Available online at www.sciencedirect.com
SCIENCE DIRECT®

Pattern Recognition 38 (2005) 2208–2212

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY
www.elsevier.com/locate/patcog

Rapid and brief communication

Universal demosaicking for imaging pipelines with an RGB color filter array

Rastislav Lukac*, Konstantinos N. Plataniotis

Multimedia Laboratory, BA 4157, The Edward S. Rogers Sr. Department of ECE, University of Toronto, 10 King's College Road, Toronto, Ont., Canada M5S 3G4

Received 12 April 2005; accepted 27 April 2005

Abstract

A universal demosaicking solution appropriate for imaging pipelines employing a red–green–blue (RGB) color filter array (CFA) is introduced. The proposed solution can demosaick a raw, grayscale, CFA sensor image captured by a digital camera equipped with any RGB-CFA currently in use. The solution utilizes a spectral model, an edge-sensing mechanism, and a postprocessor to preserve the coloration and sharpness of the captured image. The framework readably unifies existing demosaicking solutions which differ in design characteristics, performance, and computational efficiency. Simulation studies indicate that the universal demosaicking framework allows for cost-effective camera image processing and produces visually pleasing full-color digital images.

Keywords: Digital camera image processing; Color filter array; Demosaicking; CFA interpolation; Spectral interpolation; Color image enhancement; Spectral model; Edge-sensing mechanism; Postprocessing

1. Introduction

Single-sensor imaging devices are used extensively in numerous applications ranging from computer vision and multimedia systems, to sensor networks and surveillance systems [1,2]. Since the sensor, usually a charge coupled device (CCD) or complementary metal oxide semiconductor (CMOS) sensor, is a monochromatic device, a color filter array (CFA) is placed at the sensor's surface to capture all the three, red–green–blue (RGB) primary colors at the same time. The specific arrangements of color filters in the CFA vary depending on the camera manufacturer (Fig. 1) [3]. Each pixel of the raw CFA image has its own spectrally selective filter and thus, the captured CFA data constitutes a gray-scale, mosaic-like, image [2,5]. By estimating the two missing color components at each spatial location from the adjacent pixels, a full-color RGB image is obtained using a process called demosaicking [1–5].

Although demosaicking methods have been extensively researched in the last few years the main focus of attention was on demosaicking solutions for images captured using a Bayer CFA (Fig. 1a) as it can be seen from recent overviews in the area [2,4]. The Bayer CFA is of particular interest to practitioners and researchers alike due to the simplicity of the subsequent processing steps. However, when it comes to anti-aliasing, non-Bayer CFAs (Figs. 1b and c), especially those with semi-periodic, pseudo-random or random arrangements of the RGB array may be more appropriate as they lead to additional performance improvements [3].

In this paper, we introduce a solution which is capable of demosaicking a CFA image obtained using an arbitrary

* Corresponding author. Tel.: +1 416 978 6845; fax: +1 416 978 4425.
E-mail address: lukacr@dsp.utoronto.ca (R. Lukac)
URL: http://www.dsp.utoronto.ca/~lukacr (R. Lukac).

0031-3203/$30.00
doi:10.1016/j.patcog.2005.04.008

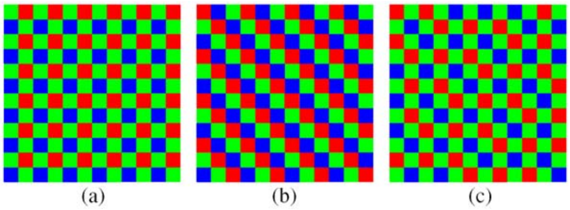

Fig. 1. Examples of the RGB CFAs: (a) Bayer pattern; (b) diagonal strip pattern; (c) pseudorandom Bayer pattern.

RGB CFA. The solution employs a spectral model, an edge-sensing mechanism, and a postprocessor to produce visually pleasing full-color images. The proposed methodology constitutes a canonical framework which readily accommodates a large number of application requirements and implementation constraints. The existence of this realistic and efficient procedure allows for demosaicking solutions to be implemented either on the camera or in a companion personal computer (PC) using raw, CFA images in the so-called tagged image file format for electronic photography (TIFF-EP). To the best knowledge of the authors, this is the first time when such a versatile and cost-effective solution is being proposed for processing single-sensor images.

2. Prior-art

Due to the monochromatic nature of the sensor the captured values create a $K_1 \times K_2$ gray-scale, mosaic-like, image $z: Z^2 \rightarrow Z$. This CFA image represents a two-dimensional matrix of integer samples $z_{(p,q)}$ with $p = 1, 2, \ldots, K_1$ and $q = 1, 2, \ldots, K_2$ denoting the image rows and columns, respectively. The demosaicking step re-arranges the acquired gray-scale sensor data to an RGB-like vectorial field, and completes missing color components using adjacent sensor data through spectral interpolation. The process produces a color (RGB) image $\mathbf{x}: Z^2 \rightarrow Z^3$ with color pixels $\mathbf{x}_{(p,q)} = [x_{(p,q)1}, x_{(p,q)2}, x_{(p,q)3}]$ represented as vectors in the RGB vectorial space. In the color vector $\mathbf{x}_{(p,q)}$ the $x_{(p,q)k}$ value, for $k = 1, 2, 3$, denotes the kth vector's spectral component. Namely, $x_{(p,q)1}$ signifies the R component, $x_{(p,q)2}$ denotes the G component, and $x_{(p,q)3}$ indicates the B component.

Since demosaicking solutions often introduce visual impairments, such as blurred edges and color shifts, demosaicked image postprocessing follows the demosaicking module in the processing pipeline in order to enhance the visual quality of the demosaicked, full-color, camera images [5]. In high-quality imaging solutions both the spatial and spectral characteristics of the acquired input should be utilized during the demosaicking and postprocessing steps. This can be done using the so-called unbiased, data-adaptive spectral estimator [1,2]

$$\mathbf{x}_{(p,q)} = \sum_{(i,j)\in\zeta} \{w'_{(i,j)} f(\mathbf{x}_{(i,j)}, \mathbf{x}_{(p,q)})\} \tag{1}$$

which operates over the pixels available in the location of interest (p, q) and the neighboring locations $(i, j) \in \zeta$ with ζ denoting the area of support. The term $f(\cdot)$ is a manifestation of the spectral model concept which indicates that in a natural image significant spectral correlation exists among the color planes [2,4]. The use of a spectral model reduces the presence of color shifts and artifacts in the outputted image. The terms $w'_{(i,j)}$ defined as

$$w'_{(i,j)} = w_{(i,j)} \Bigg/ \sum_{(g,h)\in\zeta} w_{(g,h)} \tag{2}$$

denote the normalized weighting coefficients calculated using the edge-sensing weights $w_{(i,j)}$. Weights are used to regulate the contribution to $\mathbf{x}_{(p,q)}$ of the available samples $\mathbf{x}_{(i,j)}$ inside the area ζ by [1,2,4]: (i) emphasizing inputs which are not positioned across an edge, and (ii) directing the demosaicking process along the natural edges in the true image. This ensures a sharply formed, output camera image.

The unbiased estimator in (1) generalizes numerous processing solutions, which may be derived by changing the form of the spectral model, as well as the way that the edge-sensing weights are calculated. The choice of these two elements essentially determines the characteristics and the performance of the processing solution [2,4].

3. Proposed universal demosaicking solution

The objective of this work is to device a universal, single-sensor imaging solution which can demosaick and postprocess images captured using an arbitrary CFA. Therefore, processing operations should be performed independently from the CFA structure. This can be done using the spatial location flags proposed here. Since information about the arrangements of color filters in the actual CFA is readily available either from the camera manufacturer (when demosaicking is implemented in the camera), or obtained from the raw CFA image stored in TIFF-EP format (when demosaicking is performed in a companion PC), a $K_1 \times K_2$ vectorial field $\mathbf{d}: Z^2 \rightarrow Z^3$ of the corresponding location flags $d_{(p,q)k}$ is initialized using the default value $d_{(p,q)k} = 1$ to indicate the presence of a CFA value $z_{(p,q)}$ in the color vector $\mathbf{x}_{(p,q)}$ for the proper value of k. For example, if (p, q) corresponds to a G CFA location in the image z, then $\mathbf{x}_{(p,q)} = [0, z_{(p,q)}, 0]$ and $d_{(p,q)k} = 1$ for $k = 2$ should be used. If (p, q) corresponds to a R (or B) CFA location, then $\mathbf{x}_{(p,q)} = [z_{(p,q)}, 0, 0]$ (or $\mathbf{x}_{(p,q)} = [0, 0, z_{(p,q)}]$) and $d_{(p,q)k} = 1$ for $k = 1$ (or $k = 3$) should be utilized. In all other cases, the flags are set to $d_{(p,q)k} = 0$ indicating that processing should be performed at (p, q) in the R ($k = 1$), G ($k = 2$), or B ($k = 3$) channel of the color image $\mathbf{x}$. In addition, employing postprocessing in the imaging pipeline necessitates the storage of the initial values $d_{(p,q)k}$ to flag indicators $c_{(p,q)k}$ which form a $K_1 \times K_2$ vectorial field $\mathbf{c}: Z^2 \rightarrow Z^3$.

After initialization, the proposed solution follows conventional practice [1,2,4] and starts the demosaicking process by estimating the missing G components. It is well-known that visual impairments observed in processed camera images should be attributed not only to procedural limitations but to the spatial and spectral constraints imposed during processing [2,4,5]. To prevent the demosaicking solution from operating in areas where we lack adequate input information, the following control mechanism is introduced

$$\sum_{(i,j)\in\zeta} (d_{(i,j)k} = 1) \geqslant \chi, \tag{3}$$

where χ is a design parameter denoting the minimum number of inputs values needed to be present when processing the kth color channel in the local neighborhood ζ. Using a 3×3 sliding window $\Psi_{(p,q)} = \{\mathbf{x}_{(i,j)}; (i,j) \in \zeta = \{(p-1, q-1), (p-1, q), \ldots, (p+1, q+1)\}\}$ which places the vector $\mathbf{x}_{(p,q)}$ under consideration at the center of $\Psi_{(p,q)}$, the proposed solution updates $\mathbf{x}_{(p,q)}$ when $d_{(p,q)2} = 0$ and (3) for $k = 2$ is satisfied, as follows:

$$x_{(p,q)k} = \sum_{\substack{(i,j)\in\zeta \\ d_{(i,j)k}=1}} \{w'_{(i,j)} x_{(i,j)k}\} \tag{4}$$

with $k = 2$ denoting the G plane. This processing step, which can be viewed as the componentwise variant of (1), is performed on any vectors $\mathbf{x}_{(i,j)}$ inside the neighborhood ζ which contain a true G component $x_{(i,j)2}$.

The constraint imposed through $d_{(i,j)2} = 1$ is used in the determination of edge-sensing weights $w_{(i,j)}$ in (2). Since the proposed solution is independent from the CFA structure, the traditional forms of the edge-sensing mechanism listed in Ref. [4] cannot be used. The weights $w_{(i,j)}$ in (2) must be calculated using samples obtained through $\Psi_{(p,q)}$. The objective can be achieved by tracking the structural image characteristics using aggregated absolute differences between the inputs inside the neighborhood ζ. Such an approach has been used to support demosaicking [2], demosaicked image postprocessing [5] and integrated camera image processing [1] operations. Following these design characteristics the weights in our universal imaging pipeline are determined as follows:

$$w_{(i,j)} = \left[1 + \sum_{\substack{(g,h)\in\zeta \\ d_{(g,h)k}=1}} |x_{(i,j)k} - x_{(g,h)k}| \right]^{-1}, \tag{5}$$

where $k = 2$ is used to determine structural (edge) information using the available G components.

After completing (4), the corresponding location flag is changed to $d_{(p,q)2} = 2$ and the window is centered at the next spatial location for which the corresponding flag is $d_{(\cdot,\cdot)2} = 0$. The window operator slides over the entire image successively placing each image pixel at its center. A fully populated G plane is obtained when all the location flags are set to $d_{(\cdot,\cdot)2} \neq 0$. If this is not the case upon completion of the first iteration, the location flags corresponding to already processed locations are changed from $d_{(\cdot,\cdot)2} = 2$ to $d_{(\cdot,\cdot)2} = 1$, and the algorithmic step is repeated until a fully populated G plane is generated. Thus, during the next iteration(s) the remaining G components are calculated using both the G CFA and demosaicked G components. Typically, using some random CFAs (highest complexity) and $\chi = 2$ the total number of demosaicking iterations per color channel is two, whereas the additional iteration may be required when a higher threshold, i.e. $\chi = 3$, is used.

Since: (i) the G components considerably contribute to luminance information which is essential for the perceived sharpness of the digital image [2,4], and (ii) natural images exhibit significant spectral correlation, the missing R and B components should be demosaicked by utilizing the information from the already populated G plane. By adopting the popular color difference model used in Refs. [2–5], the generalized spectral estimator in (1) is modified to demosaick the R ($k = 1$) or B ($k = 3$) component $x_{(p,q)k}$ occupying the location (p, q) with $d_{(p,q)k} = 0$ and the constraint in (3) satisfied for $k = 1$ or $k = 3$, respectively, as follows:

$$x_{(p,q)k} = x_{(p,q)2} + \sum_{\substack{(i,j)\in\zeta \\ d_{(i,j)k}=1}} \{w'_{(i,j)}(x_{(i,j)k} - x_{(i,j)2})\}, \tag{6}$$

where the normalized coefficients $w'_{(i,j)}$ are calculated in (2) using (5) with $k = 1$ (for R) or $k = 3$ (for B).

The procedure successively updates all the flags to $d_{(p,q)k} = 2$ at each location (p, q) which is being demosaicked via (6). In the sequence, flag values should be changed from $d_{(p,q)k} = 2$ to $d_{(p,q)k} = 1$ and (6) should be re-applied over an image until the R and B planes are not fully populated. After demosaicking all available locations and updating all the corresponding flags to $d_{(\cdot,\cdot)k} \neq 0$ the demosaicking step produces a reconstructed, full-color image.

It is well-known that the output image quality can be further improved by employing a postprocessor in the pipeline [5]. To guide the demosaicked image postprocessing step, the location flag values $d_{(p,q)k}$ are restored using the stored template values $c_{(p,q)k}$, for $p = 1, 2, \ldots, K_1$, $q = 1, 2, \ldots, K_2$, and $k = 1, 2, 3$. Then, postprocessing of the G color plane, in all locations (p, q) with the constraints $d_{(p,q)2} = 0$ and (3) enforced for $k = 2$, can be realized using R or B components as follows:

$$x_{(p,q)2} = x_{(p,q)k} + \sum_{\substack{(i,j)\in\zeta \\ d_{(i,j)2}=1}} \{w'_{(i,j)}(x_{(i,j)2} - x_{(i,j)k})\}, \tag{7}$$

where weights $w'_{(i,j)}$ are obtained in (2) using (5) with $k = 2$, and $d_{(p,q)2}$ are the flags to be updated similarly to (4) with

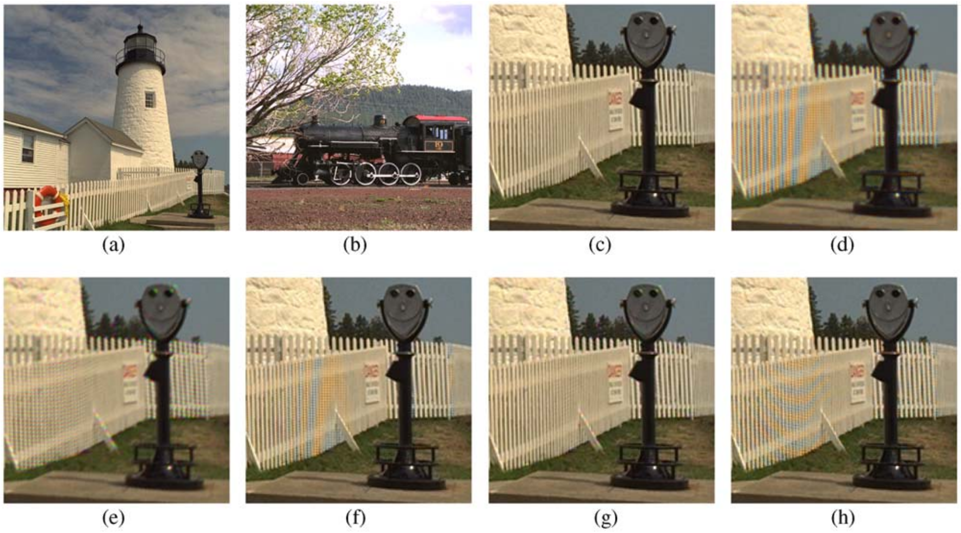

Fig. 2. Test color images: (a) Lighthouse and (b) Train, and (c)–(h) enlarged parts of: (c) original image, (d) Solution A output using Fig. 1a, (e) Solution A output using Fig. 1b, (f) Solution B output using Fig. 1a, (f) Solution B output using Fig. 1b, (f) Solution B output using Fig. 1c.

Table 1
Objective performance of the proposed universal framework applied to sensor images obtained using different CFAs

Image	Lighthouse (Fig. 2a)						Train (Fig. 2b)					
CFA	Fig. 1a		Fig. 1b		Fig. 1c		Fig. 1a		Fig. 1b		Fig. 1c	
Method	MSE	NCD	MSE	NC	MSE	NCD	MSE	NCD	MSE	NCD	MSE	NCD
Solution A	137.9	0.0579	117.8	0.0544	144.0	0.0606	473.7	0.1443	425.0	0.1418	484.3	0.1486
Solution B	13.7	0.0229	12.3	0.0247	15.4	0.0251	49.1	0.0536	78.0	0.0684	56.3	0.0569

$k=2$. Postprocessing is applied only at the spatial locations (p,q) where the G components have been obtained using (4), i.e. original CFA components are kept unchanged. If (p,q) corresponds to the R CFA location ($c_{(p,q)1}=1$), then the parameter $k=1$ should be used in (7). Otherwise, (7) is used for the B CFA location ($c_{(p,q)3}=1$) and the pertinent parameter is $k=3$. After the G plane is enhanced, the postprocessing step is completed by enhancing the demosaicked R and B components using $k=1$ and 3, respectively, in both (6) and (5) in a manner similar to the one described in Refs. [1,2,5]. This step is performed for R (or B) components only in locations corresponding to G and B (or G and R) CFA values, i.e. $c_{(p,q)1}=0$ (or $c_{(p,q)3}=0$).

Two solutions from the proposed framework are adopted in the sequence for demosaicking experimentation. Namely, the so-called *Solution A* is a simple, linear, componentwise solution which uses (4), for $k=1,2,3$, based on (2) and fixed weights $w_{(i,j)}=1$. The alternative *Solution B*, which can be seen as the generalization of Solution A as well as numerous demosaicking schemes in Refs. [2,4], uses the complete processing cycle as it was described in Section 3. Note that in this paper, at least three components from the kth channel, for $k=1,2,3$, to be demosaicked are required in Eqs. (4)–(7) and thus, the parameter $\chi=3$ is used in (3).

4. Experimental results

A number of test images, such as the 512×512 images shown in Figs. 2a and b, captured using three-sensor devices and normalized to 8-bit per channel RGB representation, have been used for demonstration purposes. The tests were performed by sampling the original images with the CFAs shown in Fig. 1, respectively, to obtain a CFA image following the standard practice previously reported in Refs. [1,2,4,5]. The used CFAs vary significantly in anti-aliasing and the complexity of the color reconstruction process [3]. Demosaicked images were obtained from applying the two universal demosaicking solutions (Solutions A and B) onto the CFA images. Performance was measured by comparing the original full-color images to the full-color demosaicked images. To facilitate the objective comparisons [2,5], the RGB color space based mean square error (MSE) criterion and the CIE-LUV color space based normalized color difference (NCD) criterion have been considered in this work.

Results presented in Table 1 and Figs. 2c–h indicate that the choice of the CFA may be one of the most important factors in designing the single-sensor imaging pipeline. Although the same camera image processing solution has been employed to process the CFA image, results corresponding to the different CFAs differ significantly. For example, visual inspection of the images obtained using a simple linear solution (Figs. 2d and e) shows substantial reduction of aliasing artifacts when the diagonal strip pattern CFA is used instead of the Bayer CFA. Additional improvements in image quality, in terms of both sharpness and coloration, as well as objective performance (Table 1) were observed when a more sophisticated solution is used (Figs. 2f–h). These results indicate that extensive research, as part of our future work, should be focused on the design, analysis and evaluation of the existing CFAs used in the single-sensor imaging pipeline.

5. Summary

A unique CFA image processing solution was introduced. The solution can process the CFA image captured using any existing RGB CFA currently in use. The framework employed an edge sensing mechanism, a spectral model, and a postprocessor to produce high-quality, full-color images.

References

[1] R. Lukac, K. Martin, K.N. Plataniotis, Digital camera zooming based on unified CFA image processing steps, IEEE Trans. Consumer Electron. 50 (1) (2004) 15–24.

[2] R. Lukac, K.N. Plataniotis, Data adaptive filters for demosaicking: a framework, IEEE Trans. Consumer Electron. 51 (2) (2005).

[3] FillFactory, Technology—image sensor: the color filter array FAQ, Available Online: http://www.fillfactory.com/htm/technology/htm/rgbfaq.htm

[4] B.K. Gunturk, J. Glotzbach, Y. Altunbasak, R.W. Schaffer, R.M. Murserau, Demosaicking: color filter array interpolation, IEEE Signal Process. Mag. 22 (1) (2005) 44–54.

[5] R. Lukac, K.N. Plataniotis, A robust, cost-effective postprocessor for enhancing demosaicked camera images, Real-Time Imaging, Special Issue Spectral Imaging 11 (2) (2005).

Rastislav Lukac received the M.S. (Ing.) and Ph.D. degrees in Telecommunications from the Technical University of Kosice, Slovak Republic in 1998 and 2001, respectively. From February 2001 to August 2002 he was an Assistant Professor at the Department of Electronics and Multimedia Communications at the Technical University of Kosice. Since August 2002 he is a Researcher in Slovak Image Processing Center in Dobsina, Slovak Republic. From January 2003 to March 2003 he was a Postdoctoral Fellow at the Artificial Intelligence & Information Analysis Lab at the Aristotle University of Thessaloniki, Greece. Since May 2003 he has been a Post-doctoral Fellow with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering at the University of Toronto in Toronto, Canada. His research interests include digital camera image processing, microarray image processing, multimedia security, and nonlinear filtering and analysis techniques for color image & video processing.

Dr. Lukac is a Member of the IEEE, a Member of the IEEE Circuits and Systems, IEEE Consumer Electronics, and IEEE Signal Processing Societies. He serves as a Technical Reviewer for various scientific journals and he participates as a Member of numerous International Conference Committees. In 2003 he was awarded the NATO/NSERC Science Award.

Konstantinos N. Plataniotis received the B. Engineering degree in Computer Engineering from the Department of Computer Engineering and Informatics, University of Patras, Patras, Greece in 1988 and the M.S and Ph.D degrees in Electrical Engineering from the Florida Institute of Technology (Florida Tech), Melbourne, Florida in 1992 and 1994 respectively. From August 1997 to June 1999 he was an Assistant Professor with the School of Computer Science at Ryerson University. He is currently an Associate Professor at the Edward S. Rogers Sr. Department of Electrical & Computer Engineering where he researches and teaches adaptive systems and multimedia signal processing. He co-authored, with A.N. Venetsanopoulos, a book on ”Color Image Processing & Applications”, Springer Verlag, May 2000, he is a contributor to seven books, and he has published more than 300 papers in refereed journals and conference proceedings in the areas of multimedia signal processing, image processing, adaptive systems, communications systems and stochastic estimation.

Dr. Plataniotis is a Senior Member of IEEE, a past member of the IEEE Technical Committee on Neural Networks for Signal Processing. He was the Technical Co-Chair of the Canadian Conference on Electrical and Computer Engineering (CCECE) 2001, and CCECE 2004. He is the Vice-chair for the 2006 IEEE Intelligent Transportation Systems Conference, and the Technical Program Co-Chair of the 2006 IEEE International Conference in Multimedia and Expo (ICME 2006).