

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.





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. Introduction

If it is true that "a picture is worth a thousand words" (Brisbane, 1911), nowadays our whole life is becoming increasingly more reliant upon images. Virtually all our visual memories are stored in real-time in our devices, uploaded in the cloud (often without the user realizing it), and possibly shared through the web. The amount of multimedia data created each day on the Internet is massive and impressive. According to (Schultz, 2017), in 2017 more than 4 million hours of content have been uploaded to Youtube, with users watching 5.97 billion hours of Youtube videos, 200,000,000 photos have been uploaded on Facebook and 67,300,000 pictures have been posted on Instagram. This enormous pervasiveness of images has a lot of consequence in all the aspects in our everyday life. How many times we form our own conviction in mind about an event, simply looking at the relative images or footages presented to us by the media. In most of the cases simply there is no time to investigate if a visual information is true or false, so people simply trust what sometimes is a fake. The danger in this behavior in that, once an opinion is rooted in this way, is really difficult to remove it (Nash, Wade and Lindsay, 2009 - Sacchi, Agnoli and Loftus, 2007).

The lack of human ability to distinguish between tampered and original images (Schetinger, Oliveira, da Silva and Carvalho, 2015), helps to increase the risk for mankind to be fooled by malicious agents. In this complex environment, the overall feeling for the Integrity Verification in Multimedia is constantly increasing (Battiato, Giudice and Paratore, 2016).

In the forensics scenario, a natural consequence of this pervasiveness of visual sources of what is called "liquid knowledge" is that very often one (or more) images or videos could become fundamental evidences in a large and heterogeneous set of legal trials. Like all the finds used as evidence, images and footage become valid, and therefore admissible, only if they had been acquired, processed and stored according with the required procedures. One of the main steps in this pipeline is devoted to ascertain the originality of the evidence. But how can we trust images or footage? How can we be sure about the source from where they are supposed to come, and most of all, how can we prove that the visual content that we would like to validate as an evidence (possibly a primary one) has not been altered?

2. The evolution of Image/Video Tampering

Despite to what one might imagine, the first documented examples of image manipulation (Figure 1) date back to 1860, only a few decades after the birth of photography. Although the reasons of this photo-editing have never been clarified, experts think that this first version of "cut-and-paste" forgery has been motivated by the best physical appearance of the politician John Calhoun compared to the one of the most famous colleague. From that episode, there had been thousands of cases in which more or less important images have undergone to substantial changes. Figures 3-6 also show some examples of "photo tampering throught history", as it is called the exhaustive and updated list of examples, spread in different historical periods, available on the famous website http://pth.izitru.com/, from which all these samples were taken.



Figure 1: The picture known as the first photographic counterfeiting on the left (circa 1860), is obtained by combining the head of American president Abraham Lincoln with the body of the southern politician John Calhoun (right).

ESSENTIAL MATERIAL

IV / MMXVIII



Figure 2: The book above (made in 1946) is one of the first known manuals about Photo Tampering in the analog era.

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

As shown in Figure 2, yet in the late 40s there were manuals (i.e. Wardell, 1946) containing technical tricks and related strategies explaining how to physically retouch images. In the old analogue world modifying the printed pictures or in some cases the negative film was in the exclusive preserve of very few experts. With the advent of digital devices and sophisticated photo editing software, the manipulation of digital images is becoming more and an easy task for common user.



Figure 3: On the left image, King George VI was removed from the original photograph, where he previously appeared together with Queen Elizabeth and Canadian Prime Minister William Lyon Mackenzie (1939).



Figure 4: TIn the upper image, the so called "Gang of Four" (Mao's wife Jiang Qing, and her associates Wang Hongwen, Yao Wenyuan, and Zhang Chunqiao) were cutted out from the original photograph of a ceremony at Tiananmen Square (1976).

During the 2004 Presidential primaries, as Senator Kerry was campaigning for the Democratic nomination, the image at the left of Figure 5 appeared, showing Senator John Kerry and Jane Fonda sharing a stage at an anti-war rally. Its caption was: "The Actress and anti-war activist Jane Fonda speaks to a crowd of Vietnam as activist and former Vietnam Vet John Kerry (LEFT) listens and prepares to speak next concerning the war in Vietnam (AP Photo)". The picture was later discovered to be a fake, composed by the picture of Senator Kerry captured in June 1971, while he was preparing to give a speech at the Register for Peace Rally in Mineola, New York, merged with the one of Jane Fonda, shooted while she was speaking at a political rally in Miami Beach, Florida in August 1972. The aftermath on the campaign of the candidate Kerry have been enormous.



Figure 5: The left image, appeared during the 2004 Presidential primaries, shows Senator John Kerry and Jane Fonda together in an anti-war event. The picture was obtained as a results of a copy-paste (also known as "splicing") between the pictures on the center and on the right, respectively shot in 1971 in 1972.

In the very first hours after Osama bin Laden was killed by US forces in Pakistan on May 2nd, 2011, the image on the right of Figure 6 was shown on Pakistani television, and immediately published by the British newspaper Mail, Times, Telegraph, Sun, and Mirror. Even in this case, the photo is a composite of two separate images, one of an alive Bin Laden, and one of another person.



Figure 6: in 2011, close to the death of Osama Bin Laden, the image of the right was spread through the media, presented as a proof of the terrorist's death. Shortly afterwards it was claimed that it was a composition obtained by joining the underlying part of the left image, made slightly darker, with the overlying part of the central image.

In a shocking TED speech titled "*Fake videos of real people* — *and how to spot them*", last April Supasorn Suwajanakorn, an American expert of Computer Vision and AI, presented his work (Supasorn, Seitz and Kemelmacher-Shlizerman, 2017) about how producing an artificial video (in their example a footage with President Obama pronouncing a speech) starting only from an audio track, and footages of the subject taken in other moments of his life. By modeling the mouth shape at each time instant, they synthesize high quality mouth texture, and composite it with proper 3D pose matching to change what he appears to be saying in a target video to match the input audio track. Their approach is then compared (see Figure 7) with the real video from where the audio was extraxted. The results are amazing.



Figure 7: Comparison between the same moment in the speech by Barak Obama. On the left the original footage, on the right the artificial video, composed starting only by the input audio track of the speech and a database of previous footages.

How different the outcomes of John Kerry's election campaign would have been without the circulation of the above Figure 5? What the social and geopolitical consequences of the relationships with Muslim people would have been without the appearance of Figure 6, which is still circulating on the web? What could happen if a fake video of the current U.S. president, artificially built with the Neural Network approach seen in Figure 7, announced an unexpected change on the U.S. economic policy? Nobody really knows the answers, but the thing that remains is that, still and more than ever, the verification of the authenticity of images and videos is an inescapable requirement. As we will see in the following, this is one of the answers given by some Image/Video Forensics approaches. ©

CITATIONS

- * Arthur Brisbane: Speakers give sound advice. Syracuse Post Standard, 18, 1911.
- Jeff Schultz (10/10/2017) How Much Data is Created on the Internet Each Day? Retrieved from https://blog.microfocus.com/how-much-data-is-created-on-the-internet-each-day/.
- Robert A.. Nash, Kimberley A. Wade, and D. Stephen Lindsay: Digitally manipulating memory: Effects of doctored videos and imagination in distorting beliefs and memories. Memory & Cognition, 37(4):414–424, 2009.
- Dario LM Sacchi, Franca Agnoli, and Elizabeth F. Loftus: Changing history: Doctored photographs affect memory for past public events. Applied Cognitive Psychology, 21(8):1005–1022, 2007.
- Victor Schetinger, Manuel M Oliveira, Roberto da Silva, and Tiago J Carvalho: Humans are easily fooled by digital images. arXiv preprint arXiv:1509.05301, 2015.
- Raymond Wardell: Short cuts to photo-retouching for commercial use ed. The House of Little Books ASIN: B0007E33M6, 1946.
- Suwajanakorn, Supasorn, Steven M. Seitz, and Ira Kemelmacher-Shlizerman: Synthesizing obama: learning lip sync from audio. ACM Transactions on Graphics (TOG) 36.4 (2017): 95.
- Sebastiano Battiato, Oliver Giudice, Antonino Paratore: Multimedia forensics: discovering the history of multimedia contents. In Proceedings of the 17th International Conference on Computer Systems and Technologies 2016, pp. 5–16, 2016.

IV / MMXVIII