SOURCE IDENTIFICATION OF CAMERA PHONES USING SVD

Ahmad Ryad Soobhany [1,2], *K. P. Lam* [1], *Peter Fletcher* [1], *David Collins* [1]

[1] Research Institute for the Environment, Physical Sciences and Applied Mathematics
Keele University, Keele, Staffordshire, ST5 5BG, United Kingdom
[2] Forensic Pathways Ltd, Tamworth, Staffordshire, B77 5ES, United Kingdom

ABSTRACT

We present a novel method to extract the sensor pattern noise (SPN) of digital images and associate them with their respective source camera phones. The method first estimates the photo response non-uniformity (PRNU) of each image by means of its energy level and then converts it to an additive noise to facilitate extraction using Singular Value Decomposition (SVD). The latter is a spectral decomposition technique that separates the PRNU from the signal subspace. The camera reference signatures of the individual cameras are computed from a sample of their respective images and compared with a mixture of image signatures from a set of known camera devices. Our studies show that it is possible to determine the source device of digital images from camera phones using such method of signature extraction, with encouraging results.

Index Terms— Source Identification, Singular Value Decomposition, Digital Image Forensics, Sensor pattern noise, PRNU

1. INTRODUCTION

Digital images that are created by digital cameras found in bespoke cameras, mobile phones, tablets or video camcorders can be used for illicit purposes and for the commission of crime. When a forensic investigator recovers images from a suspect source, for example a mobile phone or any secondary storage, s/he may want to identify the source device that created the image to be able to link the images to a suspect or find out if the content of the images has been tampered with. When an image is created in the camera pipeline, described in [1], traces or artefacts of the camera processing are left in the image. The artefacts can be used as digital fingerprints or signatures to identify the source device that created the image. Several techniques exist to use these artefacts for identifying and linking source devices, such as lens aberration, colour filter array (CFA) interpolation and demosaicing, camera response function (CRF), JPEG compression, higher order wavelet statistics and relevantly sensor pattern noise (SPN) or photo response non-uniformity (PRNU). The first and last methods are often used to identify the specific source device that generated the image.

Until recently, the most prominent method to extract the PRNU has been to use a wavelet extraction filter that extracts in the wavelet domain the medium to high frequency subbands in which the PRNU lies [2]. As such, the extracted signature contains a mixture of different types of noises including the desired SPN or PRNU, random noise, fixed pattern noise (FPN) and any high frequency scene details. The SPN consists mainly of the FPN and PRNU noise, and the latter is unique to the individual sensor of a camera. The FPN is caused by dark currents when the sensor is not exposed to light and is usually removed by some cameras. Further work has been done to reduce the FPN and other random noises to obtain a cleaner sensor pattern noise. The signature can be enhanced by reducing the high frequency scene details which pollute the signature [3].

The PRNU is a multiplicative noise whose magnitude depends on the intensity of light falling on the sensor; thus the PRNU has a relatively low energy in general. If the PRNU can be converted to an additive noise, it will therefore be much easier to extract the signature from the image. Furthermore, if the range of the energy of the PRNU can be estimated, it will be easier to separate the PRNU from other polluting noises and get a cleaner signature. This paper presents a novel method of extracting the PRNU of digital images produced by mobile phone cameras. The images are decomposed by using Singular Value Decomposition (SVD), which separates the images into ranks of descending order of energies. The extraction of the PRNU is performed in the logarithmic domain, using the homomorphic filtering technique, where the inherently nonlinear PRNU is converted to an additive noise [4]. The ranks are chosen based on the relative range of energy of the PRNU compared to the image. Experiments were performed on five models of mobile phones. Two handsets of each model, ten cameras, were used to demonstrate that the signature obtained by our method is both unique and consistent; thus enabling the differentiation between individual devices of the same camera model. Gul & Avcibas have investigated the use of SVD to identify the model of a mobile phone. Their method uses the singular values from SVD decomposition to estimate the relative linear dependency of image rows/columns, which identifies the CFA interpolation algorithm of the camera [5].

The rest of the paper is organised as follows. The background on digital image forensics and device identification is reviewed in section 2 followed by a brief description of the current PRNU extraction methods in section 3. The SVD technique is described in section 4, which leads to an elaboration of our signature extraction model in section 5. The experimental procedures are described and the results are explained in section 6 followed by a discussion of the results in section 7 and finally the conclusions in section 8.

2. DIGITAL IMAGE FORENSICS & DEVICE IDENTIFICATION

Digital image forensics can help an investigator obtain the information and knowledge about a source device that created

some suspect images. The artefacts left behind in the digital image by the camera can be from the characteristics of the imaging device itself or the processing inside the device [6]. In general, forensic investigators do not have any previous knowledge about the source of the images they recover and digital image forensics usually works as a 'blind' approach without needing à priori knowledge about the images.

The identification of the CFA interpolation and demosaicing algorithms present in digital images can be performed by calculating the correlation between the different colour channels in a colour image and estimating the demosaicing algorithm used to produce the image [7]. Lens radial distortion occurs when straight lines from the object are rendered as curved lines on the sensor of the camera and the difference between the distorted line and the straight line can be measured and used to identify the camera [8]. Quantization tables vary between camera manufacturers and different camera models from the same manufacturer [9]. Digital images are usually recompressed for storage or transmission and in these cases the original source device can be identified [10]. The CRF can be estimated by finding the mapping algorithm using a single image, and the imaging device can be identified as the source of that image [11]. Also, higher-order wavelet statistics have been applied for camera model and make identification, together with binary similarity measures and image quality measures as well as a SVM classifier [12].

The CFA, CRF, JPEG compression and statistical techniques can be used to identify a particular model or make of a camera whereas the lens aberration and SPN can be used to ascertain distinct devices of the same model. Camera lenses can be changed readily whereas the camera sensor is much harder to change as well as being more expensive and hence it is uncommon for a sensor to be changed. This makes the SPN a better camera signature than lens aberration.

3. SIGNATURE EXTRACTION OF SPN/PRNU

The SPN consists mainly of the PRNU noise that is unique to the individual sensor of a camera [2]. It can be used to identify source devices as well as to determine whether an image has been tampered with [13]. The PRNU is due to imperfections arising from the manufacturing process of the sensor and due to slight variations in conversion of light to electrical energy by individual pixel sensors [14]. A combination of the uniqueness of the imperfections in the silicon material and the different sensitivity of the pixels makes the PRNU ideal for differentiating between sensors, even if they are made from the same silicon wafer, and hence the respective cameras into which they are embedded.

Two most commonly used denoising filters for signature extraction are the Gaussian filter in the spatial domain and the wavelet domain based approach. The Gaussian filter is two dimensional where the variance of the filter can be varied to choose the cut-off frequency that will determine the level of scene content and sensor noise [15]. The second approach applies wavelet decomposition to represent the image in different levels of details. A noise filter, such as a Wiener filter, is applied to the details from which an image reconstruction is performed to obtain the noise free image. The denoising filter is described in details in [2]. The PRNU, n, which appears as a high frequency signal, can be extracted from an image, I, based on the model proposed in [2] as a high pass filter:

$$n = I - f(I) \tag{1}$$

where f is a denoising function, which acts as a low pass filter to aid the extraction of the desired spectrum of noise from the image.

The multiplicative nature of the PRNU has made it particularly susceptible to scene details. For mobile phones, in addition, it can also be contaminated by the blockiness (row/column noise) created by the JPEG compression and other processing operations performed in the camera pipeline. Consequently, further processing is often applied to facilitate the estimation of the SPN, including the attenuation of non-unique artefacts (NUA) such as the FPN, blockiness and colour interpolation [13]. The accuracy of SPN can also be improved by attenuating the interference of scene details with the enhancer described in [3], where the enhanced SPN was shown to increase the identification rate and allows the use of smaller image crop size. However, these methods did tend to decrease the overall quality and strength of the SPN, which is already a weak signal.

In general, the identification of source devices is performed by extracting the digital signatures from a number of images, say, 50 pictures from each camera. The average of these signatures is calculated to form the camera reference signature. The signatures of recovered suspect images are extracted and compared against the camera reference signature for possible matches. Two comparison methods are commonly used; namely, Peak to Correlation Energy (PCE) measure and cross-correlation coefficient (CCC). Broadly speaking, the PCE is a more stable test than correlation method, particularly when the image has undergone geometrical manipulations [14].

4. SINGULAR VALUE DECOMPOSITION

Signal decomposition is an important practical problem as the energy in most real-world signals has unevenly distributed frequency spectra [16]. Using signal compaction techniques, the energy of a signal can often be redistributed into a significantly smaller number of frequency sub-bands, allowing them to be divided into sub-spectra in order that those with more energy content will be given a significantly higher priority for further processing. By analysing and discarding signal subspaces with lower priority, a signal can be reconstructed or approximated by a decomposition-synthesis procedure that is widely adopted in such practical applications as signal compression/coding. Additionally, such a procedure also forms the mathematical basis of modern time-frequency based techniques for analysing signal subspace, providing relevantly an expansive means of spectral analysis that naturally leads to the transform coding methods representative of the eigen-decomposition approach of spectral estimations.

Most eigenvector approaches work by separating a multidimensional signal into two subspaces, which are commonly referred to the signal and noise subspaces. By convention, the ensuing transform generates eigenvalues in decreasing order and eigenvectors that are orthonormal, allowing eigenvectors that are part of the noise to be identified and eliminated. In practice, the most challenging part of eigenvector spectral analysis is to compute the appropriate dimension of the signal or noise subspace, which often resorts to a trial and error procedure. Mathematically, a matrix A with m rows and n columns with rank r, $r \leq n \leq m$, can be expanded or decomposed into:

$$A = USV^T \tag{2}$$

where U and V^T are two orthogonal matrices of size *mxm* and *nxn* respectively [17]. S is the diagonal matrix, of size *mxn*, containing r non-zero singular values. The decomposition of matrix A is known as Singular Value Decomposition (SVD). When an image is decomposed using SVD, the ranks of the image can be represented as component matrices with decreasing energy contents [18]. SVD can be used to separate the spectra of the image and the ranks can be selected in accord with the aggregated total image energy of the individual ranks; that is, sum of λ_i^2, where λ_i represents the eigenvalue associated with that eigenvector (e_i).

5. SIGNATURE EXTRACTION MODEL

The wavelet based signature extraction method described in section 3 is most widely used, as it provides better identification results than the Gaussian filtering approach [3]. The method selects a wavelet filter/family to extract the PRNU by high-pass filtering in the frequency domain, resulting in scene details being included in the extracted signature. In particular, PRNU is a multiplicative noise and consists of additive low frequency defects (e.g. dust particles on lens) and the pixel non-uniformity (PNU) [2]. In practice, the noise used for identifying the individual source camera is the PNU, since the low frequency defects are removed by averaging several images or high-pass filtering as part of the denoising process. PRNU is not a temporal noise which means it is a spatial noise only with a multiplicative spatial variance. A simplified model for a noisy image I can be represented as

$$I = I_0 + I_0.K + \eta \tag{3}$$

where I_o is the clean image (perfect absorption of light energy by pixels), K is the PRNU and η is the remaining noise, such as shot noise, dark noise and read-out noise, associated with the image. The noisy image is comprised of the clean image to which is added the product of the clean image and PRNU.

If the image model is converted from the spatial domain to the logarithmic domain, the multiplicative noise model is transformed to an additive noise model in the logarithmic domain [19]. The homomorphic filtering technique is used, where the inherently nonlinear PRNU is transformed into an additive noise. Thus, the result is an additive model consisting of the image and PRNU noise, as follows:

$$\acute{I} = \acute{I}_0 + \acute{K} \tag{4}$$

where $\acute{I} = log(I)$, $\acute{I}_0 = log(I_0)$, $\acute{K} = log(1+K)$ and the noise η is cancelled by averaging many images created by the sensor.

The energy of the PRNU in an image depends on the type of device that produced the image and is a fraction of the total energy of the image. There are two types of sensors that are primarily used in digital cameras, the CCD (Charge-Coupled Device) and the CMOS (Complementary metal–oxide–semiconductor). The CCD produces less noise but requires more power when compared to the CMOS, hence the reason CMOS is used most often in camera phones where space and battery life are crucial. The energy of the PRNU in a CMOS will be affected by other sources of noises and the average power (variance) of the PRNU can be reduced [20]. The PRNU in CCD was measured by calculating the variance (σ^2) of the noise in 100 image sets [21]. The energy of the PRNU can be estimated to be in the range of 0.01% to 1.5%, depending on the type of image and sensor.

Table 1. Mobile Phone names and aliases with maximum image resolution

Mobile Phone	Alias	Max Image resolution
nokia_C2_01_A	cam_1	1536x2048
nokia_C2_01_B	cam_2	1536x2048
nokia_E72_A	cam_3	2592x1944
nokia_E72_B	cam_4	2592x1944
nokia_N95_A	cam_5	2592x1944
nokia_N95_B	cam_6	2592x1944
samsung_galaxy_S2_A	cam_7	3264x2448
samsung_galaxy_S2_B	cam_8	3264x2448
zte_orange_sanfrancisco_A	cam_9	1536x2048
zte_orange_sanfrancisco_B	cam_10	1536x2048

The logarithmic image model is then decomposed into ranks using SVD. The ranks are in descending order of relative energy of the image. The partial logarithmic image is reconstructed using a selected range of ranks in accordance with their associated energy. The latter should be chosen to contain the PRNU of the camera that created the image. If the digital signature is converted from the logarithmic domain back to the spatial domain the original image cannot be recovered, which signifies that the signature can be stored or transferred securely. The signatures extracted by this method can be used to create the camera reference signatures of the source device for identification purposes or compared against the reference signatures of other cameras for linkage purposes.

6. EXPERIMENTS AND RESULTS

For the purpose of our experiments, a total of 1000 images were chosen evenly from 10 mobile phones; i.e., each device contributed 100 images. Most of these phones were older models of the respective make and, as such, offer a significantly lower image quality, particularly when compared with images taken from digital cameras. To demonstrate that our method can differentiate between devices of the same make and model as previously explained in section 2, there were five different makes/models, each of which has two phones.

Table 1 shows that, for each model, the two phones share the same prefix but are distinguished by different labels 'A' and 'B'. Most of the phones are products from Nokia, since it is one of the most popular makes in the low-to-medium end of the camera phone market. In addition, the inclusion of different phone models from the same make was also expected to better demonstrate the identification performance of our method. In all cases, the pictures were taken at the highest native resolution of the cameras and stored in the JPEG format, which is the de facto compression format for still images from camera phones. To ensure generality, the pictures were natural images consisting of a mixture of outdoor and indoor scenes, captured during the day and at night. Further, given the expectedly different sizes of the captured images, due to the different quality of the camera phones, they were all cropped to the same size of 512x512 pixels, consisting of the lower left corner of the image as described in [1].

The SVD-based signature extraction procedure was applied to these cropped images, allowing the creation of reference signatures for the individual cameras. This was achieved by selecting 50 images randomly out of the 100 sample and then averaging the extracted signatures as per the commonly adopted procedure
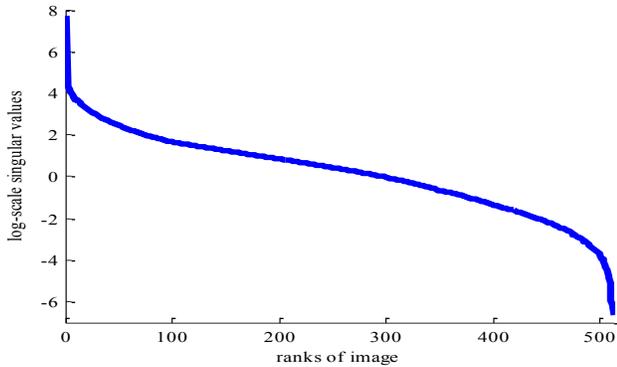
Figure 1. Plot of *log*-scaled singular values of a natural image with 512 ranks after SVD decomposition.

(described in section 3). The remaining 50 signatures from each camera dataset were then compared against the ten camera reference signature computed for the ten cameras. Peak to Correlation Energy (PCE) was used to match the signatures of the test images against the ten camera reference signatures.

The required energy range for the extraction of the PRNU was found to vary greatly depending on the amount of scene detail in the image. Images with high scene detail content will have their energy spread out more widely across the top ranks after performing SVD, whereas less 'busy' images had most of their energy concentrated among the first couple of high energy ranks. Fig. 1 shows the plot of the log-scaled singular values of a natural image with 512 ranks. There is a sharp drop after the first rank (from 7.6 to 4.4). For a blue sky image, the drop is from about 7.8 to 0.4. The first rank holds most of the scene detail energy.

The graph in Fig. 2 shows the result of comparing the camera signature of the Nokia_E72_A (*cam_3*) with the test signatures of ten images from each camera. The graph in Fig. 3 shows the result of comparing the camera signature of the Nokia_E72_B (*cam_4*) with the same test signatures as before.

7. DISCUSSION

The empirically selected range of ranks for extracting the signature was found to be between 50 and 150 inclusive, with the test images cropped at a size of 512x512 pixels [22]. It can be seen in Fig. 2 that the PCE values for the images from *cam_3* is significantly higher than the PCE values for the other cameras. The latter values are close to zero confirming the expectedly uncorrelated relationship. Furthermore, the identification results between images from *cam_4* and the reference signature of *cam_3* is similar to the results of other cameras, clearly demonstrating that our method can differentiate between two cameras of the same model.

The results in Fig. 3 corroborates the results from *cam_3* when using the camera reference signature of *cam_4*, although the PCE values for test signatures from *cam_4* are higher as shown in the figure. Preliminary studies suggest that this was largely due to the quality of the images that was selected (at random) to create the reference signature of *cam_3*; in particular, there were more saturated pixels present in these pictures, which impacts negatively on the PRNU. The results from the other cameras used in the experiment showed that it is possible to differentiate between cameras of the same model; except for cam_5 and cam_6 where the error rates were higher due to the quality of the images recovered from the cameras.
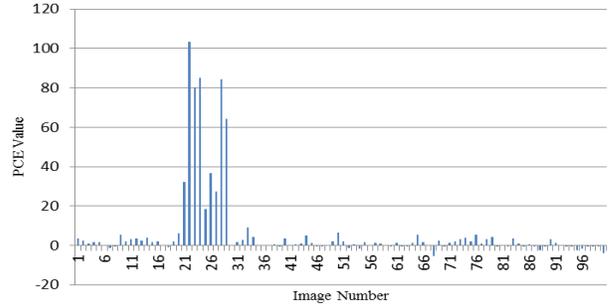

Figure 2. Nokia_E72_A camera reference signature and correlated with 100 images. Images 21 to 30 come from this camera, images 31 to 40 from *cam_4* and rest of images from the other 9 cameras.
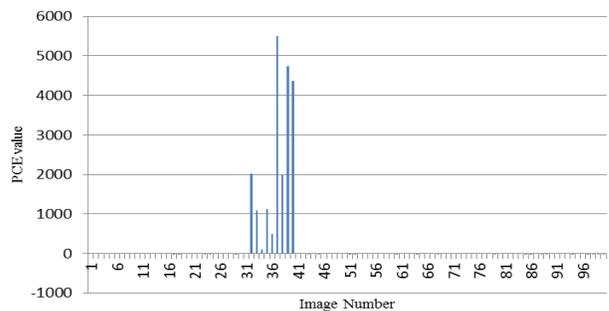

Figure 3. Nokia_E72_B camera reference signature and correlated with 100 images. Images 31 to 40 come from this camera, images 21 to 30 from *cam_3* and rest of images from the other 9 cameras.

8. CONCLUSION

A novel PRNU extraction method using SVD was introduced and demonstrated to be able to distinguish between camera phones of the same model. The extraction model described how SVD can be used as an image decomposition method for which signatures can be extracted from the individual images that can be associated with their respective source devices. The identification results of the test performed on 10 cameras showed that our method can differentiate between two cameras of the same make and model, suggesting that the signature is highly related to the SPN of the camera.

We also showed that the PRNU signature could be extracted relatively straightforwardly with most real-world/natural images. Further work, particularly on testing of automated algorithms for rank selection given the image characteristics, and classification performance including images of differing resolution and size, is underway.

REFERENCES

[1] A. R. Soobhany, R. Leary and K. P. Lam, "On the Performance of Li's Unsupervised Image Classifier and the Optimal Cropping Position of Images for Forensic Investigations," *International Journal of Digital Crime and Forensics (IJDCF),* vol. 3, pp. 1-13, 2011.

[2] J. Lukas, J. Fridrich and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security,* vol. 1, pp. 205-214, 2006.

[3] C. -T. Li, "Source Camera Identification Using Enhanced Sensor Pattern Noise," *IEEE Transactions on Information Forensics and Security,* vol. 5, pp. 280-287, 2010.

[4] R. C. Gonzalez and R. E. Woods, "Homomorphic filtering," in *Digital Image Processing*, Second ed. New Jersey: Prentice-Hall, 2002, pp. 191.

[5] G. Gul and I. Avcibas, "Source cell phone camera identification based on singular value decomposition," in *First IEEE International Workshop on Information Forensics and Security, 2009. WIFS 2009.* 2009, pp. 171-175.

[6] T. Gloe, M. Kirchner, A. Winkler and R. Böhme, "Can we trust digital image forensics?" in *Proceedings of the 15th International Conference on Multimedia,* Augsburg, Germany, 2007, pp. 78-86.

[7] A. Swaminathan, M. Wu and K. J. R. Liu, "Nonintrusive component forensics of visual sensors using output images," *IEEE Transactions on Information Forensics and Security,* vol. 2, pp. 91-106, 2007.

[8] K. San Choi, E. Y. Lam and K. K. Y. Wong, "Source camera identification using footprints from lens aberration," *Digital Photography II SPIE,* vol. 6069, pp. 172–179, 2006.

[9] H. Farid, "Digital image ballistics from JPEG quantization," Dept. Comput. Sci., Dartmouth College, Hanover, NH, Tech. Rep. TR2006-583, 2006.

[10] M. J. Sorell, "Conditions for effective detection and identification of primary quantization of re-quantized JPEG images," in *E-Forensics '08: Proceedings of the 1st International Conference on Forensic Applications and Techniques in Telecommunications, Information, and Multimedia and Workshop,* Adelaide, Australia, 2008, pp. 1-6.

[11] S. Lin, Jinwei Gu, S. Yamazaki and Heung-Yeung Shum, "Radiometric calibration from a single image," in *Proceedings of the 2004 IEEE Computer Society Conference On Computer Vision and Pattern Recognition, 2004. CVPR 2004.* 2004, pp. II-938-II-945 Vol.2.

[12] O. S. Celiktutan and B. Avcibas, "Blind identification of source cell-phone model," *IEEE Transactions on Information Forensics and Security,* vol. 3, pp. 553-566, 2008.

[13] M. Chen, J. Fridrich, M. Goljan and J. Lukas, "Determining image origin and integrity using sensor noise," *IEEE Transactions on Information Forensics and Security,* vol. 3, pp. 74-90, 2008.

[14] J. Fridrich, "Digital Image Forensic Using Sensor Noise," *IEEE Signal Process. Mag.,* vol. 26, pp. 26–37, 2009.

[15] E. J. Alles, Z. J. M. H. Geradts and C. J. Veenman, "Source camera identification for low resolution heavily compressed images," in *International Conference on Computational Sciences and its Applications, 2008. ICCSA '08.* 2008, pp. 557-567.

[16] T. Berger, *Rate Distortion Theory: A Mathematical Basis for Data Compression.* NJ: Prentice-Hall, 1971.

[17] C. Moler, "Eigenvalues and singular values," in *Numerical Computing with MATLAB* http://www.mathworks.co.uk/moler: The MathWorks, Inc, 2004, pp. 269.

[18] H. Andrews and C. Patterson, "Singular value decompositions and digital image processing," *IEEE Transactions on Acoustics, Speech and Signal Processing, ,* vol. 24, pp. 26-53, 1976.

[19] H. Xie, L. E. Pierce and F. T. Ulaby, "Statistical properties of logarithmically transformed speckle," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 40, pp. 721-727, 2002.

[20] A. El Gamal and H. Eltoukhy, "CMOS image sensors," *Circuits and Devices Magazine, IEEE,* vol. 21, pp. 6-20, 2005.

[21] K. Irie, A. E. McKinnon, K. Unsworth and I. M. Woodhead, "A model for measurement of noise in CCD digital-video cameras," *Measurement Science and Technology,* vol. 19, pp. 045207, 2008.

[22] A. R. Soobhany, "Image Source Identification and Characterisation for Forensic Analysis," *Ph.D. Thesis, Keele University, UK*, 2012.