Source Cell-phone Identification Based on Multi-feature Fusion

Xuehui Sun[1], Lisha Dong[1], Bo Wang[1], Xiangwei Kong[1] and Xin'gang You[1,2]
[1]Information Security Research Center, Dalian University of Technology, Dalian, Liaoning, China
[2]Beijing Institute of Electronic Technology and Application, Beijing, China

Abstract - *This paper proposes a new method for source cell-phone identification based on mutli-feature fusion. For a given image which is taken by cell-phone camera, three sets of features are used in this paper, they are higher-order statistics, image quality measures and CFA interpolation coefficients. The feature selection process is implemented with SFFS method to construct a vector of 19 features selected from these 3 sets of features, and the vector is then fed to the SVM classifier. The experiment shows that this algorithm gives an average accuracy of 95.94%, which is better than the existing ones.*

Keywords: source cell-phone identification; higher-order statistics; image quality measures; CFA interpolation coefficients; SFFS; SVM

1 Introduction

Nowadays, cell-phone is becoming one of the necessary electronic devices in people's daily life. Since most cell-phones have a camera function, people are likely to take photos using cell-phone rather than camera due to its convenience. Meanwhile, the image editing and processing tools are developing so fast that it makes image tampering and forgery much easier, and it is hard to tell from the real ones. Thus, the issue comes when a digital image is presented as a piece of evidence, and to identify the integrity and authenticity of a given image is the main challenge in the field of digital image forensics[1-3]. In this paper, we focus on source cell-phone identification, which is an important branch of digital image forensics.

Ismail Avcibas first proposed the issue of source cell-phone identification[4-5]. Considering that different interpolation algorithms and in-camera processing may leave certain footprints in the form of correlations across adjacent bit planes of image, they explore a set of binary similarity measures, image quality measures and higher order wavelet statistics as features to identify the originating cell-phone. The average accuracy of three different brand cell-phones is 97.82% and the result fells to 91.2% when the number of cell-phones reaches to 9. What's more, the classification result for same brand cell-phones drops greatly. Sabu Emmanuel noted that different wavelengths focusing at different positions on the sensor caused lateral aberration[6], and through the estimation of the parameters of lateral chromatic aberration by maximizing the mutual information between the corrected R and B channels with the G channel, they found an effective way to identify the source cell-phone. The average accuracy of this method is about 92.22%, but again, it drops greatly when the cell-phones are the same brand.

In order to improve the classification accuracy of the same brand cell-phones, we consider the affect of the non-linear distortion on the statistical features of the image and the differences of the in-camera processes comprehensively, and propose a new method of cell-phone identification based on feature fusion. There are three sets of features used in this paper, they are higher-order statistics of 36 dimensions, image quality measures of 13 dimensions and CFA interpolation coefficients of 336 dimensions. Then a SFFS (Sequential Floating Feature Selection) method is implemented to select the most useful features of 19 dimensions to reduce the algorithm complexity, and SVM (Support Vector Machine) is used as the classifier. The experiment result shows that the method proposed in this paper gives the best result compared to the existing ones.

This paper is organized as follow. In the next section, there is a brief introduction of the cell-phone image acquisition and post-processing. In section 3, we present the method of extracting the three sets of features, and select the set of features of the greatest contribution with SFFS method. Then we give the experiment result and analysis of the algorithm in section 4. . Final conclusions are drawn in section 6.

2 Image Capture Model

Although different cell-phones may use different image devices and different processing method, the main image capture process is the same; so firstly, we make a brief introduction of the image acquisition process and then analyze the affect that caused by different part of the model.

Fig.1 shows the image capture process. The lights from the scene pass through the optical system and are finally recorded by the sensors (most sensors are CMOS in cell-phones). Considering the high cost of the full-resolution sensors and the technique complexity, most cell-phones use a Color Filter Array (CFA) to sample the real world scene and acquire the missing color information by applying a certain CFA interpolation algorithm[7]. The most widely used CFA pattern is the Bayer pattern. After CFA interpolation, the images go through a post-processing stage. Finally, the image may be JPEG compressed in order to reduce the storage space.

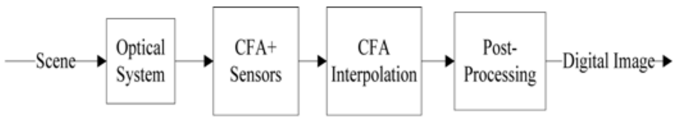

Fig.1 Image Capture Model

Although the process procedures are similar in different cell-phones, different CFA interpolation algorithms and JPEG compression methods may cause differences in the quality of the image as well as the higher-order statistic features of the image. These tiny differences may be hardly 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. In the next part, we present the specific method of feature extraction.

3 Feature Extraction

As analyzed above, the higher-order statistics and image quality and color can sever as the unique characteristics of the given image to find the source cell-phone. The wavelet features, which have been widely used in Steganalysis [8] and source camera forensics [9], can be used to analyze the multi-resolution of the image and represent the higher-order statistics of the image. In this paper, we use the subband coefficients of the wavelet to describe the affects of the in-camera processing on the frequency domain of the image. While for the image quality and color differences, we use the image quality measures proposed in [10] and the CFA interpolation coefficients proposed in [11] as features respectively.

To facilitate discussions, Let S represents the real world scene that is to be taken, p represents the CFA pattern and $S(i,j,c)$ represents a three dimensional matrix of size H×W×C, where H and W represent the height and width of the image, respectively, C=3 represent the number of color channels(e.g. red, green, blue). After the light from the scene is captured by the CFA, it is converted as a three dimensional matrix S_p in the form of

$$S_p(i,j,c) = \begin{cases} S(i,j,c) & p(i,j)=c \\ 0 & otherwise \end{cases} \quad (1)$$

Where, $S_p(i,j,c)=S(i,j,c)$ corresponds to those pixels that are captured directly and $S_p(i,j,c)=0$ corresponds to those that are to be interpolated later.

3.1 Features of higher-order statistics

Different physical imaging and software processing procedures in different cell-phones have different affects on the wavelet statistical features of the image[8-9]. Through wavelet decomposition, we can make multi-scale analysis of the image and extract more detail information about the higher-order statistics.

For the three different color channels of a given image, we do wavelet decompositions separately, and denote the vertical, horizontal and diagonal subbands as $V_k(i,j)$, $H_k(i,j)$ and $D_k(i,j)$ respectively. In this paper, the decomposition employed is based on Quadrature Mirror Filters (QMFs)[8]. This decomposition splits the frequency space into multiply scales and orientations. In each orientation, we choose the mean, variance, skewness and kurtosis of the subband coefficients, altogether we extract 36 dimensions of wavelet features. Equations (2) to (5) are the 12 dimensions in the vertical orientation, and we can get the other 24 dimensions using the similar method.

$$W_{1-3} = mean(V_k(i,j)) \quad (2)$$

$$W_{4-6} = \mathrm{var}iance(V_k(i,j)) \quad (3)$$

$$W_{7-9} = skewness(V_k(i,j)) \quad (4)$$

$$W_{10-12} = kurtosis(V_k(i,j)) \quad (5)$$

3.2 Features of image quality

Image quality measures are figures of merit used for the evaluation of imaging systems or of processing techniques, which have been used in steganalysis and compression in [10]. Since images taken with different cell-phones have differences in image quality, this can be used as another set of features to identify the source cell-phone. Take [10] as a reference, we choose 13 dimensions of the image quality measures, which are listed in Table 1.

Table 1 Image Quality Measures in Use

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 Spectral Phase-Magnitude Error；***Q10;*** Spectral Magnitude Error：***Q11；*** Spectral Phase Error：***Q12；*** Spectral Phase-Magnitude Error：***Q13***

3.3 Features of CFA interpolation coefficients

As one of the most important parts in the chain of imaging processing, the CFA (Color Filter Array) interpolation module may cause some correlations among neighboring pixels which turn out to be the statistical characteristics of the image [7]. Thus, through the estimation of the CFA interpolation algorithm and its coefficients, we can use them as another set of features to identify the source cell-phone. To simplify calculation, we suppose that the interpolation algorithm employed in the cell-phone is linear and we only consider the Bayer CFA pattern. Using the linear model and Singular Value Decomposition proposed in [11], we can estimate the interpolation coefficients. The linear equation can be formed as below:

$$y = a_1 x_1 + a_2 x_2 + \cdots + a_{48} x_{48} \quad (6)$$

Where the symbol y denotes the central pixel that is to be interpolated, $a_i (i \in [1, 48])$ specify the values of the pixels captured directly in the neighborhood, and $x_i (i \in [1, 48])$ stands for the interpolation coefficients to be estimated. Equation (6) can be simplified as

$$Ax = y \quad (7)$$

Note that the post-processing in the camera after interpolation may bring noise to the image and may change the values of the pixels more or less, That is, if let A_0, y_0 represents the ideal noiseless situation, and E、r represents the error of A、y respectively, the realistic situation can be represented as $A = A_0 - E$, $y = y_0 - r$. Take the noise into consideration, equation (7) can be rewritten as follow:

$$[A + E, y + r] \begin{bmatrix} x \\ -1 \end{bmatrix} = 0 \quad (8)$$

Since singular value decomposition works well in the situation of unstable questions, we employ this method to solve the equations, thus, the solution to this problem can be written as:

$$\begin{bmatrix} x \\ -1 \end{bmatrix} = -\frac{1}{v_{N_u+1, N_u+1}} v_{N_u+1} \quad (9)$$

Where v_{N_u+1} denotes the $(N_u+1)^{th}$ right singular vector of the combined matrix [A y], by this schematic, we can estimate the interpolation coefficients of a given image.

3.4 Feature Selection

In order to diminish the correlations among different features and reduce the feature dimensions, we implement a SFFS method to select the most useful features from the above three sets of features. When the feature dimension reaches to 19, the classification accuracy reaches to the highest value and remains stable. Thus, with this new set of 19 features, we can effectively identify the source cell-phone. The proportions of the 19 effective features are shown in Table 2.

Table 2 Proportion of Different Set of Features

Set of Features	Proportion
Higher-order statistics	52.63%
Image quality measures	26.31%
CFA interpolation coefficients	21.06%

4 Experimental result

4.1 Experiment design

To test the effectiveness of the proposed algorithm, a total of 8 different cell-phones of 3 brands are included in our experiment, the models and their parameters are listed in Table 3.

We choose 130 images from each of the 8 cell-phones. These images were taken in different light situations and have a diversity of content, and were JPEG compressed with the default quality factors.

Table 3 Models and Image Parameters Used in Experiment

Cell-Phone Model	Symbol	Maximum Resolution	Parameters of the sample images	
			Resolution	Format
Motorola_e680i	M1	640×480	640×480	JPEG
Motorola_l7	M2	640×480	640×480	JPEG
Samsung_e208	S1	1600×1200	640×480	JPEG
Samsung_d520	S2	1600×1200	640×480	JPEG
Nokia_n70	N1	1600×1200	640×480	JPEG
Nokia_n73	N2	2048×1536	640×480	JPEG
Nokia_e50	N3	1600×1200	640×480	JPEG
Nokia_7610	N4	800×1200	640×480	JPEG

In the experiment, we use SVM as the classifier [12]. There are 3 groups of experiment designed in this paper and every group has a comparison with the algorithm proposed in paper [5]. We randomly choose 80 out of the 130 images from each cell-phone as ground truth for training and the remains for testing. To make sure the effectiveness and stable of the algorithm, all the experiments below are repeated 10 times with the training sets and testing sets to be chosen randomly.

4.2 Classification of 3 cell-phones of different brands

We choose three cell-phones from the three different brands listed in Table 3. The experiment result is shown in Table 4. From Table 4, we see that the classification accuracy is 100% in our experiment, while in [5], the result is about 98.67%, so the proposed method is effective when the cell-phones in question are totally different brands.

Table 4 Classification Result of three different cell-phone brands

	Classification Result					
	Method in[5]			Method in this paper		
	M1	S1	N4	M1	S1	N4
M1	98%	2%	0%	100%	0%	0%
S1	0%	100%	0%	0%	100%	0%
N4	2%	0%	98%	0%	0%	100%

4.3 Classification of 4 cell-phones of the same brand

We choose the 4 Nokia cell-phones of different models. As is shown in the first section, most of the existing cell-phone forensic methods are not robust to this kind of situation. We can see from Table 5 that the lowest accuracy in [5] is 78%, which is far from satisfactory, but the algorithm proposed in this paper gives a high classification accuracy, with the lowest accuracy to be 94% and the average accuracy is about 95%, much higher than that of [5].

Table 5 Classification Result of four cell-phone of the same brand

	Classification Result							
	Method in[5]				Method in this paper			
	N1	N2	N3	N4	N1	N2	N3	N4
N1	88%	12%	0%	0%	94%	6%	0%	0%
N2	20%	78%	2%	0%	6%	94%	0%	0%
N3	2%	0%	98%	0%	2%	0%	98%	0%
N4	0%	0%	0%	100%	0%	0%	6%	94%

4.4 Classification of 8 different cell-phones

We use all the 8 cell-phones listed in Table 3 in this group of experiment to test the effectiveness of the proposed method when the number of classification cell-phones is much larger. The comparison result is shown in Table 6. From Table 6 we can see that the lowest accuracy in this paper is 88% with an average accuracy of 95.94%, while the lowest accuracy in [5] is 80% with an average accuracy of 95%.

Table 6 Classification Result of 8 different cell-phones

	Classification Accuracy (100%)							
	N1	N2	N3	N4	S1	S2	M1	M2
Method in [5]	98	80	94	96	94	98	100	100
Method in this paper	88	94	94	100	96	96	100	100

From the above three groups of experiment result we can see that under the same situation, the accuracy is higher in this paper than that in [5], especially when the identification cell-phones are the same brand.

5 Conclusion

In this paper, we propose an algorithm based on mutli-feature fusion. Considering the affect caused by different image acquisition techniques and the post-processing method in different cell-phones together, we extract three sets of features, namely, higher-order statistics, image quality measures and CFA interpolation coefficients, and use SFFS to select the most contributive features and SVM as the classifier. Experiment result shows that the classification accuracy can reach to 95.94% when the number of cell-phones is 8, and the method is robust to the situation of same brand cell-phone classification, which gives the best result compared to the existing ones.

6 Acknowledgement

This work was supported by the National High Technology Research and Development Program of China (863 Program, No. 2008AA01Z418) and National Natural Science Foundation of China (No. 60971095).

7 References

[1] Geradts, Z., Bijhold, J., Kieft, M., Kurosawa, K.,Kuroki, K.and Saitoh, N. Methods for Identification of Images Acquired with Digital Cameras. Proc. of SPIE, Enabling Technologies for Law Enforcement and Security[C]. Boston, MA USA, 2001(4232): 505–512.

[2] Kharrazi M., Sencar H.T., and Memon N. Blind Source Camera Identification. International Conference on Image Processing [C]. Singapore, October, 2004(1) : 709-712 .

[3] Lukáš J., Fridrich J., and Goljan M. Digital Camera Identification from Sensor Pattern Noise [J]. IEEE Transactions on Information Forensics and Security.,2006, 2(1):205-214 .

[4] Oya Çeliktutan, İsmail Avcıbaş, Bülent Sankur, Nasir Memon. Source Cell-phone Identification. 2006 IEEE 14th Signal Processing and Communications Applications[C]. Antalya Turkey, 2006(1-2) :218-221.

[5] Oya Çeliktutan, İsmail Avcıbaş, Bülent Sankur. Blind Identification of Cellular Phone Cameras. Security, Steganography, and Watermarking of Multimedia Contents [C]. San Jose, CA USA, 2007,65051H.

[6] Lanh Tran Van, Sabu Emmanuel, Mohan S Kankanhalli. Identifying Source Cell Phone Using Chromatic Aberration. 2007 International Conference on Multimedia & Expo [C]. Beijing China, 2007(1-5) :883-886.

[7] A. C. Popescu, H. Farid. Exposing digital forgeries in color filter array interpolated images[J]. IEEE Transaction on Signal Processing, 2005, 53(10): 3984-3959.

[8] H. Farid and S. Lyu. Detecting hidden messages using higher-order statistics and support vector machines. 5th International Workshop on Information Hiding[C]. Noordwikkerhout Netherlands, 2003 (2578):340-354.

[9] Siwei Lyu and H. Farid. How Realistic is Photorealistic [J]? IEEE TRANSACTIONS ON SIGNAL PROCESSING, 2005,53(2): 845-850 .

[10] Ismail Avcibas. Image Quality Statistics and Their Use in Steganalysis and Compression[D]. Boğaziçi University. 2001.

[11] A. Swaminathan, M. Wu, K. J. R. Liu. Non-intrusive component forensics of visual sensors using output images [J]. IEEE Transaction on Information Forensics and Security, 2007, 2(1): 91-106.

[12] C.C. Chang, C. J. Lin. LIBSVM: A Library for support vector machines [EB].
http://www.csie.ntu.edu.tw/~cjlin/libsvm. 2001.