple classifier scheme. The PSNR values of the multiple classifier scheme are listed in Table III.

Table III. Average PSNRs of the switched multiple classifiers.

Red (dB)	Green (dB)	Blue (dB)
35.96	39.51	36.38

In order to evaluate the visual quality of different classifiers, we present in Figs. 3 and 4 some samples (difficult cases for the existing methods) of the demosaicked color images by the PCSD algorithm using different classifiers to choose the interpolation direction. These results are compared with those of Hirakawa-Parks' algorithm [5] and Gupta-Chen's algorithm [6] (The authors of both papers gracefully sent us their software for evaluation). Hirakawa-Parks' algorithm is also based on edge-direction learning, and has visual quality comparable to ours with somewhat different artifacts (inferior to ours in Fig. 3, and slightly better in Fig. 4). The merits of the proposed classification approach to color demosaicking are evident.

V. CONCLUSION

Adaptive directional interpolation in color demosaicking is converted to a binary classification problem, and solved by Fisher's linear discriminant. Criteria for selecting the training set to design the classifier are also examined. The proposed classification technique for color demosaicking can significantly improve the performance of the existing methods in some difficult cases.

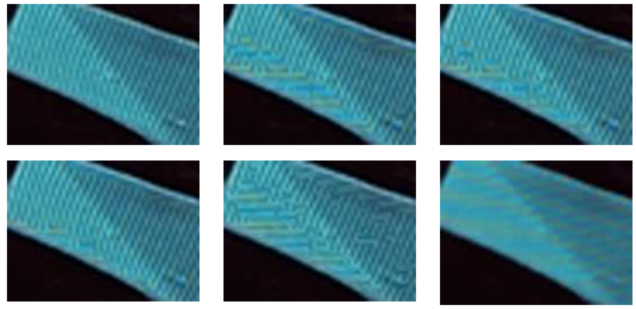

Fig. 3. Part of test image 1. Top left to right: original, empirical classifer, classifier $\mathcal{C}$; bottom left to right: classifier $\check{\mathcal{C}}$, [5], [6].

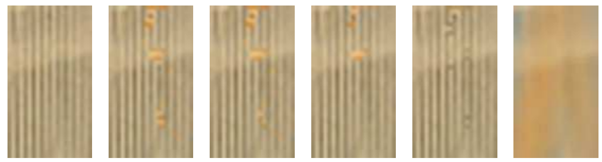

Fig. 4. Part of test image 2. From left to right: original, empirical classifer, classifier $\mathcal{C}$, classifier $\check{\mathcal{C}}$, [5], [6].

VI. REFERENCES

[1] B. K. Gunturk, Y. Altunbasak, and R. M. Mersereau, "Color plane interpolation using alternating projections," *IEEE Trans. on Image Processing*, vol. 11, no. 9, p. 997-1013, Sept. 2002.

[2] R. Ramanath, W. E. Snyder and G. L. Bilbro, "Demosaicking methods for Bayer color arrays," *J. of Electronic Imaging*, vol. 11, p.306-315, July 2002.

[3] T. Chen, $http://www-ise.stanford.edu/ tingchen/$.

[4] X. Wu and N. Zhang, "Primary-consistent soft-decision color demosaic for digital cameras," *IEEE Trans. on Image Processing* (to appear).

[5] K. Hirakawa and T. W. Parks, "Adaptive homogenity-directed demosaicing algorithm," *IEEE 2003 Int. Conf. on Image Processing*, Sept. 2003.

[6] M. Gupta and T. Chen, "Vector color filter array demosaicing," *Proceedings of SPIE Electronic Imaging Conference*, vol. 4306, 2001.