Efficient identification of arbitrary color filter array images based on the frequency domain approach

Yong-Huai Huang [a,2], Kuo-Liang Chung [b,*,1], Tseng-Jung Lin [b]

[a] Department of Electronic Engineering, Jinwen University of Science and Technology, No. 99, AnChung Rd., Xindian Dist., New Taipei City 23154, Taiwan, R.O.C.
[b] Department of Computer Science and Information Engineering, National Taiwan University of Science and Technology, No. 43, Section 4, Keelung Road, Taipei 10672, Taiwan, R.O.C.

ARTICLE INFO

ABSTRACT

For reducing the cost, most digital cameras are equipped with a CCD or CMOS sensor and a RGB color filter array (CFA) for each pixel to capture one primary color component, and hence produce a mosaic image. Suppose the input mosaic image without the CFA structure information, this paper presents a novel efficient method, consisting of a training-based scheme and an identification scheme, for identifying its CFA structure using the frequency domain approach. Initially, based on a set of training mosaic images with different CFA structures, a training-based scheme is proposed to build up the representative spectrum for every CFA structure. As the model maps, the constructed representative spectra can be reused in subsequent identification processes. The proposed identification scheme first constructs the representative spectrum of the header-less input mosaic image as the query map. Then, a matching scheme is proposed to identify the corresponding CFA structure of the query map from the model maps. Experimental results demonstrate that the proposed identification method has low computational cost and high identification accuracy merits for mosaic images without prior header information, when compared with the state-of-the-art spatial domain-based method by Chiu et al.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Nowadays, digital cameras have become increasingly popular in the consumer electronics market. To reduce hardware costs, most digital cameras use a single charge-coupled device (CCD) or a complementary metal-oxide-semiconductor (CMOS) sensor with a color filter array (CFA) to capture one primary color component for each pixel [1].

Such images are called mosaic images. Fig. 1 shows 11 typical CFA structures. The first 10 CFA structures [2] in Fig. 1 are red–green–blue (RGB) CFAs in which the Bayer CFA [3] is the most well-known structure; the last CFA structure, i.e. the Hirakawa CFA [4], is a non-RGB CFA in which six colors are considered and each one is a linear combination of red, green, and blue components.

Most of the digital cameras can produce full RGB color image in JPEG file format [5] from the captured mosaic image through a series of image processing operations, such as demosaicing, noise removal, white balance control, resizing, and image compression. The produced JPEG files are convenient for media storage. However, some of the above digital camera's operations may result in image quality degradation, the resultant images are unsuitable for professional users, e.g. photographers, artists, and graphic designers. Besides the JPEG file format, many digital cameras also offer the raw image
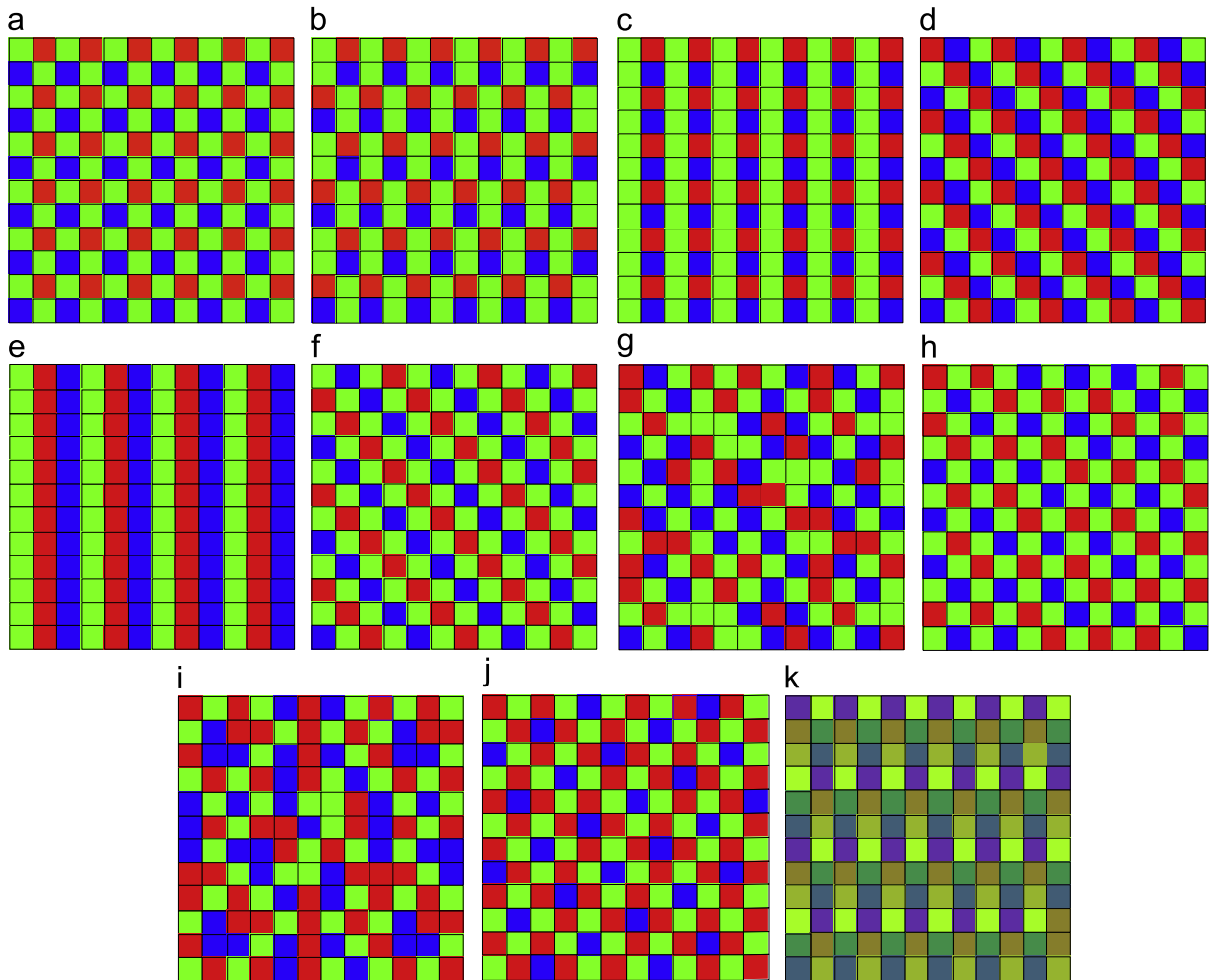
Fig. 1. Ten typical RGB CFA structures: (a) Bayer CFA, (b) Lukac and Plataniotis CFA, (c) Yamanaka CFA, (d) diagonal stripe CFA, (e) vertical stripe CFA, (f) modified Bayer CFA, (g) HVS-based CFA, (h) type I pseudo-random CFA, (i) type II pseudo-random CFA, (j) type III pseudo-random CFA, and (k) Hirakawa CFA.

format, which directly records mosaic images captured from CCD or CMOS sensor and without applying any digital camera's operations, to avoid the quality degradation. Recently, the raw image format has become popular among professional users because they can select better image algorithms instead of the ones used in digital cameras. As the popularity of the raw image format, more and more researchers focus on mosaic images and several related algorithms have been developed, such as compression [6–8], super-resolution [9,10], and demosaicing [11–14].

Given a mosaic image, its CFA structure is required to be known in advance from the header in TIFF-EP format for image manipulations. If, however, the prior header information is unknown for the input mosaic image, e.g. the header may be lost due to network packet loss in transmission or storage damage, the related image manipulations cannot work well. In such a header-less situation, Chiu et al. [15] proposed a spatial domain-based total average square difference minimization approach for identifying the CFA structure of the mosaic image. However, their method is rather time-consuming and the identification accuracy is dependent on

the window size used. Therefore, it is necessary to design a new approach to substantially reduce the computational cost and enhance the identification accuracy, leading to the main motivation of this research.

For an input mosaic image whose CFA structure is not available, this paper presents a novel efficient method for identifying its CFA structure using the frequency domain approach. The proposed method consists of two schemes: (1) the four-step training-based scheme to build up the representative spectra as the model maps for the concerned CFA structures and (2) the three-step matching scheme to identify the CFA structure of the input header-less mosaic image. In the four-step training-based scheme, suppose there are n training mosaic images for the ith CFA structure, $0 \leq i \leq 10$. In the first step, the Fourier transform is performed on every training mosaic image to obtain the high-pass spectrum map. In the second step, the high-energy blocks in each high-pass spectrum map are located and the coefficients in every high-energy block are reserved, but those in the non-high-energy block are discarded. In the third step, a successive thresholding approach is proposed

to extract the significant coefficients in every high-energy block, obtaining the significant spectrum map. In the fourth step, for the ith CFA structure, an outlier-pruning approach is proposed to build up the model map based on these n significant spectrum maps. We thus build up the 11 model maps for the 11 concerned CFA structures. In the first step of the proposed three-step matching scheme, we resize the input mosaic image to the same size as the model map and then obtain the corresponding significant spectrum map as the query map. In the second and third steps, based on the Bhattacharya distance criterion, a matching strategy is proposed to identify the corresponding CFA structure of the query map from the model maps. In addition, the computational complexity analysis of the proposed method is presented. For comparison, the computational complexity analysis of the comparative method is also exploited. Experimental results demonstrate that when compared with Chiu et al.'s spatial domain approach, the proposed frequency domain-based identification method has low computational cost and high identification accuracy merits for the input

mosaic image with an arbitrary CFA structure for which there is no prior header information.

The rest of this paper is organized as follows. In Section 2, we first present the proposed four-step training-based scheme for constructing the model map for each CFA structure, and then present the three-step matching scheme to identify the CFA structure of the input mosaic image without prior header information. Furthermore, the computational complexity analyses of the proposed method and the comparative method are presented. Section 3 gives the empirical results of the identification performance in terms of the computational cost and the identification accuracy. Concluding remarks are given in Section 4.

2. Proposed CFA identification method

For an input mosaic image whose CFA structure is not available, this section presents a novel efficient frequency domain-based method to identify its CFA structure. In the first subsection, a four-step training-based scheme is proposed to
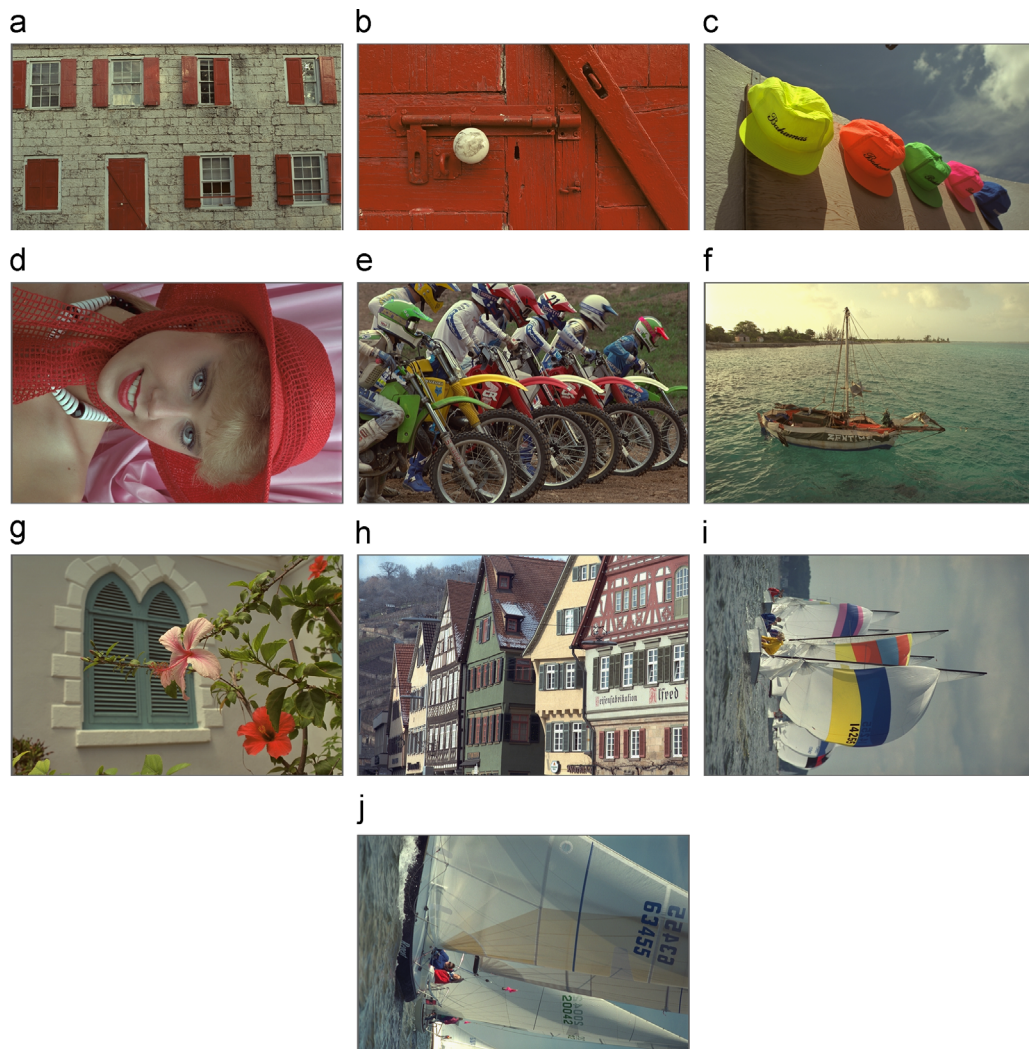

Fig. 2. Ten training color images used in the proposed four-step training scheme.

build up the model maps for the 11 concerned CFA structures. In the second subsection, the proposed three-step matching scheme is presented to identify the CFA structure of the input mosaic image. In the third subsection, the related computational complexity analyses are provided.

2.1. Proposed four-step training-based scheme to construct the model map for every CFA structure

The purpose of the proposed four-step training-based scheme is to build up the 11 model maps for the concerned CFA structures based on a set of training mosaic images. Initially, let the n RGB training full color images be denoted by $I_0, I_1,...,$ and I_{n-1}, each with size $w \times h$. According to the 11 typical CFA structures as shown in Fig. 1, the corresponding n mosaic images for each CFA structure are generated and used to construct the model map. For convenience, let I_k^p, $0 \leq p \leq 10$ and $0 \leq k \leq n$, denote the kth training mosaic image for the pth CFA structure. To explain how the proposed training scheme works, we select 10 full RGB color images from Rich Franzen's website [16] as shown in Fig. 2 to generate training mosaic images. For each image in Fig. 2, the corresponding 11 different mosaic images are generated by reserving only one color for each pixel. Specifically, for mosaic images with CFA structures in Fig. 1(a)–(j), the color of each pixel is selected from red, green, or blue component in the RGB domain; for those with the non-RGB CFA structure in Fig. 1(k), the color of each pixel is obtained from the linear combination of red, green, and blue components. Totally 10 full RGB color images can generate 110 training mosaic images for constructing the model maps.

In the first step, the training mosaic image I_k^p is transformed to the frequency domain representation F_k^p by using the Fourier transform:

$$F_k^p(u,v) = \sum_{x=0}^{w-1} \sum_{y=0}^{h-1} I_k^p(x,y) e^{-j2\pi((u/w)x + (v/h)y)} \quad (1)$$

for $0 \leq u \leq w-1$ and $0 \leq v \leq h-1$. Next, the spectrum map S_k^p is obtained by

$$S_k^p(u,v) = \log(|F_k^p(u,v)| + 1) \quad (2)$$

Consequently, the pth CFA, $0 \leq p \leq 10$, yields n spectrum maps, $S_0^p, S_1^p, ..., $ and S_{n-1}^p.

Fig. 3 shows the 11 resultant spectrum maps of Fig. 2(a). It can be observed that the spectrum distributions in the central low frequency region are similar, so that this region of these spectrum maps cannot be used to distinguish one CFA structure from the others. Therefore, the central low frequency region of S_k^p can be discarded by using the following high-pass filter:

$$S_k^p(u,v)$$
$$= \begin{cases} 0, & \sqrt{\left(u - \dfrac{w}{2}\right)^2 + \left(v - \dfrac{h}{2}\right)^2} \leq r \times \min(h,w) \\ S_k^p(u,v) & \text{otherwise} \end{cases}$$
$$(3)$$

where the value of r is in $[0,1]$ and we set it to 0.3 empirically. By Eq. (3), it yields the high-pass spectrum map which is denoted by $S_k^{p,hp}$.

In the second step, a block-based incremental approach is proposed to further locate the high-energy blocks of $S_k^{p,hp}$. The energy density value of a $b \times b$ block $B_k^{p,hp}(u_b, v_b)$, where b is set to 16 empirically and (u_b, v_b) denotes the upper-left position of that block, is in $[0,1]$ and is calculated by

$$D_e\left(B_k^{p,hp}(u_b, v_b)\right)$$
$$= \frac{1}{b^2} \sum_{k=0}^{b-1} \sum_{l=0}^{b-1} \xi\left(S_k^{p,hp}(u_b+k, v_b+l), \mu_k^{p,hp}\right) \quad (4)$$

where $\mu_k^{p,hp}$ denotes the average spectrum value of the high-pass spectrum map $S_k^{p,hp}$ and the function ξ is defined by

$$\xi(x,t) = \begin{cases} 1, & x \geq t \\ 0 & \text{otherwise} \end{cases} \quad (5)$$

By Eq. (4), $B_k^{p,hp}(u_b, v_b)$ is defined to be the high-energy block if the condition $D_e(B_k^{p,hp}(u_b, v_b)) \geq T_E$ is held, where T_E is in $[0,1]$ and is set to 0.7 empirically. To locate the high-energy blocks of $S_k^{p,hp}$, the calculation of $D_e(B_k^{p,hp}(u_b, v_b))$ is started from $u_b = v_b = 0$. If the block $B_k^{p,hp}(u_b, v_b)$ is a high-energy block, we reserve all coefficients of that block and then increase u_b by $u_b = u_b + 16$ to determine whether or not the next block is the high-energy block. Otherwise, if the block $B_k^{p,hp}(u_b, v_b)$ is not a high-energy block, we set the coefficients in the first column of the block to 0 s and then increase u_b by $u_b = u_b + 1$ to repeat the above energy density calculation process for the next block. Such an incremental calculation method can reduce the average computational cost. When the condition $u_b > w - b$ is held, we set $u_b = 0$ and $v_b = v_b + b$, and repeat the same high-energy block determination process for the first block in the next row. The above process is continuously performed until all blocks have been processed. The final resultant spectrum map, denoted by $S_k^{p,he}$, is called the high-energy spectrum map. After completing the high-energy block determination for the $10n$ high-pass spectrum maps, in the third step, the coefficients in every high-energy block will be reserved, but those in the non-high-energy block will be discarded.

In the third step, a successive thresholding approach is proposed to extract the significant coefficients in every high-energy block of $S_k^{p,he}$. Let $S_{k,i}^{p,he}$ denote the current significant spectrum map at the i-th iteration which is formulated as

$$S_{k,i+1}^{p,he}(u,v) = \begin{cases} S_{k,i}^{p,he}(u,v), & S_{k,i}^{p,hp}(u,v) \geq T_{k,i}^{p,he} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

where the initial condition is $S_{k,0}^{p,he}(u,v) = S_k^{p,he}(u,v)$ and the adaptive threshold $T_{k,i}^{p,he}$ is equal to the average value of nonzero coefficients in $S_{k,i}^{p,he}$. Empirically, nine iterations are enough to filter out those insignificant coefficients and retain the significant ones in each high-energy block of $S_k^{p,he}$. Let $S_k^{p,s}$ denote the final resultant significant spectrum map after finishing the above successive thresholding process by Eq. (6). For the pth CFA structure, $0 \leq p \leq 10$, the final n significant spectrum maps are denoted by $S_0^{p,s}$, $S_1^{p,s}, ..., $ and $S_{n-1}^{p,s}$.

In the fourth step, for the pth CFA structure, $0 \le p \le 10$, the corresponding model map will be constructed based on the n significant spectrum maps obtained from the third step. For each position (u,v), we extract the related n significant coefficients $S_0^{p,s}(u,v)$, $S_1^{p,s}(u,v)$, ..., and $S_{n-1}^{p,s}(u,v)$ from the n significant spectrum maps and these extracted n coefficients form a histogram. Totally we have $(w \times h)$ histograms for each CFA structure. Pruning the outliers of each of $(w \times h)$ histograms, we can construct the model map of the pth CFA structure by

$$M^p(u,v)$$
$$= \frac{\sum_{k=0}^{n-1} S_k^{p,s}(u,v) \cdot \xi(S_k^{p,s}(u,v), \mu^{p,s}(u,v) + \kappa\sigma^{p,s}(u,v))}{\sum_{k=0}^{n-1} \xi(S_k^{p,s}(u,v), \mu^{p,s}(u,v) + \kappa\sigma^{p,s}(u,v))}$$
$$(7)$$

where the function ξ has been defined in Eq. (5); $\mu^{p,s}(u,v)$ and $\sigma^{p,s}(u,v)$ denote, respectively, the average and the standard deviation of the n coefficients, $S_0^{p,s}(u,v)$, $S_1^{p,s}(u,v)$, ..., and

$S_{n-1}^{p,s}(u,v)$; $\kappa = 1$ empirically. After performing the fourth step for the 11 CFA structures, it yields the 11 model maps, M^0, M^1, ..., and M^{10}. Fig. 4 shows the final resultant 11 model maps based on the 10 training images as shown in Fig. 2. We observe that except for the two model maps in Fig. 4(g) and (i), the other nine model maps are quite different. In fact, the model maps in Fig. 4(g) and (i) are still slightly different. In the next subsection, a Bhattacharya distance criterion will be used to measure the difference level between all model map pairs. Then, an identification scheme is presented to identify the CFA structure of each header-less mosaic image.

2.2. Proposed three-step matching scheme for identifying arbitrary CFA structures

In this subsection, based on the constructed 11 model maps and the difference level between all model map pairs, a three-step matching scheme is proposed to identify the CFA structure of any header-less mosaic image $I_{h\ell}$ with
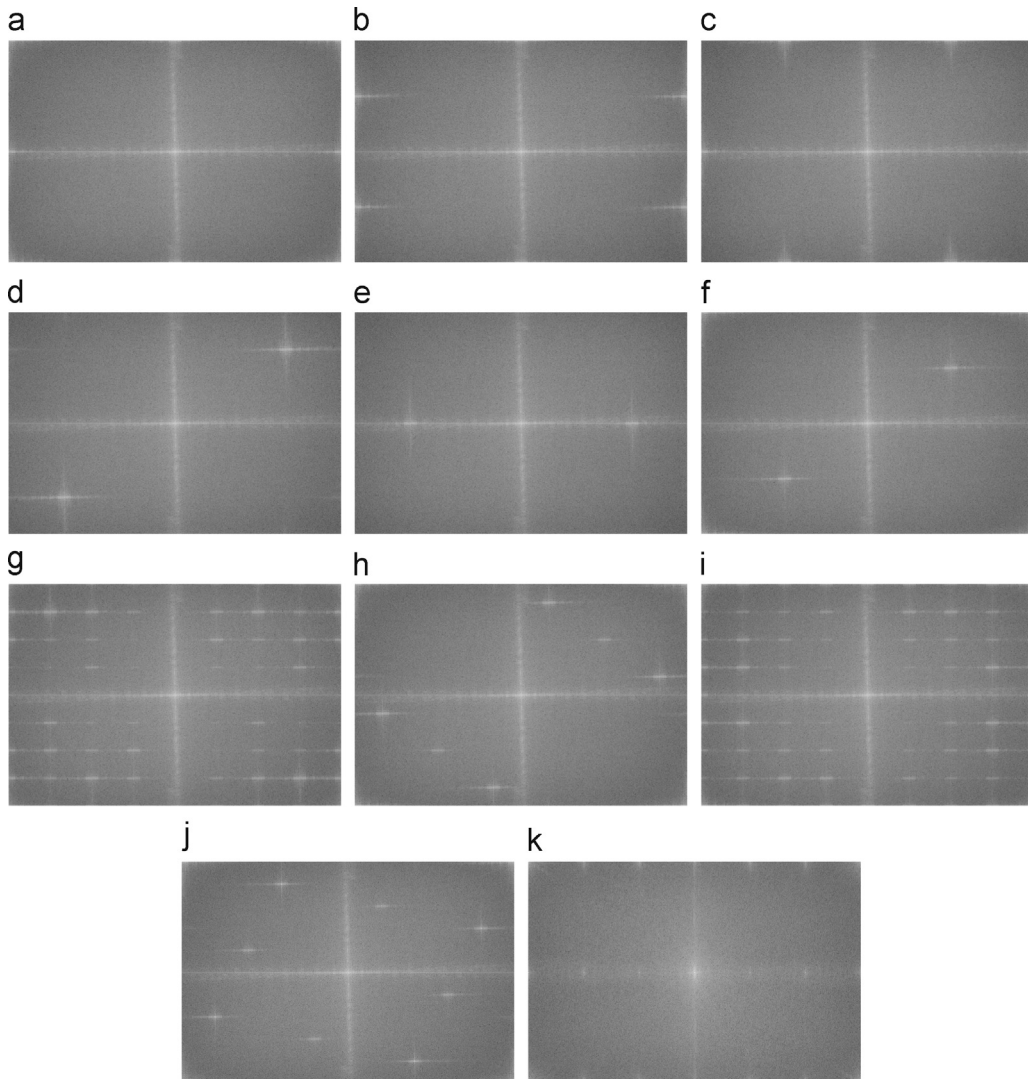

Fig. 3. Spectrum maps of the 10 mosaic images generated from Fig. 2(a) according to the 10 CFA structures as shown in Fig. 1. (a)–(k) denote the corresponding 10 spectrum maps for Fig. 1(a)–(k).

size $w' \times h'$. Before describing the details of the proposed three-step matching scheme, we first sketch the purpose of each step in the scheme. In the first step, we will first generate the significant spectrum map $S_{h\ell}^s$ of $I_{h\ell}$ and then calculate the normalized local energy of each block in the significant spectrum map to obtain the query map. According to the query map, the second step will find the best matched model map from the 11 based on the Bhattacharya distance criterion. If the CFA structure of the matched model map belongs to the CFA set $\{0, 1, 2, 3, 4, 5, 7, 9, 10\}$, then the CFA structure is reported as the output of the query map; otherwise, we proceed to the third step. In the third step, a set-based approach is presented to further determine whether the CFA structure of the query map is 6 or 8. We now describe the details of the proposed three-step matching scheme.

In the first step, we want to resize the input mosaic image $I_{h\ell}$ to the same size of the model map and transform the resized mosaic image to the query map. We first describe how to resize the input mosaic image with size $w' \times h'$ to be of size $w \times h$. If $w' > w$, we crop the last $w' - w$ columns of $I_{h\ell}$; otherwise, i.e. $w' < w$, we pad $I_{h\ell}$ with additional $w - w'$ columns produced by periodically repeating the last columns of $I_{h\ell}$. The two cases, $h' > h$ and $h' < h$, are handled in a similar way. The resized image of $I_{h\ell}$ is denoted by $I_{h\ell}^r$. Next, we perform the first three steps of the model map construction process, which has been described in the last subsection, on the resized mosaic image $I_{h\ell}^r$. After that, the significant spectrum map $S_{h\ell}^s$ of the mosaic image is obtained. Further, we partition $S_{h\ell}^s$ into a set of $b \times b$ blocks and for each block $B_{h\ell}^s(u_b, v_b)$, where (u_b, v_b) denotes the upper-left position of the block, the normalized local energy of that block is calculated by

$$E\left(B_{h\ell}^s(u_b, v_b)\right) = \frac{\sum_{k=0}^{b-1} \sum_{l=0}^{b-1} S_{h\ell}^s(u_b + k, v_b + l)}{\sum_{u=0}^{w-1} \sum_{v=0}^{h-1} S_{h\ell}^s(u, v)}. \tag{8}$$
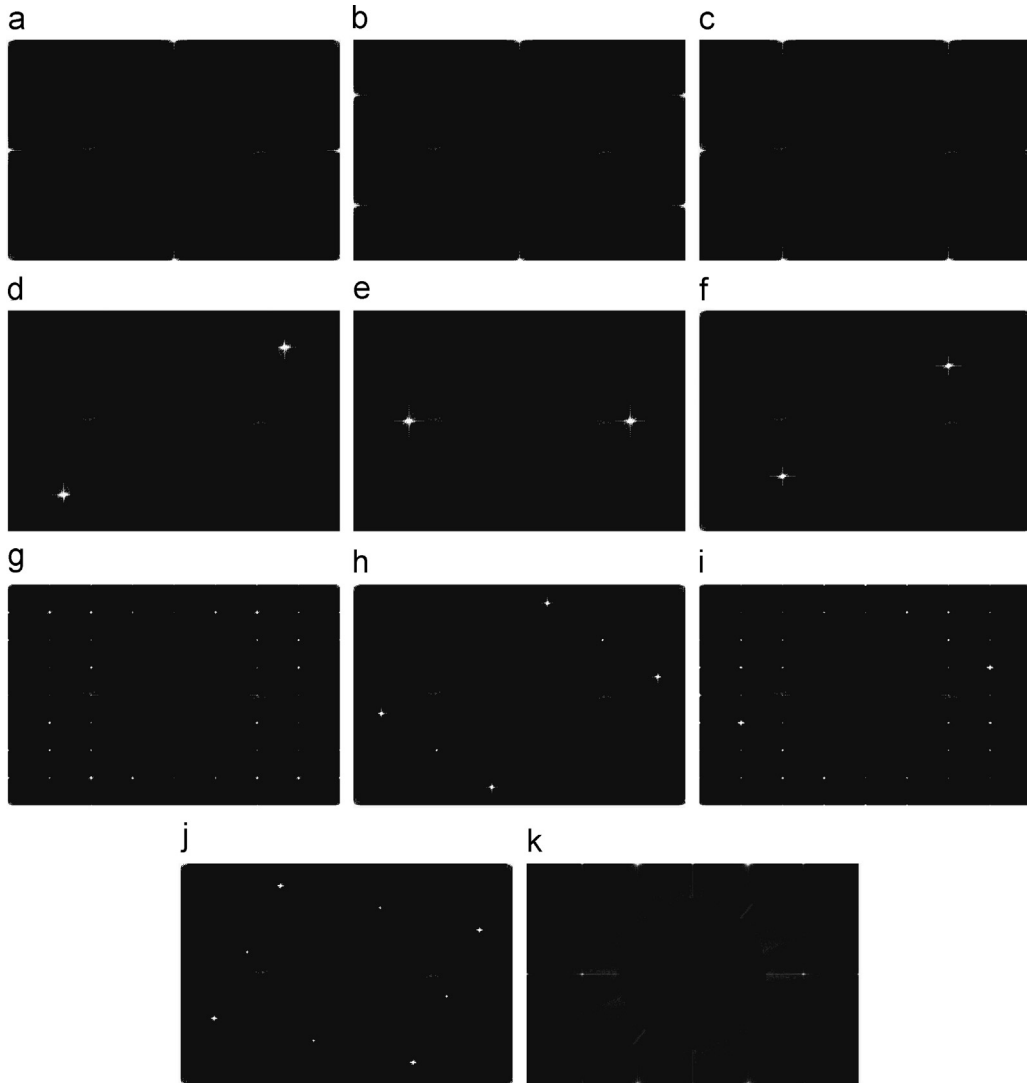

Fig. 4. The 10 constructed model maps of Fig. 2 for the 10 CFA structures: (a) Bayer CFA, (b) Lukac and Plataniotis CFA, (c) Yamanaka CFA, (d) diagonal stripe CFA, (e) vertical stripe CFA, (f) modified Bayer CFA, (g) HVS-based CFA, (h) type I pseudo-random CFA, (i) type II pseudo-random CFA, (j) type III pseudo-random CFA, and (k) Hirakawa CFA.

Because of $\sum_{(u_b,v_b)\in PO_b} E(B^s_{h\ell}(u_b,v_b)) = 1$, where PO_b denotes the set of upper-left positions of all $b \times b$ blocks, we have $0 \le E(B^s_{h\ell}(u_b,v_b)) \le 1$. After completing the computation of Eq. (8) for all blocks of the significant spectrum map, we then obtain the query map which is still denoted by $S^s_{h\ell}$ for convenience. Note that in implementation, the above normalized local energy computation process has been performed on each block for all model maps in advance and the calculated normalized local energy of each block in the model map can be reused again and again. The resultant model map is still denoted by M^p for convenience.

In the second step, we use the query map $S^s_{h\ell}$ to find the best matched model map M^j among the 11 model maps. Because $S^s_{h\ell}$ and M^j are represented by two normalized local energies of blocks which can be viewed as two normalized histograms, the similarity between $S^s_{h\ell}$ and M^j can be measured by the similarity of the two normalized histograms. In the proposed method, the Bhattacharya distance [17], which is an efficient metric for histogram matching, is used to determine the best matched model map M^j for the query map $S^s_{h\ell}$ by

$$j = \arg \min_{0 \le p \le 10} BD(S^s_{h\ell}, M^p) \qquad (9)$$

where $BD(S^s_{h\ell}, M^p)$ denotes the Bhattacharya distance between $S^s_{h\ell}$ and M^p and is defined by

$$BD(M^s_{h\ell}, M^p) = \left(1 - \sum_{(u_b,v_b)\in PO_b} \sqrt{E(B^s_{h\ell}(u_b,v_b))E(B^p(u_b,v_b))}\right)^{1/2} \qquad (10)$$

The lower the value of $BD(M^s_{h\ell}, M^p)$, the higher the similarity between the query map $S^s_{h\ell}$ and the model map M^p.

In the second step, we use the query map $S^s_{h\ell}$ to find the best matched model map M^j of the 11 model maps by

$$j = \arg \min_{0 \le p \le 10} BD(S^s_{h\ell}, M^p) \qquad (11)$$

where $BD(S^s_{h\ell}, M^p)$ denotes the Bhattacharya distance [17] between $S^s_{h\ell}$ and M^p and is defined by

$$BD(M^s_{h\ell}, M^p) = \left(1 - \sum_{(u_b,v_b)\in PO_b} \sqrt{E(B^s_{h\ell}(u_b,v_b))E(B^p(u_b,v_b))}\right)^{1/2} \qquad (12)$$

The lower the value of $BD(M^s_{h\ell}, M^p)$, the higher the similarity between the query map $S^s_{h\ell}$ and the model map M^p. After determining the best matched model map M^j by Eq. (12), two cases are needed to be considered. The first case is that $M^j \in \{M^0, M^1, M^2, M^3, M^4, M^5, M^7, M^9, M^{10}\}$ and the second is that $M^j \in \{M^6, M^8\}$. Table 1 illustrates the distinguishability level, i.e. the difference level, between all model map pairs, (M^k, M^l) for $0 \le k, l \le 10$, in terms of the Bhattacharya distance. From Table 1, we observe that when distinctive k and l are in $\{0, 1, 2, 3, 4, 5, 7, 9, 10\}$, the Bhattacharya distance value of $BD(M^k, M^l)$ is always greater than 0.75. It reveals that if the best matched model map belongs to the first case, the corresponding CFA structure can be definitely reported as the output of the query map. However, Table 1 indicates that the Bhattacharya distance value is less than 0.58, so the third step is necessary to further distinguish the HVS-based CFA structure (corresponding to M^6) from the type II pseudo-random CFA structure (corresponding to M^8). The above distinguishability analysis is coincident with the visual observation from Fig. 4.

In the third step, it is known that the best matched model map does not belong to the set for the first case, i.e. $j \notin \{0, 1, 2, 3, 4, 5, 7, 9, 10\}$, so further spectrum comparison is required to determine whether the corresponding CFA structure is the HVS-based CFA or the type II pseudo-random CFA, i.e. $j=6$ or $j=8$. Suppose the temporary CFA structure of the best matched model map obtained in the second step is $j=6$. Now we want to confirm whether j should remain as 6 or be changed to 8. Let a $b \times b$ block, where b is the same as the one used in the proposed training scheme, i.e. $b=16$, in the spectrum map be a significant block when its normalized local energy value is larger than 0. We take each significant block $B^s_{h\ell}(u_b, v_b)$ in $S^s_{h\ell}$ and examine the co-located blocks $B^6(u_b, v_b)$ and $B^8(u_b, v_b)$ in M^6 and M^8, respectively. If $B^6(u_b, v_b)$ is not a significant block but $B^8(u_b, v_b)$ is, then it offers one vote to M^8; on the contrary, if $B^6(u_b; v_b)$ is a significant block but $B^8(u_b, v_b)$ is not, then it offers one vote to M^6. In what follows, a set-based approach is proposed to realize the voting process. Let $PO^s_{hl,sb}$, PO^6_{sb}, and PO^8_{sb} denote the sets of upper-left positions of the significant blocks in $S^s_{h\ell}$, M^6, and M^8, respectively. By using $PO^{d_8}_{sb} = PO^8_{sb} (PO^6_{sb} \cap PO^8_{sb})$, where '\' and '∩' are the complement and intersection operations, respectively, we can count the number of significant blocks in $S^s_{h\ell}$ whose co-located blocks in M^8 are significant but insignificant in M^6. The favorite level of M^8 as

Table 1
The Bhattacharya distance values of all model map pairs.

Model map	M^0	M^1	M^2	M^3	M^4	M^5	M^6	M^7	M^8	M^9	M_{10}
M^0	0	0.76182	0.76547	1	0.99536	0.99343	0.97859	0.99353	0.82405	0.99195	0.8331
M^1	0.76182	0	0.99597	1	0.99621	0.99526	0.95597	0.99533	0.93641	0.99413	0.91976
M^2	0.76547	0.99597	0	1	0.99586	0.99413	0.95864	0.99422	0.91843	0.99281	0.93011
M^3	1	1	1	0	1	1	1	1	1	1	1
M^4	0.99536	0.99621	0.99586	1	0	0.99508	0.992	0.99515	0.99346	0.99387	0.82542
M^5	0.99343	0.99526	0.99413	1	0.99508	0	0.92846	0.83885	0.85246	0.99157	0.98206
M^6	0.97859	0.95597	0.95864	1	0.992	0.92846	0	0.96061	0.57329	0.98272	0.94862
M^7	0.99353	0.99533	0.99422	1	0.99515	0.83885	0.96061	0	0.94207	0.99019	0.98231
M^8	0.82405	0.93641	0.91843	1	0.99346	0.85246	0.57329	0.94207	0	0.98658	0.9153
M^9	0.99195	0.99413	0.99281	1	0.99387	0.99157	0.98272	0.99019	0.98658	0	0.9686
M^{10}	0.8331	0.91976	0.93011	1	0.82542	0.98206	0.94862	0.98231	0.91530	0.9686	0

the best matched model map can be defined by

$$FL_{hl}^8 = \frac{|PO_{hl,sb}^s \cap PO_{sb}^{d_8}|}{|PO_{sb}^{d_8}|} \qquad (13)$$

where $|PO|$ denotes the size of set PO and $0 \leq FL_{hl}^8 \leq 1$. If $FL_{hl}^8 \leq T_{FL}$, where $T_{FL} = 0.275$ empirically, the CFA structure of the query map is determined to be $j=8$; otherwise, the CFA structure of the query map remains as $j=6$. When the temporary CFA structure of the best matched model map obtained in the second step is $j=8$, it can be handled in a similar way.

Using the above three–step matching scheme, the CFA structure of the input mosaic image can be identified according to the best matched model map. Note that the proposed method can only identify the CFA structure considered in the proposed four-step training scheme. For the CFA structures which are not included in Fig. 1, we need to generate the corresponding training mosaic images from the full RGB color images and perform the proposed four-step training scheme for constructing the model maps. After combining the former 11 model maps with the newly constructed ones, we could identify all concerned CFA structures using the proposed three-step matching scheme.

2.3. Computational complexity analysis

In this subsection, the computational complexity analysis of the proposed method is offered. For comparison, the computational complexity analysis of the state-of-the-art method by Chiu et al. is also exploited.

Since the model maps constructed by the proposed four-step training-based scheme can be reused, we only focus on the computational complexity analysis in the three-step matching scheme. Without loss of generality, instead of using the image size $w \times h$ as the problem size parameter, let the image size be $N \times N$ for ease of analysis.

In the first step of the matching scheme, it takes $O(N^2)$ time to resize the input mosaic image to the same size of the model map. Then it takes $O(2N^2 \log_2 N)$ time to perform the Fourier transform on the resized mosaic image, yielding the corresponding spectrum map. Next, considering the worst case, the block-based incremental approach can be done in $O(b^2 \times N^2/b)(= O(bN^2))$, i.e. $O(16N^2)$ for $b=16$, time to locate all the high energy blocks of the spectrum map. It is obvious that the time complexity of the block-based incremental approach would be higher when $b > 16$. However, in our implementation, we observe that $b=16$ is sufficient to accurately locate the high-energy blocks in all high-pass spectrum maps, so that the corresponding time complexity can be written as $O(16N^2)$. Next, it takes $O(9N^2)$ time to extract the significant coefficients in every high-energy block of the spectrum map. It yields the significant spectrum map. Further, it takes $O(N^2)$ time

Table 2
Accuracy comparison of the proposed method and Chiu et al.'s method.

CFA structures	Proposed method			Chiu et al.'s method		
	First set (%)	Second set (%)	Third set (%)	First set (%)	Second set (%)	Third set (%)
Bayer CFA	100	100	100	100	100	97
Lukac and Plataniotis CFA	100	100	100	100	100	100
Yamanaka CFA	100	100	100	100	100	98
Diagonal stripe CFA	100	100	100	100	100	100
Vertical stripe CFA	100	100	100	100	100	100
Modified Bayer CFA	100	100	100	100	100	100
HVS-based CFA	100	100	100	100	100	100
Type I pseudo-random CFA	100	100	100	100	100	100
Type II pseudo-random CFA	100	100	100	100	100	100
Type III pseudo-random CFA	100	100	100	100	100	100
Hirakawa CFA	100	100	100	98	96	95
Average	100	100	100	99.8	99.6	99.1

Table 3
Execution-time comparison of the proposed method and Chiu et al.'s method in terms of milliseconds.

CFA structures	Proposed method			Chiu et al.'s method		
	First set	Second set	Third set	First set	Second set	Third set
Bayer CFA	210.61	193.23	194.23	22,876.21	25,854.08	85,022.56
Lukac and Plataniotis CFA	210.95	194.77	194.68	22,872.90	25,765.96	84,739.96
Yamanaka CFA	211.45	195.78	198.16	22,878.78	25,763.89	85,169.55
Diagonal stripe CFA	206.34	193.45	193.88	22,873.65	25,761.19	84,885.17
Vertical stripe CFA	206.83	191.37	192.98	22,873.88	25,832.85	85,021.59
Modified Bayer CFA	211.89	192.91	194.17	22,878.12	25,766.05	84,564.49
HVS-based CFA	209.94	193.19	199.87	22,879.31	25,766.95	84,678.81
Type I pseudo-random CFA	208.89	194.11	194.52	22,876.71	25,780.62	85,008.06
Type II pseudo-random CFA	208.11	192.99	201.12	22,875.71	25,765.93	85,893.82
Type III pseudo-random CFA	208.96	192.18	199.18	22,879.45	25,825.66	85,127.81
Hirakawa CFA	211.85	194.96	194.85	22,880.65	26,054.96	84,612.96
Average	209.62	193.54	196.15	22,876.85	25,812.56	84,974.98

to calculate the normalized local energy for all blocks, and we obtain the resultant query map.

In the second step of the matching scheme, based on the Bhattacharya distance criterion, it takes $O(11N^2/b^2)$, i.e. $O(11N^2/256)$, time to find the best matched model map of the query map. The third step of the matching scheme needs $O(3N^2/b^2)$, i.e. $O(3N^2/256)$, time to determine whether the CFA structure is the HVS-based CFA (with respect to structure 6) or the type II pseudo-random CFA (with respect to structure 8). Consequently, the overall time complexity required in the three-step matching scheme is bounded by $O(16N^2 + 2N^2 \log_2 N)$. When $N < 256$, the time bound can be rewritten as $O(16N^2)$; otherwise, the time bound is written as $O(2N^2 \log_2 N)$. Due to the fact that the size of the mosaic images captured by modern digital cameras is usually larger than 256×256, the overall computational complexity of the proposed method can be simplified to $T_{ours} = O(2N^2 \log_2 N)$. Considering the worst case for $N = 4096$, which covers nearly all possible resolutions of modern digital cameras, we thus have $T_{ours} = O(24N^2)$.

The computational complexity of Chiu et al.'s spatial domain-based method is analyzed as follows. To determine the CFA structure of the input mosaic image, Chiu et al.'s method assigns the primary color component of each pixel according to the considered CFA structure among the 11 CFA structures and then calculates the average square difference between each pixel and its neighboring pixels with the same primary color component within the $W \times W$ window. The overall average square difference can be calculated by summing up the average square differences of all pixels in the input mosaic image under the consideration of one specific CFA structure, taking $O(W^2 N^2)$ time in total. After calculating the total average square differences for the concerned 11 CFA structures, the CFA structure with the minimal total average square difference is selected as the final output. Their method takes $O(11W^2 N^2)$ time to determine the CFA structure of the input mosaic image. Because of $O(24N^2) < O(11W^2 N^2)$ for $W > 2$, it is obvious that the proposed method has lower computational complexity than Chiu et al.'s method.

3. Experimental results

For comparing the identification performance of the proposed frequency domain-based method and Chiu et al.'s spatial domain-based method, three test image sets are used to generate the test mosaic images to conduct the experiments. The first test set consists of 14 color images of size 512×768, which are adopted from the Kodak lossless true color image suite [16]. The second test set consists of 106 color images of size 576×768 adopted from [18], while the third test set consists of 100 color images of size 1200×1200 adopted from the website [19]. Totally 2420 test mosaic images were generated for the 11 CFA structures. All the experiments were implemented on an IBM compatible computer with Intel Core 2 duo E7400 CPU of 2.8 GHz, 2GB RAM, and Microsoft Windows XP operating system. The two concerned methods were realized by Borland C++ Builder 6.0.

As mentioned in the last section, among the 24 color images on the website, the first 10 images, as shown in Fig. 2, are used to generate the model maps for the 11 CFA structures, and the remaining 14 images are used as the test images. In fact, we select 5–15 images from the 24 color images for model map generation and find that the model maps generated from 10 color images are sufficient to identify the CFA structures of the 2420 test mosaic images. To determine the parameters used in the proposed method, we tune the values of r and T_E from 0 to 1 with step 0.1, the value of b from 8 to 32, and the value of κ from 0.5 to 2 with step 0.5, to find the optimal solution. Experiment results show that setting $r = 0.6$, $T_E = 0.7$, $r = 0.3$, $b = 16$, and $\kappa = 1$ can maximize the identification accuracy of the proposed method. In the third step of the proposed four-step training-based scheme, the number of iterations required in the successive thresholding approach is determined in the same way, and the number of iterations is tuned from 3 to 15. Empirically, nine iterations are enough.

After performing the proposed method and Chiu et al.'s method on the three sets of test mosaic images, for each CFA structure, the identification accuracy was measured by the ratio of the number of correctly identified test mosaic images over the number of total images. To determine the window size, $W \times W$, used in Chiu et al.'s method, we consider W from 7 to 31 and select $W = 21$ for maximizing the identification accuracy performance. The accuracy performance comparison is demonstrated in Table 2 which reveals that the proposed method can correctly identify the CFA structures for all test mosaic images. Chiu et al.s method works perfectly for the first and second test mosaic image sets; however, for the third set, their method reports few wrong outputs for the Bayer CFA and Diagonal stripe CFA. The execution-time performance comparison is shown in Table 3 and indicates that the proposed method takes much less execution-time when compared with Chiu et al.'s method. On average, the proposed method only needs 1% of execution-time required in Chiu et al.'s method to identify the CFA structure for one mosaic image.

In Table 3, the proposed method spends more execution-time on the first test mosaic image set than that on the other two sets. This is because that the images in the first set have higher saturation, i.e. there are higher difference values among the R, G, B color channels than that in the other two sets, resulting in more high-energy blocks to be processed by the proposed method. Experimental results have demonstrated the high identification accuracy and low computational cost merits of our proposed frequency domain-based identification method.

4. Conclusions

This paper has presented the proposed novel frequency domain-based method for identifying the CFA structure of the input header-less mosaic images with an arbitrary CFA structure. The proposed method first builds up the 11 model maps for the 11 concerned CFA structures by performing the four-step training-based scheme on a set of training mosaic images. According to the 11 constructed model maps, the three-step matching scheme is proposed

to identify the corresponding CFA structure of the input mosaic image. In addition, detailed computational complexity analysis of the proposed method is provided. For comparison, the computational complexity analysis of the state-of-the-art comparative method is also provided. Based on 2200 test mosaic images, the experimental results demonstrate that the proposed method can quickly and exactly identify the correct CFA structure for every test mosaic image. When compared with the state-of-the-art spatial domain-based method by Chiu et al., the proposed method has superior performance in identification accuracy and has significant execution-time improvement.

References

[1] R. Lukac, Single-Sensor Imaging: Methods and Applications for Digital Cameras, CRC Press/Taylor and Francis, Boca Raton, FL, 2008.

[2] R. Lukac, K.N. Plataniotis, Color filter arrays: design and performance analysis, IEEE Trans. Consum. Electron. 51 (April (4)) (2005) 1260–1267.

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

[4] K. Hirakawa, P.J. Wolfe, Spatio-spectral color filter array design for optimal image recovery, IEEE Trans. Image Process. 17 (October (10)) (2008) 1876–1890.

[5] W. Pennebaker, J. Mitchell, JPEG, Still Image Data Compression Standard, Van Nostrand, New York, 1993.

[6] H. Chen, M. Sun, E. Steinbach, Compression of Bayer-pattern video sequences using adjusted chroma subsampling, IEEE Trans. Circuits Syst. Video Technol. 19 (December (12)) (2009) 1891–1896.

[7] W.J. Yang, K.L. Chung, W.N. Yang, L.C. Lin, Universal chroma subsampling strategy for compressing mosaic video sequences with arbitrary RGB color filter arrays in H.264/AVC, IEEE Trans. Circuits Syst. Video Technol. 23 (April (4)) (2013) 591–606.

[8] S. Kim, N. Cho, Lossless compression of color filter array images by hierarchical prediction and context modeling, IEEE Trans. Circuits Syst. Video Technol. 24 (June (6)) (2014) 1040–1046.

[9] S. Farsiu, M. Elad, P. Milanfar, Multiframe demosaicing and super-resolution of color images, IEEE Trans. Image Process. 15 (January (1)) (2006) 141–159.

[10] B.K. Karch, R.C. Hardie, Adaptive wiener filter super-resolution of color filter array images, Opt. Express 21 (August (16)) (2013) 18820–18841.

[11] S.C. Pei, I.K. Tam, Effective color interpolation in CCD color filter arrays using signal correlation, IEEE Trans. Circuits Syst. Video Technol. 13 (June (6)) (2003) 503–513.

[12] E. Dubois, Frequency-domain methods for demosaicking of Bayer-sampled color images, IEEE Signal Process. Lett. 12 (December (12)) (2005) 847–850.

[13] B.K. Gunturk, J. Glotzbach, Y. Altunbask, R.W. Schafer, R.M. Mersereau, Demosaicking: color filter array interpolation, IEEE Signal Process. Mag. 22 (January (1)) (2005) 44–54.

[14] L. Zhang, X. Wu, A. Buades, X. Li, Color demosaicking by local directional interpolation and non-local adaptive thresholding, J. Electron. Imaging 20 (April (2)) (2011). 023016-1–023016-16.

[15] Y.H. Chiu, K.L. Chung, W.N. Yang, C.H. Lin, Y.H. Huang, Universal intra coding for arbitrary RGB color filter arrays in HEVC, J. Vis. Commun. Image Represent. 24 (October (7)) (2013) 867–884.

[16] R. Franzen, Kodak lossless true color image suite [Online]. Available: ⟨http://r0k.us/graphics/kodak/⟩.

[17] M. Mignotte, Segmentation by fusion of histogram-based k-means clusters in different color spaces, IEEE Trans. Image Process. 17 (May (5)) (2008) 780–787.

[18] [Online]. Available: ⟨http://tabby.vision.mcgill.ca/html/browsedownload.html ⟩ [Category: Landscapes].

[19] [Online]. Available: ⟨http://www.tecnick.com/public/code/cp_dpage.php?aiocp_dp=testimages⟩.