SOURCE CAMERA IDENTIFICATION BASED ON CFA INTERPOLATION

SevincBayram [a]*,HusrevT.Sencar* [b]*,NasirMemon* [b]*, IsmailAvcibas* [a]

[a] Dept. of Electrical and Computer Eng., Uludag University, Bursa, TURKEY
[b] Dept. of Computer and Information Sci., Polytechnic University, Brooklyn, NY, USA

ABSTRACT

In this work, we focus our interest on blind source camera identification problem by extending our results in the direction of [1]. The interpolation in the color surface of an image due to the use of a color filter array (CFA) forms the basis of the paper. We propose to identify the source camera of an image based on traces of the proprietary interpolation algorithm deployed by a digital camera. For this purpose, a set of image characteristics are defined and then used in conjunction with a support vector machine based multi-class classifier to determine the originating digital camera. We also provide initial results on identifying source among two and three digital cameras.

1. INTRODUCTION

The advances in digital technologies have given birth to very sophisticated and low-cost hardware and software tools that are now integral parts of our daily lives. This trend has brought with it new issues and challenges concerning he integrity and authenticity of digital images. The most challenging of these is that digital images can now be easily created, edited and manipulated without leaving any obvious traces of having been modified. This in turn undermines the credibility of digital images presented as news items or as evidence in a court of law since it may no longer be possible to distinguish whether an introduced digital image is the original or not. As a consequence, one can no longer take the authenticity of digital images for granted. Image forensics, in this context, is concerned with determining the source and potential authenticity of a digital image.

Digital watermarking has been introduced as a means for authenticating digital documents that are most likely to undergo various processing [2]. Although this approach enables the extractor to establish the degree of authenticity and integrity of a digital image, it practically requires that the watermark be embedded during the creation of the digital object. This limits watermarking to applications where the digital object generation mechanisms have built-in watermarking capabilities. Therefore, in the absence of widespread adoption of digital watermarks (which is likely to be the case in the (foreseeable future), watermarking cannot be offered as a general solution to the complex problem of authentication. Consequently, in order to determine origin, veracity and nature of digital images, alternative approaches that do not require any prior knowledge of the original digital image need to be considered (blind authentication techniques). *Atthepresenttime,however, there is a severe lack of techniques that could achieve thesegoals*

In this paper, we focus our interest on the source camera identification problem. That is, given an image can we determine the digital camera that was used in capturing the image? It should be noted that all digital cameras encode the camera model, type, date, time, and compression information in the image header; however, it is not possible to authenticate these information. In this regard, the success of blind image authentication techniques rely on the validity of assumption that all images produced by a camera will exhibit certain characteristics, regardless of the captured scene, that are unique to that camera due to its proprietary image formation pipeline. In our prior work [1], we studied the same problem and identified a set of image features by selectively combining the features based on image quality metrics [3] and higher-order statistics of images [4]. This approach essentially requires the design of a classifier that is able to capture the variations in the designated image features, introduced by different digital cameras.

Another promising approach in this area is made by Lucas *etal.* [5]. In their work, sensor's *patternnoise* is characterized via wavelet-based image denoising. The reference noise pattern for each digital camera is obtained by averaging over a number of raw (unprocessed) images, and the source camera for a given image is determined by correlating the noise pattern with the image itself. Alternately, 9n the present work, we exploit the fact that most state-of-the-art digital cameras, due to cost considerations, employ a single mosaic structured *color filter array* (CFA) rather than having different filters for each color component. As a consequence, each pixel in the image has only one color component associated with it, and each digital camera employs a proprietary interpolation algorithm in obtaining the missing color values. Our approach is inspired by the technique proposed by Popescu *etal* . for image tamper detection [6]. The rationale for their

0-7803-9134-9/05/$20.00 ©2005 IEEE

technique is that the process of image tampering very often requires up-sampling operation (which in turn introduces periodic correlations between the image pixels), and they designated statistical measures to detect such phenomena.

The rest of this paper is organized as follows. In section 2, we briefly describe the image formation process in digital cameras. The details for identifying traces of interpolation are provided in Section 3. We present our experimental results in Section 4. and conclude in Section 5.

2. IMAGE FORMATION IN DIGITAL CAMERAS

Although much of the details on the camera pipeline is considered proprietary information to each manufacturer, the general structure and sequence of stages in the camera pipeline remains to be very similar in all digital cameras. The basic structure of a digital camera pipeline is shown in Figure 1-(a) [7]. After light enters the camera through the lens, a set of filters are employed, the most important being an anti-aliasing filter. The anti-aliasing filter is needed when the spatial frequency of the scene being captured is larger than the distance between the elements (pixels) of the charge-coupled device (CCD) array.

The CCD array is the main component of a digital camera, and it's the most expensive component. Each light sensing element of CCD array integrates the incident light over the whole spectrum and obtains an electric signal representation of the scenery. Since each CCD element is essentially monochromatic, capturing color images requires separate CCD arrays for each color component. However, due to cost considerations, rather than using multiple arrays, the CCD array is arranged in a pattern by using different spectral filters, typically red, green and blue (RGB). This mask in front of the sensor is called the color filter array (CFA). Since any given CCD element only senses one band of wavelengths, the raw image collected from the array is a mosaic of red, green and blue pixels Figures 1-b and 1-c display a CFA pattern using RGB and YMCG color space respectively for a 6x6 pixel block.

Looking at the RGB values in the CFA pattern, it is evident that each sub-partition of four pixels only provides information on two green, one red, and one blue pixel values. Hence, the missing RGB values need to be interpolated for each pixel (demosaicing). The interpolation is typically carried out by applying a weighting matrix (kernel) to the neighborhood around a missing value. There are a number of different interpolation (demosaicing) algorithms and different manufactories use different interpolation techniques, i.e. kernels with different sizes and shapes. The processing block shown in the Figure 1-a produces the final image and it includes a number of operations which include color processing and compression. Although the operations and stages explained in this section are the standard section of the digital camera pipeline, the exact processing detail in each stage varies from one manufacturer to other, and even in different camera models manufactured by the same company. It should also be noted that many components in the image formation pipeline of various digital cameras, (*e.g*., lens, optical filters, sensor) are produced by a limited number of manufactures. Therefore, this should be taken into consideration in associating image features with the properties of digital cameras. However, interpolation (demosaicing) algorithm and the design of the CFA pattern remain to be proprietary to each digital camera manufacturer.

In the next section we will describe how the variations in color interpolation can be exploited to classify the images either originating from one camera or the other.

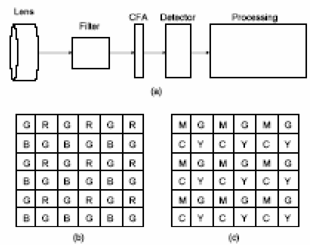


Figure 1. (a) The more important stages of a camera pipeline are shown. (b) CFA pattern using RGB values. (c) CFA pattern using YMCA values.

3. IDENTFYING TRACES OF INTERPOLATION

In [6], Popescu *et al.* employed Expectation/ Maximization (EM) algorithm to detect traces of up-sampling to identify images (or parts of images) that have undergone resizing. The EM algorithm consists of two major steps: an expectation step, followed by a maximization step. The expectation is with respect to the unknown underlying variables, using the current estimate of the parameters, and conditioned upon the observations. The maximization step then provides a new estimate of the parameters. These two steps are iterated until convergence [8]. The EM algorithm generates two outputs. One is a two-dimensional data array, called *probabilitymap* , with each entry indicating the similarity of each image pixel to one of the two groups of samples, namely, the ones correlated to their neighbors and those ones that are not, in a selected kernel. On this *map* the regions identified by the presence of periodic patterns indicate the image parts that have undergone up-sampling operation. The other output is the estimate of the *weighting (interpolation) coefficients* which designate the amount of contribution from each pixel in the interpolation kernel.

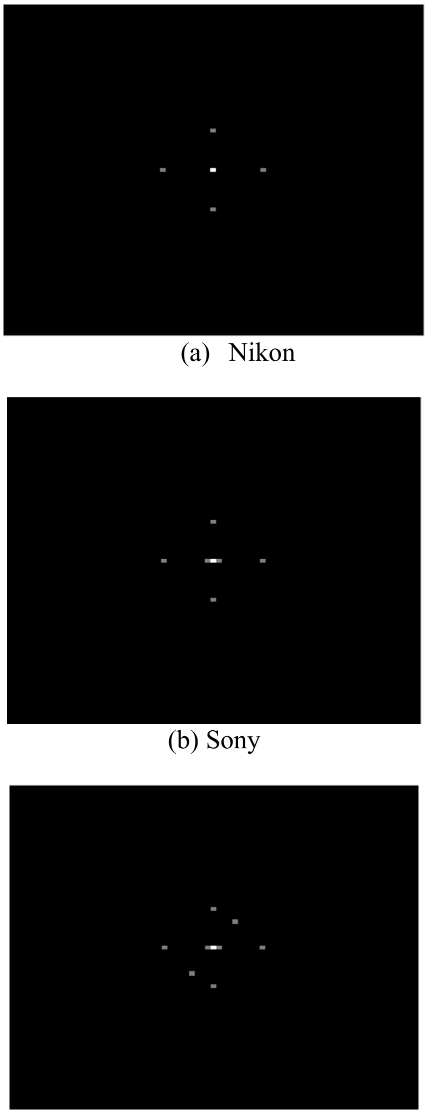
Since in a typical digital camera RGB channels are heavily interpolated, we propose to apply a similar procedure to determine the correlation structure present in each color band and classify images accordingly. Our initial experimental results [1] indicate that both the size of interpolation kernel and the demosaicing algorithm vary from camera to camera. Furthermore, the interpolation operation is highly non-linear, making it strongly dependent on the nature of the depicted scenery. In other words, these algorithms are fine-tuned to prevent visual artifacts, in forms of over-smoothed edges or poor color transitions, in busy parts of the images. On the other hand, in smooth parts of the image, these algorithms exhibit a rather linear characteristic. Therefore, in our analysis we treat smooth and non-smooth parts of images separately.

Since no *a-priori* information is assumed on the size of interpolation kernel (which designates the number of neighboring components used in estimating the value of a missing color component) probability maps are obtained for varying sizes of kernels. When observed in the frequency domain, these probability maps yield to peaks at different frequencies with varying magnitudes indicating the structure of correlation between the spatial samples. In designing our classifier we rely on two sets of features: The set of weighting coefficients obtained from an image, and the peak location and magnitudes in frequency spectrum. In Figure 2, sample magnitude responses of frequency spectrum of the probability maps for three cameras (Sony, Nikon and Canon) are given. The three responses differ in peak locations and magnitudes.

4. EXPERIMENTAL RESULTS

An SVM classifier was used to test the effectiveness of the proposed features. There are a number of SVM implementations available publicly, and we have used the LibSvm package [9]. We have also used the sequential forward floating search (SFSS) algorithm to select the best features from the given set.

In the first part of our experiments, we have used two camera models: Sony DSC-P51 and Nikon E-2100. The two cameras have both a resolution of 2 mega-pixels. The pictures were taken with maximum resolution, size of 1600x1200 pixels, auto-focus, no focusing, and other settings at default values. In order to detect properties due to the camera source not the texture of images, we used the pictures that were taken from the same scene by two cameras.



(a) Nikon

(b) Sony

(b) Canon

Figure 2. Frequency spectrum of probability maps obtained by three makes of digital cameras.

A picture data set was made by obtaining 140 pictures from each model. One fifth of these images were used for training. Then the designed classifier is used in classifying the previously unseen 4/5 of the images. We used 75x75 pixel parts of the images for experiments. First we extracted features assuming a 3x3 interpolation kernel for both Sony and Nikon cameras. The accuracy is measured as 89.3%. The corresponding confusion matrix is given in Table-1.

Table 1. The confusion table for 2 cameras assuming a 3x3 interpolation kernel

		Predicted	
		Nikon	Sony
Actual	Nikon	95.71	4.29
	Sony	17.14	82.86

Then we extract the features considering a neighborhood of 4x4. Correspondingly the accuracy in detection increased to 92.86 and the corresponding confusion matrix is in Table 2. The same experiment is repeated for 5x5 neighborhoods which lead to an accuracy of 95.71%. The corresponding confusion matrix is given in Table 3. As seen from the tables accuracy improves with larger kernel sizes. These results suggest that the actual size of the interpolation kernel used for CFA interpolation is not smaller than the considered sizes which were empirically known to be true [1].

Table 2. The confusion matrix for 2 cameras assuming a 4x4 interpolation kernel

		Predicted	
		Nikon	Sony
Actual	Nikon	91.43	8.57
	Sony	5.71	94.29

Table 3. The confusion tatrix for 2 cameras assuming a 5x5 interpolation kernel

		Predicted	
		Nikon	Sony
Actual	Nikon	94.64	5.36
	Sony	3.57	96.43

In order to see how the proposed features perform when considering three cameras, we also obtained 140 images from Canon Powershot S200. These images were acquired randomly from internet and consist of different sceneries. So we didn't know the exact setting used at the time of capture. Again we used SVM and SFSS to classify three cameras. We extract features from 5x5 neighborhoods. The accuracy was 83.33%, and corresponding confusion matrix is provided in Table 4. Larger neighborhood sizes and new features will be considered in the final version of the paper.

Table 4. The confusion table for 3 cameras assuming a 5x5 interpolation kernel

		Predicted		
		Nikon	Sony	Canon
Actual	Nikon	85.71	10.71	3.57
	Sony	10.71	75	14.28
	Canon	0	10.71	89.28

5. CONCLUSIONS AND FUTURE WORK

In this paper, we propose to identify the source camera of a digital image based on traces of color interpolation in the RGB color channels. For this, we generate a number of measures using EM algorithm. A classifier was then is designed and used to determine how reliably the selected measures could classify the images originating from the two and three cameras.

The proposed approach is another step taken in the direction of devising a set of techniques to solve blind source camera identification problem. This method is, unfortunately, limited to images that are not heavily compressed as the compression artifacts suppress and remove the spatial correlation between the pixels due to CFA interpolation.

6. REFERENCES

[1] M. Kharrazi, H. T. Sencar, N. Memon, "Digital Camera Model Identification," *Proc.ofICIP* , 2004.

[2] Special Issue on Data Hiding, *IEEETransactionson SignalProcessing* , Vol. 41, No. 6, 2003.

[3] I. Avcibas, N. Memon and B. Sankur, "Steganalysis using Image Quality Metrics," *IEEE Transactions on ImageProcessing* , Jan. 2003.

[4] S. Lyu and H. Farid, "Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines," *Proc.ofInformationHidingWorkshop* , 2002

[5] J. Lucas, J. Fridrich, and M. Goljan, "Determining Digital Image Origin Using Sensor Imperfections," *Proc. ofIS&TSPIE* , vol 5680, 2005

[6] A. Popescu and H. Farid, "Exposing Digital Forgeries by Detecting Traces of Re-sampling," *IEEE TransactionsonSignalProcessing* , 2004.

[7] J. Adams, K. Parulski and K. Sapulding, "Color Processing in Digital Cameras," *IEEE Micro* , Vol. 18, No.6, 1998.

[8] Todd Moon, 'The Expectation Maximization Algorithm', *IEEESignalProcessingMagazine, November1996.*

[9] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," 2001, Software available at http://www.csie.ntu.edu.tw/»cjlin/libsvm