Classifier Combination Based Source Identification for Cell Phone Images

Bo Wang[1], Yue Tan[1] , Meijuan Zhao[1], Yanqing Guo[1] and Xiangwei Kong[1]
[1] School of Information and Communication Engineering, Dalian University of Technology
Dalian, Liaoning, P.R. China
[e-mail: bowang@dlut.edu.cn, yue985@mail.dlut.edu.cn, zhaomeijuan@mail.dlut.edu.cn, guoyq@dlut.edu.cn, kongxw@dlut.edu.cn]
*Corresponding author: Bo Wang

Received June 9, 2015; revised September 10, 2015; accepted October 14, 2015; published December 31, 2015

Abstract

Rapid popularization of smart cell phone equipped with camera has led to a number of new legal and criminal problems related to multimedia such as digital image, which makes cell phone source identification an important branch of digital image forensics. This paper proposes a classifier combination based source identification strategy for cell phone images. To identify the outlier cell phone models of the training sets in multi-class classifier, a one-class classifier is orderly used in the framework. Feature vectors including color filter array (CFA) interpolation coefficients estimation and multi-feature fusion is employed to verify the effectiveness of the classifier combination strategy. Experimental results demonstrate that for different feature sets, our method presents high accuracy of source identification both for the cell phone in the training sets and the outliers.

Keywords: Image forensics, cell phone source identification, classifier combination, interpolation coefficient

This work is supported by the National Science Foundation of China (No. 61502076), the Scientific Research Project of Liaoning Provincial Education Department (No. L2015114), the Open Fund of Artificial Intelligence Key Laboratory of Sichuan Province (No. 2012RZJ01), and also the Fundamental Research Funds for the Central Universities (Grant No.DUT13RC201)

http://dx.doi.org/10.3837/tiis.2015.12.019 **ISSN : 1976-7277**

1. Introduction

According to the IDC's report [1], worldwide smart phone market has achieved a total of 1 billion in 2013 for the first time. The advantages of low-cost devices and easy access to amateur users have opened the smart phone floodgate. This means a new life style that people share photos with WiFi, Bluetooth etc, and send images by MMS (Multimedia Message Service). As a result, in the forensics context, the fast-growing smart phone trend has brought in an increasing number of image evidence captured by cell phone. Therefore, it is important to check the integrity and authenticity of the cell phone images presented as evidence in court. Digital image forensics which aims at ballistic analysis and exposing potential semantic manipulation of the image, has become necessary for legal purpose and security investigation [2].

In a practical blind digital forensics scenario, an analyst is assumed to gather clues and evidence from a given cell phone image without access to the device that created it [3]. An important piece of evidence is the identity of the source camera. Thus, the source identification for cell phone image becomes a branch of digital image forensics, whose task is to determine the cell phone that was used for capturing the given image.

The cell phone image source includes two different meanings. One is in term of mobile model that denotes products from different manufacturers. The other means the alternate cell phones of the same model [4-8]. In this study, we focus on the cell phone model identification.

In the area of cell phone model identification, several residual artifacts have been exploited in previous literatures. In [9], Celiktutan *et al* explored three sets of source identification features, namely binary similarity measures, image quality measures and high order wavelet statistical features. They further compared three types of decision-level fusion schemes including confidence-level fusion, rank-level fusion and abstract-level fusion in their experiments, in conjunction with SVM (Support Vector Machine) classifier [10]. By using 16 cell phones in 6 brands as experimental samples, the method received an overall average accuracy of 95.1%. Similar work is accomplished by Tsai *et al* in [11]. Also, Sun *et al* proposed a new method for source cell phone identification based on multi-feature fusion [12]. Features are selected by SFFS (Sequential Floating Feature Selection) method from three sets, which consist of higher-order statistics, image quality measures and CFA interpolation coefficients. For 8 cell phones in 3 brands, an overall average accuracy of 95% was achieved. Furthermore, they discussed the situation of classification of the different brands and the same brand. For 3 cell phones from different brands, a perfect accuracy of 100% was achieved, although the number of experimental samples seemed a little insufficient. In the more difficult scenario of classification of 4 cell phones from the same brand Nokia, the method proposed in [12] also got a good performance of 95%. Besides, the parameters of lateral chromatic aberration are also used to identify source cell phone, by maximizing the mutual information between different color components [13].

As a result of using a structured color filter array in front of sensor to obtain a mosaic image rather than full RGB color component image in the cell phone, the CFA interpolation is indispensable to recreate the missing color component for each pixel. The CFA interpolation artifacts, which are thus considered as one of the most important components in image pipeline, are widely exploited as a fingerprint for cell phone identification, as well as digital camera identification. Chuang *et al* [14] presented a study of cell phone camera model linkage

based on CFA interpolation. Furthermore, they evaluated the dependency on the content of training image collection via variance analysis. Gökhan and Ismail use SVD (Singular Value Decomposition) to obtain the micro and macro statistical feature vector introduced by CFA interpolation [15]. Most of these algorithms could achieve the classification accuracy of 90% and even higher, for several cell phone branches.

Although there are differences between cell phone and digital camera in terms of sensor, aperture, zoom and so on, the imaging pipeline is almost the same. Similar works could be found in the correlated area of digital camera identification in recent years, and most of these algorithms of digital camera identification perform well in cell phone identification [16-23]. The typical algorithm was proposed by Swaminathan *et al* [23], using a linear model to estimate the CFA coefficients. The details of the method could be found in Section 2.1.

To our best knowledge, most of the cell phone and digital camera source identification methods extract multi-dimensional features and use Fisher's linear discriminant or SVM as the classifier. As a typical multi-class classification problem in pattern recognition, this implies a tacit assumption that the given image was captured by the camera models existed in the training process because these classifiers can only distinguish the classes included in the training model. This assumption is impractical because it is impossible to train the camera models including the entire cameras in the market. In this case, the assumption means that an inevitable false classification would occur if there is an image captured by a new unknown device. In this paper, we define the device "outlier" when it is a new unknown device and out of the training model. Though the assumption is impractical, the scenario could be acceptable for digital camera source identification. The reason is that the number of mainstream digital cameras is well limited.

As for cell phone source identification, it is obvious that the assumption of traversal of all cell phone models can not be satisfied. The mainstream cell phone models are many more than those of digital cameras. Besides, various copycat cell phones increase the difficulty in the construction of training models. In this case, the previous algorithms based on traditional multi-class classification could be considered as impractical methods for real world source identification.

The proposed scheme in this paper differs from the previous works in term of unknown cell phone model identification. In this paper, we present a MC (multi-class) and OC (one-class) classifier combination method to distinguish the unknown mobile model in source identification for cell phone images.

The paper is organized as follows. In Section 2, the CFA coefficient features and multi-feature fusion consisting of image quality measure and high-order statistics extracted for classifier is described. The strategy of OC and MC classifier combination is presented and discussed in Section 3. The experiments are demonstrated in Section 4, where we indicate the performance of proposed method for 20 different cell phone models. Finally the paper is concluded in Section 5.

2. Feature Sets

Related prior studies on camera source identification have provide several efficient ways to determine the image source. These solutions can be classified into two classes: component parameter based methods and statistical characteristics based methods. Typical component parameter based methods can be found in [3,14,19-23], which widely discuss the information about CFA pattern and interpolation coefficients and present high performance in term of

identification accuracy. The statistical characteristics based methods usually use one or several sets of characteristics, such as binary similarity measures [9], image quality metrics [10], high order wavelet statistical features [12], SVD features [15] and so on. In this paper, a CFA coefficient feature set proposed in [23] and a feature set of multi-feature fusion proposed in [12] are used, separately.

2.1 CFA Coefficient Feature Set

As is known to all, the image formation pipeline of digital camera equipped on the cell phone can be described as **Fig. 1**:

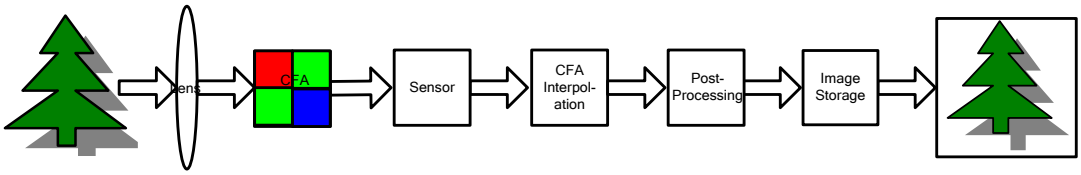


Fig. 1. Image formation pipeline

The rays from the scene of the real world first pass through the lens and a sophisticated designed filter, which is called color filter array, CFA. A typical CFA pattern called Bayer CFA consists of one red and blue color component and two green components in a 2×2 cell. The following sensor detects sampled R/G/B component at different pixel locations according to the CFA pattern. The output of sensor is considered as a mosaic image because there is only one color component at every single pixel. To rebuild the true-color image, the missing color components of each pixel are interpolated using the local area sampled data, which is called CFA interpolation. After that, a post-processing such as white balancing or gamma correction is carried out, and finaly the image is stored as pre-set format such as JPEG. Obviously, CFA interpolation is an important step to maintain the image quality in the image formation pipeline, because 2/3 of the image data is re-built by the interpolation processing. There are several different CFA interpolation algorithms with different performance [24,25]. As a unique feature set of camera brand identification, CFA interpolation coefficients are considered as an important parameter for identifying the camera source of an image.

In this paper, we use the non-intrusive algorithm, which is proposed in [23], to estimate the interpolation coefficients as the feature vector. The CFA interpolation coefficients estimation algorithm consists of two parts. First, the interpolation coefficients are preliminarily estimated with a linear model. The pixels in the image are first divided into three categories according to the texture information as following:

$$\begin{cases} Horizontal: \left(H_{x,y} - V_{x,y}\right) > T \\ Vertical: \quad \left(V_{x,y} - H_{x,y}\right) > T \\ Smooth: \quad \left|\left(H_{x,y} - V_{x,y}\right)\right| \le T \end{cases} \tag{1}$$

$H_{x,y}$ and $V_{x,y}$ respectively denote the second-order gradient values of horizontal and vertical gradients, which can be computed in equations (2) and (3), and T is a suitably chosen threshold.

$$H_{x,y} = \left| I_{x,y-2} + I_{x,y+2} - 2I_{x,y} \right| \tag{2}$$

$$V_{x,y} = \left| I_{x-2,y} + I_{x+2,y} - 2I_{x,y} \right| \tag{3}$$

where $I_{x,y}$ denotes the pixel value at the location (x, y) in the image. Without doubt, the image pixels are finally divided into nine sets according to three categories in R, G and B components. Suppose that we have a matrix of the pixel values directly captured by the cell phone, denoted by $\mathbf{A}$ of dimension $N_e \times N_u$, the linear interpolation model can be represented as following:

$$\mathbf{Ax = b} \tag{4}$$

$\mathbf{b}$ of dimension $N_e \times 1$ denotes the pixel values to be interpolated, and $\mathbf{x}$ of dimension $N_u \times 1$ stands for the interpolation coefficients to be estimated. Of course this is an idealized model for the CFA interpolation, as there is always perturbation introduced by the other image operations such as gamma correction, white balance and especially, lossy JPEG compression. Considering the perturbation, the model should be revised as:

$$(\mathbf{A} + \mathbf{E})\mathbf{x} = \mathbf{b} + \mathbf{r} \tag{5}$$

A solution for $\mathbf{x}$ with this model is to solve the minimization problem:

$$\min_{\mathbf{E},\mathbf{r}} \left\| [\mathbf{E} \quad \mathbf{r}] \right\|_F, \quad subject \quad to \quad (\mathbf{A} + \mathbf{E})\mathbf{x} - (\mathbf{b} + \mathbf{r}) = 0 \tag{6}$$

The Frobenius norm of the matrix $[\mathbf{E} \quad \mathbf{r}]$ can be computed as:

$$\left\| [\mathbf{E} \quad \mathbf{r}] \right\|_F = \left(\sum_{m=1}^{N_e} \sum_{n-1}^{N_u} \left| e(m,n) \right|^2 + \sum_{m=1}^{N_e} \left| r(m) \right|^2 \right)^{1/2} \tag{7}$$

After the CFA interpolation coefficients are preliminarily estimated, an interpolation error, which computed by a weighted sum of errors of nine pixels categories, is obtained to evaluate the veracity of the estimation. Also, detection statistics deduced by the errors are obtained as a sorting index to search different CFA patterns. Considering the high complexity, we simplify the CFA pattern process in our method. We use a typical diagonal Bayer pattern for the CFA. A full brute force search of different CFA patterns can be easily implemented in the extension.

2.2 Multi-Feature Fusion Set

As illustrated in **Fig. 1**, although the image pipeline is similar in different cell phones, the parameters in CFA interpolation and JPEG compression are different, which may cause differences in the quality of the image as well as the higher-order statistic features of the image. These tiny differences may hardly be detected by the naked eyes, but they can be used as the unique features of the image, thus provide evidences to identify the source cell-phones. A multi-feature fusion method proposed in [12] has combined the higher-order statistics and image quality measure to identify the image source of cell phone.

Image quality measure have been used for steganalysis [26] and tampering detection [27]. Typically, 13-dimonsional statistical features related to image quality measure are involved in the multi-feature fusion. **Table 1** shows the three categories of image quality measures and their corresponding detailed descriptions.

Table 1. Image Quality Measures

Category	Description
Measures based on pixel differences	Mean square error: $Q1$
	Mean absolute error: $Q2$
	Modified infinity norm: $Q3$
Measures based on correlation	Image fidelity: $Q4$
	Normalized cross-correlation: $Q5$
	Czenakowski correlation: $Q6$
	Mean angle similarity: $Q7$
Measures based on spectral distance	Block spectral magnitude error: $Q8$
	Block spectral phase error: $Q9$
	Block specrtral phase-magnitude error: $Q10$
	Spectral magnitude error: $Q11$
	Spectral phase error: $Q12$
	Spectral phase-magnitude error: $Q13$

Also, the higher-order statistics have been proved as an effective tool for steganalysis and tampering detection [28]. The statistical model for photographic images could be built upon several frequency-domain transformations. Without loss of generality, we use wavelet-like decomposition as the model. The processing of decomposition consists several separable quadrature mirror filters, which splits the frequency space of the image into multiple scales and orientations, typically a vertical, a horizontal and a diagnal subband. For full-color RGB images, the three color channels are decomposed separately. $V_k(i,j)$, $H_k(i,j)$ and $D_k(i,j)$ denote the vertical, horizontal and diagnal subbands respectively. In each orientation, the mean, variance, skewness and kurtosis of the coefficients in each subband are used to construct the feature vector, as (8) to (11) shown.

$$W_{1-3} = mean\left(V_k(i,j)\right) \tag{8}$$

$$W_{4-6} = variance\left(V_k(i,j)\right) \tag{9}$$

$$W_{7-9} = skewness\left(V_k(i,j)\right) \tag{10}$$

$$W_{10-12} = kurtosis\left(V_k\left(i, j\right)\right) \tag{11}$$

The computation are applied in three color channel of an image, and a feature vector consists of 36 features are generated.

To restrain the correlation in the feature sets, a feature selection algorithm are implemented. There are several different feature selection algorithms. A simple and effective method is SFFS, which brute force search all of the combination of the features. For a specified dimensional feature vector, SFFS selects the feature combination with the highest accuracy. For all of the dimensions, a correlation curve between feature subsets and performance could be achieved, which is further used for feature selection. More details could be found in [12] and [29]. Respect to the work in [12], we use 19 effective features to construct the feature vector.

3. Classifier Combination

The source identification of cell phone image is traditionally considered as a pattern recognition problem. The typical solution is that for several different classes with training samples as side-information, we mark the classes with different labels, and extract distinguishing feature vector. By feeding a classifier with the feature vector, a model is expected to be built to predict the best matching label for a given new sample. In this methodology, the classifier usually constructs a linear boundary or non-linear hyperplane in the two- or high- dimension space. Thus a key assumption is that the classifier must have the side-information of the training samples, as well as the class label. And also the classifier can only be assigned to the test sample with matching labels where the classifier has already known in the training process. Is this practical for the cell phone source identification in term of forensics?

Our proposed work say no unfortunately. The task of cell phone source identification is to determine the source of the image, which means we do not know how we obtain the image. However, the assumption of classification is self-contradictory because it includes an implication that the test image belongs to one of the training classes. Thus, for a more practical scenario, the forensic analyst obtains a multi-class model with training image samples, including a large cell phone model set as large as he/she could obtain. Nevertheless, the problem which he has to face with is that the test image could be captured by any cell phone in the market. If the multi-class model is directly used to predict the category of the test image, an inevitable misclassification would occur when the test image is from an outlier cell phone.

To address this issue, a combined classifier consisting of MC and OC classifiers is proposed. In the combination strategy, the multi-class classifier is supposed to provide a tool that determines the best match label in the training model, while the one-class classifier exposes the outliers of the training model. In another word, the MC classifier is used to answer the question that which cell phone captures the test image, and the OC classifier is expected to answer if the classification result of the MC classifier is correct.

The combination strategy of MC and OC classifier is illustrated as **Fig. 2**. Supposing we have an image data set consisting of image samples from *N* cell phone models, it is easy to obtain all of the OC classifier models $M_{OC1}, M_{OC2}, \cdots, M_{OCN}$. When we extract the feature vector of the test image, the MC classifier model, called M_{MC}, is first used to predict a best

matching label, denoted by C_i. Then, the corresponding OC classifier model M_{OCi} is used to identify whether the test image is captured by the specific cell phone. A positive result confirms that the test image is captured by the cell phone, and a negative result exposes an unknown cell phone source of the test image.

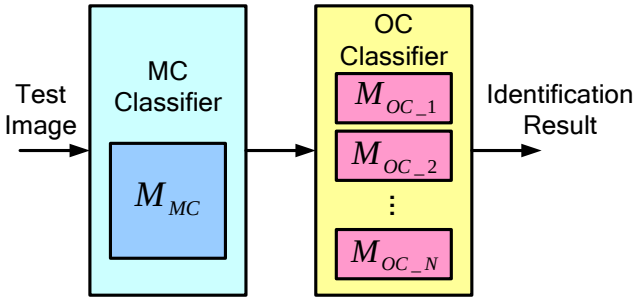


Fig. 2. Combination strategy of MC and OC classifiers

An unavoidable fact of the classifier combination is the propagation of errors. For the MC classifier, without loss of generality, we use $e_{MC}^{i}, i=[1,N]$ to denote the misclassification ratios, defined as following.

$$e_{MC}^{i}=\frac{N_{misc-mc}}{N_i} \tag{12}$$

$N_{misc-mc}$ denotes the number of misclassified samples of multi-class classifier, while N_i means the number of samples belonging to class *i*. For the OC classifier, a false positive ratio $e_{OC\_FP}^{i}$, $i=[1,N]$ and a false negative ratio $e_{OC\_FN}^{i}, i=[1,N]$ for each model are defined as following:

$$e_{OC\_FP}^{i}=\frac{N_{ci}}{N_{non-i}} \tag{13}$$

$$e_{OC\_FN}^{i}=\frac{N_{misc-oc}}{N_i} \tag{14}$$

Where N_{ci} denotes the number of samples classified as class *i*, and N_{non-i} means the number of samples NOT belonging to class *i*, while $N_{misc-oc}$ means the number of misclassified samples of one-class classifier. We evaluate and compare the performance between the strategies of traditional MC classifier and the proposed classifier combination, in the term of misclassification ratio.

For the previous work with only MC classifier, the misclassification ratio for each class is obviously $e_{MC}^{i}, i=[1,N]$, when the test image is indeed captured by some of the cell phones in the training set. Thus the average misclassification ratio is undisputed $\sum_{i=1}^{N} e_{MC}^{i} \big/ N$. Of

course, when the test image comes from an outlier cell phone, the misclassification ratio can be easily obtained as 100%, as equation (11) demonstrates.

$$\begin{cases} \dfrac{\sum_{i=1}^{N} e_{MC}^{i}}{N}, when \quad S_{Test} \in M_i, i = [1, N] \\ 1, \quad when \quad S_{Test} \notin M_i, \forall i = [1, N] \end{cases} \tag{15}$$

Then we discuss the misclassification ratio of the combined classifier strategy. When the test image source is included in the training model, we obtain an error probability of $e_{MC}^{i}, i = [1, N]$ for each model, if the MC classifier misclassifies the test image in the first step. Because if there is an error occurring in MC classifier, the output of the OC classifier, no matter it is positive result for the false cell phone model or negative result for the outlier, would be a misclassification as well. If we obtain a correct result in MC classifier with probability of $1 - e_{MC}^{i}, i = [1, N]$ for each model, the probability of misclassification will be $\left(1 - e_{MC}^{i}\right) \times e_{OC\_FN}^{i}, i = [1, N]$ according to the performance of the OC classifier. When the test image is an outlier of the training model, the ratio becomes as simple as $e_{OC\_FP}^{i}, i = [1, N]$. Finally, we get the average misclassification in the case of classifier combination in (16).

$$\begin{cases} \dfrac{\sum_{i=1}^{N} \left(e_{MC}^{i} + \left(1 - e_{MC}^{i}\right) \times e_{OC\_FN}^{i} \right)}{N}, when \quad S_{Test} \in M_i, i = [1, N] \\ e_{OC\_FP}^{i}, \qquad\qquad when \quad S_{Test} \notin M_i, \forall i = [1, N] \end{cases} \tag{16}$$

The RBF (Radial Based Function) kernel based MC SVM [30] and OC SVM [31] are adopted in this study as the specific MC and OC classifiers. Other OC and MC classifiers are also applicative in our classifier combination framework.

4. Experimental Results and Analysis

An image data set containing 24 cell phone models from 9 manufacturers is used in our experiments. The brief introduction of the image samples from these cell phones is shown in **Table 2**. For each cell phone, we collect 150 different image samples, consequently a total of 3600 samples are included. These images are collected under a variety of uncontrolled conditions, such as different resolutions, in-door/out-door scenes, natural/ artificial scenes, different compression quality factors, and so on. 17 cell phones (No. 1 to No. 17) are selected as the models and the forensic analyst can access several training samples to obtain a MC classifier model and 17 OC classifier models. 100 images from each cell phone, a total of 1700, are randomly selected as training samples. And the rest of 50 images for each of the 17 cell phones are used for test. The rest of 7 cell phones are treated as outliers, which means there is

none prior knowledge about these devices. For the outlier cell phones, all of the 150 image samples are used for test.

The experimental results are shown as following. In **Table 3**, the accuracy of source classification for all 24 cell phones in the training model and outlier is presented, compared with the CFA pattern search simplified algorithm [23] and the multi-feature fusion algorithm [12]. For the cell phones in the training model, we receive anticipative deteriorations of the results from 93.8% to 88.7% in the term of average identification accuracy using CFA pattern search simplified algorithm, because of the propagation of errors, as shown in **Table 4**. A same deterioration could be found for the multi-feature fusion algorithm. Though there is nearly 6% deduction of average accuracy of our method compared with that in [23], we consider the deduction as a small and acceptable range. The different performances for the proposed methods between different feature sets verify that the CFA coefficient features are better than multi-feature fusion sets for the term of camera source identification. Meanwhile, the method in [23] and [12] is totally invalid for the 7 outliers as we expected, because the classifier used in [23] and [12] misclassifies the outliers as the cell phones in the training model. However, we obtain an average identification accuracy of 75.3% and 66.9% for the 7 outlier cell phones, as **Table 4** shows. For all 24 cell phones, our method also achieves a higher average accuracy of 84.8% and 77.9%, compared with 66.4% and 63.8% achieved by the method in [23] and [12]. The confusion matrix shown in **Table 5** and **Table 6** describes the details of the experimental results of the method [23] and [12], which are the input of the OC classifiers in the combination strategy of the proposed method. The 17 columns corresponds to the 17 cell phones in the training model, and 24 rows corresponds to all of the cell phones. The (i, j) element in the confusion matrix gives the percentage of images from cell phone i that are classified as belonging to cell phone j. The symbol "*" denotes percentage of 0. The gray cells in Table 5 and Table 6 demonstrate the classification results of 7 outlier cell phones. For the image samples from these cell phones, the classification accuracy is 0 because the inevitable misclassifications always occurs.

Table 2. Image data sets used in the experiments

No.	Model	Resolution	Smart Phone
1	Samsung D520	1280×1024	No
2	Samsung E208	1280×1024	No
3	Motorola L7	640×480	No
4	Nokia 7610	1152×864	No
5	Nokia E50	1280×1024	No
6	Nokia N70	1600×1200	No
7	Sony Ericsson K530	1600×1200	No
8	Samsung I9000	2592×1944	YES
9	Nokia 5230	1600×1200	YES
10	HTC A3366	2048×1536	YES
11	HTC Z710e	3264×2448	YES
12	HTC G10	3264×2448	YES
13	iPhone 1	1600×1200	YES
14	Mi 1	3264×2448	YES
15	MeiZu M9	2592×1944	YES
16	Smartisan T1	4128×3096	YES
17	Mi 3	4208×3120	YES
18	Nokia N73	2048×1536	No
19	Sony Erricson T707	2048×1536	No

20	HTC G12	2592×1944	YES
21	HTC T328W	2592×1944	YES
22	iPhone 4	2592×1944	YES
23	iPhone 6	3264×2448	YES
24	RedMi 1S	3264×2448	YES

Table 3. Identification accuracy for all 24 cell phones

No.	Accuracy (%)				No.	Accuracy (%)			
	Method in [23]	Proposed Method	Method in [12]	Proposed Method		Method in [23]	Proposed Method	Method in [12]	Proposed Method
1	98	94	94	86	13	98	96	94	90
2	100	90	96	90	14	84	78	90	78
3	100	82	92	80	15	100	96	96	90
4	90	84	92	84	16	94	90	92	88
5	94	84	86	80	17	90	90	82	76
6	86	84	88	78	18	0	73	0	70
7	88	82	84	78	19	0	68	0	72
8	100	96	90	84	20	0	64	0	64
9	92	92	96	88	21	0	81	0	70
10	100	96	82	70	22	0	97	0	62
11	84	80	90	80	23	0	70	0	68
12	96	94	88	82	24	0	74	0	62

Table 4. Average accuracy comparison for 17 cell phones in the training model, 7 outlier cell phones and all 24 cell phones

Average	Method in [23]	Proposed Method	Method in [12]	Proposed Method
No.1 –No.17	93.8	88.7	90.1	82.5
No.18 –No.24	0	75.3	0	66.9
No.1 –No.24	66.4	84.8	63.8	77.9

Table 5. Confusion matrix of method in [23]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	**98**	2	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
2	*	**100**	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*
3	*	*	**100**	*	*	*	*	*	*	*	*	*	*	*	*	*	*
4	4	*	4	**90**	*	2	*	*	*	*	*	*	*	*	*	*	*
5	*	*	*	2	**94**	*	*	2	*	2	*	*	*	*	*	*	*
6	*	*	8	*	*	**86**	4	*	*	*	2	*	*	*	*	*	*
7	*	4	*	*	*	*	**88**	*	4	2	*	*	2	*	*	*	*
8	*	*	*	*	*	*	*	**100**	*	*	*	*	*	*	*	*	*
9	*	*	*	*	*	6	2	*	**92**	*	*	*	*	*	*	*	*
10	*	*	*	*	*	*	*	*	*	**100**	*	*	*	*	*	*	*
11	*	*	*	*	*	*	*	*	*	*	**84**	12	*	4	*	*	*
12	*	*	*	*	*	*	*	*	*	*	4	**96**	*	*	*	*	*
13	*	*	*	*	*	*	*	*	*	*	*	*	**98**	*	*	2	*
14	*	*	*	*	*	*	*	*	*	*	8	4	*	**84**	2	*	*
15	*	*	*	*	*	*	*	*	*	*	*	*	*	*	**100**	*	*
16	*	*	*	*	*	*	*	*	*	*	*	2	2	*	*	**94**	2
17	*	*	*	*	*	*	*	*	*	*	2	*	*	2	*	6	**90**
18	*	2	*	6	12	68	4	*	*	4	2	*	*	*	2	*	*

19	*	*	10	12	4	*	28	*	*	32	*	*	8	4	*	2	*
20	*	*	*	*	*	*	*	20	2	8	12	46	*	*	4	2	2
21	*	*	*	*	2	*	*	12	*	22	38	26	*	*	*	*	*
22	*	*	*	*	*	*	*	4	4	*	2	*	28	8	22	30	2
23	*	*	*	*	*	2	*	2	4	6	2	4	20	12	20	24	4
24	*	*	*	*	4	2	4	*	2	*	10	12	2	38	4	2	20

Table 6. Confusion matrix of method in [12]

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**
1	**94**	4	*	*	*	*	*	*	*	*	2	*	*	*	*	*	*
2	2	**96**	*	*	*	*	2	*	*	*	*	*	*	*	*	*	*
3	2	*	**92**	*	2	*	*	2	*	*	*	2	*	*	*	*	*
4	*	2	*	**92**	*	2	*	*	2	*	*	*	*	*	*	2	*
5	*	4	*	*	**86**	*	2	*	*	2	4	*	*	*	*	2	*
6	*	*	*	6	4	**88**	*	2	*	*	*	*	*	*	*	*	*
7	2	2	*	4	*	*	**84**	*	4	*	2	*	*	*	*	*	2
8	*	2	*	*	*	*	*	**90**	*	2	2	4	*	*	*	*	*
9	*	*	*	*	*	2	*	*	**96**	*	*	*	*	2	*	*	*
10	*	*	2	*	2	*	*	*	*	**82**	8	4	*	*	*	*	2
11	*	*	*	*	*	*	2	*	*	2	**90**	6	*	*	*	*	*
12	*	*	*	*	*	4	*	*	*	8	*	**88**	*	*	*	*	*
13	*	*	2	*	2	*	*	*	*	*	*	*	**94**	*	*	2	*
14	*	*	*	2	*	*	*	4	*	*	2	*	*	**90**	*	*	2
15	*	*	*	*	*	2	*	*	2	*	*	*	*	*	**96**	*	*
16	*	*	*	*	*	*	*	2	*	*	4	*	2	*	*	**92**	*
17	*	*	*	*	*	*	*	2	4	*	*	*	2	6	*	4	**82**
18	*	4	*	2	*	44	20	*	8	*	2	10	4	*	2	4	*
19	2	4	2	*	*	18	30	*	12	8	*	6	*	2	4	8	4
20	2	*	4	*	2	*	*	24	*	4	10	52	*	*	2	*	*
21	*	*	4	*	*	10	2	2	*	12	30	30	*	4	*	*	6
22	*	4	2	*	4	8	*	2	4	2	*	4	34	*	6	28	2
23	2	2	2	6	2	2	2	*	2	*	4	*	26	2	2	36	10
24	*	*	*	*	*	*	6	*	2	4	*	4	8	56	4	2	14

Considering the term of time complexity, the proposed methods is obviously higher than the baseline [12] and [23], because they combines the MC classifier with several OC classifiers in the classifier strategy. To be fair, we compare the time cost of the methods without considering the training process, because the traning process could be finished offline. That means the time cost of the proposed methods consists of three components: feature extraction, multi-class classification and one-class classification. Compared with the corresponding baseline [12] and [23], the time costs of feature extraction and multi-class classification is completely the same, while the one-class classification is the additional time complexity. The before-mentioned experiments are implemented via Matlab 2009 with a PC equipped with Intel Core i7-5960X 3.0GHz CPU and 32G Ram. **Table 7** demonstrates the segmented time costs of the proposed methods compared with the baseline [12] and [13], for all 1900 test images. The identification of all of the test images spends 442 minutes and 1030 minutes for methods in [12] and [23]. For the proposed methods, the corresponding time costs are 443 minutes and 1032 minutes, in other words, about 14 seconds and 33 seconds for each test image sample.

Table 7. Time complexity of the identification

Methods	Components	Time Costs (Min)	Total Time Costs (Min)
Method in [23]	Feature Extraction	1013	1030
	Multi-Class Classification	17	
Proposed Method	Feature Extraction	1013	1032
	Multi-Class Classification	17	
	One-Class Classification	2	
Method in [12]	Feature Extraction	430	442
	Multi-Class Classification	12	
Proposed Method	Feature Extraction	430	443
	Multi-Class Classification	12	
	One-Class Classification	1	

5. Conclusion

This paper proposed a classifier combination strategy for identifying the source cell phone of digital images. A framework of successive detections with MC classifier and OC classifier is used to obtain an acceptable average accuracy for cell phone models in the training model, and a high average identification ratio for outlier cell phones. The classifier combination strategy is implemented with two effective source camera identification algorithms, using CFA interpolation coefficients estimation and multi-feature fusion as feature vectors. Experiments indicate that the average accuracies of 88.7% and 75.3% with CFA coefficient features, 82.5% and 66.9% with multi-feature fusion, are achieved for cell phones that in and out of the training model, respectively.

In the practical scenario of image source identification for cell phones, the classification of outlier is a significant but difficult task. The classifier combination strategy is used to introduce an "outlier" label for image source identification. Though the strategy is feasible, we still plan to improve the performance of the classifier combination, and design new ingenious combination of classifiers for specific feature sets.

References

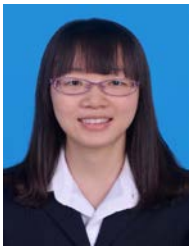
[1] http://www.idc.com/getdoc.jsp?containerId=prUS24645514
[2] H. Farid, “Digital image forensics,” *Scientific American*, vol. 6, no. 298, pp. 66-71, 2008. Article (CrossRef Link)
[3] A. Swanminathan, M. Wu, K. J. R. Liu, “Nonintrusive component forensics of visual sensors using output images,” *IEEE Transaction on Information Forensics and Security*, vol. 2, no. 1, pp. 91-106, March, 2007. Article (CrossRef Link)
[4] J. Lukáš, J Fridrich, M. Goljan, “Digital “bullet scratches” for images,” in *Proc. of IEEE International Conference on Image Processing*, pp.III-65-68, September 11-14, 2005. Article (CrossRef Link)
[5] J. Fridrich. “Digital image forensic using sensor noise,” *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 26-37, 2009. Article (CrossRef Link)
[6] M. Steinebach, M. Ouariachi, H. Liu, S. Katzenbeisser, “Cell phone camera ballistics: attacks and countermeasures,” in *Proc. of Electronic Imaging, Multimedia on Mobile Devices*, vol.7542, pp. B1-9, January, 2010. Article (CrossRef Link)

[7] M. Steinebach, M. Ouariachi, H. Liu, S. Katzenbeisser, "On the reliability of cell phone camera fingerprint recognition," *Ditital Forensics and Cyber Crime*, pp. 69-76, September 30-October 2, 2009. Article (CrossRef Link)
[8] X. Kang, Y. Li, Z. Qu, J. Huang, "Enhancing source camera identification performance with a camera reference phase sensor pattern noise," *IEEE Transaction on Information Forensics and Security*, vol. 7, no. 2, pp. 393-402, March, 2012. Article (CrossRef Link)
[9] O. Celiktutan, I. Avcibas, B. Sankur, N. P. Ayerden, C. Capar, "Source cell-phone identification," in *Proc. of IEEE 14th Signal Processing and Communications Applications*, pp. 1-3, April 17-19, 2006. Article (CrossRef Link)
[10] O. Celiktutan, B. Sankur, I. Avcibas, "Blind identification of source cell-phone model," *IEEE Transaction on Information Forensics and Security*, vol. 3, no. 3, pp. 553-566, August, 2008. Article (CrossRef Link)
[11] M. Tsai, C. Wang, J. Liu, J. Yin, "Using decision fusion of feature selection in digital forensics for camera source model identification," *Computer Standards & Interfaces*, vol. 34, no. 3 pp. 292-304, March, 2012. Article (CrossRef Link)
[12] X. Sun, L. Dong, B. Wang, X. Kong, X. You, "Source Cell-phone Identification Based on Multi-feature Fusion," in *Proc. of International Conference on Image Processing, Computer Vision, and Pattern Recognition*, pp. 590-596, 2010. Article (CrossRef Link)
[13] V. Lanh, S. Emmanuel, M. S. Kankanhalli, "Identifying source cell phone using chromatic aberration," in *Proc. of IEEE International Conference on Multimedia and Expo*, pp. 883-886, July 2-5, 2007. Article (CrossRef Link)
[14] W. Chuang, W. Min, "Semi non-intrusive training for cell-phone camera model linkage," in *Proc. of IEEE International Workshop on Information Forensics and Security*, pp. 1-6, December 12-15, 2010. Article (CrossRef Link)
[15] G. Gökhan, I. Avcibas, "Source cell phone camera identification based on singular value decomposition," in *Proc. of IEEE International Workshop on Information Forensics and Security*, pp. 171-175, December 6-9, 2009. Article (CrossRef Link)
[16] M. Kharrazi, H. Sencar, N. Memon, "Blind source camera identification," in *Proc. of IEEE International Conference on Image Processing*, October 24-27, pp. 709-712, 2004. Article (CrossRef Link)
[17] K. S. Choi, E. Y. Lam, K. Y. Wong, "Automatic source camera identification using the intrinsic lens radial distortion," *Optics Express*, vol.14, no. 24, pp. 11551-11565, 2006. Article (CrossRef Link)
[18] F. Meng, X. Kong, X. You, "A new feature-based method for source camera identification," in *Proc. of IFIP WG 11.9 International Conference on Digital Forensics*, pp.702-705, September 12-14, 2008. Article (CrossRef Link)
[19] S. Bayram, H. Sencar, N. Memon, I. Avcibas, "Source camera identification based on CFA interpolation," in *Proc. of IEEE International Conference on Image Processing*, pp. 69-72, September 11-14, 2005. Article (CrossRef Link)
[20] S. Bayram, H. Sencar, N. Memon, "Identifying digital cameras using CFA interpolation," in *Proc. of IFIP WG 11.9 International Conference on Digital Forensics*, 2006. Article (CrossRef Link)
[21] Y. Long, Y. Huang, "Image based source camera identification using demosaicking," in *Proc. of IEEE 8th Workshop on Multimedia Signal Processing*, pp.419-424, October 4-6, 2006. Article (CrossRef Link)
[22] A. Swaminathan, W. Min, K. J. R. Liu, "Non-intrusive forensic analysis of visual sensors using output images," in *Proc. of IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. V-V, May 14-19, 2006. Article (CrossRef Link)
[23] A. Swaminathan, W. Min, K.J.R. Liu, "Nonintrusive component forensics of visual sensors using output images," *IEEE Transaction on Information Forensics and Security*, vol. 2, no. 1, pp. 91-106, March, 2007. Article (CrossRef Link)
[24] J. Park, J. Chong, "Edge-preserving Demosaicing method for digital cameras with bayer-like W-RGB color filter array," *KSII Transaction on Internet and Information Systems*, vol. 8, no. 3, pp. 1011-1025, March, 2014. Article (CrossRef Link)

[25] D. Sung, H. Tsao, "Demosaicing using subband-based classifiers," *Electronics Letters*, vol. 51, no. 3, pp. 228-330, February, 2015. Article (CrossRef Link)
[26] I. Avcıbaş, N. Memon, B. Sankur, "Steganalysis using image quality metrics," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 221-229, February, 2003. Article (CrossRef Link)
[27] Y. Li, B. Wang, X. Kong, Y. Guo, "Image tampering detection using no-reference image quality metrics," *Journal of Harbin Institute of Technology*, vol. 21, no. 6, pp. 51-56, 2014. Article (CrossRef Link)
[28] S. Lyu, H. Farid, "How realistic is photorealistic?" *IEEE Transaction on Signal Processing*, vol. 53, no. 2, pp. 845-850, February, 2005. Article (CrossRef Link)
[29] P. Pudil, J. Ferri, J. Novovicova, "Floating search methods for feature selection with nonmonotonic criterion functions," in *Proc. of 12th International Conference on Pattern Recognition*, pp. 279-283, October 9-13, 1994. Article (CrossRef Link)
[30] B. E. Boser, I. Guyon, V. Vapnik, "A training algorithm for optimal margin classifiers," in *Proc. of the Fifth Annual Workshop on Computational Learning Theory*, pp. 144-152, 1992. Article (CrossRef Link)
[31] B. Schölkopf, J. C. Platt, J. S. Taylor, A. J. Smola, R. C. Williamson, "Estimating the support of a high-dimensional distribution," *Neural Computation*, vol. 13, no. 7, pp. 1443-1471, July, 2001. Article (CrossRef Link)

Bo Wang received his B.S. degree in electronic and information engineering, M.S. degree and Ph.D. degree in signal and information processing from Dalian University of Technology, Dalian, China, in 2003, 2005 and 2010, respectively. From 2010 to 2012, he was a post-doctoral research associate in faculty of management and economics in Dalian University of Technology. From 2012 up to now, he is an assistant professor in school of information and communication engineering in Dalian University of Technology. His current research interests focus on the areas of multimedia processing and security, such as digital image processing and forensics.

Yue Tan received her B.S. degree in photo communication from Jilin University, China, in 2013. She is currently pursuing the M.S. degree in information and communication engineering from Dalian University of Technology, Dalian, China. Her current research interests focus on source camera identification and tampering detection in the area of digital image forensics.



Meijuan Zhao received her B.S. degree in communication engineering from Dalian Nationalities University, Dalian, China, in 2014. She is currently pursuing the M.S. degree in information and communication engineering from Dalian University of Technology, Dalian, China. Her current research interests focus on source camera identification in the area of digital image forensics.

Yanqing Guo received his B.S. degree in electronic and information engineering, M.S. degree and Ph.D. degree in signal and information processing from Dalian University of Technology, Dalian, China, in 2002, 2004 and 2009, respectively. From 2009 up to now, he was employed by school of information and communication engineering in Dalian University of Technology, and he was an associate professor since 2012. His current research interests focus on the areas of social media computing, digital multimedia processing and computer vision.

Xiangwei Kong received her B.S. degree in electronic engineering, M.S. degree in electronic and communication engineering in 1985 and 1988 from Harbin Engineering University. And she received her Ph.D. degree in management science and engineering in 2003 from Dalian University of Technology. From 2002 up to now, she was a professor in school of information and communication engineering in Dalian University of Technology. Her current research interests focus on the areas of multimedia processing and security, content based image retrieval, knowledge management and intelligence, information fusion and so on.