

Camera model identification based machine learning approach with high order statistics features

Amel Tuama, Frédéric Comby, Marc Chaumont

▶ **To cite this version:**

Amel Tuama, Frédéric Comby, Marc Chaumont. Camera model identification based machine learning approach with high order statistics features. EUSIPCO: European Signal Processing Conference, Aug 2016, Budapest, Hungary. pp.1183-1187, 10.1109/EUSIPCO.2016.7760435 . hal-01374110

HAL Id: hal-01374110
https://hal.science/hal-01374110

Submitted on 29 Sep 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

CAMERA MODEL IDENTIFICATION BASED MACHINE LEARNING APPROACH WITH HIGH ORDER STATISTICS FEATURES

Amel TUAMA[1] *, Frédéric COMBY*[1]*, Marc CHAUMONT*[1,2]

[1] LIRMM(UMR5506), MONTPELLIER UNIVERSITY/CNRS, 34095, Montpellier Cedex 5, FRANCE
[2] NÎMES UNIVERSITY, 30021 Nîmes Cedex 1, FRANCE

ABSTRACT

Source camera identification methods aim at identifying the camera used to capture an image. In this paper we developed a method for digital camera model identification by extracting three sets of features in a machine learning scheme. These features are the co-occurrences matrix, some features related to CFA interpolation arrangement, and conditional probability statistics. These features give high order statistics which supplement and enhance the identification rate. The method is implemented with 14 camera models from Dresden database with multi class SVM classifier. A comparison is performed between our method and a camera fingerprint correlation-based method which only depends on PRNU extraction. The experiments prove the strength of our proposition since it achieves higher accuracy than the correlation-based method.

***Index Terms*—** Camera identification, Co-occurrences, CFA interpolation, Conditional Probability, SVM.

1. INTRODUCTION

Source camera identification has recently received a wide attention due to its important role in security and legal issue. It is the process of determining which camera device or model has been used to capture an image. The image acquisition process involves many steps inside camera device: lens system, filters, Color Filter Array (CFA), imaging sensor, and digital image processor. All these steps add artifacts to the image content which will provide different features for the identification process.

As a relation to prior work, researchers have proposed to use these artifacts to collect features and treat them in a machine learning approach [1]. Kharrazi et al. [2] used three sets of features to perform camera model identification. The 34 proposed features are color features, Image Quality Metrics (IQM), and wavelet domain statistics. Celiktutan et al. [3] used a subset of Kharrazi's feature sets to identify the source cell-phone camera. Then, they added the features of binary similarity measures to the previous feature sets to get 592 features.

Table 1. A comparison between feature based camera identification methods.

Camera Iden. Method	Features	No.of Features	No.of Models
Kharrazi et al. (2004) [2]	Color features, Image Quality Matrices, Wavelet features	34	5
Celiktutan et al. (2008) [3]	Image Quality matrices, Wavelet features, Binary Similarity Measures	592	16
Filler et al. (2008) [4]	Statistical moments, Block covariance, Cross-correlation of CFA, Cross-correlation of linear pattern	28	17
Gloe et al. (2012) [5]	Color features, Image Quality Matrices, Wavelet features	82	26
Xu and Shi (2012) [6]	Local Binary Patterns	354	18
Wahab et al. (2012) [7]	Conditional Probability	72	4
Marra et al. (2015) [8]	Spam of Rich models	338	10

Filler et al. [4] introduced a camera model identification method using 28 features related to statistical moments and correlations of the linear pattern. Gloe et al. [5] used Kharrazi's feature sets with extended color features to produce 82 features. Xu and Shi [6] used 354 Local Binary Patterns as features. Local binary patterns capture inter-pixel relations by thresholding a local neighborhood at the intensity value of the center pixel into a binary pattern.

Wahab et al. [7] used the conditional probability as a single feature set to classify camera models. The 72 conditional probability features were obtained using the coefficients of 8×8 DCT transform. Marra et al. [8] gathered 338 SPAM features from the rich models based on co-occurrences matrices of image residuals.

Other methods deal with camera identification but far from feature extraction and machine learning. Bayram et al. [9] explored the CFA interpolation process to determine

This work was supported and funded partially by the Ministry of Higher Education and Scientific Research in Iraq, Northern Technical University.

the correlation structure present in each color band which can be used for image classification. The main assumption is that the interpolation algorithm and the CFA filter pattern of each camera model is somewhat different from others, which will result in distinguishable correlation structures in the captured images.

A reliable one for identifying source camera based on sensor pattern noise is proposed by Lukas et al. [10]. PRNU can be used as fingerprint for uniquely identifying sensors. Choi et al. [11] proposed to use the lens radial distortion as a fingerprint to identify source camera model. Each camera model expresses a unique radial distortion pattern that helps on its identification.

Dirik et al. [12] proposed a device identification based on sensor dust in digital single lens reflex cameras (DSLR). Sensor dust patterns are used as artifacts on the captured images to identify the camera device. Table 1 shows the most known methods for camera identification based machine learning and feature extraction.

Our contribution is to use a bigger set of features in order to better describe the statistics (and that is what characterizes the proposed method). We extract a vector of 10932 features which can be considered as a huge number compared to those of the methods mentioned in Table 1.

The rest of this paper is organized as follows. Section 2 explains the correlation based method. Section 3 presents all the details of our method and feature extraction steps. In Section 4, we describe the experiments and the results. Finally, Section 5 concludes this paper and presents some perspectives.

2. CORRELATION BASED METHODS

Sensor pattern noise has drawn much attention due to its feasibility in identifying camera models of the same brand, and individual devices of the same model. The PRNU is unique to each sensor and is stable over time. By correlating the noise extracted from a query image against the known reference pattern, or PRNU, of a given camera, we can determine whether that camera was used to originally capture the query image. The reference pattern of a camera is first extracted from a series of images taken from known camera device. The reference pattern is then used to detect whether the camera used to generate the reference pattern was used to capture an unknown source image.

Generally, for each image I, the residual noise is extracted by subtracting the denoised version of the image from the image itself as follows:

$$N = I - F(I), \tag{1}$$

where $F(I)$ is the denoised image, and F is a denoising filter. Wavelet based denoising filter is recommended and it is used in most cases [13].

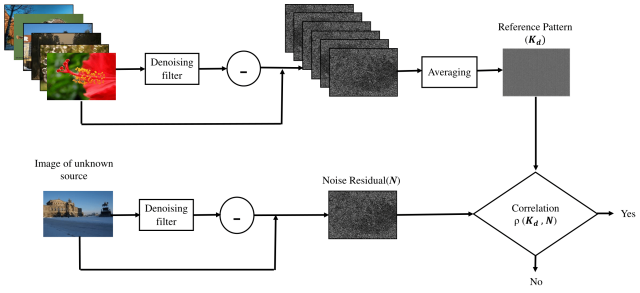


Fig. 1. The correlation based scheme

In order to extract the fingerprint of a camera, multiple images are denoised and averaged. At least 50 images are used to calculate the reference pattern K_d [13] of a known camera device D as in Equation 2.

$$K_d = \frac{\sum_{i=1}^{n}(N_i I_i)}{\sum_{i=1}^{n} I_i^2}. \tag{2}$$

A common approach to perform a comparison is to compute the Normalized Cross-Correlation which measures the similarity between the reference pattern K_d and the estimated noise N of an image under test which is of unknown source [13]. Normalized Cross-Correlation is defined as:

$$\rho(N, K_d) = \frac{(N - \overline{N}).(K_d - \overline{K_d})}{\|N - \overline{N}\|.\|K_d - \overline{K_d}\|}. \tag{3}$$

Where $\overline{N}$ and $\overline{K_d}$ are the means of N and K_d, respectively. Fig.1 illustrates the steps related to correlation based method.

3. PROPOSED FEATURE BASED METHOD

In this paper, we extract features from the residual noise N. After the image is decomposed into its three color channels, the noise residual N is extracted by subtracting the denoised version of the image I from the image itself as in Equation 1. For the denoising process, a wavelet based denoising filter, $F(I)$, is used based on a Wiener filtering of each wavelet subband for each channel as in [10].

Computing the linear pattern L, a periodic signal of the pattern noise, allows to suppress all artifacts produced by color interpolation and JPEG compression [13]. L is obtained by subtracting the average row and average column from each row and column respectively of N from each color channel separately [13]. This gives three linear patterns corresponding to each color channel, noted L_r, L_g and L_b for red, green, blue channels respectively. The three linear patterns are combined together by using the conversion formula from RGB to gray-scale as follows.

$$L_P = 0.3.L_r + 0.6.L_g + 0.1.L_b. \tag{4}$$

Three sets of features will be extracted, co-occurrences matrix, color dependencies, and conditional probability. Co-occurrences matrix will be extracted from L_P by calculating the different statistical relationships among neighboring pixels. The second features set, related to CFA arrangement, calculates the local dependencies and periodicity among neighboring pixels. The third features set is the conditional probability features which will be calculated from the original images by examining the absolute values of three selected coefficients in 8×8 DCT block. The following three sub-sections describe the theoretical part of the three features sets.

3.1. Co-occurrences Matrix

Recently rich models approach and co-occurrences matrix have been widely used in forensics applications [8, 14, 15]. The co-occurrences are a very good way to describe high order statistics of neighboring data. The co-occurrences feature vector is made of joint probability distributions of neighboring residual samples. In this work, we used the linear pattern of the noise residual obtained from Equation 4 as input to rich models. We use four-dimensional co-occurrences matrices formed by groups of four horizontally and vertically adjacent samples after they were quantized and truncated as follows:

$$R \leftarrow trunc_T(round(L_P/q)), \tag{5}$$

where $trunc_T$ is a function to minimize the residual range with $T \in \{-T, ..., T\}$, $round(x)$ gives the nearest integer value of x, L_P is the linear pattern of the noise residual, and $q \in \{1, 1.5, 2\}$ is the quantization step. The final co-occurrences matrix will be constructed from horizontal and vertical co-occurrences of four consecutive values from R. The horizontal co-occurrence matrix C_d^h is computed as follows [16]:

$$C_d^h = \frac{1}{Z}\big|\{(i,j) \mid R_{i,j} = d_1, R_{i,j+1} = d_2, R_{i,j+2} = d_3, R_{i,j+3} = d_4\}\big|, \tag{6}$$

where Z is the normalization factor, with $R_{i,j} \in \mathbb{N}$ is the coefficient from the matrix R at position $(i,j) \in \{1, ..., n\}^2$, $d = (d1, ..., d4) \in \{-T, ..., T\}^4$ with $T = 2$.

We can compute the vertical co-occurrences matrix equivalently.

3.2. Color Dependencies

The CFA pattern and the way of colors interpolation leave some periodic patterns which are important to characterize the camera models [9]. In this section, we will explain the set of features related to CFA arrangement. From the linear patterns of the noise residual L_r, L_g, and L_b, we compute local dependencies and periodicity among neighboring samples. The normalized cross-correlation, explained in Equation 7, is computed between the estimated linear pattern from the noise residual of the three color channels and their shifted version as in [4].

For each color channel pair $(A1, A2)$, $A1, A2 \in \{L_r, L_g, L_b\}$ and shift $\triangle_1 \in \{0, ..., 3\}$, $\triangle_2 \in \{0, ..., 3\}$. This step results in 96 features which are the result of six combinations of color channels by 4×4 shifts of $\triangle_1$ and $\triangle_2$.

$$\rho(A1, A2, \triangle) = \frac{\sum_{i,j}(A1_{i,j} - \overline{A1})(A2_{i-\triangle_1, j-\triangle_2} - \overline{A2})}{\sqrt{\sum_{i,j}(A1_{i,j} - \overline{A1})^2 \sum_{i,j}(A2_{i-\triangle_1, j-\triangle_2} - \overline{A2})^2}}, \tag{7}$$

where ρ is the normalized cross correlation, $\triangle = [\triangle_1 \ \triangle_2]^T$ is the 2D shift, $\overline{A1}$ and $\overline{A2}$ are sample means calculated from matrices $A1$ and $A2$ respectively.

3.3. Conditional Probability

Conditional probability features (CP) were introduced in camera identification by Wahab et al. [7]. A number of CP features can be obtained by examining the absolute values of three selected block DCT coefficients at different locations. For the usual 8×8 DCT transform, we picked three DCT coefficients from the 4×4 left upper sub-block because most non-zero coefficients are in that region. Given the three relative positions r, s, and t in a DCT block such that $\{r, s, t\} \in \{1, ..., 4\} \times \{1, ..., 4\}$, we compute the conditional probability as follows:

$$Prob(Y_i|X_i) = \frac{Prob(X_iY_i)}{Prob(X_i)}, \tag{8}$$

knowning that $X_i \in \{X_1, X_2, X_3\}$ and $Y_i \in \{Y_1, Y_2, Y_3\}$ are defined such as:

$$\begin{aligned} X_1 &= \{\text{value at position } r < \text{value at position } s\}, \\ X_2 &= \{\text{value at position } r > \text{value at position } s\}, \\ X_3 &= \{\text{value at position } r = \text{value at position } s\}, \\ \\ Y_1 &= \{\text{value at position } t < \text{value at position } s\}, \\ Y_2 &= \{\text{value at position } t > \text{value at position } s\}, \\ Y_3 &= \{\text{value at position } t = \text{value at position } s\}, \end{aligned} \tag{9}$$

Eight different arrangements of r, s, and t will be examined over nine events resulting in 72 features.

4. EXPERIMENTS AND EVALUATION

In order to assess the performance of our method, we carried out some experiments on a set of camera models from the Dresden image database [17]. It is one of the most widespread database dedicated to forensics applications. For the experimental phase, 14 camera models with single device from each model were used as in Table 3. An image is decomposed into its three color channels (R, G, B). For each camera model we used 200 full size images. Noise residual is extracted from all images by applying wavelet denoising filter.

Table 2. Identification accuracy of the proposed method and the correlation based method for 14 chosen camera models.

Camera Model	CM1	CM2	CM3	CM4	CM5	CM6	CM7	CM8	CM9	CM10	CM11	CM12	CM13	CM14
Correlation method%	98	98	100	96	99	98	97	100	96	98	97	96	97	95
Proposed method%	99.3	98.6	100	99.9	98.7	99.9	98.1	98	99.6	98.2	99.4	98.9	97.7	96.2

Table 3. Camera models used from Dresden database

Brand/Make	Model	Resolution
Agfa Photo (CM1)	DC-733s	3072x2304
Agfa Photo (CM2)	DC-830i	3264x2448
Agfa Photo (CM3)	Sensor 530s	4032x3024
Canon (CM4)	Ixus 55	2592x1944
Fujifilm (CM5)	FinePix J50	3264x2448
Kodak (CM6)	M1063	3664x2748
Nikon (CM7)	D200 Lens A/B	3872x2592
Olympus (CM8)	M1050SW	3648x2736
Panasonic (CM9)	DMC-FZ50	3648x2736
Praktica (CM10)	DCZ 5.9	2560x1920
Samsung (CM11)	L74wide	3072x2304
Samsung (CM12)	NV15	3648x2736
Sony (CM13)	DSC-H50	3456x2592
Sony (CM14)	DSC-W170	3648x2736

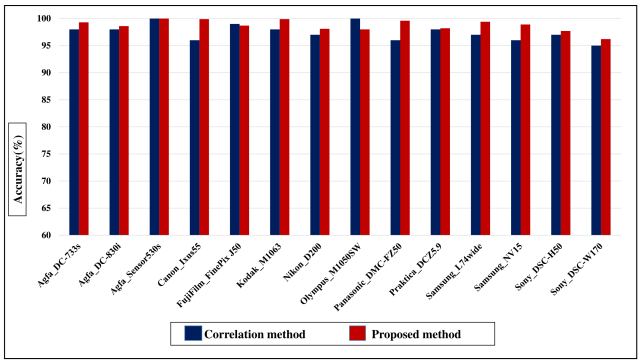


Fig. 2. Comparison of the identification results.

Table 4. Test results for images from Flickr data set.

Camera Make/Model	No.Images	Identification
Canon IXUS 55	97	99.1%
Fujifilm FinePix J50	74	98.7%

Table 5. The total accuracy for the compared methods.

Camera identification method	Accuracy
Correlation based method	97.5%
Proposed method	98.75%

SVM classifier is trained on 10932 features coming from three sets. Two feature sets are extracted from linear pattern of the noise residual for each image in the data set. The first set is the co-occurrences matrix which consists of 10764 features. While the second set consists of 96 features from CFA interpolation dependencies. The third features set is extracted by computing the conditional probability of the 8×8 DCT transform coefficients of the original images and resulting in 72 features, see Section 3.

The method of min-max scaling was used for feature normalization. In this approach, the features will be re-scaled, to a specific range [0,1]. For the classification, LIBSVM package with multi-classification was used [18] with the Radial Basis Function (RBF) and v-fold cross validation scheme. The kernel parameters values were for $\gamma = 2^{-4}$ and cost parameter $C = 2048$ after examining a grid search over a range of values. We run the training procedure 10 times then averaging the results. Each time, the training and testing data sets are selected randomly.

We used 100 images for the training and another 100 for the test. The proposed method achieved a total identification accuracy of 98.75% over 14 camera models as shown in Table 2. We recorded a perfect identification for $Agfa - Sensor530s$ and very high performance for $Canon - Ixus55$, and $Kodak - M1063$. We noticed that the two models of Sony recorded the lowest rates due to the in-camera processes they achieve.

In order to test the method against the high dimensionality problem, we performed an additional experiment. Two image subsets were downloaded from the most wide web database "$Flickr$". The images of two camera models $Canon - Ixus - 55$, and $Fujifilm - FinePix - J50$ are used only to test the network which was trained with the previous data set "$Dresden$". We achieved an identification accuracy 99.1%, 98.7% respectively as shown in Table 4. These results prove the robustness of the proposed method against the curse of dimensionality.

For comparison, we implemented the method of the correlation based sensor pattern noise for camera identification, explained in Section 2. This method depends on extracting the fingerprint of the camera which can be estimated by averaging a set of images. Normalized correlation is applied between the fingerprint and an image under test to investigate whether this image came from this camera or not. For each camera model, we used 100 images to estimate the fingerprint and we left the rest 100 images for the test. This results in 97.5% as a total identification accuracy as in Table 2. The bar chart in Figure 2 is showing the comparative accuracy for the two methods in terms of each camera model separately. In Table 5, we can see that the proposed method performs higher

than the correlation based method since it achieves 98,75% while the compared method only achieves 97.5%.

5. CONCLUSION

This paper contributes in identifying camera models based on feature extraction and machine learning. The objective in adding this big number of features is to allow enhancing the identification rate by providing strong statistic tool.

The algorithm is composed of extracting three sets of features. The noise residual is obtained by applying wavelet denoising filter. Images from 14 camera models were used from the Dresden database and classified by SVM classifier.

The experimental results show that the proposed method gives very high identification accuracy since it provides an identification rate of 98.75% in comparison with the correlation based method which achieved 97.5%. The problem of dimensionality was examined by testing images from another database.

The future work will include using a large scale database with more camera models, so as the usage of multiple devices of the same model.

REFERENCES

[1] M. Kirchner and T. Gloe, "Forensic camera model identification," in *In: Ho, T., Li, S. (eds.) Handbook of Digital Forensics of Multimedia Data and Devices.Wiley-IEEE Press,*, 2015.

[2] M. Kharrazi, H.T. Sencar, and N. Memon, "Blind source camera identification," in *IEEE International Conference on Image Processing ICIP '04.*, vol. 1, pp. 709–712, 2004.

[3] O. Celiktutan, B. Sankur, and I. Avcibas, "Blind identification of source cell-phone model.," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 553–566, 2008.

[4] T. Filler, J. Fridrich, and M. Goljan, "Using sensor pattern noise for camera model identification," in *Proc. ICIP, 15th IEEE International Conference on Image Processing, San Diego, California, October 12-15*, pp. 1296–1299, 2008.

[5] T. Gloe, "Feature-based forensic camera model identification," *Shi, Y.Q., Katzenbeisser, S. (eds.) Transactions on Data Hiding and Multimedia Security VIII. LNCS,*, vol. 7228, pp. 42–62, Springer, Heidelberg, 2012.

[6] G. Xu and Y. Q. Shi, "Camera model identification using local binary patterns," in *Proc. IEEE Int Conference on Multimedia and Expo (ICME) , Melborne, Australia*, pp. 392–397, 2012.

[7] A.W. Abdul Wahab, A.T.S. Ho, and S. Li, "Inter-camera model image source identication with conditional probability features," in *Proc. of the 3rd IIEEJ Image Electronics and Visual Computing Workshop, Paper ID 2P-2, 2012*.

[8] F. Marra, G. Poggi, C. Sansone, and L. Verdoliva, "Evaluation of residual-based local features for camera model identification," in *New Trends in Image Analysis and Processing - ICIAP Workshop : BioFor, Genoa, Italy, September 7-8*, pp. 11–18, 2015.

[9] S. Bayram, H.T. Sencar, and N. Memon, "Improvements on source camera model identification based on cfa interpolation," in *Advances in Digital Forensics II, IFIP International Conference on Digital Forensics, Orlando Florida*, pp. 289–299, 2006.

[10] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on, Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, June 2006.

[11] K. S. Choi, E. Y. Lam, and K. Wong, "Source camera identification using footprints from lens aberration," in *Proc. SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics*, vol. 6069, pp. 60690J–60690J–8, 2006.

[12] A. E. Dirik, H. T. Sencar, and N. Memon, "Source camera identification based on sensor dust characteristics," in *IEEE Workshop on Signal Processing Applications for Public Security and Forensics, SAFE '07, Washington, USA*, pp. 1–6, 2007.

[13] J. Fridrich, "Digital image forensic using sensor noise," *IEEE Signal Processing Magazine*, vol. 26, no. 2, pp. 26–37, 2009.

[14] X. Qiu, H. Li, W. Luo, and J. Huang, "A universal image forensic strategy based on steganalytic model," in *Proceedings of the second ACM Workshop on Information Hiding and Multimedia Security*, New York, NY, USA, IH&MMSec'14, pp. 165–170, 2014.

[15] L. Verdoliva, D. Cozzolino, and G. Poggi, "A feature-based approach for image tampering detection and localization," in *IEEE International Workshop on Information Forensics and Security (WIFS), Georgia, USA,*, pp. 149–154, 2014.

[16] J. Fridrich and J. Kodovsky, "Rich models for steganalysis of digital images," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 868–882, June 2012.

[17] T. Gloe and R. Böhme, "The 'Dresden Image Database' for benchmarking digital image forensics," in *Proceedings of the 25th Symposium On Applied Computing (ACM SAC 2010)*, vol. 2, pp. 1585–1591, 2010.

[18] C. Chang and C. Lin, "Libsvm: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, May 2011, Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.