A Universal Attack Against Histogram-Based Image Forensics

Mauro Barni, Department of Information Engineering and Mathematical Sciences, University of Siena, Siena, Italy

Marco Fontani, Department of Information Engineering and Mathematical Sciences, University of Siena, Siena, Italy

Benedetta Tondi, Department of Information Engineering and Mathematical Sciences, University of Siena, Siena, Italy

ABSTRACT

In this paper the authors propose a universal image counter-forensic scheme that contrasts any detector based on the analysis of the image histogram. Being universal, the scheme does not require knowledge of the detection algorithms available to the forensic analyst, and can be used to conceal traces left in the histogram of the image by any processing tool. Instead of adapting the histogram of the image to fit some statistical model, the proposed scheme makes it practically identical to the histogram of an untouched image, by solving an optimization problem. In doing this, the perceptual similarity between the processed and counter-attacked image is preserved to a large extent. The validity of the scheme in countering both contrast-enhancement and splicing-detection is assessed through experimental validation.

Keywords: Contrast Enhancement, Counter Forensics, Histogram, Image Forensics, Tamper Hiding, Universal

Images have always played an important communicative role, probably due to their immediacy and presumed objectivity. The advent of digital imaging further increased this trend, since acquiring and sharing photos is nowadays cheap and fast. However, creating forgeries and editing photos has also become much easier, thus raising some doubts about the reliability of what we see through digital images.

As an answer to this growing inconvenience, Multimedia (MM) Forensics is emerging as a discipline that aims at revealing the history of digital contents (image, video, audio) using a blind approach. Unlike active techniques like digital watermarking or fingerprinting, MM Forensics does not assume that the content is generated or controlled by the subject that will have to ensure its authenticity. Instead, the idea
at the basis of MM Forensics is that almost every step typically undergone by a digital content (e.g., acquisition, encoding, or application of processing operators) leaves a number of footprints into the media. By leveraging on these footprints, several methods have been proposed to reach some conclusions on the past history of the object under analysis: there are techniques for integrity verification, source identification or classification, analysis of near-duplicates dependencies and many others (see Redi, Taktak, & Dugelay, 2011) for a recent survey).

Together with the continuous development of new forensic techniques, however, counter-forensic (CF) methods are being developed as well. As suggested by the name, the goal of counter-forensics is to conceal the traces introduced by processing tools when the user edits/tamper a MM content, so to make his actions undetectable. As it will be clarified in the following section, existing approaches are mostly targeted at deceiving a specific detector: they exploit knowledge of the forensic algorithm to erase the traces it looks for, of course limiting the perceptual impact of the modifications. In doing so, however, they may introduce new artefacts, that could be detected using different (perhaps more sophisticated) forensic tools. This can lead to a “cat-and-mouse” game where several iterations of the forensic/counter-forensic loop are carried out. It would be interesting, instead, to devise universal CF methods that give the attacker more warranties about the undetectability of the processing operations, at least under some assumptions.

In this paper we extend our previous work in Barni, Fontani, and Tondi (2012), proposing a universal approach for concealing traces left in the image histogram by any processing operator. Compared to that work, we cast the counter-forensic scheme into the more appropriate theoretical framework provided in Barni and Tondi (2012), that better fits the realistic scenario we are considering.

The proposed tool is extremely useful whenever we can assume that the Forensic Analyst (FA) only considers first order statistics to perform its tests, as, for example, in Stamm and Liu (2008) and Stamm and Liu (2010), and that the Adversary (AD) must satisfy some requirements in terms of desired image quality. Under these assumptions we develop a counter forensic technique that is “universal” in the sense that the AD does not need to know anything about the FA detection algorithms (apart from the fact that they are based on first-order statistics), and that the AD can use the proposed technique, without any changes, to hide histogram traces introduced by any kind of processing tool. Specifically, the AD will first process (or tamper with) the image and then perform slight modifications on the resulting image so to bring the histogram as close as possible to that of another, original, unprocessed image, while respecting strict distortion constraints. Intuitively, if the AD manages to do so, the FA will be forced to classify both the original and the tampered content in the same way, thus committing either a false positive or a false negative error. Of course, this will hold only if the two images are no longer distinguishable based on the statistic the FA relies on (that is, image histogram).

The paper is organized as follows: first we give a brief overview of existing counter-forensic techniques; then we sketch and present the proposed CF approach. In the experimental result section we evaluate the performance of the method both from a universal point of view and, as a case study, in countering a specific state of the art forensic algorithm (Stamm & Liu, 2008). Finally, we also evaluate the impact of the proposed scheme in hiding traces left in the pixel histogram during the creation of realistic forgeries.

**PREVIOUS WORKS IN COUNTER FORENSICS**

Counter-forensics was firstly introduced in a seminal work by Kirchner and Böhme (2007), where the concept of fighting against image forensics was introduced together with a practical application, namely a method for resampling an image without introducing pixel correlations. Furthermore, a simple yet important taxonomy was introduced in Kirchner and Böhme (2007).
Web-Based Child Pornography: Quantification and Qualification of Demand
www.igi-global.com/article/web-based-child-pornography/37425?camid=4v1a