MULTIMEDIA FORENSIC TECHNIQUES FOR ACQUISITION DEVICE IDENTIFICATION AND DIGITAL IMAGE AUTHENTICATION

Authors: *Roberto Caldelli, Irene Amerini, Francesco Picchioni, Alessia De Rosa and Francesca Uccheddu*

Abstract

Multimedia forensics can be defined as the science that tries, by only analysing a particular digital asset, to give an assessment on such a content and to extract information that can be useful to address and support an investigation linked to the scene represented in that specific digital document. The basic idea behind multimedia forensics relies on the observation that both the acquisition process and any post-processing operation leave a distinctive imprint on the data, as a sort of digital fingerprint. The analysis of such a fingerprint may permit to determine image/video origin and to establish digital content authenticity.

Introduction

Digital crime, together with constantly emerging software technologies, is growing at a rate that far surpasses defensive measures. Sometimes a digital image or a video may be found to be incontrovertible evidence of a crime or of a malevolent action. By looking at a digital content as a digital clue, Multimedia Forensic technologies are introducing a novel methodology for supporting clue analysis and providing an aid for making a decision on a crime. Multimedia forensic researcher community aimed so far at assisting human investigators by giving instruments for the authentication and the analysis of such clues. To better comprehend such issues let firstly introduce some application scenarios. Let’s imagine a situation in which the action itself of creating a digital content (e.g. a photograph) implies an illegal action related to the content represented in the data (e.g. child pornography). In such a case, tracing the acquisition device that took that digital asset, can lead the judge to blame the owner of the "guilty" device for that action. Forensic techniques can help in establishing the origin/source of a digital media, making the "incriminated" digital content a valid, silent witness in the court. A similar approach can be used in a different circumstance, in which a forensic analysis can help the investigator to distinguish between an original multimedia content and an illegal copy of it. Different types of acquisition devices can be involved in this scenario, from digital cameras, scanners, cell-phones, PDAs and camcorders till photorealistic images or videos created with graphic rendering software. In this context, the possibility of identifying how that digital document was created may allow to detect illegal copy (e.g. digital cinema video recaptured by a camcorder). A more insidious digital crime is the one that attempts to bias the public opinion through the publication of tampered data. Motivations can spread from joking (e.g. unconvincing loving couple), to changing the context of a situation in which very important people are involved, or to exaggerating/debasing the gravity of a disaster image. Image forensic techniques can give a support in recognizing if, how and possibly where the picture has been forged.

Forensic tools work without any added information, the only features that can be evaluated are the ones intrinsically tied to the digital content. The basic idea behind multimedia forensic analysis relies on the observation that both the acquisition process and any post-processing operation leave a distinctive imprint on the data, as a sort of digital fingerprint. The estimation of such fingerprints really suggests how to evaluate the digital clue, turning it into an actual evidence.

It is the aim of this chapter to present the principles and the motivations of digital forensics (i.e. concerning images and videos), and to describe the main approaches proposed so far for facing the two basic questions: a) what is the source of a digital content? b) is such a digital content authentic or not? The chapter will be organized as it follows. The first section will introduce the reader to the basics of multimedia forensics; the different approaches for obtaining information from a digital content will be presented, as well as the diverse type of digital data that can be usually analyzed; then, the possible application scenarios that can benefit from forensic techniques will be described and an overview over the intrinsic digital fingerprints will be presented. The second and the third sections will be devoted to the analysis of the principal techniques exploited respectively for identifying the acquisition device of digital images and videos, and for assessing the authenticity of digital images. Future trends will be suggested and some conclusions will be provided in the last sections. Bibliographic references will complete the chapter.

Multimedia forensics: principles and motivations

Multimedia forensics can be defined as the science that tries, by analysing a digital asset, to give an assessment on such a content and to extract information that can be useful to address and support an investigation linked to the scene represented in that specific digital document. Multimedia forensics has to be able to develop efficient instruments to deal with the disparate digital devices that can generate images and, above all, with the different processing tools that allows also an unskilled user to manipulate digital goods. Hereafter two basic approaches are introduced, then the various kinds of data that multimedia forensic tools could have to face with are presented. After that, some possible application scenarios where these technologies could be claim to operate are described and finally a wide look to which are the possible digital fingerprints to be searched for in a multimedia content is given.

Possible approaches

When digital images (videos) had to be protected or their authenticity verified or, furthermore, their provenance tracked, the solution generally was to insert in the original data an embedded, usually unperceivable, information that permitted afterwards to determine what was happened, in which part of the content and, in particular application cases, by whom. This kind of techniques that can be grouped under the name of *digital watermarking* (Barni, 2004), follow an "active" approach, that is it is necessary to operate on the original document which has to be available from the beginning: this requirement is almost always hard to be satisfied. Embedding a watermark into an image, for instance, (see Figure 1) can be accomplished by applying some specific slight modifications to the original document I according to the information contained in the watermark W and ,often, to a private key K; after that the watermarked content I_W is obtained.

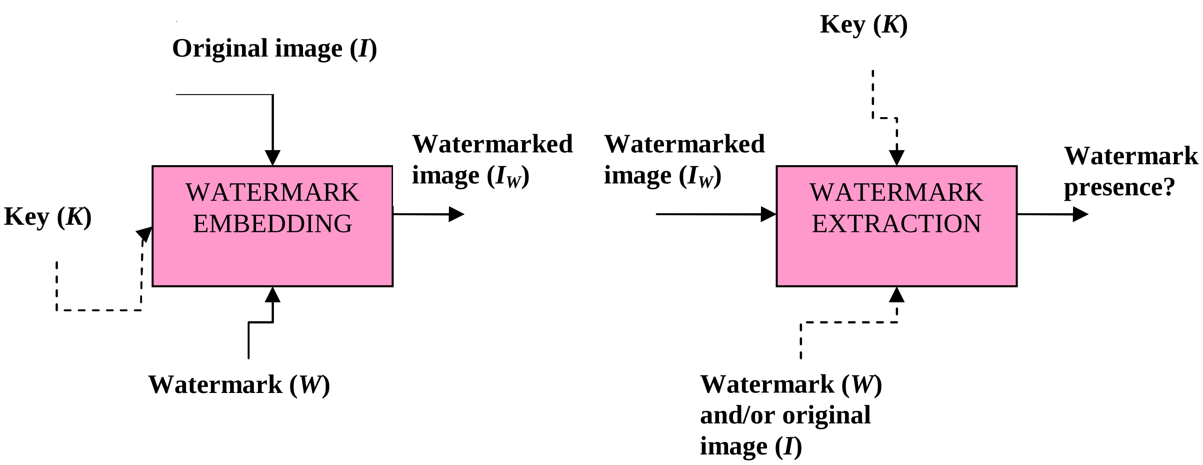

Figure 1. Watermark embedding phase (left) and watermark extraction (right).

If an assessment has to be performed to check if something has happened on the watermarked image, the detection phase is carried out by passing it, together with the private key *K* (if the algorithm is not blind the original image is needed too), to the detector that give an answer by re-extracting the watermark *W* or by comparing a verification parameter with a certain threshold.

For sake of completeness, also the cryptographic approach should be included within "active" method category. Such an approach uses digital signature for verifying author and time of signature and authenticating message contents. A digital signature is achieved by calculating a digest of the digital data by means of a hash function and encrypting it with a private key; such a signed digest is stored together with the image and can be used to prove data integrity or to trace back to its origin. There are some intrinsic weaknesses in this cryptographic approach. Firstly, the signal digest has to be tied to the content itself, e.g. by defining a proper format, and this makes impossible to use a different format, or to authenticate the data after D/A conversion. Secondly, the digest changes as soon as any modification is applied to the signal, making impossible to distinguish malicious versus innocuous modifications. Finally, cryptographic authentication usually does not allow a precise localization of tampering (Menezes,1998).

It is easy to understand that such a-posteriori evaluation can not be performed, for instance, on a common digital content obtained through the Internet (e.g. a video posted on YouTube, an image published on a newspaper web-site and so on). This kind of "active" technologies (Blythe, 2004) can be adopted to manage data in a specific application context where additional information casting is feasible but are not able to deal with an open operative environment in which only a detection step is possible.

On the contrary, in this situation a "passive" methodology would be useful; with the term "passive" an approach which tries to make an assessment only having the digital content at disposal is to be intended. It is straightforward to realize that this kind of investigation is harder and has to be founded on the thorough analysis of some intrinsic features that should have/have not been present and are not/are now recognizable inside the observed data (Popescu, 2004 a). For sake of clarity: when a photomontage, for instance, has been performed to alter the content of a digital photo, to change the meaning of the represented scene, some traces of this operation are left somehow over the "new fake" image. These traces, although unperceivable, can result in the modification of the image structure such as anomalous pixel values (e.g. sequential interpolated values or strange continuous flat values) but also in inconsistencies within the image content itself such as anomalies in the illumination direction or in the presence of slight disproportionate object size with respect to the whole context. These are only some examples of the analysis approaches to be followed; further and deeper details will be discussed in the next sections.

Kinds of digital evidence and their characterization

Digital forensic tools are asked to recover crucial information by analysing digital evidences; their intrinsic features related to the way these documents have been created, stored and managed are important elements to be considered from the very first and, particularly, can determine which investigation methodology is more appropriate.

Most of the digital data digital forensic has to deal with are images: a three-channelled bi-dimensional array (single if grey level image) is all you can get to try to give answers. First of all, if images have been originated by a digital camera framing a real scene, it follows that its content, besides presenting an intrinsic real structure, will contain all the imperfections and alterations induced by the specific acquiring sensor and by the processing block which generates the final stored file. As evidenced in Figure 2, when an image is taken from real life, light is focused by the lenses on the camera sensor which is a 2D array of CCD/CMOS which constitute the picture elements (pixels). Such elements are hit by the photons and convert them into voltage signals which are then sampled by an A/D converter.

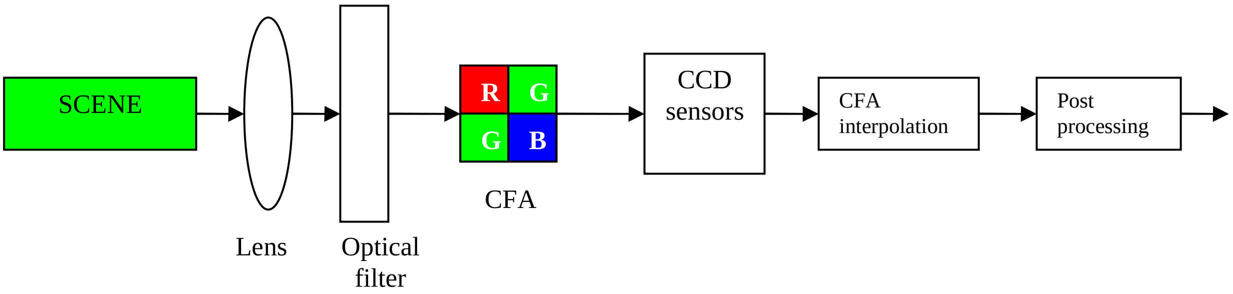

Figure 2. Acquisition process in a photo camera.

Anyway before reaching the sensor, the rays from the scene are filtered by the CFA (Colour Filter Array) which is a specific colour mosaic that permits to each pixel to gather only one particular colour. The sensor output is successively demosaicked (i.e. interpolated) to obtain all the three colours for each pixel and then this signal undergoes additional processing such as white balance, gamma correction, image enhancement and so on; after that is stored to the camera memory in a customized format, although, for commercial devices, JPEG format is usually preferred.

It is now easier to understand that the characteristics of each operation and the properties of every element, from the framed scene to the final image file, influence the digital data. In literature, in fact, there are techniques that have investigated the presence of a specific CFA (Swaminathan, 2006 a) within the image texture to go back to the brand of the camera that had taken a certain photo and other methods which have proposed to study the JPEG quantization coefficients to verify if an image had undergone a second compression thus revealing a possible tampering (Lukas, 2003). On the other side, many are the approaches based on the analysis of the anomalies left by the device over the image such as scratches on the lenses, defective pixels, etc.. In particular, attention has been paid to the sensor noise and among all, dark current, shot noise, thermal noise and so on, PRNU noise (Photo Response Non-Uniformity) is one of the most interesting for forensic applications. PRNU presence is induced by intrinsic disconformities in the manufacturing process of silicon CCD/CMOSs (Chen M., 2008). Such a noise is a 2D systematic fingerprint which characterized each single sensor, that is two cameras of the same brand and model will leave two different traces on the digital contents they acquire. So it is not properly a random noise because it is a deterministic bidimensional template which is superimposed to each taken image.

However images, needing a forensic analysis, can be, not only still, but may also be part of a video sequence; in this circumstance the data to be controlled have a temporal dimension too that has to be taken into account although most of the considerations made for digital photos regarding the presence of PRNU noise pattern and the CFA related to the acquisition phase, can be directly extended to the case of videos (Chen, 2007 b; SPIE, Mondaini, 2007). It is anyway fundamental to point out that the huge amount of available data can suffer different kinds of manipulations with respect to static ones, in particular frames can be skipped or interpolated and inserted to modify the meaning and to alter the original duration of the sequence. Furthermore a clip, coming from another recording but of similar content, could be added to the video in a not-annoying manner to change the whole represented story. Forensic analysis has to be concentrated on aspects such as inter-frame PRNU noise correlation and MPEG-X re-coding.

Another kind of images that can constitute a digital evidence to be checked, in addition to those ones acquired with a photo camera or with a camcorder, might come from a scanning operation.

This means that a printed document (e.g. the cover of a magazine or a real-life photo) located in a flatbed scanner has been illuminated row by row by a sliding mono-dimensional sensor array to originate the digital data (Khanna, 2007 a). The final file format is usually customizable but often is JPEG or PNG. In this case, due to the diversity of the device and to the digitization process, other elements, in addition to those already discussed for cameras, can be considered during the forensic analysis to highlight possible traces of digital asset misuse. For instance, the presence over the image of a 1-D noise pattern, instead of a bidimensional, could be an indicator of image origin and what's more, the direction (vertical or horizontal) of such mono-dimensional periodicity could evidence which has been the scanning manner. Another interesting aspect to control could be the existence of some pieces of dirt that were settled over the scanner plate or of small scratches over the scanner glass that during acquisition have become integral part of the image itself.

Finally it is worthy to spend some words on another type of images digital forensic tools could have to face with: these are computer-generated images. Many are the software that allow to create digital photorealistic pictures that are undistinguishable with respect to those ones acquired by a camera (http://area.autodesk.com/index.php/fakeorfoto). These systems offer the possibility to build up a completely new image or to arrange a believable photomontage merging parts of a real photo with elements synthetically generated. To do this as much actual as possible various are the instruments that are usable and the superimposition of artificial noise is only one of the shrewdness a skilled user could put in practice to develop his fake content. The basic idea to be followed when dealing with this kind of images is to extract significant features which give an indication of the intrinsic realism of the image.

Application scenarios

It is now interesting to consider which can be the possible application scenarios for digital forensic technologies and which could be the questions they can give answers to. Though in literature many have been the fields where digital forensic tools were call to operate, two are the basic categories of usage: "identification of the source" and "detection of forgeries", these two aspects will be debated in detail in the following sections of this chapter.

With the term "identification of the source" it is intended the forensic procedure to determine which is the origin where the digital image comes from. In particular, it is good to split this issue into two sub-cases. In the first sub-case the aim is to recognize which is the device that has produced that digital asset, that is if the digital content has been generated by a photo-camera (video-camera), by a scanner or was computer-generated. To achieve this target, though different approaches exist, the basic ideas are to search over the digital image for traces of the specific acquisition process and for the presence/absence of realistic characteristics within the digital data, this last mainly for distinguishing a computer generated image. On the other side, the second sub-case concerns with the individuation, within a certain set of devices, of which one has created that image. For example, taken a group of photo-cameras (scanners or video-cameras) try to discern which camera (brand and model) has taken that picture. Usually to perform this purpose is necessary to previously extract some information featuring each apparatus and this is done by constructing a sort of identifying fingerprint through the analysis of a certain number of digital contents (training set) produced by that device. Well-known procedures are based on SVM (Support Vector Machine) or on noise pattern correlation.

The second principal application scenario for digital forensic is the "detection of forgeries"; in this case it is required to establish if a certain image is authentic or has been artificially created by means of a manipulation to change its content. The aim of this modification could be very disparate ranging from commercial applications like to make an untrue journalistic scoop or to realize a pseudo-realistic advertisement clip, to some others much more crucial ones such as to alter the judgement in a trial where the image has been accepted as digital evidence or to produce satellite photos to assess that nuclear arms are stocked in a certain territory. Anyway it is important to point out that one of the main hurdles to this kind of analysis is the dimension of the forged part with respect to the whole image size. On the contrary, it is not to underestimate that a mimicking action often has to lead to a substantial alteration of the meaning of the represented scene and this is not always achievable with the exchange of a few pixels.

Intrinsic digital fingerprints

Even if forensic technologies are usually applied for different purposes (as previously described), actually it is possible to evidence how a common approach is followed by almost all the forensic algorithms proposed so far, regardless of their application for source identification or tampering detection. In particular, digital forensics is based on the idea that inherent traces (like digital fingerprints) are left behind in a digital media during both the creation phase and any other successively process (Swaminathan, 2008). By resorting only on the analyzed data, without any previously embedded information (passive approach) and without the knowledge of the related original data (blind method), forensic techniques capture a set of intrinsic information carried out by the digital asset by means of different analysis methods, i.e. statistical, geometric, etc.

Several kinds of digital fingerprints are taken into account for the forensic analysis, a possible classification of such fingerprints can be made by dividing them in three categories: digital traces left by the in-camera processing and those left by the out-camera processing and the fingerprints related to the features of the framed scene. In particular it is to be intended:

- in-camera fingerprints: each component in a digital acquisition device modifies the input and leaves intrinsic fingerprints in the final output, due to the specific optical system, color sensor and camera software; furthermore, images and in particular natural images, have general characteristics, regardless of the content, such as inherent noise or behaviour of the luminance or statistical properties that can be seen as inherent fingerprint;
- out-camera fingerprints: each processing applied to digital media modifies their properties (e.g. statistical, geometrical, etc.) leaving peculiar traces accordingly to the processing itself.

Let us note that previous fingerprints are independent off the content of the analysed data: e.g. the trace left by a given camera is the same even if different subjects have been acquired. On the contrary there is a third fingerprint category considering features related to the content of the image itself, namely:

- scene fingerprints: the world, the photo coming from, has specific properties depending on the content, like lighting properties, which characterize the reproduced scene.

After choosing the specific fingerprint, generally the procedure is to select some properties of the considered fingerprint, to explore relative parameters, and to make a decision basing on either classification or estimation procedures. In particular, in the case of source identification these traces are usually extracted and then compared with a dataset of possible fingerprints specific for each kind/model/brand of acquisition devices, in order to link the digital data to the corresponding source. On the other hand, according to the purpose of forgery detection, the idea is to detect non-uniformity or breaking of such fingerprints within the considered data; specifically, the media is usually block wise analysed and for each block the chosen fingerprints or, better, their properties or parameters, are extracted and compared each other. It is obvious that for the source identification only the first category of traces, the in-camera fingerprints, will be taken into account, whereas for integrity verification all the three categories can be exploited.

Next sections will be devoted to the two main purposes digital forensics is exploited for: acquisition device identification and integrity verification; what kind of digital fingerprint is taken into account and how it is used for the specific aim will be debated for providing a general overview of the principal approaches followed by multimedia forensics. In particular, in the next section, focused on the source identification, the so-called in-camera fingerprints are deeply analysed and their characteristics exploited for acquiring information about data origin. While the successive section focuses on tampering detection, by starting from the specific application of in-camera fingerprints to such a task and then the usage of the other two kinds of fingerprints (out-camera fingerprints and scene fingerprints) is debated.

Techniques for acquisition device identification

Techniques for device identification are focused on assessing digital data origin (images or videos). In particular two are the main aspects to be studied: the first one is to understand which kind of device has generated those digital data (e.g. a scanner, a cell-phone, a digital camera, a camcorder or they are computer-generated) and the second one is to succeed in determining the specific camera or scanner that has acquired such a content, recognizing model and brand (Figure 3).

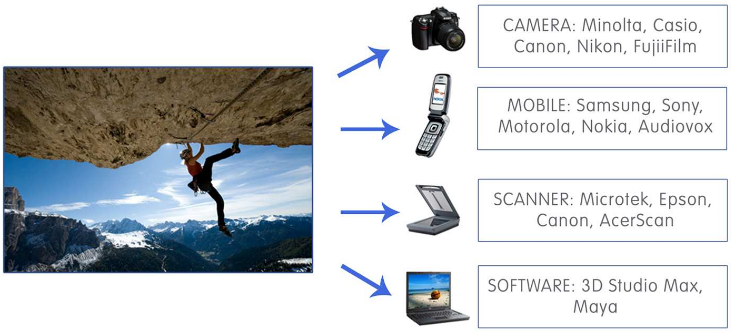
CAMERA: Minolta, Casio, Canon, Nikon, FujiiFilm
MOBILE: Samsung, Sony, Motorola, Nokia, Audiovox
SCANNER: Microtek, Epson, Canon, AcerScan
SOFTWARE: 3D Studio Max, Maya

Figure 3. The source identification problem.

Digital images, can be stored in a variety of formats, such as JPEG, GIF, PNG, TIFF, and the format can be as informative as the image. For example JPEG files contain a well-defined feature set that includes metadata, quantization tables for image compression and lossy compressed data. The metadata describe the source of the image, usually includes the camera type, resolution, focus settings, and other features (Cohen, 2007). Besides when RAW format is used, the camera creates a header file which contains all of the camera settings, including (depending on the camera) sharpening level, contrast and saturation settings, colour temperature / white balance, and so on. The image is not changed by these settings, they are simply tagged onto the raw image data. Although such metadata provide a significant amount of information it has some limitations: they can be edited, deleted and false information about the camera type and settings can be inserted. So it is important to provide a reliable source identification regardless of the type of metadata information ; such passive approach will be explored in the following.
This section will be dedicated to the analysis of the principal solutions exploited for identifying the acquisition device of digital images and videos exploring the general structure and sequence of stages of image formation pipeline, grounding on the physics and operations of the device under examination. These techniques aim at analysing those operations in order to find a fingerprint for the device (the so called in-camera fingerprints) in term of the presence of an identifying mark due to the color filter array (CFA) interpolation, sensor imperfections and lens aberration, In this section techniques based on the extraction, from images belonging to different categories (e.g. scanned images, photos, video etc.), of some robust intrinsic features that are typical of a particular devices classes will be explored. Generally these features can be used to train a classifier (e.g. SVM); when training is performed and whether features grant a good characterization, the system is able to classify the digital asset. Hereafter, it will be shown how all these techniques do not work only for

digital cameras but also for scanner and camcorder identification and also to distinguish between a photographic and a computer graphic image.

Color Filter Array and Demosaicking

In digital cameras with single imaging sensors (the most diffuse on the market) the Color Filter Array (CFA) covers the CCD sensor. Several patterns exist for the filter array (see Figure 4), the most common array is the Bayer CFA. Since the CFA allows only one color to be measured at each pixel this means that the camera must estimate the missing two color values at each pixel, this estimation process is known as “demosaicking”.

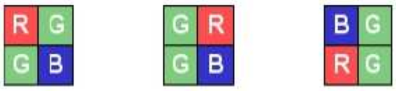

Figure 4. Examples of CFA patterns.

There are several commonly used algorithms for color interpolation and each manufacturer employs a specific algorithm for this purpose. Given an output image *I*, the techniques for acquisition device identification are focused on finding the color filter array pattern and the color interpolation algorithm employed in internal processing blocks of a digital camera that acquired image *I*.
One well-known approach (Swaminathan, 2006 a) assumes to know the CFA used in a digital camera based on the fact that most of commercial cameras use RGB type of CFA with a periodicity of 2x2.
The image *I* after the CFA sampling becomes:

$$I_s = \begin{cases} I(x,y,c), if\ t(x,y) = c \\ 0, otherwise \end{cases} \qquad (1)$$

where *t(x,y)* is the CFA pattern and *c* (colour) can be R, G and B.
Then the intermediate pixel values, corresponding to the points where $I_s(x,y,c) = 0$ in (1) are interpolated using its neighboring pixel values.
The digital forensic method proposed in (Swaminathan, 2006 a) , for every CFA pattern *t* in a search space, estimates the color interpolation coefficients in different types of texture of the image (smooth, horizontal gradient and vertical gradient image regions) through a linear approximation. Using the final camera output *I* and the assumed sample pattern *t*, it is possible to identify the set of locations in each color channel of *I* that are acquired directly from the sensor array. The remaining pixels are interpolated with a set of linear equations in terms of the colors of the pixel captured directly in each types of region. Then the algorithm reconstructs the input image *I* using the corresponding coefficients in each regions to obtain estimated final output image $\hat{I}$ for all the CFA patterns in the search space. At this point the CFA pattern that minimizes error between *I* and $\hat{I}$ is found by computing a weighted sum of the errors of the three color channels.
The color interpolation coefficients estimated from an image and the proposed CFA can be used as features to identify the camera brand utilized to capture the digital image. So a support vector machine (SVM) classifier is trained and then used to identify the interpolation method concerning different digital camera brands. The camera model is more difficult to detect because the color

interpolation coefficients are quite similar among camera models and hence it is likely that the manufacturer uses similar kinds of interpolation methods. Furthermore, others limitations to the method exist: only RGB CFA is considered and then this technique does not permit to distinguishing Super CCD cameras because those digital cameras do not employ a square CFA pattern; moreover there is a misclassification around the smooth regions of the image, in fact similar techniques, such as bicubic interpolation, around smooth region in almost all commercial cameras are used.

As explained before, at each pixel location of a CFA interpolated color image, a single color sample is captured by the camera sensor, while the other two colors are estimated from neighboring ones. As a result, a subset of samples, within a color channel, is correlated to their neighboring samples. This form of correlation is expressed by the linear model:

$$f(x,y) = \sum_{u,v=-N}^{N} \alpha_{u,v} f(x+u, y+v) \qquad (2)$$

In the above equation, $\alpha_{u,v}$ are the coefficients of the model parameters and N is the number of correlated pixel. Since the color filters in a CFA are typically arranged in a periodic pattern (see again Figure 4), then a periodic correlation is introduced.

The probability maps of the observed data obtained from the Expectation Maximization (EM) algorithm can be employed to detect if a color image is the result of a CFA interpolation algorithm and the linear coefficients, $\alpha_{u,v}$, returned by the EM algorithm, can be used to distinguish between different CFA interpolation techniques (Bayram, 2005; Bayram, 2006).

When observed in the frequency domain, these probability maps yield to peaks at different frequencies with varying magnitudes indicating the structure of correlation between the spatial samples. Then a classifier is designed on the basis of the two sets of features: the set of weighting coefficients obtained from an image, and the peak locations and magnitudes in frequency spectrum. This method does not work in case of cameras of the same model, because they share the same CFA filter and interpolation algorithm, and also for compressed image or modified image (gamma corrected, smoothed) because these artefacts suppress and remove the spatial correlation between the pixels due to CFA interpolation.

Imaging Sensor Imperfections

This class of approaches for source matching aims at identifying and extracting systematic errors due to imaging sensor, which appear on all images acquired by the sensor in a way independent by the scene content.

There are several sources of imperfections and noise that influence the image acquisition process (Healey, 1994). When the imaging sensor takes a picture of an absolutely evenly lit scene, the resulting digital image will still exhibit small changes in intensity among individual pixels.

These errors include sensor's pixel defects and pattern noise this last has two major components, namely, fixed pattern noise and photo response non-uniformity noise (PRNU).

The defective pixels can be used for camera identification as described in (Geradts, 2001). This type of noise, generated by hot or dead pixels, is typically more prevalent in cheap cameras and can be visualized by averaging multiple images from the same camera. However, many cameras post-processing remove these types of noise, then this technique cannot always be used.

So, for a reliable camera identification, the idea is to estimate the pattern noise.

The fixed pattern noise (FPN) refers to pixel-to-pixel differences when the sensor array is not exposed to light (so called dark current) and also depends on exposure and temperature. The FPN is used for source camera identification in (Kurosawa, 1999) but it is an additive noise and some middle to high-end consumer cameras suppress this noise by subtracting a dark frame from every image they take. On the basis of this consideration, photo-response non-uniformity noise (PRNU), that is the dominant part of the pattern noise in natural images, is usually searched for. The most important component of PRNU is the pixel non-uniformity (PNU), which is defined as different sensitivity of pixels to light. The PNU is caused by stochastic inhomogenities present in silicon wafers and other imperfections originated during the sensor manufacturing process. As such, it is not dependent on ambient temperature and appears to be stable over time. Light refraction on dust particles, optical surfaces and properties of the camera optics, which also contribute to the PRNU noise, are generally low spatial frequency components not characterizing the sensor and therefore not usable for source identification. Finally the noise component to be estimated and to be used as intrinsic characteristic of the sensor (fingerprint) is the PNU. It is also possible to suppress this kind of noise using a process called flat fielding (Healey, 1994), in which the pixel values are first corrected for the additive FPN and then divided by a flat field frame obtained by averaging images of a uniformly lit scene, but consumer digital cameras do not flat-field their images because it is difficult to achieve a uniform sensor illumination inside the camera.

To continue the discussion, it’s necessary to give a mathematical model of image acquisition process. The digitized output of the sensor *l* can be expressed in the following form (before any other camera processing occurs):

$$l = k(s + p) + r + d \qquad (3)$$

where *s* is the signal if no other noise sources exist, *p* is the random shot noise, *r* is the additive random noise (represented by the read-out noise, etc.) and *d* is the dark current.

The factor *k* is close to 1 and captures the PRNU noise, which is a multiplicative noise.

Because details about the processing are not always easily available (they are hard-wired or proprietary), generally is needed to use a simplified model that captures various elements of typical in-camera processing. A more accurate model tailored to a specific camera would likely produce more reliable camera identification results at the cost of increased complexity.

The simplify sensor output model described in (Lukas, 2006 a) results in the following vector form:

$$l = \sigma^{\gamma} \cdot \left[(1 + \Gamma)Y + \Pi\right]^{\gamma} + \theta_q \qquad (4)$$

In equation 4, *Y* is the incident light intensity on the sensor, σ is the color channel gain and γ is the gamma correction factor (typically, γ ≈ 0.45). The gain factor σ adjusts the pixel intensity level according to the sensitivity of the pixel in the red, green, and blue spectral bands to obtain the correct white balance. The multiplicative factor Γ is a zero-mean noise-like signal responsible for PRNU. Finally, Π is a combination of the other noise sources including the dark current, shot noise, and read-out noise, and θ_q is the quantization noise.

Assuming that either the camera that took the image is available to the forensic analyst or at least some other (non-tampered) images taken by the camera are available, the PRNU term Γ, can be estimated from a set of N images taken by the camera. To improve the SNR between the PRNU term and observed data *l*, a host signal rejection is performed by subtracting (pixel by pixel) the denoised version *(l_d)* of *l*, who can be obtained by using a denoising filter usually implemented through wavelet based algorithm (Mihcak, 1999).

$$Z = l - l_d \qquad (5)$$

Since the image content is significantly suppressed in the noise residual *Z*, the PRNU can be better estimate from *Z* than from *l*, so *Z* is designated as the reference pattern and serves as an intrinsic signature of the camera. To identify the source camera, the noise pattern from an image is correlated with the known reference patterns from a set of cameras. The camera corresponding to the reference pattern giving maximum correlation is chosen to be the source camera that acquired that image. This type of approach is used also for video source identification (Chen, 2007 b) by estimating the PRNU from a video segment and then calculating the correlation with the reference pattern from a different segment of a video clip. The method described above shows poor performance when digital image are cropped, scaled or digital magnified so an improved method for source camera identification based on joint estimation and detection of the camera photo response non uniformity has been developed in (Goljan, 2008). The detector is obtained using the generalized likelihood ratio test and has the form of a cross-correlation maximized over the parameters of the geometrical transform.

With regard to the identification between synthetic image and photographic image a method is described in (Dehnie, 2006), based on the observation that in computer generated images occurs a lack of the sensor's pattern noise artefacts due to the software generation of the image.

Furthermore a technique based on PRNU estimation, for classification of scanned and non-scanned images, is outlined in (Khanna, 2007 a; Khanna 2007 b), based on the difference in the dimension of the sensor array (scanner sensor is a one dimensional sensor array, see previous section). This technique extracts a row reference noise pattern from a single scanned image by averaging the extracted noise (via denoising) over all rows and then a procedure like (Lukas, 2006 a; Chen 2007 a) is used, based on the computation of correlation between the scanner reference pattern and the noise pattern from an image.

Lens Aberration

Due to the design and manufacturing process, lenses produce different kinds of aberrations in images. Generally two of them are investigated to solve the problem of source device identification: lens radial distortion (Choi, 2006) and chromatic aberration (Lahn, 2007).

To reduce manufacturing cost, most of digital cameras are equipped with lenses having almost spherical surfaces that introduce radial distortions.

The radial distortion causes straight lines in the object space rendered as curved lines on camera sensor and it occurs when there is a change in transverse magnification M_t with increasing distance from the optical axis. The degree and the order of compensation of such a distortion vary from one manufacturer to another or even in different camera models by the same manufacturer. As a result, lenses from different cameras leave unique imprints on the captured pictures.

The lens radial distortion can be written as:

$$r_u = r_d + k_1 r_d^3 + k_2 r_d^5 \qquad (5)$$

where r_u and r_d are the undistorted radius and distorted radius respectively. The radius is the radial distance $\sqrt{x_2 + y_2}$ of a point (*x*, *y*) from the center of distortion (the centre of an image). The goal in the method proposed in (Choi, 2006) is to find the distortion parameters k_1 and k_2 that constitute the fingerprint to identify source camera following the Devernay's straight line method. However this method fails if there are no straight lines in the image and also if two cameras of the same model are compared. Besides it is also possible to operate a software correction in order to correct the radial distortion on an image.

The second type of aberration investigated to solve the source identification problem is the chromatic aberration. Chromatic aberration is the phenomenon where light of different wavelenghts fail to converge at the same position of the focal plane. There are two kind of chromatic aberration: longitudinal aberration that causes different wavelenghts to focus at different distances from the lens, while lateral aberration is attributed at different positions on the sensor. In both cases, chromatic aberration leads to various forms of color imperfections in the image. Only lateral chromatic aberration is taken into consideration in the method described in (Lahn, 2007) for source identification. Chromatic aberration causes misalignment between the RGB channels so the task is to estimate the distorted parameters to compensate for the distortion maximizing the mutual information among the color channels. Then these parameters are used in (Lahn, 2007) to identify source cell phone through the use of a SVM classifier.

Others Approaches

There are other approaches for source identification using a set of suitable digital data intrinsic features designed to classify a device model. These features can be statistical, geometrical and color features.

In (Mehdi, 2006) a set of features are calculated, they are composed by suitably chosen image quality metrics (IQM) evaluated between an input image and its filtered version using a low-pass Gaussian filter, and integrated with color features (deviation from gray, inter-band correlation, gamma factor), and wavelet coefficient statistics. These features are used to construct multi-class classifier with images coming from different cameras, but it is demonstrated that this approach does not work well with cameras with similar CCD and it requires images of the same content and resolution.

Another group of selected features is based on the assumption that proprietary CFA interpolation algorithm leaves correlations across adjacent bit-planes of an image. Binary similarity measures (BSM) are metrics used to measure such a similarity. In (Celiktutan, 2005) the authors differentiate between cell-phone camera models by using BSM features in conjunction with IQM features. In the approach described in (Celiktutan, 2007), High-Order Wavelet Statistic (HOWS) features are added to the features used in (Celiktutan, 2005) to distinguish among various brands of cell-phone cameras.

Other techniques exist to solve the classification problem between synthetic and “real” images. The method in (Wang, 2006) proposes a wavelet based statistical model to extract features from the characteristic functions of wavelet coefficient histograms. The previous approach is then extended in (Dirik, 2007) by proposing new features to detect the use of Bayer color filter array during demosaicking (Bayram, 2005; Bayram 2006). These features are incorporated with the features in (Lyu, 2005) that capture the statistical regularities of natural images in terms of statistics of four level discrete wavelet transform coefficients.

A new set of features is taken into account for scanner identification in (Gou, 2007) because, generally, features are extracted without specifically taking the scanning process into consideration. The same features, with the addition of color interpolation coefficients, are proposed to identify images produced by cameras, cell-phone, scanners and computer graphics (McKay, 2008). These features have been chosen in particular to distinguish camera form scanner because the CCD line sensor in a scanner consists of three lines for each color (red, green, blue), so in a scanner acquisition process no color interpolation is needed.

Another set of features has been built in (Khanna, 2007 b) for classifying scanner, computer generated and digital camera due to the physical characteristic of the image sensor. In fact for a scanner, the fixed component of the noise should be nearly identical for all the rows of a scanned image due to mono dimensional image sensor, and for the same reason should be different for all the columns. Then the statistics of row correlation will differ from those of column correlation. Row correlation is defined as the correlation of each row of the image with the estimated row reference pattern calculated as average of the noise of the reference image over all rows. So the first order statistics (mean, median, mode, maximum and minimum) and the higher order statistics (variance, kurtosis and skewness) of the row correlation and column correlation are used to generate the features vector for each image and also a measure of similarity among the rows or columns of the reference pattern noise are considered (Khanna, 2007 b) to design a SVM classifier.

Techniques for assessing image integrity

Information integrity is fundamental in a trial: a verdict must be returned after considering a set of evidences and the authenticity of such proofs should be assured before making a decision. On one hand witnesses and their assertions constitute a type of evidence; on the other hand, concrete objects, e.g. a weapon, represent another type of proof, so to speak "real" evidence. In this latter category can be included all the information belonging to the crime scene, and such information have been often captured and stored by means of pictures. If pictures are just representative of the real world, then they can be considered as authentic evidences. But, it is clear that the advent of digital pictures and relative ease of digital image processing makes today this authenticity uncertain. In this scenario, an efficient assessment of the integrity of digital information, and in particular of digital images, plays a central role.

But, what does integrity mean? In a strong sense, the image must be only the outcome of an acquisition of a real world scene, without any successively processing; in a wide sense, the image must accordingly represent a real world scene and even if some processing has been probably applied, the "meaning" of the scene must not be altered.

Once evidence passes from the real world of three dimensional objects to a digital image, we lose the origin of information and we can not trust any more what we are seeing, even if the content is advertised as real. Several image processing tools are nowadays easily usable for almost everybody; let only consider that Adobe PhotoShop is already licensed to many millions of users worldwide. With such programs, a great deal of operations is allowed to affect digital photographic files: person images can be moved in different contexts; objects can be deleted from scenes; particular details can be cloned within the photograph; computer graphic objects can be added to the real scene. All these manipulations become more and more sophisticated thus making the alteration virtually imperceptible; furthermore, establishing the authenticity of images is a key point for being able to use digital images as critical evidence.

Digital forensics assume that images are intrinsically characterized by specific pattern due to the creation process and to any other process suffered after image creation. To properly individuate possible modifications, the image forensic approach considers that such intrinsic fingerprints inside images are distinguishable due to the different applied image processing, or that the original traces have been altered due to a tampering, thus losing their uniformity. So, different digital fingerprints are taken into account and studying their characteristics it is possible to verify if an image has undergone some tampering and even detect the suffered processing. Referring to the wide sense meaning of integrity (i.e. the digital photograph is a congruous representation of the captured "real" world), a lot of processing non-affecting the semantic (e.g. JPEG compression or recompression, brightness adjustment, gamma correction, etc.) can be erroneously revealed as tampering. Therefore, detection of image alteration does not necessarily prove malicious tampering, but surely questions about the content of the image and helps for further analysis.

In the following, we are going to discuss the technological approaches proposed in literature so far for verifying digital image authenticity; this discussion is structured again according to the classification of digital fingerprints previously introduced in this chapter where the three kinds of traces are categorized: in-camera fingerprints (described for their exploitation in source identification), out-camera fingerprints and scene fingerprints. Specifically, in the first and third case, forensic techniques search for some breaking or inconsistencies of such traces, whereas in the second case fingerprints are used for identifying a specific processing. As already mentioned, detection of image processing does not necessarily prove malicious tampering, but surely proves that some manipulation occurred after image creation.

Because of the great variety of existing methodologies devoted to this purpose, we have decided to provide only some hints of each analyzed technique, to allow the interested reader to get useful information and to possibly deepen his study by following the bibliographic references.

In-camera fingerprint breaking

Basically, the acquisition process is analysed and peculiarities left by some component of the chain are considered as intrinsic fingerprints (in-camera fingerprints) that characterize the kind or even the model or brand of acquisition devices. In particular, in the previous section three main components (namely color filter array, sensors and lens) are considered with their related fingerprints, that are:

- the Color Filter Array (CFA) and its related demosaicking process;
- the sensor imperfection and its related pattern noise;
- the lens aberration and its related chromatic aberration.

On the basis of the previous analysis, we now consider how the traces left by such components can be exploited for tampering detection.

In the case of *CFA* the correlations between pixels introduced by the specific algorithm for the color interpolation are analysed in order to verify if these properties are broken in certain areas, thus revealing possible tampering (Popescu, 2005 a; Swaminathan, 2008). The works in (Lukas, 2006 b, Chen M., 2008) propose a method to detect the *camera pattern noise* present in a given image: the inconsistency of camera pattern noise in some regions of digital image reveals the non integrity of the content; the proposed approach requires either the camera which produced the image or a set of images produced by the same camera, thus making such an algorithm non blind. Regarding the lens aberration, in (Johnson, 2006) the authors consider in particular the *chromatic aberration* that leads to various forms of color imperfections in the image: when these alterations fail to be consistent across the image, a tampering can be supposed to be happened.

Besides the above mentioned fingerprints, there are other in-camera traces that have been used for integrity verification. Basically, also for such algorithms a block-based analysis is computed for evidencing the coherence/incoherence of the extracted parameters on the whole image.

The image irradiance (light energy incident on the image sensors) is related to the image intensity (the final output image) by a non-linear camera response function (*CRF*), that is a characteristic of each camera. The estimation of the CRF on different regions of the analysed image and the evaluation of consistency/inconsistency between such estimated CRFs, provides a good method for deciding if the image is likely to be authentic or spliced (Ng, 2006; Lin, 2005; Hsu, 2006).

The last step of the acquisition process is usually a *JPEG compression* to reduce storage space of the output image. Such a compression leaves unique fingerprints due to the particular quantization matrix used by the specific camera, and serves as a “fragile watermark” enabling the detection of changes within the image. In (Fridrich, 2001) authors propose to detect possible manipulations by investigating the compatibility of 8×8 pixel blocks with a given quantization matrix; whereas in (He, 2006) an algorithm is developed for automatically locating the tampered regions.

The discrepancy in the signal-to-noise ratio (*SNR*) across the image can also be considered as a sign for possible tampering. Digital images have an inherent amount of noise introduced either by the imaging process or digital compression, and such a noise is typically uniform across the entire image. If two images with different noise levels are spliced together, or if small amounts of noise are locally added to conceal traces of tampering, hence changes in the SNR across the image can be used as evidence of tampering (Popescu, 2004 a).

A different in-camera fingerprint regards the luminance non-linearity, introduced during the acquisition chain in order to improve the perceptual quality of the output digital images; parameters of this non-linearity are dynamically chosen and depend on the camera and the scene, but they are typically constant on the image. The presence of several distinct non-linearities across an image can reveal the non integrity of the content. In (Popescu, 2004 a) it is described how luminance non-linearities introduce specific correlations in the Fourier domain, and how these correlations can be estimated and used for tampering detection.

Finally, another approach proposed in (Ng, 2007) consider that the camera lens often have an optical low-pass property for the purpose of anti-aliasing; hence, when an image is spliced onto another, it is likely that sharp edges are introduced into the tampered content, and that such edge transitions invalidate the low-pass behaviour. Some parameters, representing the optical low-pass property, are extracted by means of statistical methods and are used for image integrity verification.

Out-camera processing identification

A class of forensic algorithms have been proposed for identifying some processing applied after image creation, to reveal possible tampering operations.

Firstly, for generating convincing digital image forgeries, it is often necessary to resize, rotate, stretch some portions of the manipulated images, thus leading to apply a final resampling step. Although a resampling process does not typically leave perceivable artefacts, it anyway introduces specific periodic correlations between image pixels. For instance, when the image is upsampled, some of the pixel values are directly obtained from the smaller version of the image, and the remaining pixels are interpolated and, thus, they appear highly correlated with its neighbors. The authors in (Popescu, 2005 b) show how to detect a discrete approximation of the applied resampling rate in an image region. The approach relies on the detection of the introduced correlation patterns; since each pattern (based on the probability of each signal sample to be correlated to its neighboring samples) is not in a biunique relation with a resampling rate, the matching could not be uniquely identified. Another method for detecting interpolation has been proposed in (Gallagher, 2005), where authors observe a periodicity in the variance function of the interpolated signal. Authors in (Babak, 2008) analytically describe the periodic properties of an interpolated signal as well as its derivatives, thus providing also a theoretical support for the methods in (Popescu, 2005 b) and (Gallagher, 2005). The method allows the direct estimation of the resampling parameters such as the scaling factors, rotation angles and skewing factors.

Another fundamental processing to be considered is compression. Image tampering usually requires to make use of common photo-editing software: original images, often stored in JPEG format, are manipulated by the editing tools and then they are re-saved using again the JPEG format; hence the resulting tampered images have been wholly or in part, double compressed. While double compression does not necessarily prove malicious tampering, it raises suspicions that the image may be not authentic; as a matter of fact, double JPEG identification has acquired special attention in digital forensic literature, as it may serve as an useful forensics clue. Double JPEG compression often introduces specific correlations between the discrete cosine transform (DCT) coefficients of image blocks that are not present in single compressed images. These correlations can be detected and quantified by analyzing the double quantization effect of two JPEG compressions with different quality factors. Such effect is identified in the exhibition of periodic peaks and valleys in the histograms of the DCT coefficients. Not only the presence of a double compression can be estimated but also the compression quality that have been used (Lukas, 2003; Popescu, 2004 a) as well as the specific doctored parts (He, 2006). On the other hand, the works in (Luo, 2006) and (Fan, 2003) exploit the JPEG “blockiness” artefacts in order to detect a double compression. The authors in (Luo, 2006) evaluate the Blocking Artefact Characteristic Matrix (BACM) of an image which exhibits a symmetrical shape and regularity for a single JPEG compression; they show how this regularity can be destroyed by a successively non aligned compression. Fan (2003) proposes a method to determine whether a non compressed image has been previously JPEG compressed, and further to estimate which quantization matrix has been used. The original intention of such an approach was the removal of JPEG artefacts; however, it can serve as an image forensic tool by also revealing the presence of a double JPEG compression. The method assumes that if there is no compression the pixel differences across blocks should be similar to those within blocks (thus non showing any blockiness artefacts) while they should be different due to block artefacts if the image has been compressed. Finally, in (Fu, 2007) it is also found that the distribution of the first digit of the JPEG DCT coefficients can be used to distinguish a singly JPEG compressed image from a double compressed one. A single compressed image is characterized by a distribution of its DCT coefficients that follows the Benford’s law distribution; whereas, as soon as another compression is applied, the coefficients do not follow this law anymore.

One of the main common image tampering is splicing. It is defined as a simple joining of portions coming from two or more different images. In (Ng, 2004 a) some image features, particularly sensitive to splicing operations, have been extracted and used for designing a classifier. A different technique for detecting splicing searches for the presence of abrupt discontinuities in the image (Ng, 2004 b). Several other techniques estimate the camera response function from different regions of an image to detect splicing and possibly other manipulations (Hsu, 2006; Popescu, 2004 a). The authors in (Chen, 2007) observe that the spliced image may be characterized by a number of sharp transitions such as lines, edges and corners; hence, they found a parameter as a sensitive measure of these sharp transitions, and used it for splicing detection.

Another common tampering is object removal: an image region containing objects that have to be erased, is replaced by another region of the same image. This type of operation is called copy-move or region-duplication. Since there is similar information (e.g. texture, noise and color) inside the same image, it is hard to identify these forgeries via visual inspection. Furthermore, several post-processing (such as adding noise, blurring, lossy compression) may be performed on such tampered images, thus making the detection of forgery significantly harder. Works in (Fridrich, 2003; Luo, 2006; Popescu, 2004 b) are all based on block matching: firstly, the image is divided into small blocks and some features are extracted for each block; then, by comparing such features for different blocks, it is possible to identify duplicated regions.

Several works in the tampering detection literature try to define the properties of a manipulated image in terms of the distortions it goes through, and using such analysis to present methods for detecting manipulated images. In doing so, some works assume that creating a tampered image involves a series of processing operations; they propose identifying such manipulations by extracting certain salient features that would help distinguish such tampering from authentic data. Image manipulations, such as contrast changes, gamma correction, and other image nonlinearities have been modeled and used to identify them (Farid, 2001). More generally, in (Swaminathan, 2006 b), image operations, such as resampling, JPEG compression, and adding of noise, are modeled as linear operators and estimated by linear image deconvolution. In the frequency domain a “natural” signal has weak higher-order statistical correlations. The authors in (Farid, 1999) observed that “un-natural” correlations are introduced if this signal is passed through a non-linearity (which would almost surely occur in the creation of a forgery).

Scene characteristic inconsistencies

Some works have proposed to use as fingerprints the *light properties* directly derived from the scene. In particular, Johnson and Farid base their works on the idea that splicing together different images (that are the acquisition of different scenes) means likely to create a new content where light inconsistencies are present.

In (Johnson, 2005; Johnson, 2007 c) the authors consider to estimate the direction of the light source, both in a simplified case (Johnson, 2005) and in complex lighting environments (Johnson, 2007 c): if the image is supposed to be a composition of more images, hence the lighting direction is computed more than once in different positions of the image; by comparing such directions it is possible to verify whether inconsistencies are present thus revealing the suffered digital tampering.

Lighting direction can be also estimated by considering that the light source produces specular highlights on the eyes of people present in the scene. Authors in (Johnson, 2007 a) propose to compute the direction of a light source by analyzing the different highlights within an image, and by detecting inconsistencies in lighting they are able to reveal possible tampering in some part of the content. Furthermore authors evidence how it would be possible to measure from highlights also the shape and the color of the light source (besides its location), and how these parameters could help in exposing digital forgeries.

By considering specific images where eyes are present , in (Johnson, 2007 b) it is shown how to estimate the camera’s principal point (i.e. the
projection of the camera center onto the image plane) from the analysis of person’s eyes within an image. Such a principal point depends on intrinsic and extrinsic camera parameters and it is proposed to be adopted as a fingerprint, whose inconsistency across an image can be used as evidence of tampering.

Future Trends

Although many of the digital forensic techniques proposed so far are bright and groundbreaking, none of them by itself offers a stand alone solution for the considered problem (i.e. the source identification and the verification of information integrity). Furthermore, the user intervention is often desirable for validating the final results: for example, let us consider the estimation of image tampering, that without any user intervention is quite impossible, since even if an out camera processing is detected, often only a human interpreter can decide if the purpose of the modification is malicious or not.

The validation of digital forensic approaches for integrity verification, seems to be missing of a common framework, regarding both image databases and performance measures, such as accuracy, robustness, security.

An image database is fundamental for the evaluation of a proposed algorithm; furthermore, a common dataset provides an unified platform for the research community to compare various algorithms. Actually, several datasets are available for the research community (http://www.ee.columbia.edu/ln/dvmm/trustfoto/), but there are some open issues that call for a benchmark dataset. For instance, the experiments involving the camera characteristics require a dataset of images acquired by a diverse models of camera, at various acquisition settings. Furthermore, in order to facilitate the evaluation of the image forgery detection techniques using the images produced by the state-of-the-art image forgery creation techniques, a dataset of these images would be necessary. Therefore, further effort on producing and standardizing the additional benchmark dataset is needed.

Most of the proposed digital tampering forensic techniques do not provide a clear measure of the achievable performance in terms of accuracy and false-alarm rates. There is often a lack of rigorous theoretical background and concept experiments. To further refine these methods, analytical results have to be defined more clearly and appropriate test and evaluation datasets have to be designed, built and shared. The robustness to various common and malicious image processing operations is the most challenging issue that each image forensic algorithm has to face with. Proposed methods are often designed and tested to perform under limited and not general conditions, and, moreover, most techniques can be easily bypassed by a basic image processing software. Overcoming these challenges requires the development of several novel methodologies and thorough evaluation of their limitations under more general and practical settings. Alongside of robustness, a different analysis on performances of forensic algorithms comes from the security point of view. By increasing the possible solutions for forgery identification, also malevolent people, aiming at modifying digital content, increase their attention for overcoming detection of tampering processing. Hence, the analysis of forensic algorithms from the security point of view would be an interesting open issue to be addressed in the future.

Another future trend to be considered is the improvement of the use of image source imperfections as fingerprint to solve the problem of source identification. Review of the modern literature on this argument shows that good experimental results are obtained but reliable identification seems impossible if all the acquisition process and post-processing steps are not taken into account, so further investigations are necessary. Future research should focus on definition of new model for the acquisition process in order to better estimate the anomalies left by intrinsic disconformities in the manufacturing process of silicon sensor of a camera. Since this fingerprint is not a random noise but a deterministic template, which is superimposed to each taken image, should be necessary to define and use new denoising filters that grant the suppression of the image content and take into account the different kind of sensor device.

Conclusions

Nowadays, digital visual data have gained high relevance in nearly every aspect of our life and represent one of the main source of information that can bias common opinion. In particular scenarios, such as the forensic one, visual information can be used as possible evidence in a trial thus influencing the final verdict. In such a situation, it is fundamental to know the origin and the history of such data in order to be assured that opinion coming from such information has not been manipulated. In the last years, a new science, referred as multimedia forensics, has been proposed aiming at providing information on a digital asset, by means of the analysis of intrinsic fingerprints that characterize the data during its life. In particular, the analysis of these patterns may lead to identify image and video origin and to establish data integrity.

In this chapter, principles and motivations of digital forensics have been discussed and the main approaches for obtaining information from a digital content has been presented. Almost all the proposed techniques can be sketched as a forensic tool that extracts, from the considered data, some digital fingerprints, and that, by exploring some properties of such patterns, is able to make a decision based on either classification or estimation procedure. In particular, the output of such a tool can provide information on the acquisition device that has produced the visual content as well as on the possible suffered tampering.

Even though multimedia forensics is still in its infancy, the research community is showing an increasing interest for such technologies thus leading to new exciting challenges for the solution of many open issues in the next future.

References

[Babak, 2008] Babak, M., & Saic, S. (2008). Blind Authentication Using Periodic Properties of Interpolation. In *IEEE Transactions on Information Forensics and Security, vol. 3*(3), pp. 529-538.

[Barni, 2004] Barni, M., & Bartolini, F. (Ed.) (2004). Watermarking Systems Engineering: Enabling Digital Assets Security and other Applications. Marcel Dekker.

[Bayram, 2005] Bayram, S., Sencar, H., Memon, N., & Avcibas, I. (2005). Source camera identification based on CFA interpolation. In *IEEE International Conference on Image Processing*: vol.3, pp. III-69-72.

[Bayram, 2006] Bayram, S., Sencar, H.T., & Memon, N.(2006). Improvements on Source Camera-Model Identification Based on CFA Interpolation. In *WG 11.9 International Conference on Digital Forensics*.

[Blythe, 2004] Blythe, P., & Fridrich, J. (2004). *Secure Digital Camera*. Paper presented at Digital Forensic Research Workshop, Baltimore, MD.

[Celiktutan, 2005] Celiktutan, O., Avcibas, I., Sankur, B., & Memon, N. (2005). Source Cell-Phone Identification. *In International Conference on Advanced Computing & Communication*.

[Celiktutan, 2007] Celiktutan, O., Avcibas, I., & Sankur. B. (2007). Blind Identification of Cell Phone Cameras. In *SPIE: vol. 6505*, 65051H .

[Chen M., 2007 a] Chen, M., Fridrich, J., & Goljan M. (2007). Digital Imaging Sensor Identification (Further Study). In *SPIE: vol. 6505*, 65050P.

[Chen M., 2007 b] Chen, M., Fridrich, J., Goljan, M., & Lukas J. (2007). Source Digital Camcorder Identification Using Sensor Photo Response Non-Uniformity. In *SPIE: vol. 6505*, 65051G.

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

[Chen W., 2007] Chen, W., & Shi, Y. (2007) Image splicing detection using 2D phase congruency and statistical moments of characteristic function. In *SPIE: vol. 6505*, pp. 0R–0S.

[Choi, 2006] Choi, K. S., Lam, E. Y., & Wong, K. K. Y. (2006), Source Camera Identification Using Footprints from Lens Aberration, In *SPIE: vol. 6069*, 60690J.

[Cohen, 2007] Cohen, K. (2007). Digital Still Camera Forensics. *Small scale digital device forensics journal, 1*(1),1-8.

[Dehnie, 2006] Dehnie, S., Sencar, , H. T. , and Memon, N. (2006). Identification of Computer Generated and Digital Camera Images for Digital Image Forensics. In . *IEEE International Conference on Image Processing.*

[Dirik, 2007] Dirik, A.E., Bayram, S., Sencar, H.T., Memon, N. (2007). New Features to Identify Computer Generated Images. In *IEEE International Conference on Image Processing: vol.4*, no., pp.IV -433-IV -436.

[Fan, 2003] Fan, Z., de Queiroz, R. (2003). Identification of bitmap compression history: Jpeg detection and quantizer estimation. *IEEE Transaction on Image Processing, 12*(2): 230–235

[Farid, 1999] Farid, H. (1999). *Detecting Digital Forgeries Using Bispectral Analysis*. (Tech. Rep.AIM-1657). Massachusetts Institute of Technology Cambridge, MA, USA.

[Farid, 2001] Farid, H. (2001). Blind inverse gamma correction. In *IEEE Transactions on Image Processing: vol.10*, no.10, pp.1428-143.

[Fridrich, 2001] Fridrich, J., Goljan, M., & Du, R. (2001). Steganalysis based on jpeg compatibility. In *SPIE: vol. 4518*(1): 275–280

[Fridrich, 2003] Fridrich, J., Soukal, D., & Lukas, J. (2003). Detection of copy-move forgery in digital images. Paper presented at *Digital Forensic Research Workshop*, Cleveland, OH.

[Fu, 2007] Fu, D., Shi, Y. Q., & Su, W. (2007). A generalized Benford's law for JPEG coefficients and its applications in image forensics. In *SPIE: vol.6505*, 65051L

[Gallagher, 2005] Gallagher, A.C. (2005). Detection of linear and cubic interpolation in JPEG compressed images. *Proceedings of The 2nd Canadian Conference on Computer and Robot Vision:* vol., no., pp. 65-72.

[Geradts, 2001] Geradts, Z.J., Bijhold, J., Kieft, M., Kurusawa, K., Kuroki, K., & Saitoh, N. (2001). Methods for Identification of Images Acquired with Digital Cameras. In *SPIE: vol. 4232*, 505.

[Goljan, 2008] Goljan, M., & Fridrich, J. (2008). Camera Identification from Scaled and Cropped Images; In *SPIE: vol. 6819*, 68190E

[Gou, 2007] Gou, H., Swaminathan, A., & Wu, M. (2007). Robust Scanner Identification Based on Noise Features. In *SPIE: vol 6505*, 65050S.

[He, 2006] He, J., Lin, Z., Wang, L., & Tang, X. (2006). Detecting doctored JPEG images via DCT coefficient analysis. In *European Conference on Computer Vision: vol. 3953.*

[Healey, 1994] Healey, G.E., & Kondepudy, R. (1994). Radiometric CCD camera calibration and noise estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence: vol.16*, no.3, pp.267-276.

[Hsu, 2006] Hsu, Y.-F., & Chang, S.-F. (2006). Detecting image splicing using geometry invariants and camera characteristics consistency. In *Interational Conference on Multimedia and Expo*: vol., no., pp.549-552.

[Johnson, 2005] Johnson, M. K., & Farid, H.(2005). Exposing digital forgeries by detecting inconsistencies in lighting. Paper presented at *ACM Multimedia and Security Workshop,* New York, NY.

[Johnson, 2006] Johnson, M. K. , & Farid, H. (2006). Exposing digital forgeries through chromatic aberration. In *ACM Multimedia Security Workshop*, pp. 48–55.

[Johnson, 2007 a] Johnson, M.K., & Farid, H..(2007). Exposing Digital Forgeries Through Specular Highlights on the Eye. In *International Workshop on Information Hiding*.

[Johnson, 2007 b] Johnson, M.K., & Farid, H. (2007). Detecting Photographic Composites of People. In *International Workshop on Digital Watermarking.*

[Johnson, 2007 c] Johnson, M.K., Farid, H. (2007). Exposing Digital Forgeries in Complex Lighting Environments. In *IEEE Transactions on Information Forensics and Security: vol.2*, no.3, pp.450-461.

[Khanna, 2007 a] Khanna, N., Mikkilineni, A.K., Chiu, G.T.C., Allebach, J. P., & Delp, E. J. (2007). Scanner Identification Using Sensor Pattern Noise. In *SPIE: vol. 6505*, 65051K.

[Khanna, 2007 b] Khanna, N., Mikkilineni, A.K., Chiu, G.T.C., Allebach, J. P., & Delp, E. J. (2007). Forensic Classification of Imaging Sensor Types. In *SPIE: vol. 6505*, 65050U.

[Kurosawa, 1999] Kurosawa, K., & Kuroki, K., Saitoh, N. (1999). CCD fingerprint method-identification of a video camera from videotaped images, In *International Conference on Image Processing: vol.3*, no., pp.537-540.

[Lahn, 2007] Lanh, T. V., Emmanuel, S., Kankanhalli, M.S. (2007). Identifying Source Cell Phone using Chromatic Aberration. In *IEEE International Conference on Multimedia and Expo:* vol., no., pp.883-886.

[Lin, 2005] Lin, Z., Wang, R., Tang, X., & Shum H.Y. (2005). Detecting doctored images using camera response normality and consistency. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition: vol.1*, no., pp. 1087-1092.

[Lukas, 2003] Lukas, J., & Fridrich, J. (2003). Estimation of primary quantization matrix in double compressed JPEG images. Paper presented at *Digital Forensic Research Workshop*, Cleveland, OH.

[Lukas, 2006 a] Lukas, J., Fridrich, J., & Goljan, M. (2006). Digital camera identification from sensor pattern noise. *IEEE Transactions on Information Forensics and Security*: *vol.1*, no.2, pp. 205-214.

[Lukas, 2006 b] Lukas, J., Fridrich, J. & Goljan, M. (2006). Detecting digital image forgeries using sensor pattern noise. In *SPIE: vol. 6072*, pp. 0Y1–0Y11.

[Luo, 2006] Luo, W., Huang, J., & Qiu, G. (2006). Robust detection of region-duplication forgery in digital image. In *International Conference on Pattern Recognition*: *vol.4*, no., pp.746-749.

[Lyu, 2005] Lyu, S., & Farid, H. (2005). How realistic is photorealistic?. *IEEE Transactions on Signal Processing, ,* vol.53, no.2, pp. 845-850.
[McKay, 2008] McKay, C., Swaminathan, A., Hongmei Gou, & Min Wu (2008). Image acquisition forensics: Forensic analysis to identify imaging source. In *IEEE International Conference on Acoustics, Speech and Signal Processing:* vol., no., pp.1657-1660.
[Mehdi, 2006] Mehdi, K.L., Sencar, H.T., & Memon, N. (2006). Blind source camera identification. In *International Conference on Image Processing: vol.1*, no., pp. 709-712.
[Menezes, 1998] Menezes, A. , Oorschot, V., & Vanstone, S. (Ed.) (1998). *Handbook of Applied Cryptography*. Boca Raton, FL: CRC.
[Mihcak, 1999] Mihcak, M.K., Kozintsev, I., & Ramchandran K. (1999). Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising. In *IEEE Int. Conf. Acoust., Speech, Signal Processing: vol. 6*, pp. 3253--3256.
[Mondaini, 2007] Mondaini, N., Caldelli, R., Piva, A., Barni, M., & Cappellini, V. (2007). Detection of malevolent changes in digital video for forensic applications. In *SPIE: vol.6505*, 65050T.
[Ng, 2004 a] Ng, T. T., Chang, S. F (2004). *Blind detection of digital photomontage using higher order statistics.* (Tech. Rep. 201-2004-1), Columbia University, New York.
[Ng, 2004 b] Ng, T.T., & Chang, S.F. (2004). A model for image splicing. In *IEEE International Conference on Image Processing: vol.2*, no., pp. 1169-1172.
[Ng, 2006] Ng ,T. T., Chang, S. F., & Tsui, M. P. (2006). *Camera response function estimation from a single-channel image using differential invariants*. (Tech. Rep. 216-2006-2), Columbia University, New York.
[Ng, 2007] Ng, T. T. (2007). *Statistical and Geometric Methods for Passive-Blind Image Forensics*. Unpublished doctoral dissertation, Columbia University, New York.
[Popescu, 2004 a] Popescu, A. C. , & Farid, H. (2004). Statistical Tools for Digital Forensic. In *Interantional Workshop on Information Hiding: vol. 3200*, pp. 128-147.
[Popescu, 2004 b] Popescu, A., & Farid, H. (2004). *Exposing digital forgeries by detecting duplicated image regions*. (Tech. Rep. TR2004-515), Computer Science, Dartmouth College, Hanover, NH.
[Popescu, 2005 a] Popescu, A.C., & Farid, H. (2005). Exposing digital forgeries in color filter array interpolated images. *IEEE Transactions on Signal Processing, vol.53,* no.10, pp. 3948-3959.
[Popescu, 2005 b] Popescu, & A.C., Farid, H. (2005). Exposing digital forgeries by detecting traces of resampling. In *IEEE Transactions on Signal Processing, vol.53,* no.2, pp. 758-767.
[Swaminathan, 2006 a] Swaminathan, A., Min Wu, & Liu, K.J.R. (2006). Non-Intrusive Forensic Analysis of Visual Sensors Using Output Images. In *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing: vol.5*, no., pp. V V.
[Swaminathan, 2006 b] Swaminathan, A., Wu, M., & Liu, K. J. R. (2006). Image tampering identification using blind deconvolution. In *IEEE International Conference on Image Processing*: vol., no., pp.2309-2312.
[Swaminathan, 2008] Swaminathan, A., Min Wu, & Liu, K.J.R. (2008). Digital Image Forensics via Intrinsic Fingerprints. *IEEE Transactions on Information Forensics and Security, vol.3,* no.1, pp.101-117.
[Wang, 2006] Wang Y., & Moulin, P. (2006). On Discrimination between Photorealistic and Photographic Images. In *IEEE International Conference on Acoustics, Speech and Signal Processing: vol.2*, no., pp.II-II.

Key Terms and Their Definitions

multimedia forensic:

multimedia forensic can be defined as the science that tries, by only analyzing a particular digital asset, to give an assessment on such a content and to extract information that can be useful to address and support an investigation linked to the scene represented in that specific digital document.

digital evidences:

during a trial a set of evidences are considered before returning a verdict; alongside of witnesses, assertions, and concrete objects, nowadays digital data representing the acquisition and the storage of all the information belonging to the crime scene has to be considered as digital evidences.

data authenticity:

digital data can be assumed to be authentic if it is provable that it has not been corrupted after its creation. In a strong sense, any processing means corruption, that is digital data to be authentic must be only the outcome of an acquisition process of a real world scene without any successively processing; but in a wide sense, authentic data must accordingly represent a real world scene and even if some processing has been probably applied the meaning of the scene must not be modified. Data authenticity also means that a digital object is indeed what it claims to be or what it is claimed to be.

digital fingerprints:
any digital asset is characterized by inherent patterns specific of its life history; such patterns, referred as fingerprints, come from the acquisition device producing the data and/or the possible processing suffered by the data.

source identification:
given a digital asset, it is possible to trace the device that has produced the data. In particular, by focusing on visual data, source identification refers to the recovery of the type of used imaging devices between digital cameras, scanners, mobiles, computer graphic technologies, or the specific model or brand of such devices.

tampering:
a tampering operation can be defined as a particular subset of image processing, voluntarily applied, aiming at counterfeiting the meaning of the tampered data or at least at getting something to appear different from what it is really.

pattern noise:
a reference pattern noise is a particular digital fingerprint left over a digital image during acquisition. Such pattern is due to the manufacturing process and can be extracted from the images using a denoising filter.