

Smartphone image acquisition forensics using sensor fingerprint

ISSN 1751-9632

Received on 1st August 2014

Revised on 26th October 2014

Accepted on 24th November 2014

doi: 10.1049/iet-cvi.2014.0243

www.ietdl.org

Ana Lucila Sandoval Orozco¹ ✉, Luis Javier García Villalba¹, David Manuel Arenas González¹, Jocelin Rosales Corripio¹, Julio Hernandez-Castro², Stuart James Gibson³

¹Group of Analysis, Security and Systems (GASS), Department of Software Engineering and Artificial Intelligence (DISIA), School of Computer Science, Office 431, Universidad Complutense de Madrid (UCM), Calle Profesor José García Santesmases s/n, Ciudad Universitaria, 28040 Madrid, Spain

²School of Computing, University of Kent, Canterbury CT2 7NF, UK

³School of Physical Sciences, University of Kent, Canterbury, Kent, CT2 7NH, UK

✉ E-mail: asandoval@fdi.ucm.es

Abstract: The forensic analysis of digital images from mobile devices is particularly important given their quick expansion and everyday use in the society. A further consequence of digital images' widespread use is that they are used today as silent witnesses in legal proceedings, as crucial evidence of the crime. This study specifically addresses the description of a technique that allows the identification of the image source acquisition, for the specific case of mobile devices images. This approach is to extract wavelet-based features from sensor pattern noise which are then classified using a support vector machine. Moreover, there are a number of parameters that allows the authors to adapt the execution of the algorithm to specific situations desired for the forensic analyst (a variety of types and sizes of image or optimising the average accuracy rate in terms of processing time). This article describes a set of experiments with the same set of images that can obtain general conclusions for the different configurations.

1 Introduction

Owing to increasing storage capacity, usability, portability and affordability, camera-enabled mobile phones have become ubiquitous consumer electronic devices. The development of digital technologies has been advancing and continues to do so at an unstoppable rate. Every day the number of digital cameras is growing as well as the ease of access to them. Mobile digital cameras deserve special attention. According to Gartner [1], 1.745 billion handsets were sold in 2012 and it is predicted that 1.9 billion handsets will be sold in 2013. In total, according to estimates by the International Communication Union, there are 6.8 billion mobile phone subscriptions worldwide, which is a large increase from the 6 billion subscriptions in 2012 and 5.8 billion in 2011.

83% of these mobile devices have an integrated digital camera, which in contrast to conventional digital cameras are carried by their owners all the time to most places they attend and, in many cases, these devices have internet access [2]. The quality of these cameras has increased so much that many people use them as a replacement for digital still cameras (DSCs). In 2012, 31% of digital cameras sold belong to mobile phones, PCs and tablets and the forecast for 2016 according to [3] is to increase to 48%. In 2013 only 27% of market share will be from DSCs. There are also predictions that DSCs will disappear in favour of new integrated mobile device cameras [4], because the improved quality of these devices is growing at an unstoppable rate.

Having described this overview in figures on the extent of the presence of mobile devices in the world, we must not overlook the emergence in today's society of such devices in our day to day life. So much so, that according to Ahonen *et al.*, [2], a large number of people have and use more than one mobile device and a typical user turns to their mobile devices an average of 150 times a day.

The extensive use of smartphone cameras makes enforcing legal restrictions on the capture and sharing of digital photographs very difficult. Restrictions on the use of cameras include locations such

as schools, government offices and businesses. Consequently, tools which permit the identification of source devices have significant utility in various areas of law enforcement [5] such as child protection or digital rights management.

Often the pictures are considered to be real events captured by digital cameras. However, with the development of technology, powerful and sophisticated tools have emerged that facilitate the alteration of digital images in an impressive manner, even for those without technical knowledge or expertise in the area [6].

For these reasons, nowadays, digital image forensic analysis of mobile devices is very important. The study should be specific to mobile device images, because they have specific characteristics that allow for better results, not as valid digital image forensic techniques but for other kind of devices.

2 Image acquisition process in a digital camera

The first step to understanding and creating image forensic algorithms is to know in detail the image acquisition process in digital cameras. This process is summarised in Fig. 1.

Although many details of the camera pipeline belong to each manufacturer, the general structure is the same in all of them. Below is a brief description of each image acquisition phase.

When capturing an image, it is necessary to measure three or more bands for each pixel, which requires more than one sensor, and consequently it increases the cost of the camera. The most widespread and economical solution is the placement of a 'colour filter array' (CFA) in front of the sensor. There may be mechanisms interacting with the sensor to determine the exposure (aperture size, shutter speed and automatic gain control) and the focal length of the lens.

An antialiasing filter is also placed before the sensor; this filter is in charge of cleaning the signal prior to the analogue to digital conversion. This filter generates smoother contours in the image, reducing the unpleasant staggered appearance of lines.

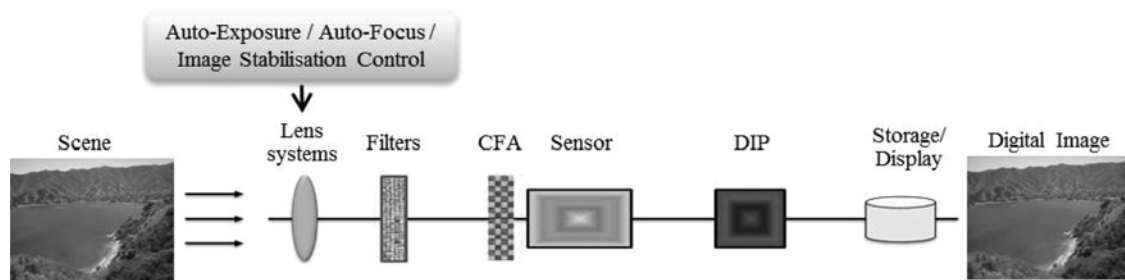


Fig. 1 Image acquisition process in a digital camera

The sensor (charge coupled device (CCD) or complementary metal oxide semiconductor) records the image converting light energy into electrical energy. The raw data obtained from the sensor needs to be processed to remove noise and other artefacts (anomalies introduced into digital signals). One of these processes is the correction of defective pixels caused by imperfections in the sensor, which corrects these pixels by interpolation. Another process is the white balance that allows for a more accurate colour reproduction without dominant colours; this effect is especially noticeable in neutral colours such as white. Demosaicing is the most complex process from the computational point of view and the techniques used are often owned by the camera manufacturer. This algorithm uses the values of the neighbouring pixels to calculate the values of the channels that have not been measured (remember that each pixel sensor detects only the channel that the array CFA allows to pass).

Another process to which the image is subjected is called gamma correction, which adjusts the intensity values of the image. Although these algorithms are in the pipeline from any camera, the exact process may vary from one manufacturer to another, and even from one camera model to another. Finally the image is compressed (mobile phone cameras typically use the algorithm joint photographic experts group (JPEG)) to save space. The compressed image is stored in the device memory with the image information in EXIF [7].

In [8] the image acquisition process in cameras of mobile devices is described, likewise a comparison of this process compared with that in DSCs and scanners is presented.

3 Source camera identification techniques

Research in this field studies the design of techniques to identify maker and model of the devices used to generate digital images. Analogously to ballistic analysis trying to relate a gun with its bullets, digital image forensics tries to identify the link between images and the digital camera which has generated them [9]. The success of these techniques depends on the assumption that the characteristics are unique to each device. The characteristics used to identify the maker and the model of digital cameras are derived from the differences between image processing techniques and technologies used in camera components [10]. The main problem with this approach is that different models of digital camera are often built using the same core components that originate from a small number of manufacturers. As a consequence it can be difficult, or in some cases impossible, to differentiate between models using such methods.

According to Van Lanh *et al.* [10], for this purpose four groups of techniques can be established depending on their base: lens system aberrations, CFA interpolation, image characteristics and sensor imperfections. The latter is the subject of this paper. In addition to the above there is another group of techniques based on metadata.

Metadata techniques are the simplest and there is plenty of research based on them. However, these techniques are highly dependent on the metadata that manufacturers decide to insert when generating images. Moreover, this method is the most vulnerable to malicious modifications or even the total elimination of metadata either intentionally or unwittingly.

During the image generation process the lens system can introduce some aberrations (spherical, coma, astigmatism, field curvature, radial distortion and chromatic aberration). The radial distortion is the one with the most impact over pictures, especially in cameras having cheap wide angle lenses. Most digital cameras use this type of lens for cost reasons. In [11] the lens radial distortion is proposed as the best technique for source identification. Radial distortion causes straight lines to appear as curves in images. The radial distortion degree of each image can be measured by a process consisting of three steps: edge detection, distorted segment extraction, and distortion error measurement. Choi experimented with three different cameras and obtained 91.28% accuracy identifying the camera source.

Some authors consider that CFA choice and the interpolation algorithm specifications generate some of the most striking differences between different camera models.

In [12] an algorithm for identifying and classifying colour interpolation operations is presented. This proposal is based on two methods to perform the classification process: first using an algorithm to analyse the correlation of each pixel value with its neighbours' values, and secondly an analysis of the differences between pixels independently. The accuracy for the source camera identification with images from four to five different models were of 88% and 84.8%, respectively.

In [13] correlations between pixels are used for the source identification, obtaining a coefficient matrix for each colour channel while defining a pixel quadratic correlation model. Neutral networks are used for classification. The method was tested with cartoon images from four cameras. The success rate obtained was 98.6%. This approach is not efficient at differentiating between different models from the same maker.

In [14] a set of binary similarity measures is used as metrics to estimate the similarity between image bit planes. The fundamental assumption of this work is that CFA interpolation algorithms from each maker leave correlations along image bit planes and can be represented by a set of 108 binary similarity measures for classification. The success rate of their experiments was between 81 and 98% to classify three cameras and decreased to 62% to identify between nine cameras.

The techniques based on image features use a set of features extracted from image content to identify the source. These features are divided into three groups: colour characteristics, image quality metrics and wavelet domain statistics.

In [8], the authors extend the source identification to different devices such as mobiles, phones, digital cameras, scanners and computers. In this proposal, colour interpolation coefficients and noise characteristics are used to classify. Their experiments showed an overall result of 93.75% accuracy. Identifying the maker and model of five mobile phone models, the accuracy obtained was 97.7%.

In [15], a method based on the bi-coherence statistics phases and magnitudes along with the wavelet coefficients is used for the identification. This method captures the unique nonlinear distortions in the wavelet domain produced by the cameras when performing processing operations over images. As a result an accuracy of 97% in the identification was obtained in distinguishing different models from the same manufacturer.

In [16], a technique to differentiate images using the wavelet family transforms is explained. Ridgelets and contourlets subbands

statistical models are proposed to extract the representative features from images. Experiments were conducted to identify three different cameras obtaining accuracies of 93.3% with wavelet-based approach, 96.7% using ridgelets, and 99.7% with contourlets.

In [17], a method using the marginal density discrete cosine transform (DCT) coefficients in low-frequency coordinates and neighbouring joint density features on both intra-block and inter-block from the DCT domain is proposed. In experiments with images of different scale factors from five smartphone models of four makers, an accuracy of between 86.36% and 99.91% was obtained.

The techniques based on sensor noise study the traces left by sensor defects in images. These techniques are mainly divided into two branches: pixel defects and sensor pattern noise (SPN). The first branch studies pixel defects, hot pixels, dead pixels, row or column defects, and group defects. In the second branch a pattern is constructed by averaging multiple residual noises computed by any noise removal filter; The presence of the pattern is determined using a correlation method or machine classification support vector machine (SVM).

In [18], pixel defects of CCD sensors are studied, focusing on different features to analyse images and then identify their source: CCD sensor defects, the file format used, noise introduced in the image and watermarking introduced by makers. Among the CCD sensor defects are considered hot spots, dead pixels, group defects, and row/column defects. Results indicate that each camera has a different defect pattern. Nevertheless, it is also noted that the number of pixel defects for images from the same camera is different and varies greatly depending in the image content. Likewise, it was revealed that the number of defects varies with temperature. Finally, the study found that high quality CCD cameras do not have this kind of problem. When considering only defective CCD sensors this study is not applicable to the analysis of images generated by mobile devices.

In [19], the authors analyse the SPN from a set of cameras, which functions as a fingerprint allowing the unique identification of each

camera. This pattern noise is obtained by averaging the sensor noise extracted from different images with a noise removal filter. To identify the camera from a given image, the reference pattern is considered as a watermark in the image and its presence is established by a correlation detector. It was found that this method is affected by processing algorithms such as image JPEG compression and gamma correction. The results for pictures with different sizes were unsatisfactory [10].

In [20], an approach to source camera identification in open set scenarios is proposed, where unlike closed scenarios it is not assumed to have access to all the possible image source cameras. This approach, in contrast to others, considers nine different regions of interest (ROIs) located in the corners and the centre of the images (not only the central region of the image). Using these ROIs, it is possible to work with different resolution images without requiring zero padding or colour interpolating. The SPN is computed for each colour channel generating a total of 36 representative features for each image. Then, the features of images taken by the camera under investigation are labelled as positive class and features from images made by other cameras as negative classes. After the SVM training phase, in which the hyper-plane that separates the positive and negative classes is estimated, this hyperplane is moved by a given value either inward (for positive classes) or outward (for negative classes) for the purpose of considering the open scenario unknown classes. The results had an accuracy of 94.49, 96.77 and 98.10%.

In [21], the sensor noise is extracted by calculating similarities as a classification method on the basis of [19]. The authors state that the sensor noise can be highly contaminated by the scenario details, and they propose that the stronger a component of the sensor noise is, the less reliable it is and therefore it should be attenuated. They performed experiments with six different DSCs. For images of 1536×2048 pixels, they obtained an accuracy of 38.5% with the implementation without the improvement and 80.8% with the proposed improvement. For images of 512×512 pixels, they obtained an accuracy of 21.8% without improvement and 78.7% with the proposed improvement.

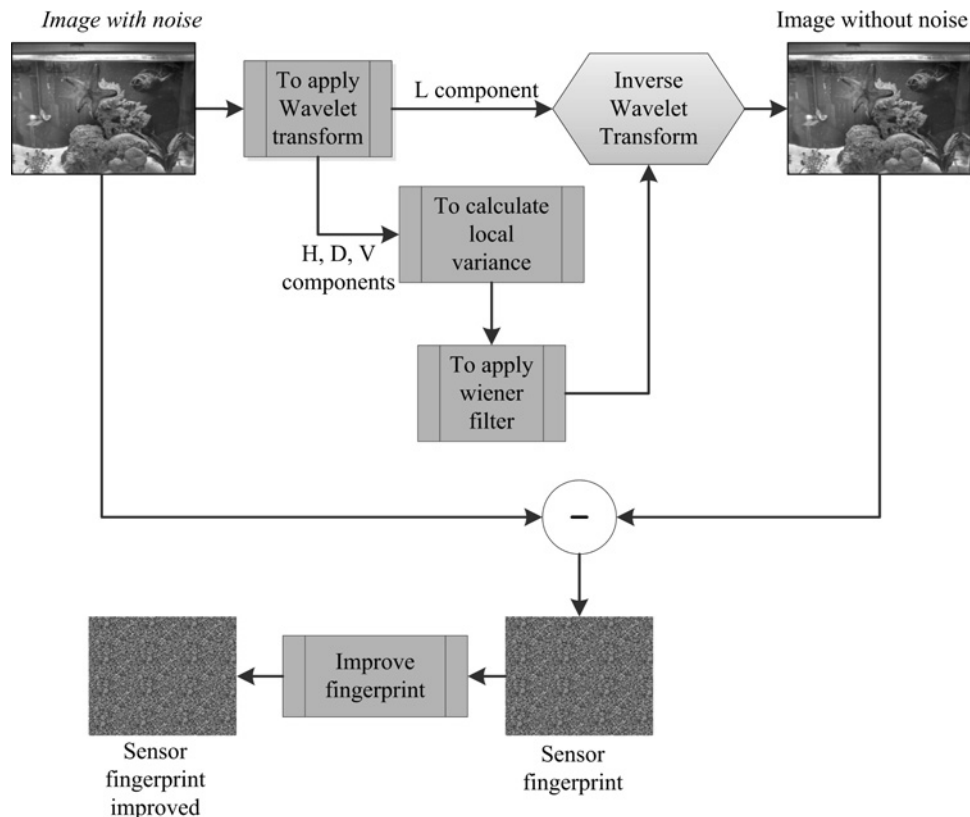


Fig. 2 Scheme functional

A detailed comparison of different source identification techniques is presented in [22].

4 Source identification algorithm

Previous work has shown SPN [18, 22, 19] and wavelet transform [15, 16] to be an effective method for source camera identification. However, almost all studies have focused only on traditional cameras, excluding mobile cameras. This makes it an area of study that requires attention especially with mobile devices. Using a biometric analogy, we consider each noise pattern to be a fingerprint of its source camera's sensor. In our study, SPN is used to classify images captured by camera-enabled smartphones. Our approach characterises the fingerprints using wavelet-based feature vectors. The scheme presented in Fig. 2 shows the functional diagram of our proposal.

Noise images were obtained using the method previously described by [19] and also summarised by Fig. 3 as follows.

To extract its noise pattern, an image is decomposed into its red, green and blue colour channels. Then, a four-level wavelet decomposition of each colour channel is calculated using the Daubechies, 8-tap, separable quadrate mirror filters. The number of decomposition levels can be increased to improve accuracy or to reduce processing time.

Horizontal H , vertical V and diagonal D high-frequency images are obtained for each level of decomposition. For each detail image, the local scene variance in a $W \times W$ window is estimated. Four estimates are obtained with window sizes corresponding to $W \in \{3, 5, 7, 9\}$. Finally, we choose the estimate which maximises the *a posteriori* probability

$$\hat{\sigma}^2(i, j) = \max\left(0, \frac{1}{W^2} \sum_{(i,j) \in N} c^2(i, j) - \sigma_0^2\right), \quad (i, j) \in J \quad (1)$$

Algorithm 1:

Input: Image I
Variance estimation: adaptive or non-adaptive
Result: Sensor fingerprint I_{noise}

- ① **procedure** EXTRACTPRNU(I)
- ② Apply a wavelet decomposition in 4 levels to I ;
- ③ **foreach** wavelet decomposition level **do**
- ④ **foreach** component $c \in \{H, V, D\}$ **do**
- ⑤ Compute the local variance;
- ⑥ **if** adaptive variance **then**
- ⑦ Compute 4 variances with windows of size: 3, 5, 7 and 9 respectively;
- ⑧ Select the minimum variance;
- ⑨ **else**
- ⑩ Compute the variance with a window of size 3;
- ⑪ Compute noiseless wavelet components applying the Wiener filter to the variance;
- ⑫ Obtain I_{clean} by applying the inverse wavelet transform with clean components calculated;
- ⑬ Obtain the sensor noise with
 $I_{noise} = I - I_{clean}$;
- ⑭ Apply zero-meaning to I_{noise} ;
- ⑮ Increase the green channel weight with
 $I_{noise} = 0.3 \cdot I_{noise_R} + 0.6 \cdot I_{noise_G} + 0.1 \cdot I_{noise_B}$;
- ⑯ **end procedure**

Fig. 3 Extracting PRNU

Algorithm 2:

Input: Sensor fingerprint I_{noise}
Result: 81 features

- ① **procedure** EXTRACTFEATURES(I)
- ② Separate R, G and B color channels of I_{noise} ;
- ③ **foreach** color channel **do**
- ④ Apply a wavelet decomposition in 1 level;
- ⑤ **foreach** component $c \in \{H, V, D\}$ **do**
- ⑥ Compute k central moments with
$$m_k = \frac{1}{n} \sum_{i=1}^n |c_i - \bar{c}|^k;$$
- ⑦ **end procedure**

Fig. 4 Extracting features

where, $c(i, j)$ is the high-frequency component and $c \in \{H, V, D\}$; σ_0 controls the degree of noise suppression.

The minimum of four variances is chosen as the best estimate

$$\hat{\sigma}^2(i, j) = \min\left(\begin{matrix} \sigma_3^2(i, j), \sigma_5^2(i, j), \\ \sigma_7^2(i, j), \sigma_9^2(i, j) \end{matrix}\right), \quad (i, j) \in J \quad (2)$$

An alternative and less accurate method is to simply use $W = 3$ as the estimated local variance.

The denoised wavelet coefficients are defined by the Wiener filter as follows

$$c_{clean}(i, j) = c(i, j) \frac{\hat{\sigma}^2(i, j)}{\hat{\sigma}^2(i, j) + \sigma_0^2} \quad (3)$$

The noise residual is obtained by calculating the inverse transform and subtracting the denoised image from the original image. JPEG and demosaicing artefacts, presented in the noise image, are suppressed by subtracting the mean column and row values [23]. Greater weight is given to the green channel since the configuration of the colour matrix this channel contains more information about the image [24–26].

The next step is to obtain features that characterise the sensor fingerprint for the purpose of classification. A total of 81 features (3 channels \times 3 wavelet components \times 9 central moments) is extracted using the Fig. 4.

Classification was performed using a SVM with RBF kernel. We used the LibSVM package in which the SVM is extended to multiple classes yielding class probability estimates [27]. A grid search was used to obtain the best kernel parameters (γ and C). The classifier was trained and tested with feature vectors extracted from randomly selected images.

5 Experiments and results

To assess the effectiveness of the proposed algorithms, a set of experiments have been made with a variety of configuration parameters. Table 1 summarises the parameters used and their possible values.

The PRNU extraction algorithm and feature extraction algorithm are implemented in Python 2.7 and C language. In a Intel Core i7 Q720 1.6 GHz and 8GB of RAM it takes approximately 20 s to extract the PRNU and compute the features for a 1024×1024 crop of an image and 5 s for a 512×512 crop of an image using adaptive variance estimation and zero meaning. The same case with no adaptive variance takes approximately 5 s and 1.5 s for 1024×1024 and 512×512 crops, respectively. Training the SVM classifier and testing for 600 images is realised in one minute and a fraction of a second, respectively. A random sample of 100 images was used for training and a different random sample of

Table 1 Parameters used in the proposed algorithm and its possible values

Parameter	Possible values
number of training photos by camera	100
number of testing photos by camera	100
image crop	centre: 1024 × 1024 or 512 × 512
variance estimation	adaptative (steps 7 and 8 of Fig. 4) or non-adaptive (step 9 of Fig. 4)
zero-meaning	ysed or not used (step 13 of Fig. 4)

Table 2 Configurations used in mobile device digital cameras

Brand	Model	Resolution	Taking Conditions
Apple	iPhone 3G (A1)	2 MP (1600 × 1200)	scene type: any orientation: vertical flash: disabled light: natural white balance: auto digital zoom ratio: 0 exposure time: 0 seg ISO speed: automatic
	iPhone 4S (A2)	8 MP (3264 × 2448)	
	iPhone 3 (A3)	2 MP (1600 × 1200)	
	iPhone 5 (A4)	8 MP (3264 × 2448)	
Black Berry	8520 (B1)	2 MP (1600 × 1200)	
Sony	UST25a (SE1)	5 MP (2592 × 1944)	
Ericsson	U51 (SE2)	8 MP (3264 × 2448)	
Samsung	GT-I9100 (S1)	8 MP (3264 × 2448)	
	GT-S5830 (S2)	5 MP (2592 × 1944)	
	GT-S5830M (S3)	5 MP (2592 × 1944)	
	EK-GC101 (S4)	16.3 MP (4608 × 3456)	
LG	E400 (L1)	3.2 MP (2048 × 1536)	
	P760 (L2)	5 MP (2592 × 1944)	
HTC	Desire HD (H1)	8 MP (3264 × 2448)	
	Desire (H2)	5 MP (2592 × 1944)	
Nokia	E611 (N1)	2 MP (1600 × 1200)	
	800-Lumia (N2)	8 MP (3264 × 2448)	
Zopo	ZP979 (Z1)	12.6 MP (4096 × 3072)	

100 images was used for testing. However, we used EOLO the HPC of Climate Change of the International Campus of Excellence of Moncloa for computing.

The first experiment of [28] shows that the performance changed only slightly in different experiment runs, which indicates stability over different training and testing image sets.

In experiments 1 to 8 we used the same number of phones and the same brands and models. This allows us to perform a comparative study and to obtain conclusions about what parameters can be favourable or optimal in different situations.

All the mobile devices used are shown in Table 2.

Table 3 Parameter configuration of experiments

Experiment	Resolution	Number of devices	Multiple neighbour	Zero mean required	Average accuracy
test 1	1024 × 1024	6	t	t	96.33
test 2	1024 × 1024	6	t	f	98
test 3	1024 × 1024	6	f	t	97.5
test 4	1024 × 1024	6	f	f	97.83
test 5	512 × 512	6	t	t	73.76
test 6	512 × 512	6	t	f	93.17
test 7	512 × 512	6	f	t	92.5
test 8	512 × 512	6	f	f	91.67
test 9	1024 × 1024	14	f	f	87.21

Table 4 Experiment 1

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony Ericsson ST25A
Apple iPhone 5	96	0	2	0	2	0
Samsung EK-GC101	5	88	2	2	3	0
Nokia 800-Lumia	0	0	100	0	0	0
Zopo ZP979	0	0	2	98	0	0
LG P760	0	0	0	0	100	0
Sony Ericsson ST25A	0	0	3	0	1	96

Once they have been presented with configuration parameters and cameras, the experiments with their corresponding parameters are shown in Table 3.

5.1 Experiment 1

The parameters chosen for this experiment are: crop centre 1024 × 1024, variance estimation adaptative and zero meaning.

The confusion table from six cameras is showed in Table 4. The average accuracy rate for correctly identifying camera make and model for this experiment was 96.33%.

5.2 Experiment 2

The parameters chosen for this experiment are: crop centre 1024 × 1024, variance estimation adaptative and no zero-meaning. That is, the same parameters as in Experiment 1, except that in this experiment the zero-meaning does not apply. Given that the images used for all experiments are the same we will be able to check the impact of this change to the results.

The confusion table from six cameras is showed in Table 5. The average accuracy rate for correctly identifying camera make and model for this experiment was 98%.

It is noted that the zero-meaning gets worse the average accuracy rate (1.67% from Experiment 1), although the difference is not very significant to obtain definitive conclusions. It can also be noted that except for the model LG P760 (passing from 100 to 99%) the rest of the mobile devices increases the hit rate.

5.3 Experiment 3

The parameters chosen for this experiment are: crop centre 1024 × 1024, variance estimation non-adaptative and zero-meaning. That is, the same parameters as in Experiment 1, except that in this experiment the variance estimation adaptative does not apply. Among others, the main objective of this experiment is to check if the chosen type of variance estimation is determinant in the results of the algorithm. It is also important to note that the use of adaptive or non-adaptive variance has important effects on the execution time of the algorithm, because algorithm execution time with non-adaptative variance is approximately four times faster.

The confusion table from six cameras is showed in Table 6. The average accuracy rate for correctly identifying camera make and model for this experiment was 97.5%. At first it was expected that non-adaptive variance estimation would produce worse results, but

Table 5 Experiment 2

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony Ericsson ST25A
Apple iPhone 5	97	0	1	0	2	0
Samsung EK-GC101	1	95	0	3	1	0
Nokia 800-Lumia	0	0	100	0	0	0
Zopo ZP979	0	0	2	98	0	0
LG P760	0	0	0	0	99	1
Sony Ericsson ST25A	0	0	0	0	1	99

Table 6 Experiment 3

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony Ericsson ST25A
Apple iPhone 5	95	0	2	0	3	0
Samsung EK-GC101	1	95	0	3	1	0
Nokia 800-Lumia	0	0	100	0	0	0
Zopo ZP979	0	1	1	98	0	0
LG P760	0	1	0	0	99	1
Sony Ericsson ST25A	0	0	1	0	1	98

it is observed that the results of the above experiments do not differ by far.

5.4 Experiment 4

The parameters chosen for this experiment are: crop centre 1024×1024 , non-adaptive variance estimation and no zero-meaning. That is, the same parameters as in Experiment 2, except that in this experiment we apply the non-adaptive variance estimation. Similar to the previous experiment one of the objectives of this experiment is to check if the chosen type of variance estimation has effects in the results. Besides we can watch the behaviour of zero-meaning for non-adaptive variance estimation.

The confusion table from six cameras is showed in Table 7. The average accuracy rate for correctly identifying camera make and model for this experiment was 97.83%.

In contrast to what occurs between Experiments 1 and 3, in this experiment a small worsening on the average accuracy rate of Experiment 2 is observed. Therefore, it can be concluded that in the case of 1024×1024 crop using adaptive variance estimation does not improve significantly the results, because the results are almost the same with minor improvements or deteriorations. Moreover it is observed that the use of zero-meaning with non-adaptive variance estimation does not significantly improve the results.

Table 7 Experiment 4

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony EricssonST25A
Apple iPhone 5	96	2	0	0	2	0
Samsung EK-GC101	1	95	0	3	1	0
Nokia 800-Lumia	0	0	100	0	0	0
Zopo ZP979	0	0	2	98	0	0
LG P760	0	2	0	0	99	0
Sony Ericsson ST25A	0	0	1	0	0	100

Table 8 Experiment 5

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony Ericsson ST25A
Apple iPhone 5	93	3	2	0	2	0
Samsung EK-GC101	16	76	2	0	6	0
Nokia 800-Lumia	0	0	86	0	2	12
Zopo ZP979	0	19	2	0	79	0
LG P760	2	0	1	0	93	4
Sony Ericsson ST25A	0	0	4	0	2	94

5.5 Experiment 5

The parameters chosen for this experiment are: crop centre 512×512 , variance estimation adaptive and zero-meaning. That is, the same parameters as in Experiment 1, except that in this experiment the crop size is reduced. One of the aims of this experiment and the following three is to check the influence of the crop sizes in the results with different parameters.

The confusion table from six cameras is shown in Table 8. The average accuracy rate for correctly identifying camera make and model for this experiment was 89.33%. As expected, the average accuracy rate is down considerably (by 7%) relative to Experiment 1, because the amount of information used to obtain the image features is considerably less.

5.6 Experiment 6

The parameters chosen for this experiment are: crop centre 512×512 , variance estimation adaptive and no zero-meaning. That is, the same parameters as in Experiment 5, except that in this experiment zero-meaning does not apply. This experiment has among others aims seeing the influence of zero-meaning in small crops using adaptive variance estimation.

The confusion table from six cameras is shown in Table 9. The average accuracy rate for correctly identifying camera make and model for this experiment was 93.17%. As expected, the average

Table 9 Experiment 6

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony Ericsson ST25A
Apple iPhone 5	94	2	3	0	1	0
Samsung EK-GC101	6	91	1	1	1	0
Nokia 800-Lumia	0	0	93	0	0	7
Zopo ZP979	0	5	2	92	1	0
LG P760	2	0	0	0	95	3
Sony Ericsson ST25A	0	0	4	0	2	94

Table 10 Experiment 7

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony Ericsson ST25A
Apple iPhone 5	95	0	2	0	3	0
Samsung EK-GC101	5	89	0	3	3	0
Nokia 800-Lumia	0	0	85	0	1	14
Zopo ZP979	0	1	2	97	0	0
LG P760	2	0	2	0	93	3
Sony Ericsson ST25A	0	0	4	0	0	96

Table 11 Experiment 8

Camera	Apple iPhone 5	Samsung EK-GC101	Nokia 800-Lumia	Zopo ZP979	LG P760	Sony Ericsson ST25A
Apple iPhone 5	95	0	2	0	3	0
Samsung EK-GC101	5	89	0	3	3	0
Nokia 800-Lumia	0	0	85	0	1	14
Zopo ZP979	0	1	2	97	0	0
LG P760	2	0	2	0	93	3
Sony Ericsson ST25A	0	0	4	0	0	96

accuracy rate is down (4.83%) relative to Experiment 2, because of the reduction of crop size. In the case of smaller crop not using zero-meaning, the success rate increases compared with the previous experiment (3.84%), although this increase is not a significant improvement.

5.7 Experiment 7

The parameters chosen for this experiment are: crop centre 512×512 , non-adaptive variance estimation and zero-meaning. That is, the same parameters as in Experiment 5, except that in this experiment we apply non-adaptive variance estimation. One of the aims of this experiment is to see the influence of the adaptive variance estimation using small crops.

The confusion table from six cameras is shown in Table 10. The average accuracy rate for correctly identifying camera make and model for this experiment was 92.50%. As expected, the average accuracy rate is down (5%) relative to Experiment 3, because of the reduction of crop size. Relative to the comparison with experiment 5, it can be seen that the results are better with

non-adaptive variance (3.17%). Moreover, relative to the comparison with Experiment 6 which also uses adaptive variance estimation the impact in results is minimal.

5.8 Experiment 8

The parameters chosen for this experiment are: crop centre 512×512 , variance estimation non-adaptive and no zero-meaning. That is, the same parameters as in Experiment 4, except that in this experiment the crop size is reduced. One of the aims of this experiment is to see the influence of zero-meaning using small crops and adaptive variance estimation.

The confusion table from six cameras is shown in Table 11. The average accuracy rate for correctly identifying camera make and model for this experiment was 91.67%. As expected, the average accuracy rate is down (6.16%) relative to Experiment 4, because of the reduction of crop size. It is confirmed that the results obtained in this experiment and the results obtained between the comparison of the results of Experiments 6 and 8 show that the use of adaptive variance estimation does not significantly improve the results.

Table 12 Confusion matrix of Experiment 9

Camera	A1	A2	A3	A4	B1	SE1	SE1	S1	S1	S3	L1	H1	H2	N1
A1	90	0	0	2	0	0	0	0	7	0	1	0	0	0
A2	0	91	0	3	0	0	0	3	0	0	0	1	2	0
A3	0	0	98	0	0	0	0	2	0	0	0	0	0	0
A4	0	0	1	88	0	0	0	0	0	0	3	6	0	2
B1	0	0	0	2	73	0	0	0	4	0	0	1	0	20
SE1	7	0	0	0	0	80	0	0	0	0	1	12	0	0
SE2	1	0	0	2	2	0	86	1	2	5	1	0	0	0
S1	4	5	0	4	0	0	1	83	0	0	1	0	2	0
S2	0	0	0	0	0	0	0	0	100	0	0	0	0	0
S3	0	0	1	0	0	0	8	0	0	85	0	1	0	5
L1	0	0	0	9	0	6	0	0	2	0	70	13	0	0
H1	2	0	0	0	0	11	0	0	1	0	1	85	0	0
H2	0	6	0	0	0	0	0	0	0	0	0	0	94	0
N1	0	0	0	0	2	0	0	0	0	0	0	0	0	98

5.9 Experiment 9

To evaluate the scalability of the method to a larger number of classes, a group of 14 mobile device digital cameras from seven different manufacturers was used. The average classification rate dropped to 87.21% as shown in the confusion matrix of Table 12 indicating a small loss in performance when the number of classes (cameras) is increased.

Remember that in all of this work, 100 images were employed for training and 100 for testing.

6 Conclusions

According to the structure and operation of mobile device digital cameras, the most appropriate techniques for forensic analysis are those based on sensor noise and wavelet transforms. In this paper, an algorithm was proposed for identifying the mobile source combining techniques based on sensor fingerprint and the wavelet transforms. The algorithm is mainly composed of two phases: the first is dedicated to extracting the sensor fingerprint, and the second to extracting features from this fingerprint that will serve as input to the SVM used as classification method.

A method for source camera identification, based on wavelet features of image noise residuals and SVM classification, was tested on photographs acquired from a range of smartphones. Eight experiments have been made with the same pictures, for the purpose of analysing the different configuration parameters and improvements in the used algorithm, which allow it to adapt to different situations. First, in general, note that the best results obtained have an average accuracy rate of 98% and the worst of 89.33%. This wide range implies that the possibility exists to set parameters to improve the algorithm for each situation.

Then, the general conclusions are presented after the previous analysis of the experiments.

The first expected conclusion is that regardless of the parameters used in the algorithm, we obtain worse results as the used crop is smaller. There is not a case in the experiments that the average accuracy rate with a small crop exceeds the worst results with a big crop for the same number of devices. Obviously, the processing in terms of execution time increases as higher crop is used.

The second general conclusion is that there are not clearly defined configuration parameters for the algorithm for each crop size that allows the best results to be obtained. Any obtained combination of parameters has similar results, although it is noteworthy that there are parameters that optimise the average accuracy rate to a greater extent. It is the responsibility of the forensic analysts to achieve greater results optimisation at the expense of a longer execution time or otherwise. Moreover, it can be concluded that none of the parameters used are superfluous because none of them independently weaken the results for all possible combinations.

A third general conclusion is that for both large and small crops there is a common configuration that gets the best results: adaptive variance estimation and no zero-meaning.

Focusing on the case of each crop size, the conclusions are shown below.

For the case of large crops (1024×1024) it can be concluded that the use of different configuration parameters does not clearly generate better results compared with the other options (the largest difference between all the results is 1.67%). The best option is to use adaptive variance estimation and not zero-meaning and the second best option does not use zero-meaning either. Hence, we can conclude that for large crops the zero-meaning does not provide any improvement and it makes the results slightly worse. Regarding the type of variance to use, it can be concluded that taking into account the processing time using adaptive variance it takes a long time. For large crops and a large number of images to be analysed it is better not to use it (in the worst case the results worsen by 0.5%), unless there are not time restrictions or we have high throughput.

In the case of small crops (512×512), there are no significant differences with respect to the use of different configuration parameters. The worst case is the one that uses the adaptive variance estimation and zero-meaning; in small crops we conclude that it is a bad choice because it gets far worse results than the other options (2.34% in the best case).

Concerning the use of various types of variance estimation and zero-meaning conclusions are similar to the case of large crops.

To evaluate the scalability of the approach, we repeated the experiment using 14 models from seven manufactures and achieved an average success rate of 87.21%.

Depending on the number and the type of images that have to be analysed and maximising the success rate depending on the desire processing time, the forensic analyst has the possibility of setting certain parameters in the algorithm of identifying the source acquisition. This will allow the analyst to obtain results closer to their needs and processing constraints.

Our results, tentatively, suggest that the method is applicable to datasets containing images from a large number of different cameras and therefore the method promises potential uses for digital forensics and data mining applications.

7 Acknowledgments

Part of the computations of this work were performed in EOLO, the HPC of Climate Change of the International Campus of Excellence of Moncloa, funded by MECD and MICINN. This is a contribution to CEI Moncloa.

8 References

- 1 'Gartner says smartphone sales grew 46.5 percent in second quarter of 2013 and exceeded feature phone sales for first time', <http://www.gartner.com/newsroom/id/2573415>, 2013
- 2 Ahonen, T., Moore, A., Almanac, T.A.: 'Mobile telecoms industry annual review', 2012
- 3 Embedded imaging takes off as stand-alone digital cameras stall, <http://www.icinsights.com/news/bulletins/Embedded-Imaging-Takes-Off-As-Standalone-Digital-Cameras-Stall/>, 2013
- 4 Baer, R.: 'Resolution limits in digital photography: the looming end of the pixel wars – OSA technical digest (CD)'. Proc. Imaging Systems, Optical Society of America, June 2010, p. ITuB3
- 5 Al-Zarouni, M.: 'Mobile handset forensic evidence: a challenge for law enforcement'. Proc. Fourth Australian Digital Forensics Conference, School of Computer and Information Science, Edith Cowan University, December 2006
- 6 Gloe, T., Kirchner, M., Winkler, A., Bohme, R.: 'Can we trust digital image forensics?'. Proc. 15th Int. Conf. on Multimedia, September 2007, pp. 78–86
- 7 Ramanath, R., Snyder, W.E., Yoo, Y., Drew, M.S.: 'Color image processing pipeline', *IEEE Signal Process. Mag.*, 2005, 22, (1), pp. 34–43
- 8 McKay, C., Swaminathan, A., Gou, H., Wu, M.: 'Image acquisition forensics: forensic analysis to identify imaging source'. Proc. IEEE Int. Conf. on Acoustics, Speech and Signal Processing, June 2008, pp. 1657–1660
- 9 Wang, B., Guo, Y., Kong, X., Meng, F.: 'Source camera identification forensics based on wavelet features'. Proc. Int. Conf. on Intelligent Information Hiding and Multimedia Signal Processing, IEEE Computer Society, September 2009, pp. 702–705
- 10 Van Lanh, T., Chong, K.S., Emmanuel, S., Kankanhalli, M.S.: 'A survey on digital camera image forensic methods'. Proc. IEEE Int. Conf. on Multimedia and Expo, July 2007, pp. 16–19
- 11 Choi, K.S.: 'Source camera identification using footprints from lens aberration'. Proc. Digital Photography II, number 852 in 6069, SPIE Int. Society For Optical Engineering, February 2006, pp. 60690J–60690J-8
- 12 Bayram, S., Sencar, H.T., Memon, N.: 'Classification of digital camera-models based on demosaicing artifacts', *Digit. Invest.*, 2008, 5, (1–2), pp. 49–59
- 13 Long, Y., Huang, Y.: 'Image based source camera identification using demosaicing'. Proc. IEEE Eighth Workshop on Multimedia Signal Processing, October 2006, pp. 419–424
- 14 Celiktutan, O., Avcibas, I., Sankur, B., Ayerden, N.P., Capar, C.: 'Source cell-phone identification'. Proc. IEEE 14th Signal Processing and Communications Applications, April 2006, pp. 1–3
- 15 Meng, F.J., Kong, X.W., You, X.G.: 'Source camera identification based on image bi-coherence and wavelet features'. Proc. Fourth Annual IFIP WG 11.9 Int. Conf. on Digital Forensics, Kyoto, Japan, January 2008
- 16 Ozparlak, L., Avcibas, I.: 'Differentiating between images using wavelet-based transforms: A comparative study', *IEEE Trans. Inf. Forensics Sec.*, 2011, 6, (4), pp. 1418–1431
- 17 Liu, Q., Li, X., Chen, L., et al.: 'Identification of smartphone-image source and manipulation', in Jiang, H., Ding, W., Ali, M., Wu, X. (Eds.): 'Advanced

- research in applied artificial intelligence' (Springer, Berlin Heidelberg, Dalian, China, 2012), pp. 262–271
- 18 Geradts, Z.J., Bijhold, J., Kieft, M., Kurosawa, K., Kuroki, K., Saitoh, N.: 'Methods for identification of images acquired with digital cameras'. Proc. Enabling Technologies for Law Enforcement and Security, SPIE-Int. Society for Optical Engine, February 2001, vol. 4232, pp. 505–512
 - 19 Lukas, J., Fridrich, J., Goljan, M.: 'Digital camera identification from sensor pattern noise', *IEEE Trans. Inf. Forensics Sec.*, 2006, **1**, (2), pp. 205–214
 - 20 Costa, F.D.O., Eckmann, M., Scheirer, W.J., Rocha, A.: 'Open set source camera attribution'. Proc. 25th Conf. on Graphics, Patterns and Images, IEEE, August 2012, pp. 71–78
 - 21 Li, C.T.: 'Source camera linking using enhanced sensor pattern noise extracted from images'. Proc. Third Int. Conf. on Crime Detection and Prevention (ICDP 2009), Curran Associates, Inc., December 2009, pp. 1–6
 - 22 Sandoval Orozco, A.L., Arenas González, D.M., Corripio, J.R., García Villalba, L. J., Hernandez-Castro, J.C.: 'Techniques for source camera identification'. Proc. Sixth Int. Conf. on Information Technology, May 2013, pp. 1–9
 - 23 Chen, M., Fridrich, J., Goljan, M., Lukas, J.: 'Determining image origin and integrity using sensor noise', *IEEE Trans. Inf. Forensics Sec.*, 2008, **3**, (1), pp. 74–90
 - 24 Celiktutan, O., Sankur, B., Avcibas, I.: 'Blind identification of source cell-phone model', *IEEE Trans. Inf. Forensics Sec.*, 2008, **3**, (3), pp. 553–566
 - 25 McKay, C.: 'Forensic analysis of digital imaging devices'. Technical report, University of Maryland, 2007
 - 26 Adams, J., Parulski, K., Spaulding, K.: 'Color processing in digital cameras', *IEEE Micro*, 1998, **18**, (6), pp. 20–30
 - 27 Chang, C.C., Lin, C.J.: 'LIBSVM: A library for support vector machines', Version 3.17, 26 April 2013, <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
 - 28 Corripio, J.R., Arenas González, D.M., Sandoval Orozco, A.L., García Villalba, L. J., Hernandez-Castro, J.C., Gibson, S.J.: 'Source smartphone identification using sensor pattern noise and wavelet transform'. Proc. Fifth Int. Conf. on Imaging for Crime Detection and Prevention (ICDP 2013), 16–17 December 2013, pp. 1–6