Fingerprint Recognition of Image Forensics Using Random Projections

Karthika A [1], Prema S [2], Divya E [3], Chitra V [4], Ramya E [5]
Assistant Professor of ECE, SNS College of Technology, Coimbatore, Tamil Nadu, India.

Abstract – Sensor imperfections in the form of photoresponsenonuniformity (PRNU) patterns are a well-established fingerprinting technique to link pictures to the camera sensors that acquired them. The fingerprint image has to be compressed and encrypted by the authentication and matching purposes in the forensic tasks by using the PRNU values to be reconstructed the original image quality.The digital image processing is to be used for compressed camera fingerprint matching via random projections. Fingerprints are one of those irregular twists of nature. The fingerprints are to be used authentication and identification processes in forensic tasks such as detection of digital forgeries. Forensic tasks can to be performed in device identification problem, device linking problem, fingerprint matching problem. For random projections the compression technique is to be required with no information loss and to be measured by PRNUvalues.

Index Terms – Random projections, PRNU, image forensics.

1. INTRODUCTION

IMAGING sensor imperfections can be considered as a unique fingerprint identifying a specific acquisition device, enabling various important forensic tasks, such as device identification, device linking, recovery of processing history, detection of digital forgeries [1]. The most common camera fingerprint is the photo-response nonuniformity (PRNU) of the digital imaging sensor [2]. The PRNU is due to slight variations in the properties of individual pixels, which producea noise-like, yet deterministic pattern affecting every image taken bysensor. Several works demonstrate that the PRNU is a robust fingerprint, usually surviving processing with the same size as the imaging sensor is due to thewide availability of sensors counting tens of millions of pixels, a realistic database of a few thousand sensors will require to store more than 1010 individual pixel values in uncompressed format. In addition, the complexity of looking for a particular fingerprint in a large database is also very high, typically requiring the computation of a correlation with each fingerprint in the database. The issue of compression of PRNU patterns does not arise when the results of device identification have to be used as evidence in the court of law, because that case typically involves small databases and requires the highest accuracy. Instead, large scale problems, such as image classification, clustering or image retrieval problems based on camera identities, involve a huge number of PRNU patterns. Hence, these problems call for techniques to efficiently store and query such databases. Another problem with PRNU fingerprints is that the test image should be geometrically aligned with the fingerprint in the database. A possible solution is to provide several versions of the same fingerprint with different scale and/or cropping factors [5], however at the cost of managing an even larger database.

The fingerprint image has to be compressed and encrypted by the authentication and matching purposes in the forensic tasks by using the PSNR values to be reconstructed the original image quality. The digital image processing is to be used for compressed camera fingerprint matching via random projections. Fingerprints are one of those irregular twists of nature. The fingerprints are to be used authentication and identification processes in forensic tasks such as detection of digital forgeries. Forensic tasks can to be performed in device identification problem, device linking problem, fingerprint matching problem. For random projections the compression technique is to be required with no information loss and to be measured by PSNR values.

Recently, several authors [6] started to address the problems related with the management of a large database of camera fingerprints. In [7] and [8], the authors propose a so-called *fingerprint digest*, which works by keeping only a fixed number of the largest fingerprint values and their positions, so that the resulting database is independent of the sensor resolution. An improved search strategy based on fingerprint digest is proposed in [9] and [10]. Fingerprint digests can also be used to ease fingerprint registration in case of geometrically distorted images, as shown in [11].

An alternative solution is to represent sensor fingerprints in binary-quantized form [12]: even though the size of binary fingerprints scales with sensor resolution, binarization can considerably speed-up the fingerprint matching process.In the case of PRNU fingerprints, it is easy to show that preserving the distance between two fingerprints is equivalent to preserving the angle between them. Since PRNU fingerprints of different sensors are known to be highly uncorrelated, and thus to form wide angles, we can expect that also the angles between compressed fingerprints obtained by random projections will be wide. As a consequence, in this paper we adapt the standard correlation detector [1] to solve fingerprint matching and camera identification problems in the compressed domain.

As to practical issues, the complexity of randomly projecting a large fingerprint is greatly reduced by employing partial circulant matrices [15], which are known to be almost as good as fully random matrices. Moreover, inspired both by the work of [12] and by recent results in compressed sensing literature [16], we propose a binary version of the compressed fingerprint that further reduces storage and computational requirements. In Section II, we provide notations and definitions and we briefly review forensic tasks based on PRNU and random projections. The proposed compressive PRNU forensic systems are described in Section III, while theoretical performance is analyzed in Section IV. Extensive numerical results on different datasets are presented and discussed in Section V. Finally, in Section VI we draw some conclusions.

2. BACKGROUND

1. Notation and Definitions

This is a real-world problem: the Federal Bureau of Investigation (FBI) maintains a large database of fingerprints. The FBI uses eight bits per pixel to define the shade of gray and stores 500 pixels per inch, which works out to about 700 000 pixels and 0.7 megabytes per finger to store finger prints in electronic form. By turning to wavelets, the FBI has achieved a 15:1 compression ratio. In this application, wavelet compression is better than the more traditional JPEG compression, as it avoids small square artifacts and is particularly well suited to detect discontinuities (lines) in the fingerprint. Note that the international standard JPEG 2000 includes the wavelets as a part of the compression and quantization process. This points out the present strength of the wavelets.

We denote (column-) vectors and matrices by lowercase and uppercase boldface characters, respectively. The l-th element of column vector $\mathbf{v}$ is $v\ell$. The i-th column of the matrix $\mathbf{A}$ is $\boldsymbol{a_i}$. The notation $\mathbf{A}\cdot\mathbf{B}$ denotes the elementwise product between matrices $\mathbf{A}$ and $\mathbf{B}$, while $\mathbf{A}/\mathbf{B}$ denotes element wise division.

The notation $\langle(a,b)\rangle$denotes the scalar product between vectors $\mathbf{a}$ and $\mathbf{b}$, and $\|a\|_2 = \sqrt{(a,a)}$.

The notation dH($\mathbf{a}$, $\mathbf{b}$)denotes the Hamming distance between $\mathbf{a}, \mathbf{b} \in \{0,1\}m$, where dH($\mathbf{a}$, $\mathbf{b}$)$=\frac{1}{m}\sum_{i=1}^{m} a_i \oplus b_i$ and $\oplus$ denotes the XOR operator.

The notation $\mathbf{a} \sim \mathcal{N}(\mu, \Sigma)$means that the random vector $\mathbf{a}$ is Gaussian distributed, its mean is $\boldsymbol{\mu}$, and its covariance matrix is Σ .

2. PRNU Forensics

PRNU [1], [2] of imaging sensors is a property unique to each sensor array due to the different ability of each individual optical sensor to convert photons to electrons. This difference is mainly caused by impurities in silicon wafers and its effect is a noise pattern affecting every image taken by that specific sensor. Hence, the PRNU can be thought of as a spread–spectrum *fingerprint* of the sensor used to take a specific picture or a set of pictures. The PRNU is multiplicative, *i.e.*, if an imaging sensor is illuminated ideally with a uniform intensity $\mathbf{i}$,[1]neglecting other sources of noise, the output of the sensor will be $\mathbf{o} = \mathbf{i} + \mathbf{I} \cdot \mathbf{k}$ where $\mathbf{k}$ represents the matrix characterizing the PRNU values.

k exhibits the following properties. It has the same pixel size as the sensor, and carries enough information to make it unique to each sensor. It is universal in the sense that every optical sensor exhibits PRNU. It is present in each picture taken by a sensor except from completely dark ones (due to its multiplicative nature). It is stable under different environmental conditions and is robust to several signal processing operations.

The PRNU characterizing one sensor can be extracted from a set of images (tipically, 20 to 50 smooth images are enough). The procedure to extract the fingerprint $\mathbf{k}$ of a sensor from a set of pictures depends on the model used to characterize the optical sensor. Denoting with $\mathbf{i}$the incident light intensity, the sensor output $\mathbf{o}$ can be modelled as expressed as given below,

Where as gamma correlation is,

$$\mathbf{o} = g^{\gamma} \cdot [(1 + \mathbf{k}) \cdot \mathbf{i} + \mathbf{e}]^{\gamma} + \mathbf{q}, \qquad (1)$$

where$g\gamma$is the gamma correction (g is different for each color channel and γ is usually close to 0.45), $\mathbf{e}$ accounts for other noise sources internal to the sensor while $\mathbf{q}$ models external noise (*e.g.* quantization). The goal is to extract $\mathbf{k}$, so, after keeping the first order term in the Taylor expansion of [*(*1 + $\mathbf{k}$*)* · $\mathbf{i}$+ $\mathbf{e}$]$^{\gamma}$, the output image can be factorized as

$$\mathbf{o} = \mathbf{o}^{\mathrm{id}} + \mathbf{o}^{\mathrm{id}} \cdot \mathbf{k} + \tilde{\mathbf{e}}, \qquad (2)$$

Where $\mathbf{o}^{\mathrm{id}} = (g\mathbf{i})^{\gamma}$ is the ideal sensor output, $\mathbf{o}$id· $\mathbf{k}$ is the PRNU term and ollects other sources of noise.Assuming to be able to obtain through proper $\tilde{\mathbf{e}} = \gamma\,\mathbf{o}^{\mathrm{id}} \cdot \mathbf{e}/\mathbf{i} + \mathbf{q}$ filtering a denoised version of $\mathbf{o}$, referred to as $\mathbf{o}$dn, then this can be used as an approximation of the ideal sensor output and subtracted from each side of (2) to obtain the so-called *noise residual*, which can be modeled as:

$$\mathbf{w} = \mathbf{o} - \mathbf{o}^{\mathrm{dn}} = \mathbf{o} \cdot \mathbf{k} + \tilde{\mathbf{q}}, \qquad (3)$$

where˜$\mathbf{q}$ accounts for ˜$\mathbf{e}$ and for the non-idealities of the model [1]. Suppose now that a certain number $C \geq 1$ of images is available. Considering the pixels of the noise term ˜$\mathbf{q}$ as zero–mean Gaussian noise with variance σ2 and independent from the signal $\mathbf{o} \cdot \mathbf{k}$, for each image l,l= 1, . . . ,C, it can be written

$$\mathbf{w}^{(\ell)}/\mathbf{o}^{(\ell)} = \mathbf{k} + \tilde{\mathbf{q}}/\mathbf{o}^{(\ell)}, \text{ where } \mathbf{w}^{(\ell)} = \mathbf{o}^{(\ell)} - \mathbf{o}^{(\ell)\mathsf{dn}}. \quad (4)$$

Under the above assumptions, the log–likelihood of $\mathbf{w}^{(l)}/\mathbf{o}^{(l)}n$given$\mathbf{k}$ satisfies

$$L(\mathbf{k}) = -\frac{C}{2}\sum_{\ell=1}^{C}\log\left(2\pi\sigma^2/(\mathbf{o}^{(\ell)})^2\right) \quad (5)$$

$$+\sum_{\ell=1}^{C}\left(\mathbf{w}^{(\ell)}/\mathbf{o}^{(\ell)} - \mathbf{k}\right)^2 \Big/ \left(2\sigma^2/(\mathbf{o}^{(\ell)})^2\right) \quad (6)$$

from which the maximum likelihood estimate ˆ**k** can be obtained as

$$\hat{\mathbf{k}} = \sum_{\ell=1}^{C}\left(\mathbf{w}^{(\ell)}\cdot\mathbf{o}^{(\ell)}\right) \Big/ \sum_{\ell=1}^{C}(\mathbf{o}^{(\ell)})^2 \quad (7)$$

From the Cramer–Rao bound, the variance of the estimator can be estimated as from which we can notice that good photos for fingerprint evaluation are photos with high luminance (but not saturated) and smooth content (which lowers $\sigma 2$). To improve further the quality of the estimation, artifacts shared among cameras of the same brand or model can be removed by subtracting row and column averages. In the case of color images, the estimation must be performed separately on each color channel, *i.e.*, we must obtain ˆ**k** R, ˆ**k** G and ˆ**k**B. After that, a "global" grayscale PRNU fingerprint will be obtained applying the usual RGB–to–gray conversion.

Several forensic tasks can be performed using the aforementioned model for camera sensors.

$$\sigma^2_{\hat{\mathbf{k}}} = \sigma^2 \Big/ \sum_{\ell=1}^{C}(\mathbf{o}^{(\ell)})^2, \quad (8)$$

$$\hat{\mathbf{k}} = 0.3\hat{\mathbf{k}}_{\mathsf{R}} + 0.6\hat{\mathbf{k}}_{\mathsf{G}} + 0.1\hat{\mathbf{k}}_{\mathsf{B}}. \quad (9)$$

• The *device identification* problem [3] (also known in the biometrics field as *verification*) tests whether a given picture was taken by a specific device. An estimate of the fingerprint of the device has been extracted in advance from a set of training pictures and stored in a database. The noise residual or a single-image fingerprint estimate is extracted from the query image and correlated with the fingerprint in the database. The original detector presented in [4] correlates the noise residual of the query image with the database fingerprint modulated by the query image intensity, denoted as corr*(***w**, **o** · ˆ**k** *)*.* • The *device linking* problem [17] is presented with two images and must determine whether they have been acquired by the same device. The noise residuals of he two photos are correlated, namely corr*(***w**1,**w**2*)*. We ill not discuss this usage case in the remainder of the paper.

• The *fingerprint matching* problem (also known in the\ biometrics field as *identification*) is presented with a database of fingerprint estimates and a set of pictures acquired by the same camera, which can be used to extract a fingerprint estimate. The goal is determine which device in the database (if present) has acquired the given pictures. Essentially, for all fingerprints, and if one fingerpring yields a correlation that is large enough, it is declared to be correct.

3. Random Projections

As will be explained in detail in Section III, PRNU databases can rapidly grow in size. For this reason, a method to "compress" them is required, with slight or ideally no information loss. One possible option is represented by *Random Projections* (RP), a low–complexity and yet powerful method for dimensionality reduction. The idea of RP is to project the original *n*–dimensional data to an *m*–dimensional subspace, with $m < n$, using a random matrix $\_ \in \mathrm{R}m\times n$. Hence, a collection of *N n* dimensional data $\mathbf{D} \in \mathrm{R}n\times N$is reduced to an *m*–dimensional subspace $\mathbf{A} \in \mathrm{R}m\times N$by

$$\mathbf{A} = \mathbf{\Phi D}. \quad (10)$$

The key property behind RP is the Johnson–Lindenstrauss lemma [13], concerning low–distortion embeddings of points from high–dimensional into low-dimensional Euclideanspace. The lemma states that a small set of points in a high-dimensional space can be embedded into a space of much lower dimension in such a way that distances between the points are nearly preserved.

Lemma 1 (Johnson–Lindenstrauss): Let $\varepsilon \in (0, 1)$*. For every set Q of* $|Q|$ *points in* R*n, if m is a positive integer such that*

$$(1-\varepsilon)\|\mathbf{u}-\mathbf{v}\|_2^2 \le \|f(\mathbf{u}) - f(\mathbf{v})\|_2^2 \le (1+\varepsilon)\|\mathbf{u}-\mathbf{v}\|_2^2$$

It has been shown that *f* can be taken as a linear mapping represented by a random matrix $\_ \in \mathrm{R}m\times n$, whose entries are randomly drawn from certain probability distributions [14], like the Gaussian or Rademacherdistributions.The properties of RP are strictly related to the field of Compressed Sensing [18], [19], and in particular to the Restricted Isometry Property (RIP) of the sensing matrices [20]. In particular, in [20] it is shown that sensing matrices whose elements follow the aforementioned distributions respect the RIP as well as the JL lemma. One can think of the RIP as a JL lemma specific for sparse vectors. In fact, a matrix R$m\times n$is said to satisfy the RIP with constant $\delta\kappa$if there exists a constant $\delta\kappa$such that

$$(1-\delta_\kappa)\|\mathbf{u}\|_2^2 \le \|\mathbf{\Phi}_\kappa\mathbf{u}\|_2^2 \le (1+\delta_\kappa)\|\mathbf{u}\|_2^2, \quad (11)$$

whereis every possiblesubmatrix obtained by keeping columns of and the techniques presented in this paper bear some similarity with techniques used in Locality Sensitive Hashing

(LSH) [21], [22]. Unlike standard hashing techniques, where the aim of the hashing function is to avoid collisions of hashes of different objects, LSH is a hashing technique for large databases using hashing functions whose aim is to maximize the probability of collision for objects close to each other rather than far apart. Then, the gap between the probability of collision of the hashes of similar objects and the probability of collision of the hashes of different objects is further amplified by concatenating several hashing functions. This allows one to perform, for example, a nearest–neighbor research in a large database using the hash of the query point retrieving elements stored in buckets containing that point. Several LSH families have been discovered in literature, each of them allowing a random choice of hashing functions. Among them, one, dubbed *arccos*, bears some similarity with 1-bit Compressed Sensing [16] and with the techniquesexplained in this paper. In words, the hashing function consists in the sign of the random projections, obtained with a sensing matrix with independent and identically distributed entries.

$$\rho(\hat{\mathbf{k}}, \mathbf{d}_i) = \frac{\langle \hat{\mathbf{k}}, \mathbf{d}_i \rangle}{\|\hat{\mathbf{k}}\|_2 \|\mathbf{d}_i\|_2}, \quad i = 1, \ldots, N \tag{12}$$

We propose to compress the database and test fingerprint representing them through a small number of random projections. This operation can be seen as the product times an $m \times n$ sensing matrix ϕ

$$\mathbf{A} = \mathbf{\Phi D} \tag{13}$$

$$\mathbf{y} = \mathbf{\Phi}\hat{\mathbf{k}} \tag{14}$$

Algorithm 1 Dictionary Creation

Require: $\mathbf{D}$, ϕ
Ensure: $\mathbf{A}$
  for $i = 1, \ldots, N$ **do**
    $\mathbf{a}_i \leftarrow \text{IFFT}\,[\text{FFT}[\mathbf{d}_i] \cdot \text{FFT}[\phi]]$
    $\mathbf{a}_i \leftarrow$ first m entries of $\mathbf{a}_i$
  end for

Algorithm 2 Matching

Require: $\mathbf{A}$, $\hat{\mathbf{k}}$, ϕ
  $\mathbf{y} \leftarrow \text{IFFT}\left[\text{FFT}[\hat{\mathbf{k}}] \cdot \text{FFT}[\phi]\right]$
  $\mathbf{y} \leftarrow$ first m entries of $\mathbf{y}$
  for $i = 1, \ldots, N$ **do**
    if $\rho(\mathbf{y}, \mathbf{a}_i) > \tau$ **then**
      Declare a match
    end if
  end for

Random projections can effectively reduce the dimension of the space the fingerprints live in thanks to the fact that they approximately preserve the geometry of the point cloud composed of the fingerprints. Since random projections approximately preserve the angle between any two fingerprints and since this angle is wide thanks to their incoherent nature, we can expect a compressive system to exhibit robust performance, while dramatically reducing the problem size. The system has to store the compressed dictionary **A** and a way to generate the compressed fingerprint whenever a test pattern is presented, using the same _ (typically the seed of a pseudorandom number generator is stored).

Since the preservation of the angles is the main interest for the matching problem, we will also consider the case of binary random measurements obtained as:

$$\mathbf{A} = \text{sign}(\mathbf{\Phi D}) \tag{15}$$

In the case of binary measurements the correlation coefficient is replaced by the Hamming distance as test metric.

$$d_H(\mathbf{y}, \mathbf{a}_i), \quad i = 1, \ldots, N \tag{16}$$

In Sec. IV we discuss how the Hamming distance tendsto be concentrated around $dS(\hat{\mathbf{k}}, \mathbf{di}) = \pi^{-1}$arccosbeing arcosthe angle between two uncompressed fingerprints. The higher the correlation between fingerprints, the narrower the angle between them. Hence, the angle between two matching fingerprints is typically narrower than the angle between non-matching fingerprints. This is reflected on the binary random projections, where the Hamming distance between matching fingerprints is typically smaller than that between non-matching fingerprints. Binaryrandom projections allow to compress significantly, while the performance degradation is limited. As we will show in Sec. V, the degradation due to binarization is small but it allows to obtain a significant gain in terms of space. Moreover, computing the Hamming distance is a very fast and efficient operation. Binarization of the fingerprints was considered by Bayram*et al.* [12] as an effective method to reduce storage requirements. We go one step further by showing that binarization of random projections is effective as well, while further reducing the storage and computational requirements and providing additional flexibility by modulating the number of random measurements. Binarization of the fingerprints themselves can be seen as a special case of the presented framework, in which the sensing matrix is the identity.

A.Camera Identification

The camera identification problem is conceptually very similar to the fingerprint matching scenario. The main difference is that a single test image is available instead of a set of them. Chen *et al.* [4] showed that the optimal detector for this problem correlates the noise residual of the image with a modulated version of the fingerprint stored in the database, where the modulating term is the test image. Extending this detector to the compressed domain is not possible because of the elementwise product between test image and the fingerprint in the database. Instead, we investigate the performance of two

simplified detectors that can be readily mapped to the compressed domain. The first simplified detector correlates the noise residual **w** of the test image with the fingerprint stored in the database. Essentially this system eliminates the modulating effect of the test image, thus it will be sub-optimal unless the test image is a constant pattern. It is sufficient to apply the sensing matrix to both noise residual and fingerprint to translate this detector to the compressed domain.

$$\rho(\mathbf{w}, \mathbf{d}_i) \longmapsto \rho(\mathbf{\Phi w}, \mathbf{\Phi d}_i) \tag{17}$$

$$\rho(\hat{\mathbf{k}}, \mathbf{d}_i) \longmapsto \rho(\mathbf{\Phi}\hat{\mathbf{k}}, \mathbf{\Phi d}_i) \tag{18}$$

B.Detection Metrics

The matching problem is concerned with finding the column of the dictionary that best matches a test compressed pattern. The test compressed fingerprint undergoes a binary hypothesism test for each column of the compressed dictionary. The two hypotheses are defined as:

$$\hat{\mathbf{k}} = \mathbf{d}_i + \mathbf{z},$$

H0: the compressed test fingerprint and the reference are not from the same camera

H1: the compressed test fingerprint and the reference are from the same camera

• **False Alarm**: the null hypothesis was incorrectly rejected.

• **Detection**: the null hypothesis was correctly rejected.

• **True Detection**: the null hypothesis was rejected only for the correct camera.

False acceptance corresponds to the case in which all the columns of the dictionary are tested, and at least one column containing the compressed fingerprint of a different camera with respect to the compressed fingerprint under test is declared as a match. On the other hand, *true detection* occurs when all the columns of the dictionary are tested, and a match is declared *only* for the column corresponding to the same camera of the compressed fingerprint under test.

3. EXPERIMENTAL RESULTS

We tested the performance of the compressed system under various conditions.We used two datasets of actual photographs to obtain the receiver operating characteristic (ROC) of the system under different scenarios. We constructed the first dataset (PoliTO database) by shooting photographs of walls with 8 different cameras. The uniform subject and the control over light conditions make those photos nearly ideal for the extraction of camera fingerprints. The second database is the publicly available Dresden image database [29]. Each database is constructed from a number of training photos, while *T* additional photos are used for testing. Extraction of the camera fingerprints is performed using the Camera Fingerprint toolbox [30], [31].

Referring to the events described in Sec. III-D and the probabilites defined in Section IV, we estimate the detection probability *PD(i)*, averaged over all the cameras *i*= 1, . . . , *N*, with the *true positive rate* as

$$\text{True Positive Rate} = \frac{\#\text{ of detections}}{T},$$

while the false alarm probability *PFA(i, j)*, averaged over all the cameras *i*= 1, . . . , *N* and *j* _= *i*, is estimated with the *false positive rate* as

$$\text{False Positive Rate} = \frac{\#\text{ of false alarms}}{(N-1)T}.$$

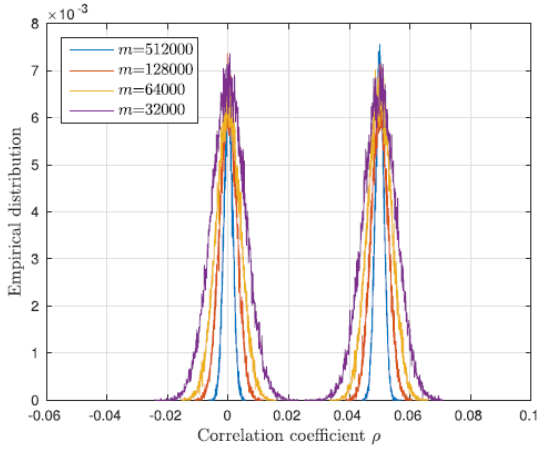
Next, we estimate the true detection probability *PT (i)*, averaged over all the cameras *i*= 1, . . . ,*N*, with the *true detection rate*, as

$$\text{True Detection Rate} = \frac{\#\text{ of true detections}}{T},$$

while the false acceptance probability *PF(i)*, averaged over all the cameras *i*= 1, . . . , *N*, is estimated with the *false acceptance rate* as

$$\text{False Acceptance Rate} = \frac{\#\text{ of false acceptances}}{T}.$$

A second ROC plots the True Detection Rate vs. the False Acceptance Rate. The ideal curve is the top-left–bottom-right diagonal.



Empirical distribution of correlation between matching andnon-matching compressed fingerprints.

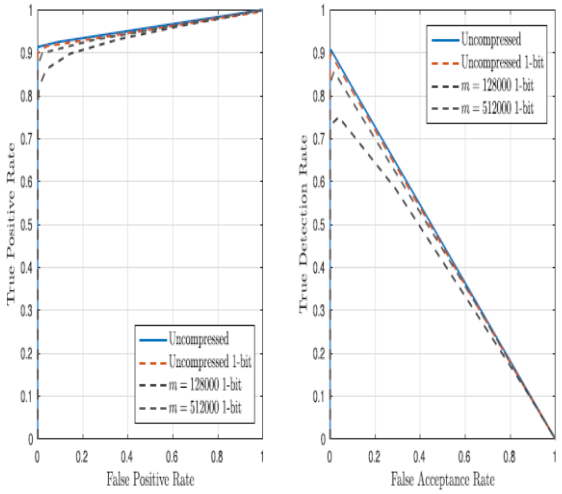
A.PoliTO Database

The PoliTO database is composed of pictures from 8 different consumer cameras. The pictures are defocused photos of walls under good illumination conditions. Each camera has at least 100 photos, all in landscape format, shot at the full resolution and maximum quality JPEG compression. We use60 photos of each camera to extract the ground truth fingerprint to be stored in the database, while the remaining ones are used for testing purposes.

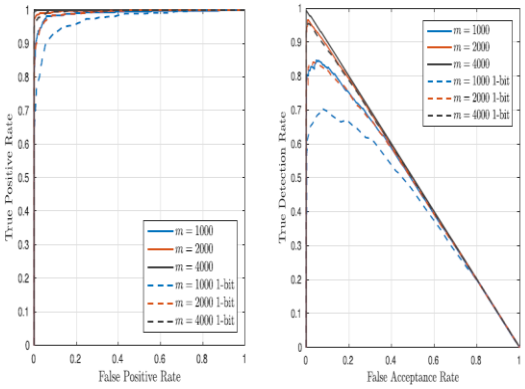
Each ROC curve is obtained by sweeping the threshold parameter τ . Test images are presented to the system one at a time, the noise residual is extracted and then compressed using the same sensing matrix used to compress the database. ROCs parametrized by the number of measurements. It can be noticed that a very small number of random measurements is enough to get almost indistinguishable performance from a perfect detector, while saving a considerable amount of storage space. Table I shows some actual figures for the space needed to store the dictionary of fingerprints on disk (without any additional form of lossless compression, which is anyway highly ineffective due to the high entropy of the PRNU and of the random measurements).

B.Dresden Database

The database assembled in [29] is composed of both flatfield images and scenes from indoor and outdoor environments. We selected 53 cameras having both flatfield and natural photos. The database is created from the flatfield images in order to have high quality fingerprints, while the test images are taken from the natural scenes. The natural photos present varying amounts of details and illumination conditions,thus making this dataset much more challenging than the PoliTO database. All photos are registered to the same sensor orientation.



ROC curves for the Dresden database. Binarization of the fingerprints is compared against binarized random projections.



ROC curves for the PoliTO database. Binarization of the fingerprints is compared against binarized random projections

4. OUTPUT FOR THE FINGERPRINT IMAGES

Fig a: Input image of the given fingerprint.en input

The output for the fingerprint images with compressed output it contains for no.of retained energy and no.ofzerosvalues for given input It has 96.0% for its retained values.No.of.zeros had 86.05%.

Fig b: Thinned image for given fingerprint

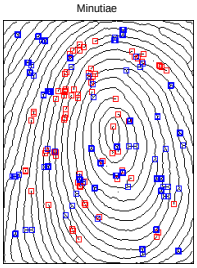


Fig c: Minutiae values for the given fingerprint.

Fig d: compressed image for a given fingerprint.

We compare the ROC obtained on the PoliTO database for Gaussian and circulant matrices, having the first row drawn as Gaussian i.i.d.. Experimental results shown in Fig. 8 confirm that circulant constructions perform very closely to the fully random ones, though they provide enormous advantages in terms of memory and computational requirements.

5. CONCLUSIONS

This paper proposed a technique to address the issues of storage and matching complexity in camera fingerprint databases, by using random projections. Motivated by the incoherent nature of fingerprints based on PRNU patterns of camera sensors, we showed that random projections can effectively preserve the geometry of the database and significantly reduce the dimension of the problem with small penalties. We characterized the usage of real-valued and binary random measurements from a theoretical point of view in terms of the detection and false alarm probabilities.

Experimental tests have confirmed the validity of the proposed method on two databases of actual photographs. Practical issues such as the complexity of calculating random projections are of significant importance when dealing with million-pixel images, but we solved them by using circulant sensing matrices. The use of random projections forcompression of camera fingerprints paves the way to many interesting applications involving maintaining large databases of fingerprints or applications requiring transmission of fingerprints over bandlimited channels. From this perspective, random projections are significantly better than the other existing methods discussed in this paper because they can provide higher compression ratios and improved scalability, *i.e.*, a fine-grained control over the compression/performance tradeoff by modulating the number of projections according to the specific needs, and an embedded representation where acompressed version of the fingerprint already embeds versions at higher compression ratios (fewer measurements used).

REFERENCES

[1] Hong cao and alex C. Kot , "accurate detection of demosaicing regularity for digital image forensics", *IEEE trans.* Information forensics and security., Vol. 4, no. 4,pp.899-910, dec. 2009.

[2] W. Sabrina lin, k. Tjoa, h. Vicky zhao , k. J. Ray liu, "digital image source coder forensics via intrinsic fingerprints", *IEEE trans*, information forensics and security, vol. 4, no. 3, pp-460-475,sep. 2009.

[3] AshwinSwaminathan,Hongmei Gouand Min Wu, "Intrinsic Sensor Noise Features for Forensic Analysison Scanners and Scanned Images", *IEEE Trans*, Information Forensics and security, vol. 4, no. 3, pp.476-491, Sep. 2009 .

[4] Yanmei Fang , Ahmet Emir Dirik , Xiaoxi Sun , Nasir Memon ,"Source Class Identification for DSLR and Compact Cameras", *IEEE Trans*, Information Forensics and security,vol.4,no.3,pp .4244-4464,Oct.2009.

[5] Daniel Garcia-Romero and Carol Y. Espy-Wilson,"Automatic Acquisition Device Identification From Speech Recordings", *IEEE Trans*, Information Forensics and security,vol.4,no.3,pp. 4244-4296,Sep.2010

[6] Matthew C. Stamm and K. J. Ray Liu,"Forensic Detection of Image Manipulation UsingStatistical Intrinsic Fingerprints", *IEEE Trans*, Information Forensics and security,vol.5,no.3,pp .492-506,Sep.2010.

[7] WeiqiLuo, ,Jiwu Huangand GuopingQiu,"JPEG Error Analysis and Its Applications to Digital Image Forensics", *IEEE Trans*, Information Forensics and security,vol.5,no.3,pp .492-506,Sep.2010.

[8] Matthew C. Stamm and K. J. Ray Liu,"Anti-Forensics of Digital Image Compression", *IEEE Trans*, Information Forensics and security, vol.5,no.3,pp .1050-1065,Sep.2011.

[9] Hung-Min Sun, Chi-Yao Weng, Chin-Feng Lee, and Cheng-HsingYang,"Anti-Forensics with Steganographic Data Embedding in Digital Images" *IEEE Trans*, Information Forensics and security,vol.29,no.7,pp .1392-1403,Aug.2011.

[10] Mani MalekEsmaeili, MehrdadFatourechi, and RababKreidieh Ward, "A Robust and Fast Video Copy Detection SystemUsing Content-Based Fingerprinting", *IEEE Trans*, Information Forensics and security,vol.5,no.3,pp .213-226,Mar.2011.

[11] Hai-Dong Yuan,"Blind Forensics of Median Filtering in Digital Images", *IEEE Trans*, Information Forensics and security,vol.6,no.4,pp .1335-1345,Dec.2011.

[12] Arun Rossand AsemOthman,"Visual Cryptography for Biometric Privacy",*IEEE Trans*, Information Forensics and security,vol.6,no.1,pp .70-81,Mar.2011.

[13] Chang-Tsun Li, and YueLi,"Color-Decoupled Photo Response Non-Uniformity for Digital Image Forensics", *IEEE Trans*, Information Forensics and security,vol.22,no.2,pp .260-271,Feb.2012.

[14] Ming Yan, Yi Yang and Stanley Osher, "Robust 1-bit Compressive Sensing using Adaptive Outlier Pursuit" ,*IEEE Trans*, Information Forensics and security,vol.6,no.1,pp .1-8,Mar.2012.

[15] Xiangui Kang, Yinxiang Li, ZhenhuaQu, and Jiwu Huang, "Enhancing Source Camera Identification PerformanceWith a Camera Reference Phase Sensor Pattern Noise", *IEEE Trans*, Information Forensics and security,vol.7,no.2,pp .393-402,Apr.2012.

[16] Hafiz Malik, *Member*,"Acoustic Environment Identification and Its Applications to Audio Forensics", *IEEE Trans*, Information Forensics and security,vol.8,no.11,pp .1827-1837,Nov.2013.

[17] Ahmed F.Shosha, Lee Tobin and PavelGladyshev,"Digital Forensic Reconstruction of A Program Actions", *IEEE Trans*, Information Forensics and security,vol.7,no.2,pp .119-122,Nov.2013.

[18] Eryun Liu, Anil K. Jain and JieTian,"A Coarse to Fine Minutiae-Based Latent Palmprint Matching" , *IEEE Trans*, Information Forensics and security,vol.35,no.10,pp .2307-2322,Oct.2013.

[19] Giuseppe Valenzise, Marco Tagliasacchi, and Stefano Tubaro, "Revealing the Traces of JPEGCompression Anti-Forensics", *IEEE Trans*, Information Forensics and security,vol.8,no.2,pp .355-349,Feb.2013.

[20] Ahmed F.Shosha, Lee Tobin and PavelGladyshev ,"Digital Forensic Reconstruction of A Program Actions" *IEEE Trans*, Information Forensics and security,vol.8,no.2,pp .355-349,Feb.2013.

[21] Eryun Liu, Anil K. Jain, and JieTian, "A Coarse to Fine Minutiae-Based Latent Palmprint Matching" *IEEE Trans*, Information Forensics and security,vol.8,no.2,pp .355-349,Feb.2013.

[22] Yoichi Tomioka, and Hitoshi Kitazawa, "Robust Digital Camera Identification Based onPairwise Magnitude Relations of ClusteredSensor Pattern Noise", *IEEE Trans*, Information Forensics and security,vol.8,no.12,pp .1986-1995,Dec.2013.

[23] ThanhHai Thai, RémiCogranne and FlorentRetraint, "Camera Model Identification Based on the Heteroscedastic Noise Model", *IEEE Trans*, Information Forensics and security,vol.8,no.2,pp .355-349,Feb.2013.

[24] Giovanni Chierchiay, DavideCozzolino, Giovanni Poggi, Carlo Sansone, Luisa Verdoliva,"Guided Filtering For Prnu-Based Localization Of Small-Size Image Forgeries", *IEEE Trans*, Information Forensics and security,vol.22,no.1,pp .250-263,Jan.2014.

[25] Gang Cao, Yao Zhao, Rongrong Ni and Xuelong Li, "Contrast Enhancement-Based Forensics in Digital Images", *IEEE Trans*, Information Forensics and security,vol.9,no.3,pp .515-525,Mar.2014.