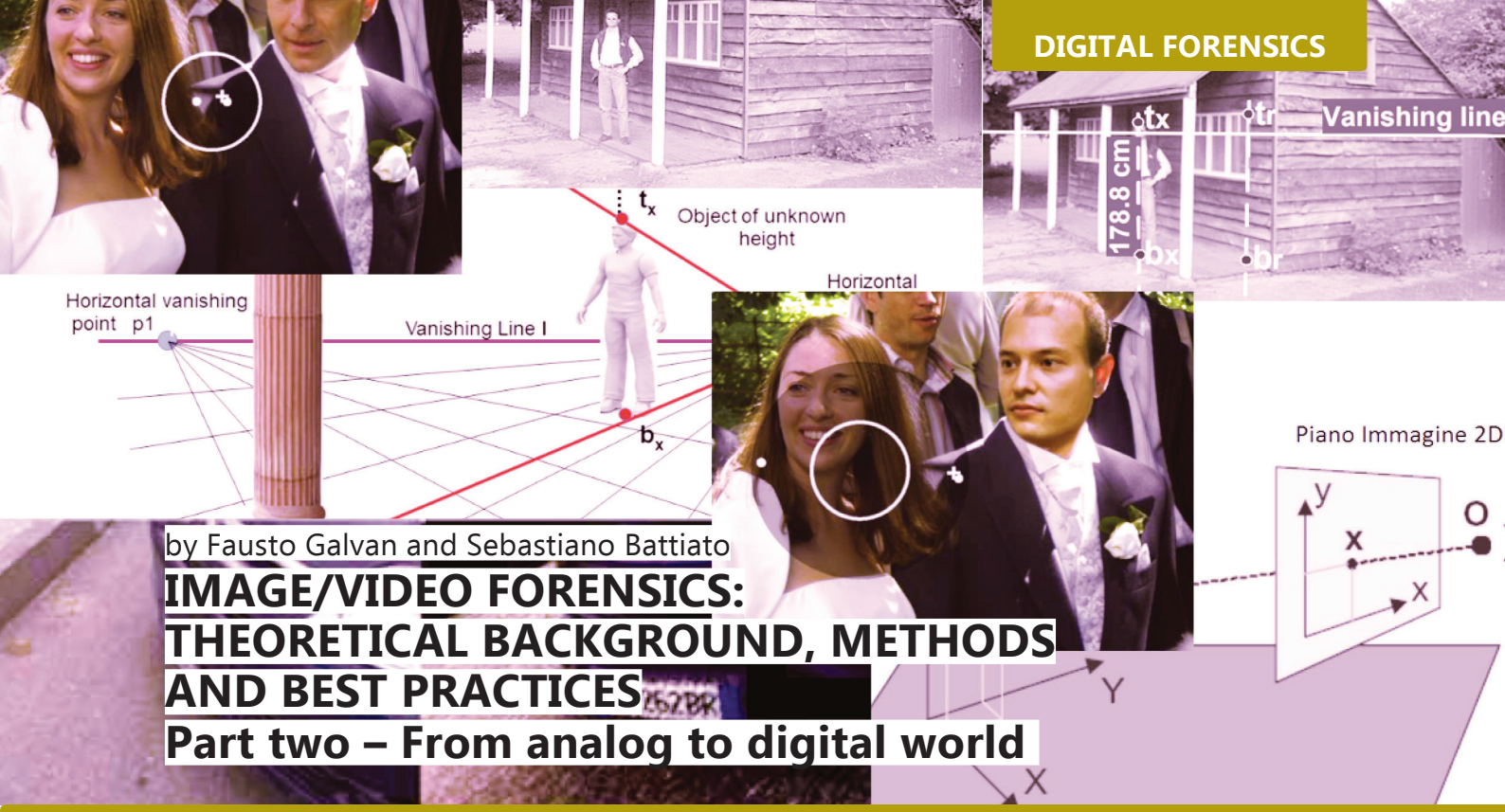

by Fausto Galvan and Sebastiano Battiato

IMAGE/VIDEO FORENSICS: THEORETICAL BACKGROUND, METHODS AND BEST PRACTICES

Part two – From analog to digital world

Fausto GALVAN is a Warrant Officer at Arma dei Carabinieri, where he has been working since 1991. He received his degree in Mathematics in 2002 and his PhD in Computer Science in 2016, both at the University of Udine. His research area is Image/Video Forensics. He co-authored more than a dozen publications for scientific journals and international conferences, chapters of books and national magazines. He has been a member of the scientific and organizer committee, and he gave speeches, in national and international seminars and conferences regarding Computer Forensics issues. At the present time he works at the Public Prosecutor's office at the Court of Udine.

Sebastiano BATTIATO is a full professor of Computer Science He is currently the Scientific Coordinator of the PhD Program in Computer Science at the University of Catania. He is involved in research and directorship of the IPLab research lab (http://iplab.dmi.unict.it). He coordinates IPLab's participation on large scale projects funded by national and international funding bodies, as well as by private companies. His research interests include Computer Vision, Imaging technology and Multimedia Forensics. He is Director (and Co-Founder) of the International Computer Vision Summer School (ICVSS), Sicily, Italy. He is the recipient of the 2017 PAMI Mark Everingham Prize for the series of annual ICVSS schools. In 2016 he founded iCTlab (www.ictlab.srl) a spin-off company working on the field of Digital Evidence.

From the beginning of this century, Image/Video Forensics experts faced the need to extract the largest number of information from a digital visual content, developing a plethora of methods and algorithms. These approaches, which may concern the authentication of images or videos, the identification of the device in which the visual data was originated, or the alterations to which the document has been subjected, find applications both in the civil and criminal context. In a series of three papers, we provide first an introductory part about the powerful impact of images and videos in today's reality, followed by a section where we highlight the differences between the analog and digital age in the formation of an image. Then we will define what is a digital evidence, and we will introduce Image/Video Forensics as a branch of the forensic sciences, highlighting its potential and limits. In the following, we will examine in detail some methods allowing to retrieve information from images when they are not readily available, and finally will provided a list of free and non-free software to face the daily challenges coming from processing images and videos for forensic purposes. The work ends with a list of publications containing the Best Practices in the field.

1. *Images from analog to digital world*

Before the advent of digital photography, the authenticity of an image presented as evidence in a trial was very rarely questioned. Nowadays, on the contrary, the risk of dealing with manipulated images is very high (Battiato and Galvan, 2013), and image forgeries can be classified into three categories:

- Image processing using computer graphics (GC) methods (e.g. artificially generated / modified objects or details);
- Alteration of the image meaning, without modifying its content (e.g. color variations and / or brightness, resizing);
- Altering the image content, inserting (e.g. copying and pasting) or eliminating (e.g. cropping, deleting) significant parts.

Before examining in detail the various approaches used to check the originality of an image, is important to understand the reason why, at the present time, our confidence in this kind of documents is so diminished. To this aim, it's useful to highlight the differences in the image formation pipeline between the analogical and current ages.

In analog cameras, the image was formed on a thin strip of plastic on which other layers of different materials are superimposed, in quantities and composition depending on technical choices. The copies generated starting from the impressed film (which can be thought as a sort of "matrix" of the image), were identical to each other, excepted for slight chromatic variations determined by the dosages of some reagents. Everytime it was necessary joining the file of an investigation with the related photographs, it was used providing the entire original film, the so-called "negatives" of the images. This last measure was considered enough to avoid any complaints about the originality of that kind of proof. Possessing the entire original film not only assured of having the original images, but furthermore preventing the risk that some "inconvenient" clues could be erased, since every image in the film was marked in chronological order. Actually, even negatives could be altered, as we saw in Figure 1 in the first part of this set of three papers. The approaches to do this were mainly two: acting physically on the film, removing or adding some parts and then extracting the image from the modified negative, or duplicating the negative with a special instrumentation, applying appropriate masks suitable for hide or insert the desired details. In both methods, however, the changes were easily detectable by an expert eye: in the first case examining the modified negative, in the latter leveraging the different characteristics (grain, thickness) of the negative-copy, which for technical reasons were never the same as those of camera rollers.

In today's cameras, from a functional point of view, the analogue of the thin strip of plastic where the image was impressed, is the sensor. In fact, on this thin slice of silicon the information upon the brightness of the scene is stored as captured by the corresponding photoreceptors. In detail, the image formation pipeline in modern devices follows the path exposed in Figure 1: the light coming from the physical world passes through a (more or less complex) system of lens, then through the Color Filter Array, and in the electronic sensor, where the conversion of light to pixel values succeeds. Before this step is completed, it has to undergone to some regulation and clearing inner software, which varies upon the various models. On the sensor, therefore, the image is present for a very short period of time, before being stored in the device's memory.

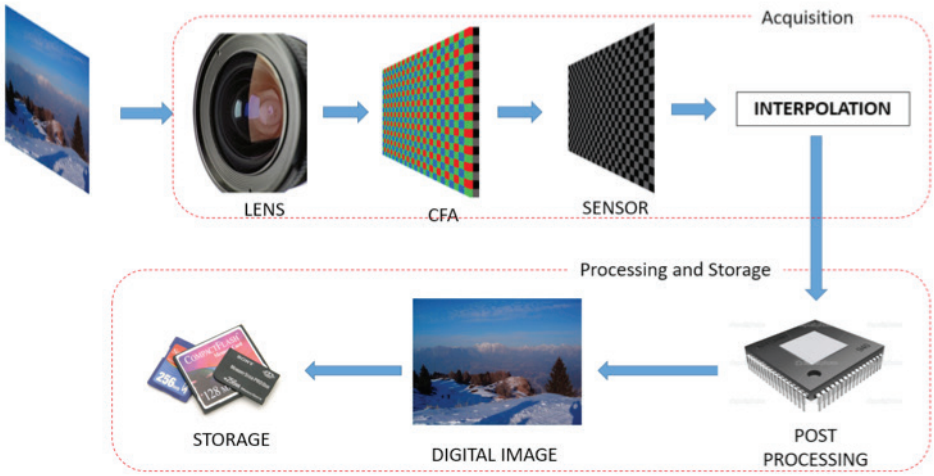

Figure 1: The pipeline of the formation of an image inside a camera: the light coming from the physical world passes through a system of lens, the Color Filter Array, and then in the electronic sensor, where the conversion of light to pixel values succeeds. Before the conversion of the physical signal into the digital image file is completed, it has to undergone to some regulation and clearing inner software, which varies upon the various models.

2. *Digital footprint as a source of evidence*

Digital evidence belongs to the group of the scientific evidences, defined as: *evidences that are provided by some scientific-technical tool with the addition of specific technical skills, possibly with the intervention of an expert in the specific field* (AA.VV., 2008), in particular to the scientific area denoted as **Digital Forensics**. This science includes many subareas, as exposed in Figure 2, and can in turn be described in two ways, with respect to different important aspects:

- **The purposes**: the use of scientifically derived and proven methods toward the preservation, collection, validation, identification, analysis, interpretation, documentation and presentation of digital evidence derived from digital sources for the purpose of facilitation or furthering the reconstruction of events found to be criminal, or helping to anticipate unauthorized actions shown to be disruptive to planned operations (Beebe, 2009).
- **The intervention of an expert**: the science of locating; extracting and analyzing types of data from different devices, which are interpreted by specialists in order to be used as legal evidence (Fenu and Solinas, 2013).

The following formal definition of digital evidence (Carrier, 2003) is widely accepted: *digital data that establish that a crime has been committed can provide a link between a crime and its victim, or between a crime and the perpetrator*. From a practical point of view, a digital evidence is composed by a digital content, often stored as a file of whatever format, and characterized by the following attributes:

- **Volatility**, such as residues of gunpowder. Think about a chat with an internet browser setted in private browsing;
- **Latency**, such as fingerprints or a DNA evidence. This is the case of data that have been erased or hidden (like steganography);
- **Easy to modify or to spoil** since reliability of digital data are intrinsically fragile. Indeed, with a simple copy-paste operation could affect the strength of a digital evidence.

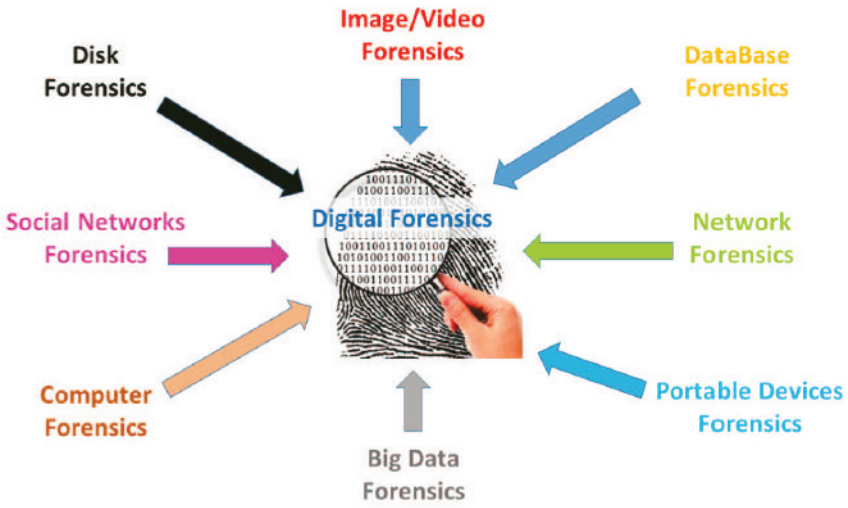

Figure 2: An example of possible categorization of Digital Forensics in sub areas. It's very difficult to separate these components, since often they are closely connected. As an example, if we have to extract a chat session made using WhatsApp stored inside a smartphone (a Portable Devices Forensics's issue), we surely must use Computer Forensics's rules, adding some attentions related to the fact that information in that kind of device may changes during the analysis (if the cell phone is connected to the net). The extracted information is then examined according to DataBase Forensics.

3. *Limits and potentialities of Image / Video Forensics*

As a part of Digital Forensics, Forensic Analysis of Images , also known as Image Forensics, is a forensic science carried out since the very first photos were made. From FBI web site (www.fbi.gov), we could read that "*Forensic Image Analysis is the application of image science and domain expertise to interpret the content of an image or the image itself in legal matters*". From this definition, we can identify the main points that should characterize every forensics approach to the analysis of an image: it must be provided by someone with **adequate technical skills**, but also able **to interpret the extracted information** inside a legal and judicial framework. Image/Video Forensics methods are grouped in six main categories (Redi, Taktak and Dugelay, 2011 - Piva, 2013):

- **Image forgery identification**: Modeling the path of light during image creation reveals physical, geometric, and statistical regularities that are destroyed during the creation of a fake. Various forensic techniques exploit these irregularities to detect traces of tampering. These methods are further divided in Pixel Based (to detect cloning, resampling and splicing operations), Statistical Based, Format Based (to detect, for example, single or multiple JPEG compressions), Camera Based (that specifically model artifacts introduced by various stages of the imaging process), Physics Based (that leverage the inconsistencies introduced in the tampered image when its parts are coming by different environments), and finally Geometric Based (aimed to find inconsistencies connected to the formation of the image inside the camera) (Farid, 2016).
- **Source camera identification**: identification of the device (hopefully the exact one, more often the brand of the device) that generated the image. Sometimes the first step is devoted to discriminate between natural or artificial (also known as Computer Generated) images. In general, the methodology switches to the identification of the source that generated the image, ascertaining the type of device (scanner, camera, copier, and printer) and then trying to determine the particular device.
- **Image reconstruction/restoration/enhancement**: restoration and improvement of the quality of deteriorated images in order to identify, even partially, the original content and/or retrieve useful information (Gonzalez and Woods, 2002).
- **Image/video analysis**: dynamic or behavioral analysis, for example with the aim to identify the *consecutio temporum* of an event of interest.
- **3D reconstruction and comparison**: bi/three-dimensional information extraction to derive measures or reference values (for example the height of an individual) and for the comparison between images (for example to compare the identity of a subject with the known offender from a footage taken by a video surveillance system).
- **Steganalysis**: detection of hidden information within an image with steganographic techniques, for example by changing the least significant bit in the number that defines the color of a pixel (LSB approach).

The first two of the above list are the one more closely connected to forensics issues, whereas the others are general-purposes approaches. Nonetheless, the other areas are equally relevant both in the field of security (e.g. video surveillance) and for more traditional investigative purposes, when simply it is necessary to highlight details and information contained in images. However, we want to point out that such methodologies of analysis can be used to extract the relevant forensic evidence only if the information is present (although apparently in a few amount, as in Figure 4). It may seem trivial to highlight this aspect, but very often this kind of analysis is asked by someone who does not accept the fact that information is actually absent. An example in this sense (unfortunately still very common) concerns images acquired by video surveillance devices, which, despite recording the criminal event, are unusable due to the poor quality of the recovery system. In certain cases, data simply doesn't exist, and can't be "invented". Clear examples of what an Image/Video Forensics will not never be able to do is the so-called "CSI effects", well represented online by video as the one available at: https://www.youtube.com/watch?v=Vxq9yj2pVWk: impossible zooms, 3D reconstructions without any scientific basis, and so on.

In the following, we listed some typical examples of questions that may be posed by the prosecutors to law enforcement agencies in the various steps of an investigation:

- *Image Enhancement*: Is it possible to improve the image/video in order to extract the license plate?
- *Image Autentication*: Did the images/videos undergone to some alterations or are they authentic, with reference to the time of shooting?
- *Source Identification - 1*: Can we state that the images/videos come from the seized device?
- *Source Identification - 2*: Do the images / videos come from a device as camera or videorecorder (which means, shooting a real scene), or are they produced using Computer Graphics (CG) methods?

3. *Extraction of plates and heights: reconstruction of 2D and 3D scenes from images*

Using proper algorithms, which leverage some mathematical tools, it is possible to extract details of an image which are not available at a first sight, and, under certain conditions, even reconstructing with great precision the 3D representation of the scene from where the two-dimensional image came from.

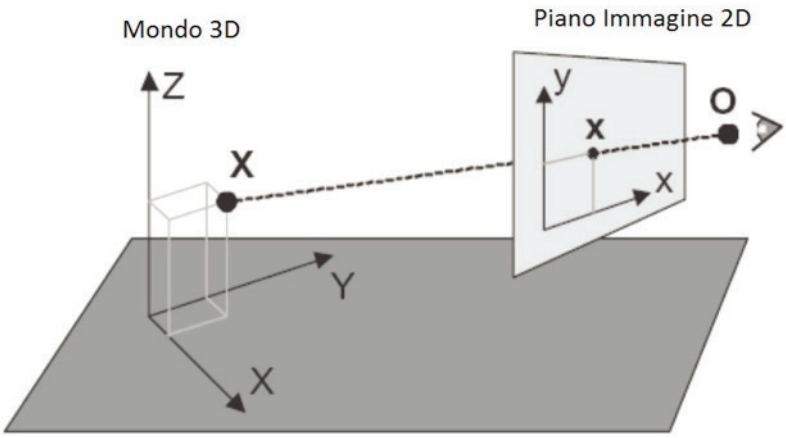

Figure 3: The formation of an image inside a camera can be mathematically modeled as an operation that transports (maps) 3D real-point points into points belonging to a 2D plane (the camera sensor) that matches the image (Criminisi, 2002)

According to the classic Pinhole Camera model, (Criminisi, 2002) an X point in the 3D world is projected (mathematically, "mapped") in the two-dimensional plane of the image in a corresponding point x. This is the intersection of the image plane with the line segment that joins the optical center O of the camera and the X point of the shot scene (Figure 3). The algebraic interpretation of this projection is summarized by the following simple equation:

$$x=\mathbf{P}X \quad (1)$$

where **P** is a mathematical operator called "*projection matrix*", a table whose coefficients define the rules for the transformation of the real points into image points. If we know (or we are able to reconstruct in any way) the matrix **P**, we have the possibility of reverse this transformation. In other words, **starting from a point in the image, we can determine its position within the real scene that the image reproduces**.

In practice, it is not always possible completely rebuilding the projection matrix, but under certain conditions, or when information on the camera that shot the image (step known as *Calibration*) are avoidable, and/or some measures of the objects portrayed in the image, the number of unknowns needed to define the matrix is greatly reduced. An useful example of this last condition is in case of the so called *Rectification*, when the object being photographed is itself 2D. In Figure 4 an inversion of this type has been applied to an image in which we want to highlight the numbers of a car license plate.

Figure 4: Example of rectification. It can be appreciated how this perspective transformation highlights the information relative to the car license plate, otherwise not visible, although present in the original data.

Working on three-dimensional objects, if we need to estimate the height of a subject, it is necessary first of all retrieving the real measurements of objects or elements present in the image (which can be recovered also afterwards). For example, Law Enforcement operators can return to the place where the photo was taken to manually detect the dimensions of a door, a window, or the wall of a house.

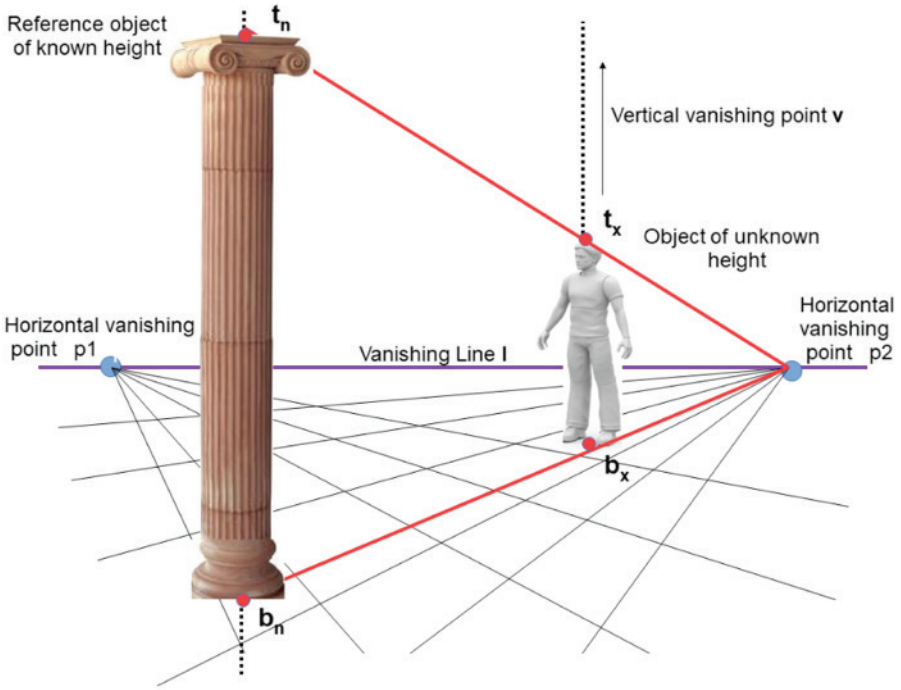

Figure 5: Estimating the size of a subject in an image or footage, requires the calculation of horizontal and vertical vanishing points. In the figure, the height of the person is obtained using a known dimension (the height of the column), the horizontal vanishing points (p1 and p2) and vertical one (v) as well as the vanishing line (or horizon line).

Otherwise, this information can be derived from appropriate documentation, if present. The second step to be taken, consists in extracting from the image a series of characteristics known as vanishing points and lines (Figure 5). In literature there are a lot of algorithms allowing to estimate these geometric entities directly from the image, without knowing the intrinsic parameters (e.g. concerning the internal settings of the camera) or extrinsic ones (concerning the positioning of the camera relative to the scene) (Szeliski, 2010).

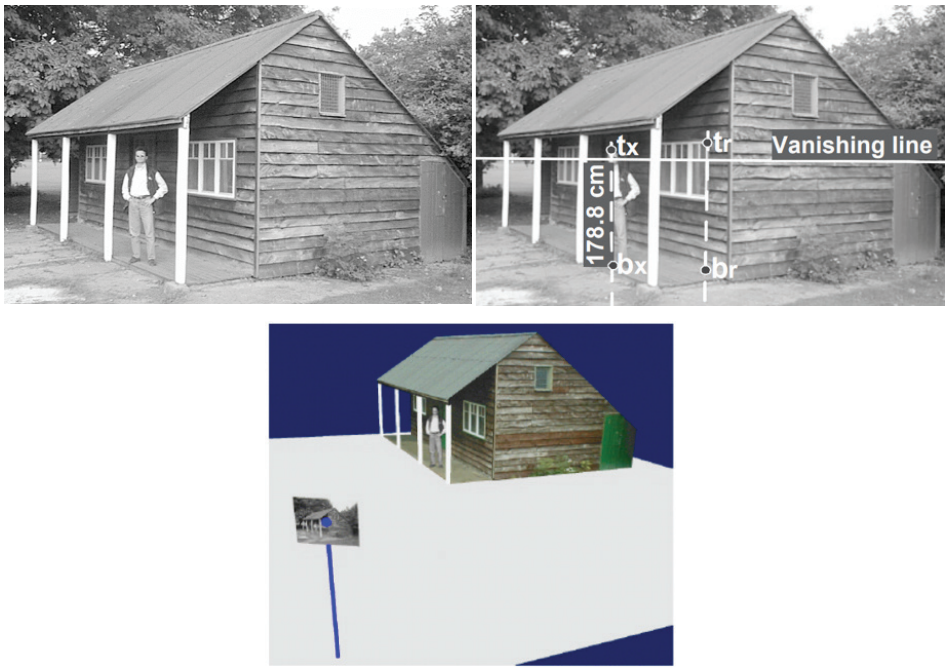

Figure 6: Example of reconstruction of a real scene from a single image: on the left of the top row the original image, on the right of the top row the estimate of the height of the subject, obtained after having retrieved the horizontal vanishing line and having measured the distance between the points br and tr. On the bottom row the 3D reconstruction of the real scene, together with the estimated position of the camera at the time of shooting (Criminisi, 2002).

Searching for information in images, sometimes we can go far beyond the simple extrapolation of single measures, like heights of people or objects. Indeed, increasing the number of "real" measures detected, it is possible to extend the above methods up to reconstructing the entire 3D scene from which the image is taken, as well as the position of the camera at the time of shooting. An example is shown in Figure 6, where the heights of the main window frame, of a column, and the dimensions of the two sides of the porch base were used as reference dimensions (Criminisi, 2002).
We want to point out that, like all estimation methods, the processes described above are subject to errors that may derive from multiple sources: the incorrect collection of reference measurements, an image affected by distortions (e.g. lens distortion, blurred, poor definition), subject not perfectly vertical, etc. In general, however, it is possible to obtain a good estimate, together with a known and measurable error level.

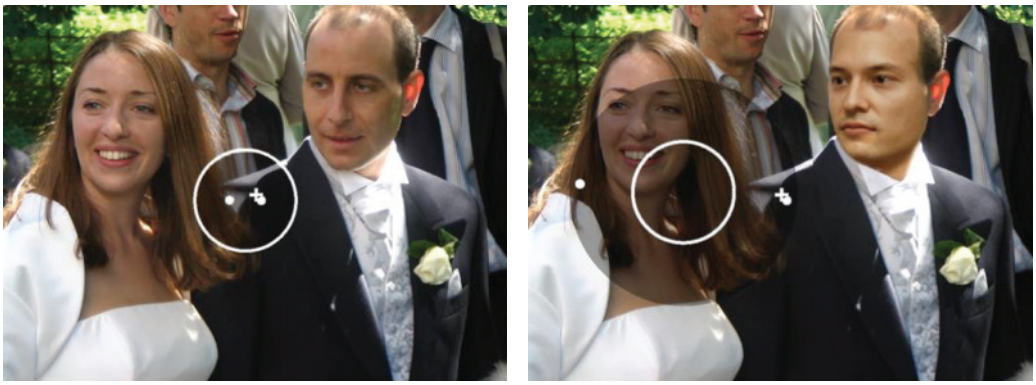

Figure 7: In an authentic image (on the left), the principal point of any subject is close to the center of the photo. In case of a forgery, obtained by the artificial insertion of the face of the man in the image (on the right), its principal point results away from the same points relative to the other subjects of the scene (Johnson and Farid, 2007).

Beyond the application to biometry, in a forensics scenario these approaches may be useful also for Image Forgery Detection, since the discrepancies between geometric clues are sought in different parts of the photo to highlight possible manipulations (Johnson and Farid, 2007 - Wu and Wang, 2011). In the paper, leveraging the estimate of the *principal point* of the camera (the projection of the optical center on the image plane), authors are able to highlight the non-originality of the photo. Indeed, in case of an authentic image this point is close to the center of the photo image for every subject, whereas in case of a forgery made by image splycing, as can be appreciated in Figure 7, the principal point of a subject artificially inserted in the image would result away from the same points relative to the other subjects of the scene. ©

REFERENCES

- S. Battiato, F. Galvan: La validità probatoria di immagini e video. Sicurezza e Giustizia – II, pp. 30:31, 2013.
- AA.VV.: Enciclopedia del Diritto. Giuffrè Editore, 2008.
- N. Beebe: Digital forensic research: The good, the bad and the unaddressed. Advances in Digital Forensics V, pages 17–36. Springer, 2009.
- G. Fenu and F. Solinas: Computer forensics between the italian legislation and pragmatic questions. International Journal of Cyber-Security and Digital Forensics (IJCSDF), 2(1):9–24, 2013.
- B. Carrier, E. H. Spafford: Getting physical with the digital investigation process. International Journal of digital evidence, 2(2):1–20, 2003.
- A. Piva: An overview on image forensics. ISRN Signal Processing, 2013.
- J.A. Redi, W.Taktak, J.L. Dugelay: Digital image forensics: a booklet for beginners. Multimedia Tools Application - 51:133-162, 2011;
- H. Farid: Photo forensics. MIT Press, 2016.
- R. C. Gonzalez, E. R. Wood: Digital Image Processing. Prentice Hall Press, 4th Edition, 2018.
- A. Criminisi: Single-view metrology: Algorithms and applications. Joint Pattern Recognition Symposium. Springer, Berlin, Heidelberg, 2002.
- R. Szeliski: Computer Vision. Algorithms and Applications, Springer, 2010.
- M.K. Johnson, H. Farid: Detecting Photographic Composites of People, 6th International Workshop on Digital Watermarking, Guangzhou, China, 2007.
- L. Wu,Y. Wang: Detecting Image Forgeries using Geometric Cues. Chapter in Computer Vision for Multimedia Applications: Methods and Solutions, 2011.