Source Cell-phone Identification

Oya Çeliktutan[a], İsmail Avcıbaş[a], Bülent Sankur[b], Nasir Memon[c]
[a]*Uludag University, Electronics Engineering Department, Bursa, Turkey*
[b]*Bogazici University, Electrical-Electronics Engineering Department, Istanbul, Turkey*
[c]*Dept. of Comp. and Inf. Science, Polytechnic University, Brooklyn, NY, USA*
oyaxceliktutan@hotmail.com, avcibas@uludag.edu.tr, bulent.sankur@boun.edu.tr, memon@poly.edu

Abstract

The techniques to validate the authenticity of digital images are rather limited. In this paper, we focus on blind source cell-phone identification problem. The main idea is that proprietary interpolation algorithm (involved due to the structure of color filter array [CFA]) leaves footprints in the form of correlations across adjacent bit planes of the image. For this purpose, we define a set of binary similarity measures and image quality measures in conjunction with a KNN classifier to identify the originating cell-phone. We provide results on identifying source among three cell-phones.

1. Introduction

Image forensics is a new emerging field concerned with determining the source and potential authenticity of a digital objects and possibly reconstructing the history of manipulations effected. In this sense image forensics tries to meet the new challenge of safeguarding the authenticity of digital image and to enable their continued usefulness as trustworthy documents and legal evidence. Digital images can obviously be easily created, edited and manipulated with increasingly more sophisticated tools, which do not leave much of any perceptible trace.

Digital watermarking falls short to meet all desiderata of this particular problem [1]. On the one hand, watermarking requires that imaging devices be equipped with built-in watermarking capabilities; on the other hand, watermarks may not be able to classify all types of attack. Forensic tools, however, can be envisaged to identify the nature of the manipulation. Finally, forensic tools can be used concomitantly with watermarking in decision fusion schemes.

In this paper, we focus on the identification of source cell-phones. In other words, the problem is to determine the make and the brand of the camera with which the given image was captured. The camera brand/made identification is based on the telltale effects due the proprietary image formation pipeline. In fact, the main difference between cameras originates from the color filter array that is used to interpolate between color pixels. In prior works [2], [3], source camera identification problem was studied using feature sets based on image quality metrics [4] and higher-order statistics [5].

All camera identification techniques exploit the fact that state-of-the-art cell-phone cameras, due to cost considerations, employ a single mosaic structured *color filter array* (CFA) rather than having different filters for each color component [6]. This process is illustrated in Fig. 1. As a consequence each pixel in the image has only one color component associated with it, and each digital camera employs a proprietary interpolation algorithm in obtaining the missing color values. This very proprietary interpolation algorithm leaves footprint like correlations between contiguous bit planes of an image.

In this work we use binary symmetry features, which directly address correlation properties within and between planes. We consider also mixtures of other categories of features, such Image Quality Measures (IQM) [4]. The rest of this paper is organized as follows. In section 2, we briefly describe the similarity measures used in the classifier design, which were selected from a set of measures described in [4], [7]. The details of the technique and experimental results are provided in Section 3. We discuss future work and present our conclusions in Section 4

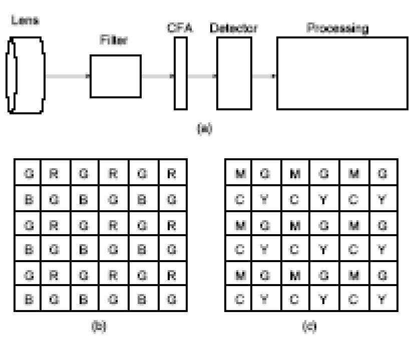

Fig. 1. (a) The more important stages of a camera pipeline are shown [6]. (b) CFA pattern using RGB values. (c) CFA pattern using YMCA values.

2. Similarity Measures

Since each bit plane is also a binary image, we start by considering similarity measures between two binary images, that is, between quantal bit planes of images. The binary similarity measures were extensively studied in [7]. We discuss here two of them for illustrative purposes.

Let's consider the 5-point stencil function and apply it in the bit plane *b*:

$$a_c^n(k,b) = \begin{cases} 1 & if \quad x_c = 0 \quad and \quad x_n = 0 \\ 2 & if \quad x_c = 0 \quad and \quad x_n = 1 \\ 3 & if \quad x_c = 1 \quad and \quad x_n = 0 \\ 4 & if \quad x_c = 1 \quad and \quad x_n = 1 \end{cases}$$

where the four arguments are defined as follows: The subscript *c* defines some central pixel and the superscript *n* denotes one of the possible four neighbor pixels. We sum $a_c^n(k,b)$ over its four neighbors (i.e. n runs over East, West, South and North neighbors) as well as over all the pixels (i.e., c runs over the MxN pixels). After the summations the sub- and superscripts can be omitted. The first argument *k indicates* one of the four agreement scores {1,2,3,4} and the second argument indicates the bit plane in which this computation is being done. Obviously $\{a(k,b), k = 1,..,4\}$ variables, that is, the agreement scores the central pixel – neighbor pixel transition types in a particular bit plane. Normalizing the agreement scores we obtain the score pdf's: $p_k^b = \alpha(k,b) / \sum_k \alpha(k,b)$. Based on these normalized *four-bin* histograms, we define binary Kullback Leibler distance as:

$$m_1 = -\sum_{n=1}^{4} p_n^7 \log \frac{p_n^7}{p_n^8} .$$

The second measure m_2 is somewhat different in that we use the neighborhood-weighting mask proposed by Ojala [8]. The histogram of the Ojala moments for different cameras is plotted in Fig. 4. For each binary image we obtain a 512-bin histogram based on the weighted neighborhood, where the score is given by: $S = \sum_{i=0}^{7} x_i 2^i$ by weighting the eight directional neighbors as shown in Fig. 2. Defining S_n^7 the count of the n^{th} histogram bin in the 7^{th} bit plane and S_n^8 the corresponding one in the 8^{th} plane, after normalizing these 512-*bin* histograms, we can define absolute histogram difference as:

$$m_2 = \sum_{n=0}^{511} \left| S_n^7 - S_n^8 \right| .$$

1	2	4
128	256	8
64	32	16

Fig. 2: The weighting pattern of the neighbors in the computation of Ojala score. For example, the score becomes S=2+ 4+8=14 in the example where E, N, NE bits are 1 and all other bits are 0.

The image quality measures were extensively studied in [4]. We discuss here one of them for illustrative purposes. The Czenakowski distance gives a metric useful to compare vectors with strictly non-negative components, as in the case of color images:

$$m_3 = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} \left(1 - \frac{2\sum_{k=1}^{3} \min\left(C_k(i,j), \hat{C}_k(i,j)\right)}{\sum_{k=1}^{3} \left(C_k(i,j) + \hat{C}_k(i,j)\right)} \right),$$

where $C_k(i,j)$ is $(i,j)^{\text{th}}$ pixel of the k^{th} band of a color image and $\hat{C}_k$ is the denoised version of the corresponding k^{th} band color image. Denoising is employed on the image to obtain a reference image to calculate the metric.

In Fig. 3 we give the scatter plot of three cell-phone cameras for three features, namely, m1, m2, m3 measures. As can be seen the used features cluster well enough for a successful classification.

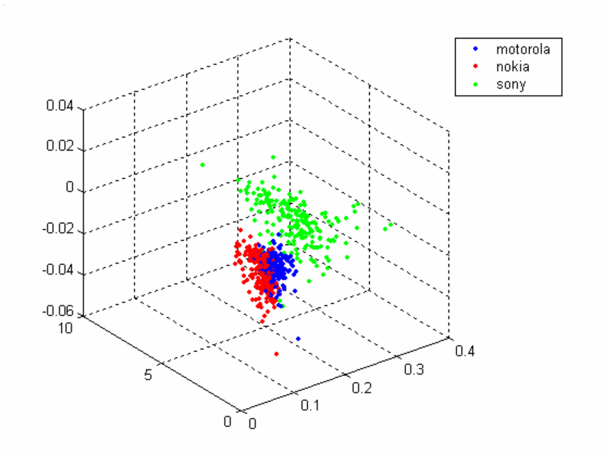

Fig. 3. The scatter plot of three cell-phone cameras for three similarity measures.

Overall we considered 108 BSM features and 10 IQM features. The BSM features consisted of the 7-8, 6-7, 5-6, 4-5, 3-4 bit planes of the red channel and of the 5th bit plane of the remaining blue and green channels. These features were then selected using the Sequential Forward Feature Selection (SFFS) algorithm.

3. Experimental Results

We have considered nine makes and/or brands of cell phone cameras, as detailed in Table 1:

Table 1. Types of cameras tested and their display characteristics**.**

Acronym	**Make/Brand**	**Colors**	**Resolution (pixel)**
M1	MotorolaV3	260K	176 x 220
M2	MotorolaV500	65K	176 x 220
N1	Nokia5140	65,536	128 x 128
N2	Nokia6230	65,536	208 x 208
N3	Nokia6600	65,536	176 x 208
N4	Nokia7270	65,536	128 x 160
S1	SonyK700	65,536	176 x 220
S2	SonyK750	262,144	176 x 220
L1	LG5600	65K	128 x 160

We collected 200 images from each one of them with maximum resolution, size of 640X480 pixels, at day light and auto-focus mode. Half of the 1800 images are used for training and the designed classifier is tested with the other unseen half set of images. The images were typical shots varying from nature scenes to close-ups of people. We experimented with the KNN classifier (K=5) as well SVM algorithm of RBF variety (γ=2.0,ε=0.001,C=8.0,cachesize=40). Sample images of outdoors scenes in the image database are shown in Fig. 4.

In a first exploratory experiment, we grouped cameras in three-tuples and ran SFFS algorithm for each combination for the best selection of features. Sample confusion tables from these three-cameras groups of are given below (best, middle, worst case tables given):

Table 2a. Confusion matrix for the SonyK700, MotorolaV3, Nokia6230 group. SFFS resulted in 5 features. Overall performance = 98.7%.

	SonyK700	MotorolaV3	Nokia6230
SonyK700	100	0	0
MotorolaV3	0	100	0
Nokia6230	0	4	96

Table 2b. Confusion matrix for the SonyK750, MotorolaV3, Nokia6600 group. SFFS resulted in 3 features. Overall performance = 90.0%.

	SonyK750	MotorolaV3	Nokia6600
SonyK750	92	8	0
MotorolaV3	8	87	5
Nokia6600	1	8	91

Table 2c. Confusion matrix for the SonyK750, MotorolaV3, Nokia7270 group. SFFS resulted in 7 features. Overall performance = 81.3%.

	SonyK750	MotorolaV3	Nokia7270
SonyK750	71	6	23
MotorolaV3	1	97	2
Nokia7270	18	6	76

The average performance of all 16 different three-tuple experiments was 93.4%.

In a more challenging experiment we tried to classify the pool of nine camera types. Again the SFFS algorithm was run for the ensemble of camera classes with the SVM classifier. The results are given in Table 3. The overall accuracy fell to 62.3, which is still considerably better than the random guess results of 11.1%.

Table 3: Recognition performance of camera brands in pairwise comparisons.

	S1	S2	M1	M2	N1	N2	N3	N4	L1
S1	92	4	0	0	0	0	0	3	0
S1	5	63	1	0	4	0	0	32	3
M1	0	3	60	11	3	1	3	3	5
M2	0	0	22	67	4	9	5	0	5
N1	0	6	2	3	57	3	6	7	18
N2	0	1	2	6	3	68	22	0	0
N3	0	1	7	10	1	18	62	1	1
N4	3	16	2	0	7	0	0	36	12
L1	0	6	4	3	21	1	2	18	56

4. Conclusions and Future Work

In this work, we proposed to identify the originating cell-phone of a digital image based on the combination of binary similarity measures, computed across contiguous bit planes of an image, and of image quality measures. The apposite features were selected based on the SFFS algorithm and an SVM classifier was trained for classification.

The performance of small groups of camera makes/brands is very satisfactory. The classification performance in groups of two is close to 100% and in groups of three it scores around 93%. For the larger group of 9 cameras, a classifier attains 62.3% correct classification.

This study can be advanced along several avenues. A larger set of features, including the so-called Higher Order Statistical measures as in [5,9]. We have just considered the red channel in this work. The perturbation of the correlation structure across color channels as well as within the blue and green channels remains to be investigated. Fusion techniques, especially the sum or product rule variety, has been shown to be very effective in improving classifier performance [10]. Another aspect of our study is to measure the drop in the performance when the technique is applied on medium- and low-quality (better compressed) images. Another direction will be to utilize binary similarities on all color channels and use more sophisticated classifiers like support vector machines.

Acknowledgement: This work has been supported in part by TÜBİTAK under the research grant number 104E056.

5. References

[1] Special Issue on Data Hiding, *IEEE Transactions on Signal Processing*, Vol. 41, No. 6, 2003.

[2] M. Kharrazi, H. T. Sencar, N. Memon, "Digital Camera Model Identification," *Proc. of ICIP*, 2004.

[3] S. Bayram, H. T. Sencar, N. Memon, I. Avcibas, "Source Camera Identification Based on CFA Interpolation", ICIP 2005.

[4] I. Avcibas, N. Memon and B. Sankur, "Steganalysis using Image Quality Metrics," *IEEE Transactions on Image Processing*, Jan. 2003.

[5] S. Lyu and H. Farid, "Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines," *Proc. of Information Hiding Workshop*, 2002.

[6] J. Adams, K. Parulski and K. Sapulding, "Color processing in digital cameras," *IEEE Micro*, Vol. 18, No.6, Jun. 1998.

[7] I. Avcibas, M. Kharrazi, N. Memon, B. Sankur, "Image Steganalysis with Binary Similarity Measures", *Journal of Applied Signal Processing*, in press, 2005.

[8] T. Ojala, M. Pietikainen, D. Harwood, A Comparative Study of Texture Measures with Classification Based on Feature distributions, Pattern Recognition, vol. 29, pp. 51-59.

[9] Popescu and H. Farid ,"Exposing Digital Forgeries by Detecting Traces Of Re-sampling", IEEE Transactions on Signal Processing ,2004.

[10] J. Kittler, F.M. Alkoot, Sum Versus Vote Fusion in Multiple Classifier Systems, IEEE Trans. Patttern Recognition and Machine Intelligence**,** 25, 110–115, 2003.

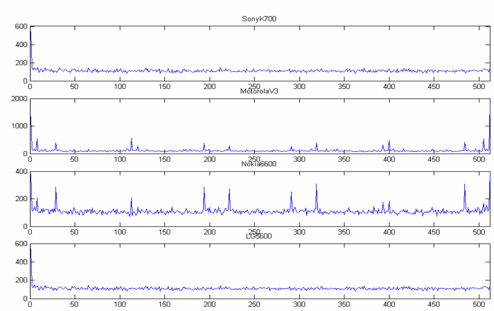

Fig. 4: Plot of Ojala histograms for different cameras.