Digital Image Forensics:
a booklet for beginners

Judith A. Redi

Eurecom, 2229 Route des Crêtes, BP 193 - 06560 Sophia-Antipolis, France

Phone: +33 (0)4 93.00.82.15

Fax: +33 (0)4 93.00.82.00

E-mail: Judith-Alice.Redi@eurecom.fr,

Website: http://image.eurecom.fr/

Wiem Taktak

Eurecom, 2229 Route des Crêtes, BP 193 - 06560 Sophia-Antipolis, France

E-mail: Wiem.Taktak@eurecom.fr,

Jean-Luc Dugelay

Eurecom, 2229 Route des Crêtes, BP 193 - 06560 Sophia-Antipolis, France

E-mail: Jean-Luc.Dugelay@eurecom.fr,

Abstract. Digital visual media represent nowadays one of the principal means for communication. Lately, the reliability of digital visual information has been questioned, due to the ease in counterfeiting both its origin and content. Digital image forensics is a brand new research field which aims at validating the authenticity of images by recovering information about their history. Two main problems are addressed: the identification of the imaging device that captured the image, and the detection of traces of forgeries. Nowadays, thanks to the promising results attained by early studies and to the always growing number of applications, digital image forensics represents an appealing investigation domain for many researchers. This survey is designed for scholars and IT professionals approaching this field, reviewing existing tools and providing a view on the past, the present and the future of digital image forensics.

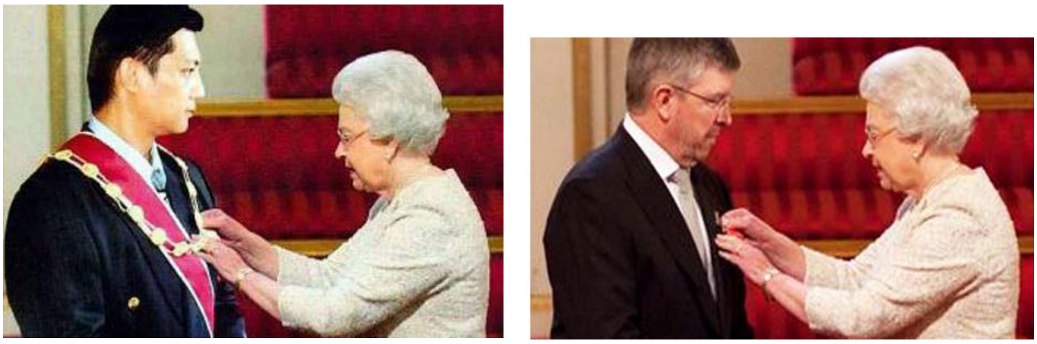
Keywords: Digital Image forensics, Multimedia security, Image tampering detection, Image source authentication, counter-forensics

1. Introduction

Images and videos have become the main information carriers in the digital era. The expressive potential of visual media and the ease in their acquisition, distribution and storage is such that they are more and more exploited to convey information, even sensible. As a consequence, today images and videos represent a common source of evidence, both in every-day life controversies and in trials. The simplest video in TV news is commonly accepted as a certification of the truthfulness of the reported news. In a similar way, video-surveillance recordings can constitute fundamental probationary material in a court of law.

Together with undoubted benefits, the accessibility of digital visual media brings a major drawback. Image processing experts can easily access and modify image content, and therefore its meaning, without leaving visually detectable traces. Moreover, with the spread of low-cost, user friendly editing tools, the art of tampering and counterfeiting visual content is no more restricted to experts. As a consequence, the modification of images for malicious purposes is now more common than ever. Digital Image Forensics is that branch of multimedia security that, together with Digital Watermarking, aims at contrasting and exposing malicious image manipulation.

In July 2010 Malaysian politician Jeffrey Wong Su En claimed to have been knighted by the Queen Elizabeth II, as recognition for his contribution to the international aid organization Médecins Sans Frontières. A picture of him being awarded by the Queen of England accompanied his statement, diffused in local media (fig. 1.a). When questioned about the award though, the British High



(a) *(b)*

Figure 1 – The doctored image depicting Jeffrey Wong Su En while receiving the award from Queen Elizabeth II, published in Malaysian dailies, and the original picture of Ross Brawn receiving the Order of the British Empire from the Queen (b).

Commission in Kuala Lumpur made clear that the name of Mr. Wong was not included in the official knighthood recipients lists, and that the picture was inconsistent with the usual protocol adopted for knighthood ceremonies. The image was finally shown to be a splicing between an original ceremony photo (fig. 1.b) and Mr. Wong’s face, built to increase his popularity.

This kind of episodes [1] contributed in making more and more questionable the use of digital images as evidence (for an updated and detailed archive, cfr. [2]). A confirmation of their authenticity is needed, before further relying on their content. For this reason, two questions about the history of the image have to be answered:

a) Was the image captured by the device it is claimed to be acquired with?

b) Is the image still depicting its original content?

The first question is of major interest when the source of the image is the evidence itself, i.e. when the ownership of the capturing camera is compromising, or when an accusatory content is such only if it was recorded by a specific device (e.g. video surveillance). The second question is of more general interest, and can be directly applied to the fake knighthood picture case. Answering to those questions is relatively easy when the original image is known. In the case of the fake knighthood, the simple availability of the original image was sufficient to expose the forgery. In practical cases, though, almost no information can be assumed to be known a priori about the original image. Investigators need therefore to authenticate the image history in a blind way.

Digital image forensics (DIF) aims at providing tools to support blind investigation. This brand new discipline stems from existing multimedia security-related research domains (e.g. Watermarking and Steganography) and exploits image processing and analysis tools to recover information about the history of an image. Two principal research paths evolve under the name of Digital Image Forensics. The first one includes methods that attempt at answering question a), by performing some kind of ballistic analysis to identify the device that captured the image, or at least to determine which devices did not capture it. These methods will be collected in the following under the common name of image source device identification techniques. The second group of methods aims instead at exposing traces of semantic manipulation (i.e. forgeries) by studying

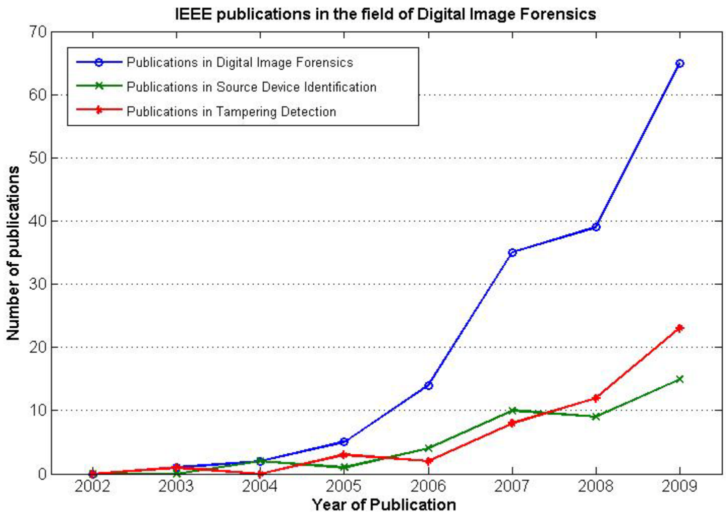


Figure 2 - Number of IEEE publications in the field of Image forensics over the last 8 years. Data were retrieved from the IEEE explore website, http://ieeexplore.ieee.org by submitting the queries “Digital Image Forensics”, “Source Device Identification” and “Tampering Detection”

inconsistencies in natural image statistics. We will refer to these methods as tampering detection techniques.

Both these fields and the DIF domain in general are attracting a growing interest from the scientific community. Figure 2 reports the number of papers concerning DIF that have been published in IEEE conferences and journals over the last 8 years. The first publications in the topic date back to 2003, although previous work might have been made public a few years earlier, such as Hani Farid’s investigation on bicoherence features for tampering detection [3]. In fact, figure 2 is not aimed at reporting a precise quantification of the publications in the field, but rather at highlighting a trend: the intensity of research activity in the DIF domain started to increase about five years ago. Nowadays, the domain is so mature that researchers are starting to propose techniques to contrast forensics methods, exploiting their weaknesses to better hide or counterfeit the manipulation. These studies represent a precious contribution to the development of image forensics, pushing experts to create more and more robust tools.

All this considered, Digital Image Forensics represents an appealing topic for many researchers. In this overview, we aim at providing useful tools to scholars and IT professionals who, coming from related fields, want to start research in Digital Image Forensics. We assume to deal with researchers who have some background either in digital forensics or signal processing. When approaching DIF for the first time, these people might have in mind questions such as:

- What is digital image forensics and why is it important?
- Which background should I have to start research in this field, and how could I apply my know-how to it?
- Which is the state-of-the-art of DIF tools?
- How is DIF evolving and which are the future challenges?

In the following we will try to provide enough useful information to answer these questions. In section 2 we put digital Image Forensics in perspective with other multimedia security-related disciplines. In section 3 we report on methods dealing with the authentication of the acquisition device of a particular image. In section 4 we propose an overview of tampering detection methods. Section 5 presents a view on the evolution of DIF: counter-forensics techniques, or how to fool existing detection and authentication methods. Finally, with section 6 we point out achievements and limitations of the current approaches, and try to outline possible, new research paths in the Digital Image Forensics domain.

2. The role of Digital Image Forensics in multimedia security

Digital Image Forensics is a quite recent discipline, as pointed out with fig. 2; nonetheless, it is tightly connected with a number of different research fields. DIF inherits its goals and attitude from classical (analog) forensic science and from the more recent field of computer forensics (for a complete dissertation on similarities and differences between these domains, see [4]). Forensic disciplines in general aim at exposing evidence of crimes; to do so, they have to deal with the burglars' ability in either hiding or possibly counterfeiting their traces.

In digital imaging both the acquisition process and the tampering techniques are likely to leave subtle traces. The task of forensics experts is to expose these traces by exploiting existing knowledge on digital imaging mechanisms, being aided by consolidated results in multimedia security research.

For better grasping the mission of image forensics investigators, it might be useful to explore the relationships between DIF and other multimedia security-oriented disciplines. Image processing for forensics shares indeed several technical challenges and similar techniques with digital watermarking and steganography [5].

Digital watermarking consists in hiding a mark or a message in a picture in order to protect its copyright. There exists a complex trade-off between three conflicting parameters (payload, robustness, and invisibility). In short, a good algorithm in digital watermarking should hide enough bits without modifying significantly the cover and should be able to recover the message even if the picture underwent some modifications between the signer and the retriever. As a first approximation, a watermark can be seen as a specific noise.

A particular application for watermarking is image integrity protection [6]. In this case, a fragile watermark is applied to the cover image so that it gets destroyed at a tampering attempt. This ensures some control on the image content manipulation. One major difficulty here is to make the distinction between malicious and naïve modifications (tampering versus fair compression for example). Within this context, digital watermarking can be seen as an active protection, whereas tools in digital image forensics can be seen as passive ones. However, in many scenarios it is not realistic to envision that images and videos have been protected by a fragile watermark prior to dissemination. This gives a plus in favor of image forensic tools for the detection of tampering. Anyway, in both cases, image processing tools are analyzing high frequencies either to recover the watermark or to detect inconsistencies in noise and expected patterns, e.g. related to the acquisition process.

Steganography consists in communicating secretly via some media (in particular images and videos). The choice of the cover is not really important here. Also, one can assume that the stego-picture will not undergo photometric or geometric attacks among the transmission. The main point for two persons who communicate some information using this technology is to be not detected by a third party. To make the message not detectable, algorithms mix secret information in high frequencies with other existing noises, e.g. related to sensors.

We can observe that, to some extent, all of these techniques are working in high frequencies to add, to recover, to detect, or more generally to analyze “noises”, or characteristic patterns.

However, Digital Image Forensics has a very precise role among multimedia security disciplines: authenticating images for which no reference is known and no previous integrity protection has been set. For this reason, they are often referred to as *passive* (as opposed to the active protection provided by e.g. the insertion of a fragile watermark), and *blind*. This makes DIF techniques the only applicable tools in a large set of practical situations.

3. Image source device identification

In tracing the history of an image, identifying the device used for its acquisition is of major interest. In a court of law, the origin of a particular image can represent crucial evidence; the validity of this evidence might be compromised by the (reasonable) doubt that the image has not been captured from the device it's claimed/supposed to be acquired with, as in the case of video-surveillance material or covert videos.

Helpful clues on the source imaging device might be simply found in the file's header (EXIF), or by checking (if present) a watermark consistency [5, 7]. However, since this information can be easily modified or removed, it cannot always be used for forensics purposes. As a consequence, blind techniques are preferred for the acquisition device identification.

Blind image forensics techniques take advantage of the traces left by the different processing steps in the image acquisition and storage phases. These traces mark the image with some kind of camera fingerprint, which can be used for authentication. The techniques presented in the following retrieve information on the source device at two different levels. As a first attempt, they try to distinguish between different camera models. On a second, more informative although more challenging level, the goal is to distinguish between single devices, even different exemplars of the same camera model. To provide the reader with a better understanding of these techniques, we start with an illustration of the most common steps in the image acquisition and storage processes, in order to describe the origin of image artifacts and how to exploit them in image authentication.

In reviewing existing techniques, we would like to warn the reader that no direct performance comparison is available between different methods. This is mainly due to the lack of established benchmarks, and, in particular, of a reference test dataset. Recently, the *Dresden Image Database for Benchmarking Digital Image Forensics* [8] has been made available for forensic investigators, containing over 14,000 images with various content, captured by 73 different camera models. To the best of the author's knowledge, no study besides [8] has been benchmarked on this dataset so far. It is our belief, though, that this useful tool will be soon exploited by the community to have a clearer view on the state of the art of acquisition device identification methods.

3.1. Image acquisition and storage

When capturing a digital image, multiple processing steps are performed prior to the storage. Figure 3 shows the typical image acquisition pipeline [9]. The Light enters the imaging device through a system of optical lenses, which conveys it towards the *imaging sensor*. The imaging sensor is the heart of every digital camera, and it is composed of an array of photo detectors, each corresponding to a pixel of the final image, which transform the incoming light intensity into a proportional voltage. Most cameras use CCD (Charged Coupled Device) sensors, but CMOS (Complementary Metal Oxide Semiconductor) imagers can also be found. To render color, before reaching the sensor the light is filtered by the Color Filter Array (CFA), a specific color mosaic that permits to each pixel to gather only one particular light wavelength (i.e. color). The CFA pattern arrangement depends on the manufacturer, although Bayer's filter mosaic is often preferred. As a result, the sensor output is a mosaic of e.g. red, green and blue pixels arranged on a single layer. To obtain the canonical 3-channels representation, the signal needs to be interpolated. *Demosaicing* algorithms are applied to this purpose; the

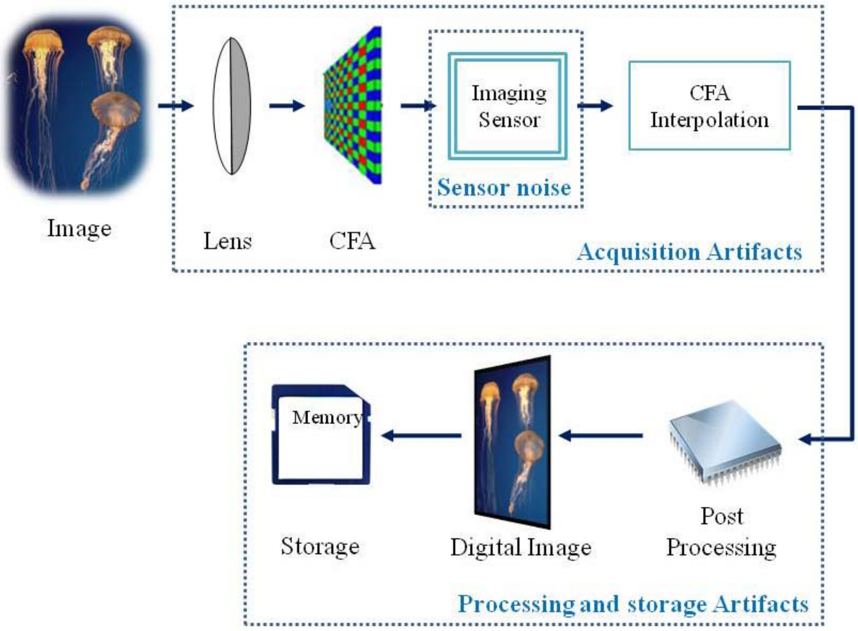


Figure 3 - A schematic view of a standard digital image acquisition pipeline

missing pixel values in each layer are estimated based on the values of existing neighbors. Before the eventual storage, additional processing is performed, such as white balance, gamma correction, and image enhancement. Finally, the image is recorded in the memory device. Also in this case the format can vary, but a common choice is JPEG.

The described image acquisition pipeline is common for most of the commercially available devices; nonetheless, each step is performed according to specific manufacturer choices, and hence might depend on the camera brand and model. This variation can be used to determine the type of camera from which a specific image was obtained. Indeed, each stage in the pipeline can introduce imperfections in the final image or characteristic traits: lens distortion, chromatic aberration, pixel defects or CCD sensor imperfections, statistical dependencies related to proprietary CFA interpolation algorithms and other intrinsic image regularities which leave tell-tale footprints. These artifacts are statistically stable and can be considered as a signature of the camera type or even of the individual device. In addition, some techniques focus on statistical regularities in the image such as color or compression features.

Each of the techniques presented in the following section has been proved to be effective to some extent in identifying the acquisition device. However, the scope of most of them is limited to the discrimination between different camera models. Sensor imperfections seem to be at this stage the only traits able to distinguish between different exemplars of the same camera model. Therefore, recent studies mostly concentrate on exploiting sensor imperfection and in extracting pattern noise as a sensor biometric trait. Yet, advances were made also with the other approaches, especially using demoisaicing regularity [10].

3.2. Forensics methods for image source identification

In the following, image source identification methods will be presented according to the kind of cues they explore, following their location along the pipeline. Section 3.2.1 will introduce those methods which use acquisition artifacts, produced by lenses or CFA interpolation. Section 3.2.2 reports on sensor imperfections; finally, section 3.2.3 describes methods which analyze properties of the output image generated in the last stages of the pipeline.

3.2.1. Identification through artifacts produced in the acquisition phase

Due to each lens and image geometries and to the design of the camera, lenses produce distortions (aberrations) in the captured images. Radial distortion for example deforms the image so that straight lines in object space appear as curved lines. Choi et al. [11] propose to analyze this kind of lens aberration as a fingerprint to identify the source camera. The authors find the distortion parameters of a camera using Devernay's line extraction method [12], to then measure the error between the distorted line segment and the corresponding straight lines. The estimated parameters can be used to train a classifier to distinguish among images captured by different cameras. The method achieves good discrimination rates among different models of cameras but no proof of robustness is given for distinct exemplars of the same model. Additionally, this method will fail to measure radial distortion in absence of straight lines to be extracted from the picture. On a similar path, Van and others [13] propose to help the correct identification of the source camera using lateral chromatic aberration analysis. Chromatic aberration derives from imperfection in the lens, which provokes a discrepancy in the location in which the sensor receives light of different wavelengths. This phenomenon is modeled in [14] through three parameters, regulating the distance from the center of distortion (optical center) and its magnitude. The model is applied to compute the misalignment between pairs of color channels (i.e., red-green and blue-green), which is formulated as a registration problem. The model parameters, estimated by maximizing the mutual information between color channels, eventually feed a SVM, trained to recognize the source camera. Again, this method is shown to provide accurate identification of different camera models, but not of single devices. It might be interesting, for future research, to attempt at integrating these methods with features describing other kinds of lens aberration (e.g. spherical, longitudinal aberration).

Proceeding on the image pipeline, the demoisaicing step provides important information on the acquisition device, being strictly dependent on the manufacturer (or even on each camera model). Demosaicing introduces a specific type of correlation between the color value of one pixel and its neighboring samples in the same color channel. Therefore, different demosaicing algorithms can be described based on these dependencies, as shown in [15] by Bayram et al.

Their method is inspired by the technique proposed by Farid et al. intended for image tampering detection [16]. The authors propose to detect traces of CFA interpolation in color bands by using the expectation/maximization (EM) algorithm. The analysis outputs a Probability Map which exposes peaks at different frequencies indicating structural correlation between spatial samples; weighting interpolation coefficients are also estimated to distinguish between different CFA interpolations. In their experiments, the authors use 140 smooth pictures from two camera models, obtaining over 95% accuracy in classification, supported by a SVM. In a later work [17], the same authors improve their approach by designing new features based on image quality metrics [18] and higher-order statistics of the image [19]. Good performance is obtained among up to three digital cameras. In a different approach [20], Long et al. use a quadratic pixel correlation model to obtain a coefficient matrix for each color channel, whose principal components are extracted to then feed a neural network. The method is shown to effectively identify four different camera models, although its performance decreases with the increase of compression. Celiktutan et al. also propose a CFA interpolation based technique [21] applied to cell-phone cameras. The authors extract a set of binary similarity measures between the bit-planes, assuming that the correlation across adjacent bit-planes can be representative of a specific camera CFA interpolation algorithm. In their experiments, the authors use these features in conjunction with image quality measures as inputs to a classifier. 200 images from each of 9 cameras are used in the test. The method performs a successful classification when discriminating among a small number of cameras, but for a larger number, the accuracy drops.

3.2.2. Identification through sensor imperfections

Imaging sensors have been shown to introduce various defects and to create noise in the pixel values [22]. The *sensor noise* is the result of three main components, i.e. pixel defects, fixed pattern noise (FPN), and Photo Response Non Uniformity (PRNU).

Pixel defects include point defects, hot point defects, dead pixels, pixel traps, and cluster defects, which reasonably vary across different sensors, independent on the specific camera model. Geradts et al in [23] attempt at reconstructing pixel defects patterns. The authors propose to determine pixel noise by taking images with black or green background with 12 different cameras and then comparing the defect points which appeared as white. Their experiments show that each camera has distinct patterns of defect pixels also across the same model; nonetheless, the impact of defect pixels closely depends on the content of the image. Furthermore, some camera models do not contain any defectives pixels or they eliminate it. Therefore, this method is not applicable to every digital cameras.

FPN and PNRU are the two components of the so-called *pattern noise*, and depend on dark currents in the sensor and pixel non-uniformities, respectively. Hence, they are independent on the image content but closely related to the physical characteristics of each single sensor. Lukas and others [24], propose to analyze pattern noise for camera identification, as it is a unique stochastic characteristic for both CCD and CMOS sensors. As such, the pattern noise extracted from images taken by the same camera should be more correlated than those extracted from different cameras. Pattern noise can be estimated by taking the difference between an image **I** and its denoised version:

$$\mathbf{W} = \mathbf{I} - F(\mathbf{I}) \qquad (1)$$

Where **W** is the residual noise and *F* is a denoising filter. As random components might appear in the residual noise due to the image content, the reference pattern noise for a particular camera C is obtained by averaging the noise residual computed for a sufficiently large number of images *m*:

$$\mathbf{W}_{ref}^{C} = \frac{1}{m}\sum_{i=1}^{m}\mathbf{W}_{i}^{C} \qquad (2)$$

To establish whether a given image $\mathbf{I}_\mathrm{N}$ was captured by a digital camera Γ, the correlation between the pattern noise estimated for the individual image $\mathbf{W}_\mathrm{N}$ and the reference pattern $\mathbf{W}_{ref}^{\Gamma}$ is computed, and compared to an empirically determined threshold, following a procedure largely resembling to that applied for people authentication in biometry. Indeed, we can easily associate pattern noise to a biometric trait, which can be used to determine whether the camera that took the picture is the claimed one (client) or another one (impostor). In this sense, this approach could be defined as "Hardwermetry". The method proves to attain high accuracy with low false rejection rates even for cameras of the same model, as also validated by an independent study [25]. Appreciable results are shown also in presence of re-sampling and JPEG compression. A possible weakness is that the authors used the same image set both to calculate the camera reference pattern and then to test the method. As a result, the effectiveness of this technique has not been proven to be effective independently on the training set.

In a later study, Chen and others [26, 27] propose to refine the previous method by estimating PRNU information from pattern noise **W**. It is observed that pattern noise can be expressed as the joint contribution of two terms:

$$\mathbf{W} = \mathbf{IK} + \mathbf{\Xi} \tag{3}$$

Where **IK** is the PRNU signal and **Ξ** is a term reflecting the contribution of all the other sources of noise that can intervene in the acquisition process, e.g. CFA interpolation or quantization due to compression. Given *d* images $\mathbf{I}_k$, *k* =1, …, *d*, taken from the same camera C the reference PNRU factor $\hat{\mathbf{K}}^C$ can be estimated with through Maximum Likelihood

$$\hat{\mathbf{K}}^C = \frac{\sum_{k=1}^{d} \mathbf{W}_k \mathbf{I}_k}{\sum_{k=1}^{d} (\mathbf{I}_k)^2} \tag{4}$$

Where $\mathbf{W}_i$ is computed as per eq. (1). Detection of PRNU for single images is then formulated as a Neyman-Pearson hypothesis testing problem. In its refined version, this authentication technique gains both in computational efficiency and in error analysis accuracy. One of its main limitations, i.e. the lack of robustness to geometric transformations, is partially addressed in [28].

Recently, C. Li [29] observed that the extraction of sensor noise from a single image might produce a pattern contaminated by the fine details and the structure of the depicted scene, and that this deviation might reduce the probabilities of matching with a reference. The author proposes to enhance the fingerprint estimation by weighting noise components in a way inversely proportional to their magnitude, in order to suppress information derived from non-smooth image parts. As a result, high classification accuracy is obtained also on small-sized image regions.

3.2.3. Source identification using properties of the imaging device

Exploiting the digital image acquisition process is not the only way to identify the source of an image: post-processing performed in the storage phase can also produce interesting cues.

Kharrazi et *al.* [30] propose to use a set of image features to characterize a specific digital camera, assuming that the image can be affected by color processing and transformations operated by the camera prior to the storage. The authors study statistical properties of the image organized into two groups: color-related measurements, such as average pixel value, RGB pairs correlation, neighbor distribution center of mass, energy ratio and wavelet domains statistics, and image quality features. Supported by a SVM classifier, this approach shows an effective result on low compressed images taken by different camera models. However, this technique can only be applied on images depicting similar content. In a later Work, Tsai et al. [31] use similar image features to construct multi-class classifiers. They show that independent of the image content, using the same number of camera and training images as in [30], their classifier can distinguish between source cameras with good precision.

After the processing phase, many digital camera models encode images in JPEG format before storing them. The well known JPEG compression standard transforms 8x8 pixels blocks of the image in the DCT domain. DCT coefficients are quantized by a given amount *q*, according to a pre-defined quantization table. Quantized coefficients are then entropy-coded into a bit-stream, and the information lost in the quantization step cannot be recovered in the de-compression phase. Based on this lossy compression scheme, the identification of the quantization table can be a useful hint for source camera identification, as it usually varies among manufacturers. Farid [32] analyzes the quantization tables for ballistic fingerprinting. The author proves how this technique can provide useful information on the source device, by testing it on 204 digital camera models. However, it is likely that the quantization table alone will not be sufficiently unique to allow the discrimination between all possible camera models; therefore, the author advises its usage in conjunction with other forensic tools (e.g., using it to exclude from further analysis incompatible manufacturers/models).

Image thumbnails were also shown to include some fingerprinting information. A thumbnail is a reduced version of a full resolution image, often stored together with it to allow users a quick preview. The process that transforms an image into its thumbnail can involve a series of filtering operations, contrast adjustment and JPEG compression. Each operation implementation depends on the imaging device manufacturer as well as on the device model. E. Kee and H. Farid [33] propose to use the formation and storage parameters of an embedded image thumbnail for digital source camera identification. The authors describe a set of techniques to estimate the parameters adopted by the specific device to perform each of the thumbnail creation steps, namely cropping, pre-filtering, down-sampling and post-filtering operations, contrast adjustment, and JPEG compression. These parameters can actually characterize a single device and how they differ e.g. from camera models to photo-editors. The authors remark that, to facilitate the application of the method, a database including the characterizing parameters of a wide range of devices should be made available.

4. Tampering Detection

According to the Oxford dictionary [34], the verb *to tamper* literally means “to interfere with something in order to cause damage or make unauthorized alterations”. In the context of digital imaging, tampering recalls the intentional manipulation of images for malicious purposes: as images are natural carriers of information, image manipulation is denoted as tampering when it explicitly aims at modifying the semantic meaning of the visual message.

The story of image forgery dates back to the early twentieth century to support political propaganda actions [2]. With the increasing use of visual content as a message conveyer, tampering techniques developed accordingly. Furthermore, with the advent of digital imaging and photo-editing software, image manipulation became affordable also for non-darkroom experts, resulting in a general lack of reliability of digital image authenticity, not only in investigative activities, but, more in general, in the media and information world.

The studies reported in this section respond to the need of the society of being able to identify tampered images and, more importantly, to detect the forged regions. In this overview of tampering detection techniques, we first provide a short summary of the most frequently used manipulation techniques, then, for each of them, we analyze the state of the art of forgery exposal methods.

4.1. A short summary of the most common tampering techniques

Image semantic content can be altered in two major ways: by removing information or by adding it. To remove information, usually forgers do not need to access the content of another image, as shown in figure 4, top row. Conversely, one typical way to add extraneous information to an image is to use material extracted from one or more distinct images. It should be noticed, however, that this operation is not always necessary, as the simple processing of an image can convey relevant alteration of both semantics and pragmatics of an image, as in the well-known case of the TIME magazine cover depicting a darkened O. J. Simpson portrait (fig 5). We analyze tampering methods and their malicious intentions by dividing them in two categories: those which produce the forgery working on a single image, and those that access the content of more than one image (i.e. composites).

4.1.1. Forgeries involving a single image

Deleting undesired objects from an image is one of the most straightforward methods to alter its meaning. In such circumstances, forgers need to “fill” the region of the image from which the object has been removed. A typical solution in this case is to copy a portion of the same image and replace with it the void left from the deletion (copy-move technique). Of course, the same approach can be used to replicate objects instead of deleting them, as shown in the bottom images of fig. 4. To better hide this operation to the human eye, the forger can perform geometric transforms on the region to be copied, such as rotation or scaling. Furthermore, to produce a smooth transition between the (original) surround and the object to be pasted, matting and blending techniques can be exploited [35, 36].

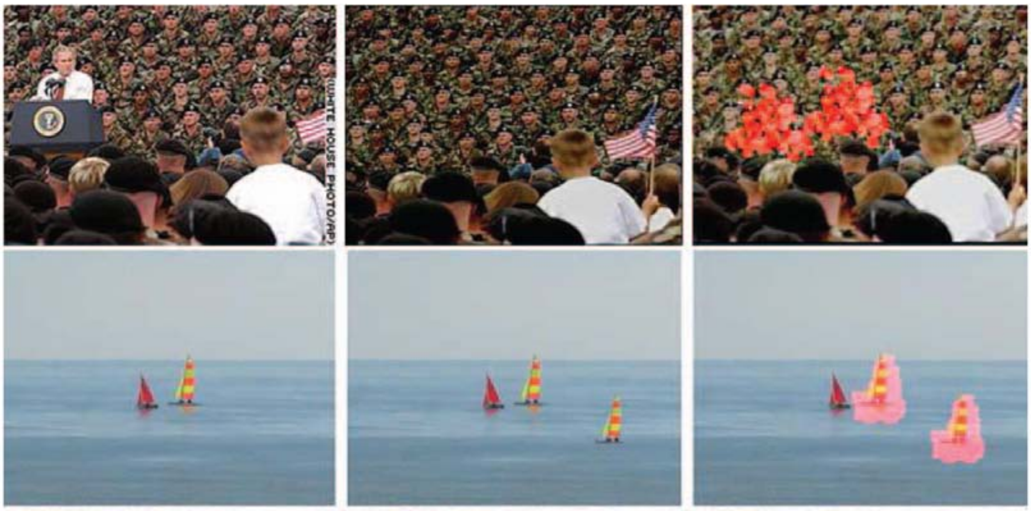


Figure 4 Examples of copy-move attacks for object removal (above) and duplication (below). Images were taken from [50]

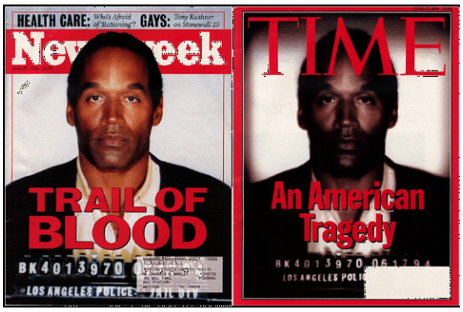


Figure 5 The Newsweek and Time covers reporting on the O.J. Simpson case. Time magazine was accused to have darkened the image with racist purposes

Object removal can be also achieved by means of in-painting techniques [37]. Inspired by real techniques for painting restoration, in-painting methods fill the holes left by object removal by exploiting the information preserved in the regions surrounding the gaps. In particular, in-painting is based on an iterative process of smooth information propagation from the image to the region to be filled. The gap is gradually filled from the periphery to the center, resulting in a perceived continuity in the final image. However, the algorithm struggles with filling highly textured areas.

Recently, the seam carving method was shown to be a powerful tool for object removal [38]. Initially proposed for content-aware image resizing [39], the algorithm is based on the notion of seam. A seam is defined to be a monotonic and connected path of pixels including one pixel per row (column) and traversing the image from the top to the bottom (left to the right). Seams are iteratively removed based on a minimum energy criterion; each seam deletion corresponds to a horizontal (vertical) resizing of one pixel, leading to a final result perceptually more coherent than what could be obtained by simple re-sampling of the image. When targeting the algorithm on a specific region or object, the same methodology becomes a very accurate tool for object removal, which is achieved by iteratively eliminating all the seams traversing the selected region.

Forgeries can be performed on a single image also without recurring to object removal. Image semantics can be modified by applying simple image processing techniques, such as histogram manipulation or contrast enhancement (fig. 5). Additionally, brightness adjustment and scaling can be of appropriate application to better mask copy-move forgeries: more in general, the application of filters or simple geometric transforms can compromise the forensics analysis, by covering or deleting traces of further tampering, as shown in [40] for scientific images.

4.1.2. Forgeries using multiple images as source for tampering

The insertion in an image of material originally coming from another source is one of the most powerful tools to overturn the message contained in visual media. Modern techniques and editing software allow easy creations of image composites (e.g. through layers superposition) obtaining results that are hardly detectable by the human eye (see, e.g. figure 1 and 6). Blending and matting techniques are again applicable to mask the boundaries of the spliced regions and to give the image a more uniform aspect. Also, the creation of image composites might require geometric transformation. Rotation, scaling and translation are often needed to make sure that the spliced object respects the original image perspective and scale. Geometric transforms typically involve re-sampling, which in turn calls for interpolation (e.g. nearest neighbor, bilinear, bicubic). The re-sampling process produces artifacts in the image histogram, and hence provides a useful cue for compositing detection.

It should be also taken into account that inserted material does not necessarily have to come from natural images. As computer graphics evolves, more and more realistic 3D objects can be modeled and rendered to be eventually spliced into an image composite. Furthermore, the extraction of 3D scene structure from images allows manipulating objects by morphing them: in this case, the splicing involves a remodeled (i.e. artificial) version of a part of the original image. This technique is applied when the forger aims e.g. to modify a facial expression as experimented in video rewriting [41].

4.2. Methods for forgery detection

In the remainder of this section we review some of the most effective techniques for tampering detection. The methods will be presented according to their target: forgeries produced from a single image (sec. 4.2.1) and composites (sec 4.2.2). Techniques able to detect both kinds of tampering will eventually presented in subsection 4.2.3.

It should be noticed that also in the case of tampering detection no well-established benchmark exists. However, researchers from Columbia University made available in 2004 the “Dataset of authentic and spliced image blocks” [42], including 933 authentic and 912 spliced image blocks of size 128 x128 pixels extracted from the CalPhotos image set [43]. This dataset represents a valuable support for researchers dealing with image composite detection. As it will be shown in section 4.2.2, fair comparison of different splicing detection methods was already possible thanks to this useful tool.

4.2.1. Detecting tampering performed on a single image

One of the most popular methods for modifying image semantics is the copy-move of an image region, due to its simplicity and effectiveness. When region copy-move is performed without further retouching, the tampered area still shares most of its intrinsic properties (e.g. pattern noise or color palette) with the rest of the image. On the other hand, a structural analysis of image segments might reveal high similarities between distant regions. Based on this idea, a first attempt in retrieving tampered areas was described in [44]. The authors propose to look for matches among DCT representations of image segments. To avoid the computational burden of a brute force comparison, DCT coefficients are lexicographically sorted and adjacent identical pairs are considered as potentially tampered regions. A refinement of this selection is based on a spatial criterion. A histogram is built counting the number of matching segments which are separated by the same distance. The higher the number of pairs located at the same distance, the higher is the probability that those pairs belong to copy-moved regions. Eventually, bins of the histogram including a number of segments higher than a predefined threshold are selected, and the corresponding regions are marked as tampered. This strategy avoids the selection of isolated segments, and is based on the assumption that likely the copy-move forgery is performed on relatively extended regions.

On a similar path, Popescu and Farid [45] propose to perform a principal component analysis (PCA) for the description of image segments (in particular, overlapping square blocks). PCA-based description is shown to bring higher discriminative power, together with a reduction of the computational cost achieved through the truncation of less significant eigenvectors. The resulting complexity of the algorithm is O (N_t N log$N$), being N_t the dimensionality of the truncated PCA representation and N the number of image pixels. The approach is shown to produce a low number of false positives and to be resistant to image degradation due to noise or lossy JPEG compression. Further research on robust features for segments matching is reported in [46]. To deal with computational complexity, the use of kd-trees is proposed in [47] and exploited in [48, 49]. In the latter, Dybala and others try to tackle the problem of region copy-paste masking through re-touching. As well as in [40], the method is able to detect traces of the use of Adobe Photoshop healing brush and Poisson cloning.

A main limitation of the previous methods is their low robustness to the likely scaling and rotation operations that tampered regions might undergo before being pasted. In this regard, the work of Huang et al [50] exploits SIFT features to obtain robust tampering detection even under geometric transforms. In this case, only matching SIFT key-points are retrieved, by means of the best-bin-first nearest neighbor identification. Hence, it is not possible to define the exact boundaries of the tampered regions. Dealing with the same problem, in [51] Bayram and others represent image segments in a Fourier-Mellin Transform domain, known for its invariance to geometric transformations. Further, the method interestingly proposes the use of bloom-filters to reduce the computational complexity of the segments matching phase. This aspect still has to be refined, as it proved to be less robust than the lexicographic sorting, since it requires the exact segment matching rather than just a high similarity among them. The SIFT-based approach was recently further developed in two independent works [52, 53]. In both studies, not only matching key-points are detected, but also the geometric transforms parameters (e.g. rotation angle and scaling factor) are estimated with acceptable accuracy. Method [53] also exploits these parameters to recover an approximated boundary of the tampered area.

With the introduction of in-painting and content-aware resizing techniques, image forgers found new efficient tools to perform object removal. To contrast this tendency, Sarkar et al. in [54] proposed a method for detecting seam insertions and carving in images. The authors observe that seam carving introduces inconsistencies in the image high frequency DCT components; such inconsistencies can be modeled with a Markov random process, whose parameters are exploited to feed an SVM, trained to recognize tampered images. Although not treating malicious applications of the seam insertion/removal, the paper represents a first attempt to deal with these new techniques. More recently, the study by Fillion and Sharma [55] treated the problem by extracting energy-bias-based features and wavelet absolute moments. The algorithm is able to identify images which have been deliberately resized to hide information for malicious purposes. Nonetheless, both methods are only able to define whether an image has been tampered by seam carving or not. To the best of the authors' knowledge, no method has been proposed so far able to also detect the tampered areas or explicitly point out object removal. The in-painting attack was analyzed in [56] by Wu et al; the authors use zero-connectivity features and fuzzy membership to detect such doctoring.

The detection of image processing based tampering is not always straightforward. Nonetheless, several methods have been proposed in the literature to identify it. Avcıbaş and others [57] designed a classifier able to discriminate between different kinds of manipulation (i.e. rotation, scaling, brightness adjustment and histogram equalization). Image quality metrics are adopted as discriminative features, as they reflect alterations in the image through a measure of its quality. Hani Farid in [40] explores tampering of scientific images, which often is performed by removing salient objects and by replacing them with a uniform region with the mean intensity of the background. The author proposes to detect these modifications by considering the image as a graph, where each pixel is a vertex, and to divide it in intensity segments by performing normalized cuts. Traces of manipulation such as intensity and texture differences in the background are exposed by appropriately weighting the edges of the graph prior to the cut.

A more complete treatment of image processing for malicious tampering is discussed in [58]. The authors observe that most of the image processing operations can be viewed as pixel mappings which leave statistical traces; as such, every tampered image carries some kind of "fingerprint" describing the image processing history. Given a mapping (processing algorithm) *m*, each pixel **w** in the tampered image **J** is related to its corresponding pixel **x** in the original image **I** by the relationship **w** = *m*(**x**). Therefore, the original image histogram $H_{\mathbf{I}}$ and the histogram $H_{\mathbf{J}}$ corresponding to **J** are related by:

$$H_{\mathbf{J}}(l) = \sum_{t=0, m(t)=l}^{255} H_{\mathbf{I}}(t) \tag{5}$$

The authors define the intrinsic fingerprint of the tampering *m* as $f_m(l) = H_{\mathbf{J}}(l) - H_{\mathbf{I}}(l)$, which describes the changes in the image histogram after the application of *m*. The fingerprint is modeled processing-dependent and analyzed to discern between tampered and genuine images. Contrast enhancement, for example, is observed to produce an increase in high frequency components of the original histogram. This cue is exploited for tampering identification, allowing also to extend the approach to the detection of circumstantiate contrast enhancement, typically applied to mask copy-paste forgeries.

4.2.2. Detecting image composition

Image composites result from the splicing of parts of different images in a single one. The simple splicing operation itself, even when visually masked with blending or matting techniques, leaves traces in image statistics. Bicoherence features were proposed by Farid in [3] to highlight these traces and later successfully applied in [59] for splicing detection. Bicoherence is a normalized bispectrum, i.e., the third order correlation of three harmonically related Fourier frequencies of a signal, which appears to capture quite well the discontinuities introduced in the image after splicing. Methods in [60, 61, 62] propose alternative techniques based on the Hilbert-Huang transform, on statistics of 2-D phase congruency and on wavelet sub-bands features together with Markov transition probabilities of difference JPEG 2-D arrays, respectively. The latter, in particular, outperforms the other techniques when applied on the Columbia Image Splicing Detection Evaluation Dataset [42].

An interesting research area in image composition exposure involves image appearance cues. Even if modern editing tools allow covering the traces of splicing in a convincing way, it is not always possible for the forger to match the lighting conditions of the regions that make up the composite, as in the well-known case of the Kerry and Fonda photomontage (fig. 6). Several studies have

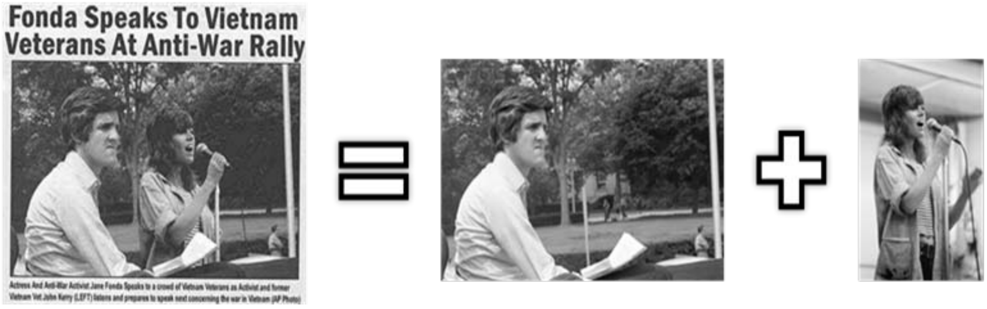


Figure 6 - The famous photomontage of John Kerry and Jane Fonda.

been dedicated to forgery detection through the scene illumination analysis. A first attempt was proposed by M. Johnson and H. Farid in [63], in which they estimate the incident light direction for different objects in order to highlight mismatches. Assuming to deal with lambertian surfaces and an infinitely far away point light source, the authors express the image intensity I for pixel $\mathbf{x}$ as:

$$I(\mathbf{x}) = \mathrm{R}(\mathbf{N}(\mathbf{x}) \cdot \mathbf{L}) + \mathrm{A} \tag{6}$$

Where R is the (assumed to be) constant reflectance term, $\mathbf{N} = (\mathrm{N}_x\ \mathrm{N}_y\ \mathrm{N}_z)$ is the normal to the illuminated object, **L** is the 3D incident light direction and A is a constant ambient light term. By assuming constant reflectance, R can be considered as a scale factor and the illumination conditions (**L** and A) can be computed independently from it. Problem (6) can be solved with standard Least Squares, having defined the quadratic cost function:

$$E(\mathbf{L},A) = \left\| \mathbf{M}\begin{pmatrix} L_x \\ L_y \\ L_z \\ A \end{pmatrix} - \begin{pmatrix} I(\mathbf{x}_1) \\ I(\mathbf{x}_2) \\ \dots \\ I(\mathbf{x}_p) \end{pmatrix} \right\|^2 = \left\|\mathbf{Ml} - \mathbf{i}\right\|^2, \quad \mathbf{M} = \begin{pmatrix} N_x(\mathbf{x}_1) & N_y(\mathbf{x}_1) & N_z(\mathbf{x}_1) & 1 \\ N_x(\mathbf{x}_2) & N_y(\mathbf{x}_2) & N_z(\mathbf{x}_2) & 1 \\ \dots & \dots & \dots & \dots \\ N_x(\mathbf{x}_{p1}) & N_y(\mathbf{x}_p) & N_z(\mathbf{x}_p) & 1 \end{pmatrix}$$

$$\Rightarrow \qquad \mathbf{l} = \begin{pmatrix} L_x & L_y & L_z & A \end{pmatrix}^T = \left(M^T M\right)^{-1} M^T \mathbf{i} \tag{7}$$

As in this formulation the light estimation problem requires the knowledge of the surface normal in at least four different points, which is rarely given unless the image geometry is known, the authors propose to perform light direction estimation along occluding contours, reducing to three the number of unknowns. Occluding contours consist in those points where an object hides itself from the viewing direction (e.g. the horizon for the earth). Along occluding contours the surface normal is perpendicular to the viewing angle, and hence $\mathrm{N}_z = 0$. In such situation, problem (6) can be rewritten as:

$$I(\mathbf{x}) = \mathrm{R}\ (\ \mathrm{N}_x(\mathbf{x})\ \ \mathrm{N}_y(\mathbf{x})\ )\ (\ \mathrm{L}_x\ \ \mathrm{L}_y\ )^{\mathrm{T}} + \mathrm{A} \tag{8}$$

Since the surface normal can be computed analytically from the expression of the curve best approximating the occluding contour, and since the intensity values along the contour can be extracted through interpolation, the components L_x and L_y can be estimated by using the least squares solution in (7). The authors also propose a way to regularize the least squares problem in order to relax the constant reflectance assumption and consider local instead of infinitely far away light source, more appropriate for e.g. indoor images.

In a later work [64] the same authors present a more complex lighting environment model, exploiting spherical harmonics for its representation. Assuming that the objects in the scene are bounded by convex and lambertian surfaces, the light environment is expressed as a function $L(\mathbf{V})$ on a sphere, being $\mathbf{V}$ a unit vector in cartesian coordinates. The light received by a surface at any point (irradiance $E(\mathbf{N})$) corresponds to the convolution of the reflectance function of the surface $R(\mathbf{V},\mathbf{N})$ and the lighting environment function, which in turn can be expressed in terms of spherical harmonics. Assuming a linear camera response, image intensity I at point $\mathbf{x}$ can be expressed as:

$$I(\mathbf{x}) = E(\mathbf{N}(\mathbf{x})) = \int_{\Omega} L(\mathbf{V}) R(\mathbf{V},\mathbf{N}) d\Omega \approx \sum_{n=0}^{\infty} \sum_{m=-n}^{n} \hat{r}_n l_{n,m} Y_{n,m}(\mathbf{N}) \quad \textit{with} \quad \hat{r}_n = \sqrt{\frac{4\pi}{2n+1}} r_n \qquad (9)$$

Where r_n is a constant related to the reflectivity, $Y_{n,m}$ is the m$^{\text{th}}$ spherical harmonic of order n. The coefficients $l_{n,m}$ describing the lighting environment can be eventually estimated by truncating the spherical harmonics expansion to e.g. $n = 2$, and solving a least squares problem in 9 unknowns. To further simplify the estimation, again the model can be applied along occluding contours.

Although these approaches provide accurate light direction estimation, the strong assumptions and the so far unresolved problem of automatic occluding contour extraction prevent them from being used in other applicative domains, where large amounts of real-world images are involved.

Johnson and Farid further explored inconsistencies in lighting by detecting composites of people from specular highlights on the eyes [65]. Zhang and others [66] proposed instead a detection technique based on the analysis of shadows geometrical and photometrical properties. Observing that an object illuminated by a point light source and its shadow on a ground plane are related by a planar homology, the authors propose to verify whether all objects in a scene respect the homology constraints. To describe the shadow photometric properties the authors use the histograms of shadow matte values computed along the shadow boundary points. The technique achieves a remarkable accuracy, nonetheless its applicability is limited to images respecting the infinitely far away light source assumption (i.e. either outdoors or artificially rendered), and require manual selection of shadows boundaries.

Section 3.1 reported in details how different steps in image acquisition can leave on the image some sort of fingerprint, which is related to the camera model and to the single camera sensor. These characteristics are permanent: when a splicing of images acquired with two different cameras is performed, the final composite will likely present inconsistent patterns. Lukas et al. [67] propose to detect and localize tampering by analyzing the inconsistencies in the sensor pattern noise extracted from an image. The noise patterns computed for various regions of the image are correlated with the corresponding regions in the camera's reference pattern. A decision is then made based on the comparison of correlation results of the region of interest with those of other regions. The authors conclude that it is possible to obtain reliable forgery detection even after subsequent JPEG compression.

4.2.3. Tampering detection independent on the type of forgery

Most of the techniques described above take advantage of the knowledge on the kind of forgery that compromised the image. A more general approach is presented in this section, which can expose tampering on both single and composite images. The techniques presented below analyze the effects of operations performed independently on the kind of tampering, but which likely take place when images are manipulated. As an example, section 4.2.1 treated malicious image processing, which can be either exploited for explicit tampering purposes (fig.5) or to cover forgery traces [58, 68]. Those techniques are usually applied on a whole image, which could be, in turn, already a composite: for this reason, methods targeting image processing detection can be assimilated to those that will be described in the following.

Three different types of artifacts can be exploited to detect general tampering: traces of re-sampling, compression artifacts and inconsistencies in acquisition device fingerprints.

As already pointed out in section 4.2.1, to render copy-move forgeries more natural, often geometric transforms are applied to the copied regions. This is even truer when splicing is performed, as the geometry of different images can be consistently different. Geometric transforms often involve a re-sampling process, which in turn produces correlation among the image pixels. Popescu and Farid [69] use the Expectation/Maximization algorithm to estimate the probability of each pixel to be correlated with its neighbors: they show how for re-sampled images, these probabilities are arranged in periodic patterns. The detection of such patterns is eventually considered as an evidence of forgery. Detector [69] is further improved by Kirchner in [70] from the computational point of view. Although its effectiveness is proven on regular images, it is observed in [69] that the method is weaker when dealing with heavily compressed images, where further periodicity is caused by quantization effects. In [71], Nataraj et Al. propose the addition of Gaussian noise to JPEG images to suppress JPEG periodic artifacts and enable the detection process. On a similar path, Kirchner and Gloe [72] analyzed the problem of re-sampling detection in case of double compression, often occurring in case of image tampering, showing how actually JPEG blocking artifacts can bring benefit to the eventual tampering detection. It should be noticed that most of the techniques discussed above are weaker in dealing with down-sampling than with up-sampling.

Digital images need to be stored in one of the available image formats for further usage. In real world applications, most of images are saved in JPEG format, being it one of the most widespread compression techniques. As already pointed out in section 3.2.3, the quantization table is strongly dependent both on the device from which the compression is performed and on the quality factor. As a result, it is likely that if an image has been originated e.g. from two different JPEG images, traces of different compressions could be exposed. To this purpose Farid et al. [73] showed how to reveal whether different zones of the image have different “compression histories”. The method is based on the comparison (i.e. the difference) of differently compressed version of the image to the original (possibly tampered) one. When the tampered region is present and has a lower quality factor than its surrounding, the method allows localizing it by detecting spatial local minima, the JPEG ghosts, which appear in correspondence of the forgery.

In a more general scenario, forged images likely undergo a double compression, the first proper of the un-tampered image, and the second applied at the moment in which the result of the forgery is saved. Although double compression is not to be systematically considered as a symptom of forgery, it introduces specific artifacts in the DCT coefficient histograms that represent important cues in tampering detection. Popescu [74] proposed the first theoretical analysis on the formation of these artifacts, mainly due to the effects of double quantization. He observed that the quantization of signal $s[t]$ with step a produces a signal $s_a[t] = q_a(s[t])$, with $q_a(u) = \left\lfloor \frac{u}{a} \right\rfloor$, and that the histogram $H_a(v)$ of $s_a[t]$ is related to that of the original signal $H(u)$ by $H_a(v) = \sum_{k=0}^{a-1} H(av+k)$. This indicates that

- every bin in the range [*av*, *av* + (*a*-1)] contributes to the v^{th} bin of $H_a(v)$
- exactly *a* bins of the original histogram contribute to each bin *v* of the quantized signal.

In the case of double quantization, with step *b* first and step *a* in a second stage, the quantized signal can be expressed as:

$$s_{ab}[t] = q_{ab}(s[t]), \qquad q_{ab}(u) = \left\lfloor \left\lfloor \frac{u}{b} \right\rfloor \frac{b}{a} \right\rfloor \tag{10}$$

The number of bins $n(v)$ of $H(u)$ contributing to the v^{th} bin v of the histogram of the double quantized signal $H_{ab}(v)$ depends now on v itself. Denoting with u_{min}, u_{max} the smallest and the largest values of u that map to v, the author finds:

$$H_{ab}(v) = \sum_{u=u_{\min}}^{u_{\max}} H(u) \quad \text{and} \quad n(v) = u_{\max} - u_{\min} + 1 \tag{11}$$

u_{min} and u_{max} depend on the quantization factors a and b:

$$u_{\min} = \left\lceil \frac{a}{b} v \right\rceil b \qquad u_{\max} = \left\lceil \frac{a}{b}(v+1) \right\rceil b - 1 \tag{12}$$

When $a > b$, typical empty bins are produced in the histogram (fig. 7.C). Conversely, if $b > a$, characteristic periodic patterns are created (fig. 7.D). In both cases, recovering such characteristic patterns is a strong indication of double compression, and of possible tampering. Chen et Al. [74] further improved Popescu's work, exploiting JPEG 2-D array features and a Support Vector Machine to increase the discrimination rates. Popescu's study is also the basis on which Lin and others [76] design their method for tampering localization, analyzing to what extent DCT coefficients randomize the double quantization effect. It is worth to notice that tampering localization represents a major advance in this research area, since simple detection of double quantization does not necessarily imply malicious manipulation. Further methods for double

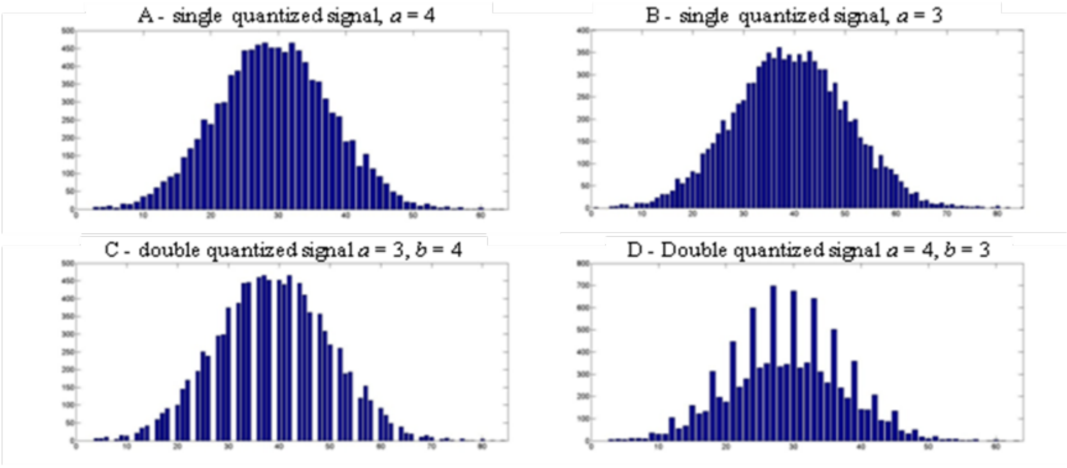


Figure 7 - effects of double quantization on signals

compression recognition are proposed e.g. in [77]. Feng and Doerr extract DCT histogram features to feed both an LDA classifier and an SVM. The system is shown to outperform the Generalized Benford's law descriptors proposed in [78].

From a different perspective, information about an image compression history can be obtained by estimating its primary quantization matrix, to be matched then with the one included in the image header. Lukas and Fridrich in [79] propose a technique to estimate low-frequency coefficients of the primary quantization table. Each image is decompressed and cropped by a few pixels; then, it is re-compressed with a predefined set of candidate matrices. The DCT coefficient histograms of the re-compressed images are compared to that of the original image; the primary quantization table is selected as the one having minimum distance from the original DCT coefficient histogram. Following this seminal work, Luo and others [80] propose another similarity criteria based on the analysis of JPEG rounding errors. The method is shown to be more accurate than that in [79]. However, in both cases it is assumed that the primary quantization matrix belongs to a set of known matrices. Although this is reasonable for most applications, sophisticated forgeries involving non-standard recompression matrices would fool these techniques. This lack is addressed by Pevný and Fridrich in [81].

Tampering also produces inconsistencies in regular patterns created in the image during the acquisition process (see sec. 3.1). Chromatic aberration inconsistencies are examined in [82, 14]. Farid and Johnson [14] propose a computational technique for automatically estimating lateral chromatic aberration based on the model described in section 3.2.1, maximizing the mutual information between color channels, and show its efficacy in detecting digital tampering. The method works well for un-compressed non-uniform parts of the image, but for the uniform regions or typical JPEG images, the results are weaker. Alternatively, Popescu and Farid [16] argue that tampering would likely break the periodicity produced by CFA interpolation. Based on the expectation/maximization (EM) algorithm, they analyze the CFA patterns of separate regions of the image, showing the soundness of their method. In [83], Dirk et al. use two features to analyze traces of CFA. This method is successful for both tampering detection and discrimination between real and synthetic images, although partially failing for stronger JPEG compression. On the other hand, the method is able to not only detect but also to localize the doctored parts of the image with a good precision.

5. A new Phase: Counter-Forensics

The methods discussed in the previous sections testify the effort and the indubitable high level of expertise reached by researchers in blind detection of malicious image manipulations. However, as well as for other security-related disciplines, improvements in defense and investigation techniques stimulate the burglars in the design of new, powerful attacks. As a consequence, lately a growing interest has been shown around the creation of undetectable forgeries and manipulations.

In the field of forensic sciences, countermeasures to the investigation activities are known under the name of counter-forensics or anti-forensics. Harris [84] defines anti-forensics techniques as “any attempt to compromise the availability or usefulness of evidence to the forensic process”. Under this interpretation, the simple wiping-off of fingerprints from a crime scene can be considered as a counter-forensic act. In a similar way, multimedia counter-forensics involves all those means that allow covering traces of image manipulation, or, more precisely, to make manipulation invisible to the existing detection methods. As pointed out by Böhme and others [4], the public availability of most of the research in the field gives burglars a valid support to develop attacks targeted on a specific tool. Since it cannot be excluded that forgers are also expert in digital image processing, digital image forensics reliability is now becoming questionable.

Although it is common practice of authors to remark the limitations of their methods [45, 76], a new series of studies was recently published pointing out the weaknesses of existing detectors. As an example, in [85] Wang and others propose a probabilistic analysis of Generalized Benford's Law, which was previously proposed for tampering detection in [78] (cfr. Sec. 4.2.3). The authors use histogram equalization in the log-domain to restore the First Significant Digit distribution so that it follows Benford's law also after manipulation. It is then shown how the discriminative power of the Benford's law-based detector decreases when simple histogram manipulation is performed as a compensation for tampering. This study represents a first attempt of tamper hiding. More sophisticated techniques have been published lately, targeted either to specific detectors or more general detection methods. In a similar way, a few attempts have been made to counterfeit the image source device fingerprint. It should be noticed that the terms hiding and counterfeiting reflect Harris' classification of counter-forensic measures. As image tampering traces cannot be destroyed or removed, but at most covered, we propose to classify the techniques described in section 5.1 as evidence hiding techniques. Section 5.2 reports instead on methods that suppress image fingerprints to replace them with other, synthesized ones. This again fits with Harris' definition of counterfeiting counter-forensics, where evidence is replaced with a misleading one. Finally, section 5.3 reports on the first examples of counter-measures to counter-forensics recently introduced to the community.

5.1. Tamper hiding

In [86], Kirchner and Bohme introduce a technique to hide traces of region re-sampling. The study is targeted on Popescu and Farid's tampering detector (see sec. 4.2.3, [69]), which exposes forgeries by detecting the linear dependencies that re-sampling typically induces among pixels. To break these linear dependencies, post-processing the image with a median (i.e. non linear) filter is shown to be effective. The authors propose to use it in conjunction with geometric distortions, to be performed on high frequency components, for a further decrease of the forensic detectability of re-sampling.

An interesting point raised by Kirchner and Bohme is that concerning the visual noticeability of tamper hiding. Both median filtering and geometric attacks can bring about visual artifacts and compromise the eventual image quality. As a consequence, counter-forensic techniques should also be evaluated according to their impact on visual quality. To this purpose, an adaptive method to reduce jittering in the final image is proposed, which reduces the strength of geometric attacks along edges, obtaining a good trade-off between low visual and forensic detectability.

In a later study [87], the same authors target the CFA-based tampering detector of Popescu and Farid [16]. The method aims at restoring, after tampering, those periodic patterns that detector [16] seeks to establish the genuineness of an image. The authors formulate this problem as a Least Squares one, assuming bilinear interpolation. CFA artifacts in the image channel $\hat{\mathbf{y}}$ can be described as being originated through the linear equation $\hat{\mathbf{y}} = \mathbf{Hc}$, where $\mathbf{H}$ is the matrix of interpolation weights and $\mathbf{x}$ is the 2D intensity lattice. In tampered images the typical CFA pattern will deviate from the original one according to the model $\mathbf{y} = \mathbf{Hc} + \xi$. Hence, to obtain minimal distortion in the CFA synthesis, the signal $\mathbf{x}$ should be chosen so that it minimizes $\|\xi\| = \|\mathbf{y} - \hat{\mathbf{y}}\|$. The solution of this standard least squares problem is given by: $\mathbf{c} = (\mathbf{H'H})^{-1}\mathbf{H'y}$, and the restored pattern can be derived by the distorted one as a pre-filtering operation:

$$\hat{\mathbf{y}} = \mathbf{H}\left((\mathbf{H'H})^{-1}\mathbf{H'y}\right) \tag{12}$$

The authors provide computationally feasible methods to estimate the needed pre-filtering coefficients to restore $\hat{\mathbf{y}}$. The approach is proven to be effective, also in terms of quality preservation; nonetheless, the authors point out that CFA interpolation is not the last step in the image acquisition process, and further processing might leave further traces. An investigator retrieving convincing CFA artifacts in contrast with other tampering traces might eventually assume it as further proof of manipulation.

Stamm and others in [88] attempt at covering traces of previous compressions by modifying the distribution of the DCT coefficients before saving the tampering results. After compression, DCT coefficients are clustered around the integer multiples of the quantization step, as discussed in sec. 4.2.3. Masking previous compression does not aim at recovering the exact original distribution of these coefficient, but rather attempts at restoring a plausible distribution for an uncompressed image. The authors accomplish this task by perturbing quantized coefficient with additive white noise. The noise perturbation depends upon the value of the DCT coefficient and is designed to eventually convey the estimated un-quantized DCT coefficient distribution. The method is shown to defeat the technique proposed in [89], which attempts at detecting previous compression in bitmap images, but no validation was performed in case of double compression, e.g. dealing with methods [78] or [81].

5.2. Image source counterfeiting

The first attempt at fooling source device identification techniques was proposed by Gloe and others in [90]. The study proposes an attack to the identification method [24], which is based on the extraction and the analysis of the camera pattern noise (cfr. Sec. 3.2.2). As a countermeasure to it, the authors propose to use flat-fielding to estimate both the FPN and the PRNU for the device of interest. The FPN can be estimated as a dark frame **d** by averaging a sufficiently large number of images captured in a dark environment. The PNRU (**K**) is estimated from a set of images captured in a homogeneously illuminated environment, from which **d** is subtracted to suppress FPN. Since FPN is additive and PNRU is multiplicative, pattern noise can be eventually suppressed from image **x** by:

$$\mathbf{J} = \frac{\mathbf{I} - \mathbf{d}}{\mathbf{K}} \tag{13}$$

In a similar way, pattern noise can be estimated from another device in terms of ($\mathbf{d}_e$, $\mathbf{K}_e$) and forged into the polished images to counterfeit the camera signature (inverse flat-fielding): $\mathbf{J}_{cont} = \mathbf{J} \cdot \mathbf{K}_e + \mathbf{d}_e$. Inverse flat-fielding is shown to suppress quite well the original pattern noise, when experimented on different acquisition device models. On the other hand, no verification has been made using distinct exemplars of the same camera model.

A different method for counterfeiting image source fingerprints is proposed in [91], and is based on the manipulation of JPEG artifacts. The authors present a technique complimentary to [88] for the suppression of traces of compression and then re-forge the image by compressing it with a different quantization table, e.g. characteristic of another device. While method [88] restores a plausible histogram for unquantized images, it does not eliminate visual compression artifacts, i.e. blockiness. For this reason, technique [91] cleans blocking artifacts by applying first a median filter and then adding low-power white noise to the image. It is shown that this operation outperforms existing de-blocking techniques when coupled to method [88] for compression hiding. “Polished” images are then recompressed with quantization tables typical of other cameras, and their quantization matrices estimated through method [89]. Almost no original quantization table is shown to be recovered. Although this technique seems to be promising, again it was not tested on more recent methods dealing with double compression, or with the device identification method proposed by Farid in [30].

5.3. Countering counter-forensics

Even if someone could think that the development of counter-forensics is potentially dangerous for multimedia security, the examples that we explored in the previous sections are not meant to defeat existing forensics tools. Rather, researchers attempted to point out the weaknesses of their counterparts to promote further improvements in the field, and to encourage the community to keep on developing this discipline, as existing tool cannot be considered completely safe. This is indeed a common practice in security-related disciplines, e.g. watermarking or steganography.

The efforts served the purpose: in very recent days, the first examples of counter-counter-forensics have been proposed to the community, as direct answers to specific counter-forensics techniques. Kirchner and Fridrich [68] targeted the median filter attack perpetrated in [86], also arguing that since many detectors rely on the detection of linearities, the use of median filtering can be a general symptom of tampering. The authors propose to inspect image by measuring streaking artifacts, typically originated from the application of median filters. The method is shown to be successful for uncompressed images, but fails when images are compressed after filtering. Therefore, the authors propose the use of *SPAM* features to strengthen the analysis and feed a SVM, which reliably detects median filtering also at medium compression qualities.

A second study [92] addresses the counterfeiting pattern noise attack [91]. In this case, plausible scenarios are studied, assuming that for the burglar to synthesize some specific camera pattern noise, he/she should have access at least to some images that are also accessible to the investigator. The authors propose therefore a “triangle test” which takes advantage of the availability of both the (possibly) counterfeit image and of the knowledge of the images from which the camera fingerprint might have been forged. Results indicate that traces of the original pattern noise remain in the forged image, and can still be exposed given the availability of the mentioned material.

6. Conclusions

The techniques that we reviewed in this survey represent important results for multimedia security, especially considering that the problems they tackle were previously (almost) unexplored. A large set of tools is now available to investigate on image sources and to authenticate acquisition devices. Among them, tools that analyze pattern noise were proven to be promising for identifying even different exemplars of the same device model. An even larger number of techniques have been developed to detect image tampering, some of which are also able to localize the forged areas.

Despite these achievements, major challenges remain still open for Digital Image Forensics. A first main concern is the robustness of the existing tools. Apart from [62], no real comparative study exists which evaluates the actual accuracy of DIF methods. This is mainly due to the lack of established benchmarks and of public testing databases. Few efforts in this direction [8, 42] have been already done; however, a large dataset comprehensive of different scenes, illumination and environmental conditions and attacks is still missing for, e.g., tampering detection. Given the maturity of the domain, it is reasonable to assume that soon new databases and comparative studies will appear, as well as public competitions, similar to the BOSS challenge for steganalysis [93]. Such an evolution is desirable both for improving communication between researchers and for better establish the actual state-of-the-art in DIF.

Also, the rise of counter-forensics contributes in exposing the limitations of DIF tools. Hence, confirming or strengthening the robustness of DIF techniques is a present priority for DIF experts. Furthermore, like every other security-related discipline, image forensics evolves based on the attacks perpetrated to the existing techniques, as discussed in section 5. Therefore, the development of counter-forensics is to be encouraged, and can represent an appealing domain for new researchers.

From another perspective, a future challenge for Digital image Forensics is the extension to other media, and in particular to video. Videos are even more powerful vectors than images in communication. Moreover, advances in video surveillance-related technologies (such as camera networks or covert video recordings) will likely enable a massive usage of digital video evidence in trials. More and more sophisticated forgery techniques will threaten the reliability of this material. As a consequence, future efforts in DIF should be also addressed towards video authentication.

Finally, perhaps the major challenge in the future of image forensics consists in integrating it with visual perception. This is not only about the joint use of human and automatic inspection of visual media. On a broader perspective, understanding the perception of visual semantics might lead to the solution of one of the main limitations of current DIF techniques: the distinction between malicious tampering and “innocent” retouching, such as red-eye correction or artistic manipulation. If in the case of the fake knighthood of Jeffrey Wong Su En (fig. 1) the malicious purpose was evident, in many occasions manipulation can be performed to e.g. augment the beauty of an image, as in the case of models re-touching in advertisement. The same case of the TIME magazine cover (fig. 5) was claimed to be an erroneous interpretation of an artistic editing. The line between evil and naïve purposes is often fuzzy. However, some manipulations bring a more critical impact on the semantic content of the image, and therefore on the viewer.

Most of the existing tools are only able to establish the non-authenticity of the image. No investigation about the kind of alteration of the message conveyed by the medium is performed, not to mention the motivation of the tampering. Of course, the human factor can in many cases provide a correct interpretation of the detected forgery. Nonetheless, a joint effort of researchers in media semantics, visual perception and media security fields might produce interesting results, not only from the strict forensic investigation point of view. To some extent, the work of De Rosa and others [94] can be considered a first step in this direction. The authors propose a theoretical framework to retrieve (semantic) dependencies among groups of images. The work is so far limited to detecting whether images have been generated from others. However, in a long term perspective, image dependencies and the analysis of the mutation of their semantic content across these dependencies might bring useful information about e.g. their owner or the role of different websites and the habits of their visitors. In a similar way, we can envision extensive forensics activity on social networks or video databases.

References

[1]. Farid H (2006) Digital Doctoring: how to tell the real from the fake. Significance: statistics making sense, vol. 3, no. 4 pp.162-166

[2]. Photo tampering throughout history. http://www.cs.dartmouth.edu/farid/research/digitaltampering/. Accessed 3 July 2010

[3]. Farid H (1999) Detecting Digital Forgeries Using Bispectral Analysis. Technical Report, AIM-1657, MIT AI Memo

[4]. Böhme R, Freiling F, Gloe T, and Kirchner M (2009) Multimedia forensics is not computer forensics. In Proc. International workshop on Computational Forensics, IWCF 2009, LNCS 5718

[5]. Katzenbeisser S, Petitcolas F (1999) Information hiding techniques for steganography and digital watermarking. Artech House 1st ed. , 10/1999, p. 240 - ISBN : 1580530354

[6]. Rey C, Dugelay J.-L (2002) A survey of watermarking algorithms for image authentication. EURASIP Journal on applied Signal Processing, Volume N°6 - June 2002, pp 613-621

[7]. Cox I-J, Miller M-L, Bloom J-A (2002) Digital Watermarking. San Francisco, CA: Morgan Kaufmann

[8]. Gloe T and Bohme R (2010) The *Dresden Image Database* for Benchmarking Digital Image Forensics. SAC 10 March 22-26, 2010, Sierre, Switzerland

[9]. Adams J, Parulski K, Sapulding K (1998) Color Processing in Digital Cameras. IEEE Micro, Vol. 18, No. 6

[10]. Cao H, Kot A. C (2009) accurate Detection of Demoisaicing Regularity for digital Image Forensics, IEEE transaction on information forensics and security, val 4

[11]. Choi K-S, Lam E-Y, Wong K. K. Y (2006) Source Camera Identification Using Footprints from Lens Aberration. Proc. of SPIE

[12]. Devernay F, Faugeras O (1995) Automatic calibration and removal of distortion from scenes of structured environments. In Proc. SPIE Investigative and Trial Image Processing, vol. 2567, pp. 62–67.

[13]. Van L. T, Emmanuel S, Kankanhalli M (2007) Identifying Source Cell Phone using Chromatic Aberration. In Proc. IEEE International Workshop on Multimedia & Expo, ICME 2007

[14]. Johnson M, Farid H (2006) Exposing digital forgeries through chromatic aberration. ACM Multimedia and Security Workshop, Geneva, Switzerland

[15]. Bayram S, Sencar H. T, Memon N, Avcıbaş I (2005) Source camera identification based on CFA interpolation. In Proc. of ICIP, vol. 3, pp. III-69-72

[16]. Popescu A. C, Farid H (2005) Exposing digital forgeries in color filter array interpolated images. IEEE Trans. on Signal Processing, vol. 53 (10), pp. 3948-3959

[17]. Bayram S, Sencar H-T, Memon N (2006) Improvements on source camera-model identification based on CFA interpolation. In Proc. WG 11.9 Int. Conf. on Digital Forensics

[18]. Avcıbaş I, Memon N, Sankur B (2003) Steganalysis using Image Quality Metrics. IEEE Transactions on Image Processing

[19]. Lyu, Farid H (2002) Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines. Proc. of Information Hiding Workshop

[20]. Long Y, Huang Y (2006) Image based source camera identification using demosaicing. In Proc. of MSP, pp. 419-424

[21]. Celiktutan O, Avcıbaş I, Sankur B, Memon N (2005) Source Cell-Phone Identification. Proc. of ADCOM

[22]. Holst G-C (1998) CCD Arrays, Cameras, and Displays, Second Edition. JCD Publishing & SPIE Press, USA

[23]. Geradts, Z, Bijhold, J, Kieft, M, Kurosawa K, Kuroki K, Saitoh, N (2001) Methods for identification of Images Acquired with Digital Cameras. Proc. of SPIE, Enabling Technologies for Law Enforcement and Security, vol. 4232, pp. 505–512

[24]. Lukáš J, Fridrich J, Goljan M (2006) Digital camera identification from sensor pattern noise. IEEE Transactions on Information Forensics and Security, 1(2):205–214

[25]. Delp E.J, Khanna N, Mikkilineni A.K (2009) Forensic Camera Classification: Verification of Sensor Pattern Noise Approach. Proceeding of the SPIE international Conference on Security

[26]. Chen M, Fridrich J, Goljan M, Lukas J (2008) Determining Image Origin and Integrity Using Sensor Noise. IEEE Transactions on Information Forensics and Security, volume 3, no.1, pp.74-90

[27]. Chen M, Fridrich J, Goljan M (2007) Digital imaging sensor identification (further study). In Proc. SPIE International Conference on Security, Steganography, and Watermarking of Multimedia Contents IX, vol. 6505, no. 1

[28]. Goljan, M, Fridrich, J (2008) Camera Identification from Cropped and Scaled Images. Proc. SPIE Electronic Imaging, Forensics, Security, Steganography, and Watermarking of Multimedia Contents X, vol. 6819

[29]. Li C. T (2010) Source camera identification using enhanced sensor pattern noise. IEEE Transactions on Information Forensics and Security , volume 5, no.2

[30]. Kharrazi m, Sencar H. T, Memon N. D (2004) Blind source camera identification. In Proc. ICIP’04, pages 24–27

[31]. Tsai M.-J, Wu G-H (2006) Using Image Features to Identify Camera Sources. Proc. Of IEEE ICASSP

[32]. Farid H (2006) Digital image ballistics from JPEG quantization. Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2006-583

[33]. Kee E, Farid H (2010) Digital Image Authentication from Thumbnails. SPIE Symposium on Electronic Imaging, San Jose, CA

[34]. The Oxford dictionary online. http://oxforddictionaries.com/ Accessed 7 July 2010

[35]. Wang J, Cohen M-F (2007) Image and video matting: a survey. Found. Trends. Comput. Graph. Vis., 3(2):97–175

[36]. Pérez P, Gangnet M, Blake A (2003) Poisson image editing. ACM Transactions on Graphics (SIGGRAPH’03), 22(3):313-318

[37]. Bertalmio M, Sapiro G, Ballester C, Caselles V (2000) Image inpainting. In Proc. Computer Graphics, SIGGRAPH 2000, 417–424

[38]. Rubinstein M, Shamir A, Avidan S (2008) Improved Seam Carving for Video Retargeting. SIGGRAPH 2008

[39]. Avidan S, Shamir A (2007) Seam carving for content-aware image resizing. ACM Transactions on Graphics (TOG), v.26 n.3

[40]. Farid H (2006) Exposing digital forgeries in scientific images. In Proc. ACM workshop on Multimedia and security, Geneva, Switzerland, pp. 29-36

[41]. Bregler C, Covell M, and Stanley M (1997) Video rewrite: Driving visual speech with audio. In Computer Graphics Proceedings, Annual Conference Series. ACM SIGGRAPH, 1997

[42]. Ng T-T, Chang S-F (2004) A data set of authentic and spliced image blocks. Columbia University technical report, Available: http://www.ee.columbia.edu/trustfoto

[43]. Calphoto (2000) A database of photos of plants, animals, habitats and other natural history subjects. University of Berkeley, Available: http://elib.cs.berkeley.edu/photos/

[44]. Fridrich J, Soukal D, Lukas J (2003) Detection of Copy-Move Forgery in Digital Images. In Proceedings of Digital Forensic Research Workshop

[45]. Popescu C and Farid H (2004) Exposing Digital Forgeries by Detecting Duplicated Image Regions. Technical Report, TR2004-515, Department of Computer Science, Dartmouth College

[46]. Li G, Wu Q, Tu D, Sun S (2007) A Sorted Neighborhood Approach for Detecting Duplicated Regions in Image Forgeries based on DWT and SVD. In Proceedings of IEEE ICME, Beijing China, pp. 1750-1753.

[47]. Langille A, Gong M (2006) An Efficient Match-based Duplication Detection Algorithm. Proc. CRV

[48]. Mahdian B, Saic S (2010) Blind Methods for Detecting Image Fakery. IEEE Aerospace and Electronic Systems Magazine, 25:4(18-24)

[49]. Dybala B, Jennings B, Letscher D (2007) Detecting filtered cloning in digital images. Proceedings of the 9th workshop on Multimedia & security, Dallas, Texas, USA

[50]. Huang H, Guo W, Zhang Y (2008) Detection of copy-move forgery in digital images using SIFT algorithm. IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application

[51]. Bayram S, Sencar T, Memon N (2009) An efficient and robust method for detecting copy-move forgery. In Proc. IEEE ICASSP

[52]. Amerini I, Ballan L, Caldelli R, Del Bimbo A, Serra G (2010) Geometric tampering estimation by means of a SIFT based forensic analysis. In Proc. IEEE ICASSP

[53]. Pan X, Lyu S (2010) Detecting image duplication using SIFT features. In. Proc. IEEE ICASSP

[54]. Sarkar A, Nataraj L, Manjunath B-S (2009) Detection of Seam Carving and Localization of Seam Insertions in Digital Images. Proc. of 11th ACM Workshop on Multimedia and Security, Princeton, New Jersey

[55]. Fillion B, Sharma G (2010) Detecting content adaptive scaling of images for forensic applications. In Proc. SPIE: Media Forensics and Security XII, vol. 7541, San Jose, CA, pp. 7541–36

[56]. Wu Q, Sun S-J, Zhu W, Li G-H, Tu D (2008) Detection of digital doctoring in exemplar-based inpainted images. In Machine Learning and Cybernetics, 2008 International Conference on, volume 3

[57]. Avcıbaş, Bayram S, Memon N, Ramkumar M, Sankur B (2004) A classifier design for detecting image manipulation. Proc. of IEEE ICIP

[58]. Stamm M, Liu J-R (2010) Forensic Detection of Image Manipulation Using Statistical Intrinsic Fingerprints. IEEE Trans. On Information Security and Forensics, vol. PP:99

[59]. Ng T-T, Chang S-F, Sun Q (2004) Blind Detection of Photomontage Using Higher Order Statistics. IEEE International Symposium on Circuits and Systems, Vancouver, Canada

[60]. Fu D, Shi Y. Q, Su W (2006) Detection of image splicing based on Hilbert– Huang transform and moments of characteristic functions with wavelet decomposition, in: International Workshop on Digital Watermarking, Jeju, Korea, pp. 177–187.

[61]. Chen W, Shi Y.Q, Su W (2007) Image splicing detection using 2-d phase congruency and statistical moments of characteristic function. SPIE Electronic Imaging: Security, Steganography, and Watermarking of Multimedia Contents, San Jose, CA, USA

[62]. Shi Y.Q, Chen C, Chen W (2007) A natural image model approach to splicing detection. ACM Workshop on Multimedia and Security (ACM MMSEC07), ACM, New York, NY, USA, pp. 51–62.

[63]. Johnson M and Farid H (2005) Exposing digital forgeries by detecting inconsistencies in lighting. Proc. ACM Multimedia and Security Workshop, New York, NY

[64]. Johnson M, Farid H (2007) Exposing digital forgeries in complex lighting environments. IEEE Transactions on Information Forensics and Security 3 (2) 450–461.

[65]. Johnson M, Farid H (2007) Exposing digital forgeries through specular highlights on the eye. 9th International Workshop on Information Hiding, Saint Malo, France

[66]. Zhang W, Cao X, Zhang J, Zhu J, Wang P (2009) Detecting photographic composites using shadows. IEEE International Conference on Multimedia and Expo, pp. 1042–1045.

[67]. Lukas J, Fridrich J, Goljan M (2006) Detecting Digital Image Forgeries Using Sensor Pattern Noise. Proc. of Security, Steganography, and Watermarking of Multimedia Contents VIII, part of EI SPIE 2006, San Jose, CA, USA

[68]. Kirchner M, Fridrich J (2010) On Detection of Median Filtering in Images. Proc. SPIE, Electronic Imaging, Media Forensics and Security XII, San Jose, CA, January 17–21, pp. 10-1 - 10-12, 2010

[69]. Popescu A, Farid H (2005) Exposing digital forgeries by detecting traces of re-sampling, IEEE Transactions on Signal Processing 53 (2) 758–767.

[70]. Kirchner M (2008) Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue. Proceedings of the 10th ACM workshop on Multimedia and security, ACM, New York, NY, USA,pp. 11–20.

[71]. Nataraj L, Sarkar A, Manjunath B-S (2010) Improving Re-sampling Detection by Adding Noise. Proc. SPIE Electronic Imaging, Media Security and Forensics

[72]. Kirchner M, Gloe T (2009) On Resampling Detection in Re-compressed Images. First IEEE Workshop on Information Forensics and Security (WIFS 2009), pp. 21–25

[73]. Farid H (2009) Exposing digital forgeries from jpeg ghosts, IEEE Transactions on Information Forensics and Security

[74]. Popescu A. C (2005) Statistical tools for digital image forensics, Ph.D. Thesis, Department of Computer Science, Dartmouth College, Hanover, NH

[75]. Chen C, Shi Y. Q, Su W (2008) A machine learning based scheme for double JPEG compression detection. Proc. IEEE ICPR

[76]. Lin Z, He J, Tang X, Tang C.K (2009) Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis. Pattern Recognition

[77]. Feng X, Doerr G (2010) Jpeg recompression detection, in: SPIE Conference on Media Forensics and Security

[78]. Fu D, Shi Y. Q, Su W (2007) A generalized Benford's law for jpeg coefficients and its applications in image forensics. SPIE Electronic Imaging: Security, Steganography, and Watermarking of Multimedia Contents, San Jose, CA, USA

[79]. Lukas J, Fridrich J (2003) Estimation of primary quantization matrix in double compressed JPEG images. In Proc. of DFRWS

[80]. Luo W, Huang J, Qiu G (2010) JPEG Error Analysis and Its Applications to Digital Image Forensics. IEEE Transactions on Information Forensics and Security

[81]. Pevný T, Fridrich J (2008) Detection of Double-Compression for Applications in Steganography. *IEEE Transactions on Information Security and Forensics*, **3**(2), pp. 247-258

[82]. Gloe T, Winkler A, Borowka K (2010) Efficient estimation and large-scale evaluation of lateral chromatic aberration for digital image forensics. SPIE Conference on Media Forensics and Security

[83]. Dirik A. E, Memon N (2009) Image tamper detection based on demosaicing artifacts. ICIP (09), Cairo, Egypt, pp. 429–432

[84]. Harris R (2006) Arriving at an anti-forensics consensus: Examining how to define and control the anti-forensics problem," Digital Investigation 3(Supplement 1), pp. 44–49

[85]. Wang J, Cha B, Cho S, Jay Kuo C.C (2009) Understanding Benford's law and its vulnerability in image forensics, in Proc. ICME

[86]. Kirchner M, Böhme R (2008) Hiding traces of resampling in digital images. IEEE Transactions on Information Forensics and Security 3(4), 582–592

[87]. Kirchner M, Böhme R (2009) Synthesis of color filter array pattern in digital images" In Proceedings of SPIE-IS&T Electronic Imaging: Media Forensics and Security XI, vol. 7254, p. 725421

[88]. Stamm M. C, Tjoa S. K, Lin W, Liu K.L (2010) Anti-forensics of JPEG compression. Proc. IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP)

[89]. Neelamani R, de Queiroz R.L, Fan Z, Dash S, Baraniuk R.G (2006) Jpeg compression history estimation for color images, IEEE Transactions on Image Processing 15 (6)1365–1378

[90]. Gloe, T, Kirchner M, Winkler A, Böhme R. (2007) Can we trust digital image forensics? Proc. 15th international conference on Multimedia, September 24–29, 2007, pp. 78–86. ACM Press, New York

[91]. Stamm M.C, Tjoa S.K, Lin W. S, Liu K. J. R (2010) Undetectable Image Tampering Through JPEG Compression Anti-Forensics. IEEE Int'l Conf. Image Processing (ICIP), Hong Kong, China

[92]. Goljan M, Fridrich J, Chen M (2010) Sensor Noise Camera Identification: Countering Counter-Forensics. Proc. SPIE, Electronic Imaging, Media Forensics and Security XII, San Jose, CA, January 17–21, pp. 0S-01 - 0S-12, 2010.

[93]. BOSS – Break Our Steganographic System, http://boss.gipsa-lab.grenoble-inp.fr/Warming/

[94]. De Rosa A, Uccheddu F, Costanzo A, Piva A, Barni M (2010) Exploring image dependencies: a new challenge in image forensics. Proc. SPIE 7541