2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing

# AN IMPROVED ALGORITHM FOR CAMERA MODEL IDENTIFICATION USING INTER-CHANNEL DEMOSAICKING TRACES

*Yongjian Hu[1,2], Chang-Tsun Li[1],*
[1]Department of Computer Science
University of Warwick
Coventry CV4 7AL, UK
{Yongjian.Hu, ctli}@dcs.warwick.ac.uk

*Xufeng Lin[2], Bei-bei Liu[2]*
[2]School of Electronic and Information Engineering
South China University of Technology
Guangzhou 510640, China
eeyjhu@scut.edu.cn

***Abstract*—Most CFA (color filter array) interpolation-based digital image forensic methods characterize inter-pixel relationship with a linear model and use the estimated interpolation coefficients as features for image source camera identification. However, various CFA models and interpolation algorithms must be tried for coefficient estimation during the detection process in that the CFA pattern of an image is often unknown at the receiver's end. This incurs high computational complexity. Instead of using inter-pixel correlations, Ho *et al.* proposed to use inter-channel demosaicking/color interpolation traces for identifying the source camera model of a test image. In this work, we propose an improved algorithm. We first extract two variance maps by estimating the variances of each component of the green-to-red and green-to-blue spectrum differences, respectively, and then take the shape and texture features of these two maps for camera model identification. Experimental results show that our method achieves better overall performance.**

***Keywords-device identification; CFA; demosaicking; color interpolation; image forensics***

## I. INTRODUCTION

In the imaging process of most digital cameras, a color filter array (CFA) is placed before the sensor to capture one of the three primary colors for each pixel while the other two color components are interpolated with a specific demosaicking algorithm. Many interpolation algorithms utilize relationship between neighboring pixels to estimate the two missing color components, which inevitably introduces correlations among adjacent pixels or different color channels. As a result, the traces left by the color interpolation algorithm can be used as evidence for digital image source identification.

In order to determine the interpolation algorithm employed by a camera, Popescu and Farid [1] estimated the probability map and linear interpolation coefficients using an EM (Expectation Maximization) algorithm. The traces of the interpolation algorithm for source camera identification were also studied by Bayram *et al.* in [2]. They detected cyclical peaks in the spectrum caused by the periodicity of CFA interpolation, and fed the features into a support vector machine (SVM) for image classification (i.e., camera device identification). Swaminathan *et al.* [3] established linear equations for nine types of pixels based on the linear models of CFA interpolation. Since the equations were solved by searching through all CFA models and the interpolation algorithms, that scheme is time consuming. Wang *et al.* [4] proposed a blind detection method, which employs a covariance matrix to set up and solve the interpolation coefficient equations.

Different from the above inter-pixel correlations-based methods, Ho *et al.* [5] introduced an inter-channel correlations-based algorithm by assuming that the Fourier spectral coefficients of different color channels have the same variance. The variances of the coefficients of the color spectrum difference were used to distinguish different interpolation algorithms. With a simple and unified model, their method achieves relatively high accuracy. However, this method assumes that the red and blue channels work in an identical fashion, which is inconsistent with the way some types of cameras actually work. In addition, it often requires a large amount of computational load due to the EM estimation step for the calculation of variances and the use of K-Nearest Neighbors classifier for final classification.

In this paper, we proposed an improved algorithm that aims at overcoming the weaknesses of Ho *et al.*'s work. We obtain the variance map that corresponds to different color channels by calculating the variance of spectrum differences. Both shape and texture features are then extracted from the variance maps for classification. Compared with the algorithm in [5], our algorithm achieves better overall performance.

The paper is organized as follows. We first introduce the motivation in Section 2, and then describe the proposed algorithm in Section 3. Experimental results and conclusion will be given in Section 4 and 5, respectively.

## II. PROBLEM STATEMENT

It is commonly known that human eyes are more sensitive to the green component of visible light. For this reason, most CFAs tend to sample the G channel at a higher rate than those of the R and B channels. In the well-known Bayer CFA sampling pattern, G samples are twice as many as R and B samples. The spectrum of G channel therefore has less aliasing and its high frequency components are better preserved. Most of the color interpolation algorithms are based on the assumption that different color channels have similar high frequency components. They use the high frequency components of G channel to replace those of R and B channels (e.g., [7]-[10]). Taking account of this fact, Ho *et al.* [5] proposed to identify the camera model based on the interpolation algorithm employed because the interpolation

This work was partially supported by the EU FP7 Digital Image and Video Forensics project (Grant Agreement No. 251677), NSF of China 60772115, and the Fundamental Research Funds for the Central Universities, SCUT (No. 2012ZM0027).

978-0-7695-4712-1/12 $26.00 © 2012 IEEE
DOI 10.1109/IIH-MSP.2012.85

process would leave inter-channel demosaicking traces on the images. Below we briefly review [5] and point out its shortcomings. Since the elements of R and B channels appear at the same frequency in the Bayer CFA sampling pattern, we only use R channel as an example in the narration.

Let $R_s$ and $G_s$ be the color planes sampled by the CFA. For the Bayer CFA sampling pattern, the size of $R_s$ is only 1/4 of the image size while $G_s$ is 1/2 of the image size. To obtain $R_s$ and $G_s$ with the full image size, zeros are filled in at non-sample locations. Assume that $R$ and $G$ are color planes reconstructed from sample values $R_s$ and $G_s$, respectively. Due to the reason in the last paragraph, $G$ is simply reconstructed from $G_s$. The reconstruction of $R$ can then be helped with $G$. Let $G_{sr}$ be the color plane that is produced by sampling $G$ at red sample locations and filling in zeros at other color sample locations. $R$ can be reconstructed by

$$\begin{aligned} & R = \psi\{R_s - G_{sr}\} + G \\ \Rightarrow\ & R = R_l + (G - G_l) \\ \Rightarrow\ & R = R_l + G_h \end{aligned} \tag{1}$$

where $\psi\{\bullet\}$ denotes interpolation algorithm. $R_l$ and $G_l$ denote the low-frequency bands of $R$ and $G$, respectively. $G_h$ denotes the high-frequency band of $G$. Likewise, one can reconstruct the blue color plane $B$.

Ho *et al.* investigated the relationship between the interpolation algorithm and (1). They performed the Fourier transform on both sides of (1):

$$\widetilde{G} - \widetilde{R} = \alpha\left(\widetilde{G}_{sr} - \widetilde{R}_s\right) \tag{2}$$

where $\alpha$ is the frequency response of the interpolation algorithm. $\widetilde{G}$, $\widetilde{G}_{sr}$, $\widetilde{R}$ and $\widetilde{R}_s$ are the spectra of $G$, $G_{sr}$, $R$ and $R_s$, respectively. So the distinction between different interpolation algorithms can be reflected by the frequency response $\alpha$. Ho *et al.* further calculated the variance of the spectrum difference $\Delta = \widetilde{G} - \widetilde{R}$ as:

$$\begin{aligned} & Var\{\Delta_{uv}\} = \alpha_{uv}^2 Var\left\{\left(\widetilde{G}_{sr} - \widetilde{R}_s\right)_{uv}\right\} \\ & = \alpha_{uv}^2\left(\sigma_{uv(G_{sr})}^2 + \sigma_{uv(R_s)}^2 - 2\sigma_{uv(G_{sr})}\sigma_{uv(R_s)}\rho_{uv}\right) \end{aligned} \tag{3}$$

where $\Delta_{uv}$ is the $uv^{th}$ entry of $\Delta$. $\sigma_{uv(G_{sr})}^2$ and $\sigma_{uv(R_s)}^2$ are the variances of $\widetilde{G}_{sr}$ and $\widetilde{R}_s$ at $(u,v)$, respectively. $\rho_{uv}$ is the correlation coefficients between $\widetilde{R}_{s(uv)}$ and $\widetilde{G}_{sr(uv)}$. Assuming that the components in $\widetilde{R}$ and $\widetilde{G}$ have the same variance, Ho *et al.* let $\sigma_{uv(G_{sr})}^2 = \sigma_{uv(R_s)}^2 = \sigma_{uv}^2$. By applying these equations to (3), they obtained

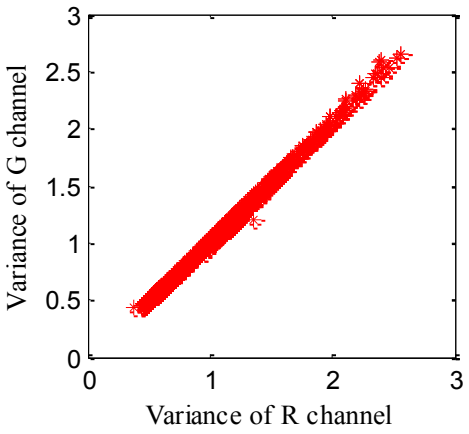
$$\begin{aligned} Var\{\Delta_{uv}\} &= \alpha_{uv}^2 Var\left\{\left(\widetilde{G}_{sr} - \widetilde{R}_s\right)_{uv}\right\} \\ &= 2\alpha_{uv}^2(1 - \rho_{uv})\sigma_{uv}^2 \end{aligned} \tag{4}$$

Because $\alpha$ corresponds to a low-pass filter, it leaves a salient trace on the variance of the color spectrum difference [5]. The change of $Var\{\Delta_{uv}\}$ can reflect the change of $\alpha_{uv}^2$. So $Var\{\Delta_{uv}\}$ was used in [5] to distinguish different interpolation algorithms. During camera model identification, Ho *et al.* did not distinguish the variances of $\Delta = \widetilde{G} - \widetilde{R}$ and those of $\Delta = \widetilde{G} - \widetilde{B}$. Below we make some analysis of [5].

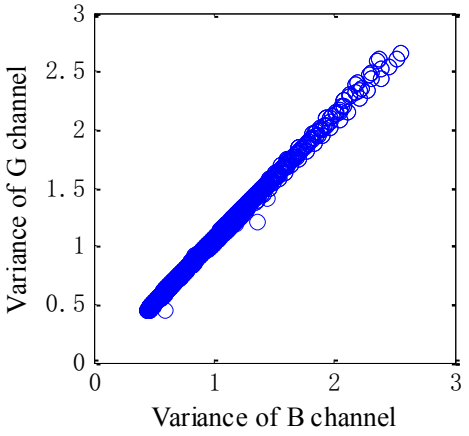
Let us first investigate the assumption about $\sigma_{uv(G_{sr})}^2 = \sigma_{uv(R_s)}^2 = \sigma_{uv}^2$. We calculate the variances from 100 images of size $4272 \times 2848$, each image being split into non-overlapping $100 \times 100$ blocks. The relationship of the variance of the spectrum coefficients between different color channels is shown in Fig 1. It can be seen that almost all points are scattered along the line y = x. This verifies that the assumption of [5] about $\sigma_{uv(G_{sr})}^2 = \sigma_{uv(R_s)}^2 = \sigma_{uv}^2$ is true.

We then investigate feasibility of Ho *et al*'s assumption that the variances of $\Delta = \widetilde{G} - \widetilde{R}$ and those of $\Delta = \widetilde{G} - \widetilde{B}$ are the same. We give an example in Fig. 2. It clearly demonstrates the difference between R and B channels of Panasonic DMC-LX2. $\rho_{uv}^2$ of G-R channels and G-B channels have quite different patterns. So we say that some types of cameras treat the R and B channels differently. In order to design an algorithm for identifying different camera models, we should treat the variances of $\Delta = \widetilde{G} - \widetilde{R}$ and those of $\Delta = \widetilde{G} - \widetilde{B}$ separately.

Other apparent weaknesses of the algorithm in [5] include the use of the EM algorithm to calculate the variances and the Nearest Neighbors classifier for image classification. Both would incur high computational complexity. To tackle these problems, we propose an improved algorithm.

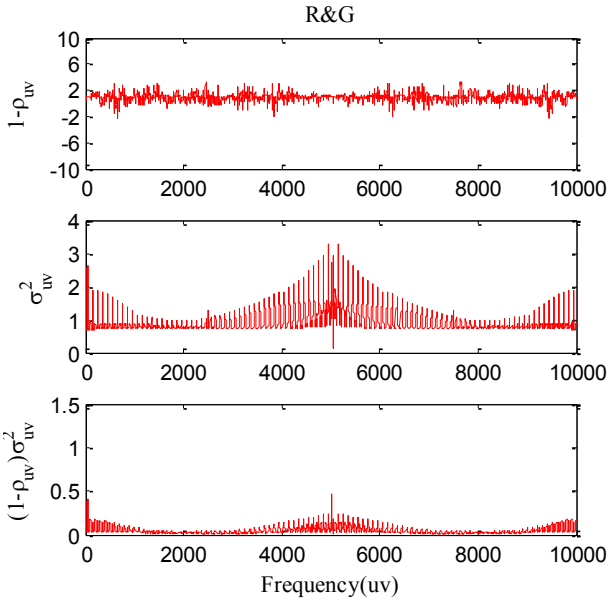



(a)

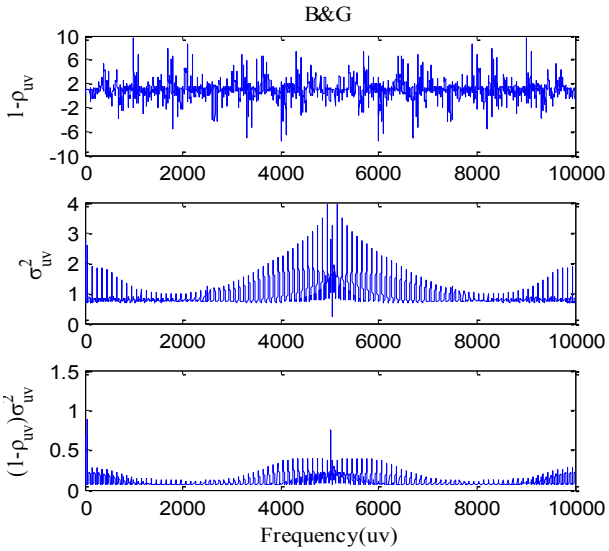



(b)

**Fig. 1.** Relationship of variances of the spectrum coefficients between R and G channel (a) and between B and G (b).




(a)




(b)

**Fig. 2.** Difference between R channel (a) and B channel (b). The two-dimensional coordinate of 100 × 100 frequencies have been converted into 1 × 10000 one-dimensional frequency coordinate. R&G refers to the pair of color channels R and G, while B&G refers to the pair of color channels B and G.

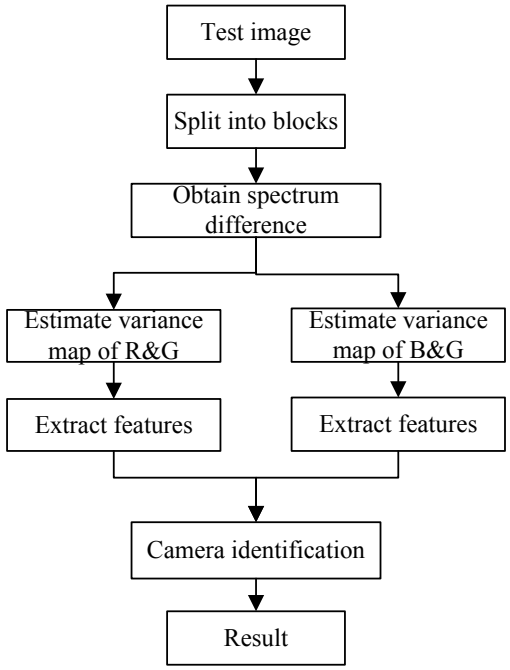



**Fig. 3.** Flow chart of the proposed algorithm.

## III. PROPOSED ALGORITHM

We propose an improved algorithm, as shown in Fig. 3. The steps and reasons are given as follows.

1) Given a test image or a number of test images, we divide

each image into non-overlapping blocks to obtain sufficient samples for variance estimation. If the number or the size of test images is small, we use small-size blocks, and vice versa. Note that if samples are insufficient, the estimation of variance is inaccurate; however, if the block size is too small, the statistical significance of the extracted spectrums would be reduced.

2) DFT is applied to the three color planes for each block to get $\widetilde{R}$, $\widetilde{G}$, and $\widetilde{B}$. The spectrum difference $\widetilde{G}-\widetilde{R}$ and $\widetilde{G}-\widetilde{B}$ are then calculated. For notational convenience, we denote the $M$ samples at location $(u,v)$ of each block as $\Omega_i(u,v), i=1,2,...,M$. We can get the variance maps $\sum_{RG}$ and $\sum_{BG}$ corresponding to different color channels, where $\sum(u,v)=Var\{\Omega_i(u,v)\}$.

3) We then extract the image features from the obtained variance map. These features will drive the AdaBoost classifier. We have two major reasons to use classifiers like Stumps AdaBoost classifier rather than K-Nearest Neighbors classifier. Firstly, the latter would become more computationally inefficient with the increase of dimensionality of the variance map. Secondly, it is easier to control the set of features driving the classifiers like Stumps AdaBoost classifier. For example, we can add features into or subtract features from the set to optimize the classifier. Besides, we can easily borrow features originally derived for other applications.

In this work, we propose to use the shape and texture of the variance maps because we observed that they are very similar for the same camera model and are quite different for different camera models (i.e., low intra-model and high inter-model variations). We select 4-dimensional shape features from [11] and 8-dimensional texture features from [12] for each color channel. To make the shape and texture more prominent, we equalize the histogram of the variance map prior to the feature extraction. The shape features employed are

$$f_1=\eta_{20}+\eta_{02} \tag{5}$$

$$f_2=(\eta_{20}-\eta_{02})^2+4\eta_{11}^2 \tag{6}$$

$$f_3=(\eta_{30}-3\eta_{12})^2+(3\eta_{21}-\eta_{03})^2 \tag{7}$$

$$f_4=(\eta_{30}+\eta_{12})^2+(\eta_{21}+\eta_{03})^2 \tag{8}$$

where $\eta_{pq}$ is the scaled $(p+q)$ order central moments of $f(x,y)$. The above 4-dimensional invariant moments have been proved to be invariant to the translation, rotation, and scale transform [11].

We use the statistical distribution of pixel values to reflect the characteristics of image textures. This work extracts Gray Level Co-occurrence Matrixes (GLCM)-based features [12], which is defined as the joint probability matrix of two fixed-interval pixel values. To describe texture features more comprehensively, the textures in four different directions (i.e., 0°, 45°, 90°, and 135°) are considered. The four groups of texture features are calculated as follows.

The first group is Texture Energy, which is a measure of homogeneity of the image.

$$f_5=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{mean}\left(\sum_{g_1}\sum_{g_2}p(g_1,g_2)^2\right) \tag{9}$$

$$f_6=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{STD}\left(\sum_{g_1}\sum_{g_2}p(g_1,g_2)^2\right) \tag{10}$$

The second group is Texture Contrast, which is a measure of contrast or the amount of local variations present in an image.

$$f_7=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{mean}\left(\sum_{g_1}\sum_{g_2}k^2p(g_1,g_2)^2\right) \tag{11}$$

$$f_8=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{STD}\left(\sum_{g_1}\sum_{g_2}k^2p(g_1,g_2)^2\right) \tag{12}$$

The third group is Texture Correlation, which is a measure of gray-tone linear dependencies across the image.

$$f_9=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{mean}\left(\left(\sum_{g_1}\sum_{g_2}g_1g_2p(g_1,g_2)-\mu_x\mu_y\right)/\sigma_x\sigma_y\right) \tag{13}$$

$$f_{10}=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{STD}\left(\left(\sum_{g_1}\sum_{g_2}g_1g_2p(g_1,g_2)-\mu_x\mu_y\right)/\sigma_x\sigma_y\right) \tag{14}$$

The fourth group is Texture Entropy, which is a measure of the amount of information contained within an image.

$$f_{11}=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{mean}\left(-\sum_{g_1}\sum_{g_2}p(g_1,g_2)\ln p(g_1,g_2)\right) \tag{15}$$

$$f_{12}=\underset{\{0^\circ,45^\circ,90^\circ,135^\circ\}}{STD}\left(-\sum_{g_1}\sum_{g_2}p(g_1,g_2)\ln p(g_1,g_2)\right) \tag{16}$$

where

$$k=|g_1-g_2|,\ \mu_x=\sum_{g_1}g_1\sum_{g_2}p(g_1,g_2)$$

$$\mu_y=\sum_{g_2}g_2\sum_{g_1}p(g_1,g_2)$$

$$\sigma_x^2=\sum_{g_1}(g_1-\mu_x)^2\sum_{g_2}p(g_1,g_2)$$

$$\sigma_y^2=\sum_{g_2}(g_2-\mu_y)^2\sum_{g_1}p(g_1,g_2)$$

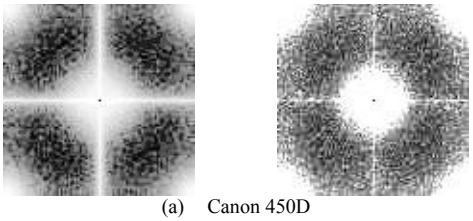
For (9)-(16), $g_1$ and $g_2$ are values of the two fixed-interval pixel locations, and $p(g_1, g_2)$ represents GLCM of the variance map.

In total, we get 24 image features from the variance maps, that is, 12 image features for each pair of color channels (R and G, and B and G).

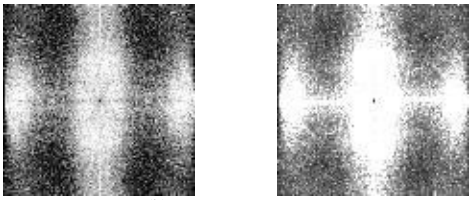
4) We choose Stumps AdaBoost in [13] for classification because of its simplicity and high speed. This algorithm can achieve high accuracy by using misclassified samples repeatedly.

## IV. Experiment and Discussion

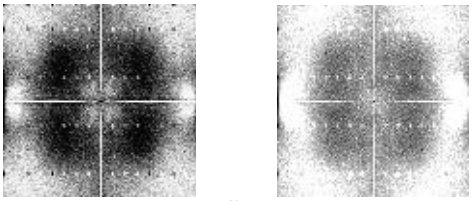
To evaluate the performance of the proposed algorithm, images from seven popular consumer digital cameras including Canon 450D (DSLR), Canon IXUS870, Nikon D40 (DSLR), Nikon D90 (DSLR), Nikon E3200, Olympus C3100Z and Panasonic DMC-LX2), are used in our experiment. For each camera, 50 of 100 images are used for the training set while the other 50 are used for the test set. The image block size is $100 \times 100$ and the number of weak classifiers of AdaBoost is 20.



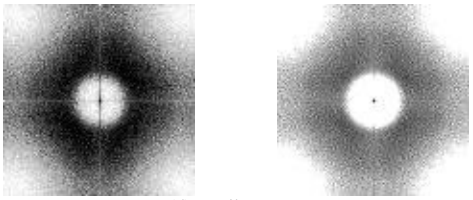
(a) Canon 450D



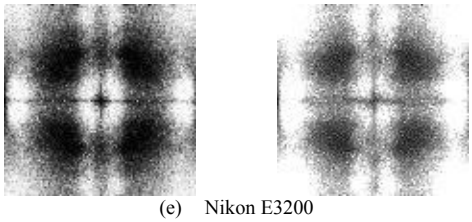
(b) Canon IXUS870



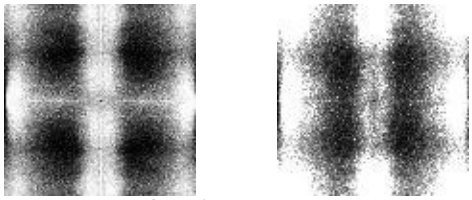
(c) Nikon D40



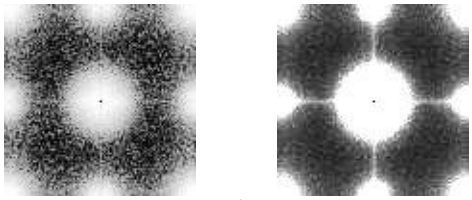
(d) Nikon D90



(e) Nikon E3200



(f) Olympus C3100Z



(g) Panasonic DMC-LX2

**Fig. 4.** Sample variance maps. The variance maps extracted from G-R spectrum difference (left) and from G-B spectrum difference (right).

It can be observed from Fig. 4 that the variance maps for different camera models usually possess different patterns. This fact justifies our algorithm. On the other hand, there are distinguishable differences between the variance maps from G-R spectrum difference and those from G-B spectrum difference. This fact verifies that the manner we treat the variances extracted from G-R spectrum difference and those extracted from G-B spectrum difference separately is sound.

**Table 1.** Comparison of two algorithms. QF (quality factor)

| algorithms / JPEG QF | Ours | [5] |
|---|---|---|
| uncompressed | 91.90% | 97.00% |
| 100 | 95.67% | 94.95% |
| 90 | 94.29% | 92.90% |
| 80 | 92.33% | 91.76% |
| 70 | 91.90% | 90.29% |

To investigate the detection accuracy and robustness of the proposed algorithm, we conduct pair-wise classification among these seven cameras. We compare our algorithm with the algorithm in [5] to demonstrate the improvement our method has made. It can be seen from Table 1 that in the case of JPEG compression, our algorithm outperforms the algorithm in [5] in both detection accuracy and robustness. Note that the detection accuracy in this work is the average value. Although the algorithm in [5] outperforms ours for uncompressed images, JPEG compressed images are more common for real-world applications. Therefore, the advantage

of our algorithm is apparent. The reason why the algorithm in [5] outperforms ours for uncompressed images is yet to be studied. However, it seems that slight compression can remove high frequency noise and thus benefits the extraction of our features. In other words, our shape and texture features are robust against compression to some degree.

We then discuss two cases for the treatment of R and B channels. In Case 1, we obtain two variance maps, one from G-R spectrum difference, and the other from G-B spectrum difference. We separately calculate the shape and texture features from these two maps and obtain 24 features. In Case 2, we do not distinguish R and B channels and treat them equally. So we only obtain 12 features. Table 2 shows the results achieved in these two cases. Apparently, the accuracy for Case 1 is higher than that in Case 2. This result proves that the way we advocate to treat the variance map from G-R spectrum difference and that from G-B spectrum difference separately is more advantageous.

**Table 2.** Comparison of two different approaches

| approach / JPEG QF | Case 1 | Case 2 |
|---|---|---|
| uncompressed | 91.90% | 88.52% |
| 100 | 95.67% | 94.67% |
| 90 | 94.29% | 93.00% |
| 80 | 92.33% | 91.43% |
| 70 | 91.90% | 89.43% |

## V. CONCLUSION

The use of inter-channel demosaicking traces for source camera identification has seldom been studied in literature. In this work, we have proposed a new algorithm of using the features of inter-channel demosaicking traces for determining the source camera models. Our major contributions include the separate treatment of G-R and G-B spectrum differences and the use of a feature-driven classifier. Experimental results have confirmed the superiority of our algorithm in terms of both detection accuracy as well as robustness to JPEG compression. Our algorithm is also adaptable to different features. Good features would enhance its performance. This paper only uses simple shape and texture features to demonstrate the feasibility of our method. The selection of more suitable features is our future work.

## REFERENCES


[1] A.C. Popescu and H. Farid, "Statistical Tools for Digital Forensics," in *Proc. of 6th International Workshop on Information Hiding*, pp.128-147, 2004.

[2] S. Bayram, H. Sencar, and N. Memon, "Source camera identification based on CFA interpolation," in *Proc. of IEEE International Conference on Image Processing*, Genova, Italy, vol. 3, no. III, pp. 69-72, Sep. 2005.

[3] A. Swaminathan, M. Wu, and K.J.Ray Liu, "Nonintrusive component forensics of visual sensors using output images," *IEEE Transactions on Information Forensics and Security*, vol. 2, no. 1, pp. 91-106, Mar. 2007.

[4] B. Wang, X. Kong, X. You, H. Fu, "Blind CFA Interpolation Detection Based on Covariance Matrix," *Transactions on Journal of Electronics & Information Technology (Chinese)*, vol. 31, pp. 1175-1179, May. 2009.

[5] John S. Ho, Oscar C. Au, J. Zhou, and Y. Guo, "Inter-channel demosaicking traces for digital image forensics," in *Proc. of IEEE International Conference on Multimedia and Expo*, Suntec City, pp. 1475-1480, July 2010.

[6] Y. Fang, A.E. Dirik, X. Sun, N. Memon, "Source Class Identification for DSLR and Compact Cameras," in *Proc. of IEEE International Workshop on Multimedia Signal Processing*, 2009.

[7] I. Pekkucuksen, Y. Altunbasak, "Edge Strength Filter Based Color Filter Array Interpolation," *IEEE Transactions on Image Processing*, 2011. To appear.

[8] I. Pekkucuksen, Y. Altunbasak, "Gradient based threshold free color filter array interpolation," in *Proc. of IEEE International Conference on Image Processing*, pp. 137-140, Sept. 2010.

[9] J. F. Hamilton and J. E. Adams, "Adaptive color plan interpolation in single sensor color electronic camera," US Patent, 5629734, 1997.

[10] B. K. Gunturk, J. Glotzbach, Y. Altunbasak, R. W. Schafer, and R. M. Mersereau, "Demosaicking: color filter array interpolation," *IEEE Signal Processing Magazine*, vol. 22, no. 1, pp. 44-54, Mar. 2005.

[11] M-K Hu. "Visual pattern recognition by moment invariants," *IEEE Transactions on Information Theory*, vol. 8, no. 2, pp. 179-187, Feb. 1962.

[12] R.M. Haralick, K. Shanmugam, and I. Dinstein, "Texture features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, no. 6, pp. 610-621, Nov. 1973.

[13] J. Rennie, "Boosting with decision stumps and binary features," *Massachusetts Inst. Technol.*, Cambridge, MA, Tech. Rep., Apr. 2003.