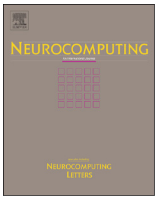
Neurocomputing 189 (2016) 117–122

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom



Brief papers

A Multi-purpose countermeasure against image anti-forensics using autoregressive model

Hui Zeng [a,b], Xiangui Kang [a,*], Anjie Peng [a]

[a] Guangdong Key Laboratory of Information Security Technology, School of Data and Computer Science, Sun Yat-sen University, Guangzhou 510006, China
[b] Jiangsu Engineering Center of Network Monitoring, College of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China

ARTICLE INFO

Article history:
Received 23 October 2015
Received in revised form
17 December 2015
Accepted 24 December 2015
Available online 6 January 2016

Keywords:
Autoregressive model
Counter anti-forensics
C-SVM
Image forensics

ABSTRACT

Image anti-forensics, which aims to remove or forge traces upon which image forensics is based, has made rapid progress recently. To rebuild the credibility of forensics, many countermeasures have been proposed for detecting different anti-forensics. However, most existing countermeasures just target only one type of anti-forensics and are difficult to extend to counter other anti-forensics. In this paper, a multi-purpose countermeasure using autoregressive (AR) model is proposed for detecting various anti-forensics. Experimental results demonstrate that the proposed countermeasure achieves satisfactory performance in detecting all of the five well-known anti-forensic methods discussed in this paper. Even compared to the state-of-art specific counter-measures, our proposed countermeasure achieves similar or better performance.

1. Introduction

Image forensics, which aims to establish trust in images, has been widely applied over the last decade [1]. For example, in law enforcement, the judges resort to image forensics to identify the authenticity of an image before accept it as physical evidence [2]. Owing to its important influence, some farsighted forgers use their knowledge about forensic tools to remove or forge traces upon which forensics is based. The research field that challenges forensics is called anti-forensics [3–8], and its development urges researchers to find countermeasures to rebuild the credibility of forensics.

Up to now, many countermeasures have been proposed for detecting different anti-forensics, such as revealing the traces of JPEG anti-forensics [9–11] and countering median filtering anti-forensics [12]. However, specific countermeasure requires the knowledge of the specific anti-forensics, and most existing methods just target only one type of anti-forensics and are difficult to extend to counter others. For example, the countermeasure to median filtering anti-forensics is not suitable for detecting resampling anti-forensics. As a result, the forensic analyst needs to master as many countermeasures as possible, which is impractical as the rapid development of anti-forensics. Therefore, a multi-purpose countermeasure to various anti-forensics is needed. This work makes a first attempt in this direction.

We begin with analyzing the common traces left by five well-known anti-forensics, and find that all of these methods destroy some inherent local correlation with an original image. Based on this, we propose a multi-purpose countermeasure using autoregressive (AR) model. Experimental results have shown the versatility of this countermeasure. Even compared with some state-of-art specific countermeasures, the proposed countermeasure achieves comparable or better performance.

The rest of this paper is organized as follows. Section 2 analyzes the common traces left by five well-know anti-forensic methods. Our proposed countermeasure is described in detail in Section 3. Section 4 shows the experimental results and the conclusion is made in Section 5.

2. Traces left by anti-forensics

To reveal the common traces left by anti-forensics, we first make a review of five well-know anti-forensics and existing countermeasures in three forensic scenarios.

2.1. JPEG anti-forensics

The first scenario we considered is the case of JPEG forensics. It is well known that JPEG compression would introduce quantization artifact and blocking artifact [13]. To remove the

[*] Corresponding author.
E-mail address: isskxg@mail.sysu.edu.cn (X. Kang).

http://dx.doi.org/10.1016/j.neucom.2015.12.089
0925-2312/© 2016 Elsevier B.V. All rights reserved.

quantization artifact, a typical anti-forensics was proposed in [3]. The authors introduced dither into the DCT coefficients to approximately restore the histogram of each subband. Fig. 1a–c show the histogram of coefficient values in the (2, 2) DCT subband from an uncompressed image, the JPEG compressed image with quality factor (QF) 75 and the anti-forensically modified image respectively. It is observed that the quantization artifact is successfully removed with anti-forensics. This forgery attracted much attention in forensic area and several countermeasures have been proposed [9,10].

As an extension work of [3], the same authors combined two post-processing to simultaneously remove the quantization artifact and blocking artifact in [4]. That is, boundary blurring on the result image after add dithering. This operation not only removes the traces of blocking artifact, but also makes the aforementioned countermeasures [9,10] lose effectiveness. To counter the extended anti-forensics, a specific countermeasure based on noise level estimation was proposed in [11].

2.2. *Median filtering anti-forensics*

The second scenario we considered is the median filtering detection. Most existing median filtering detectors are operating on the statistics of pixel value differences of an image [14,15]. Subsequently, a target attack to these detectors was proposed by modifying pixel difference distribution with adding anti-forensic noise [5]. Fig. 2a–c show the pixel difference distribution for an original image, the median filtered image and the anti-forensically modified one respectively. It is observed that the pixel difference distribution is successfully restored by the anti-forensics.

To counter such anti-forensics, a specific countermeasure [12] was proposed by analyzing the periodicity of the noise adding strategy of the forger. However, it becomes useless facing another median filtering anti-forensic method [6]. The basic idea of [6] is to add perturbation in highly textured areas to interfere with footprint left by median filtering. This method would not leave periodical trace as that in the method [5], so it cannot be detected by the specific countermeasure [12].

2.3. *Resampling anti-forensics*

The final scenario we considered is the resampling detection. Resampling is a common operation involved in image tampering and the detection of resampling can provide important information for forensics [16]. The first popular resampling detector is based on the periodic pattern in the residual signal of local linear predictors in the spatial domain [17]. To avoid the periodic pattern, a so called *dual-path* anti-forensics was proposed by introducing edge-modulated geometric distortion during resampling [7]. Fig. 3 illustrates the detection process. From top to bottom: an original image, the same image upsampled by 20%, and the forged image. The estimated probability maps [17] are displayed in the middle column and the Fourier transform of these maps are displayed in the right column (For display purpose, each Fourier transform was high-pass filtered and independently auto-scaled). For the resampled image, there is periodic nature in the probability map and obvious peaks in corresponding Fourier transform, whereas for the anti-forensically modified image, such fingerprints are successfully removed.

2.4. *Common traces of anti-forensics*

All of the anti-forensic methods mentioned above include some certain disturbance in spatial domain or frequency domain, *e.g.*, the dither to DCT coefficients in [3] or Gaussian noise in [4]. As we know, there is some inherent local correlation with an original

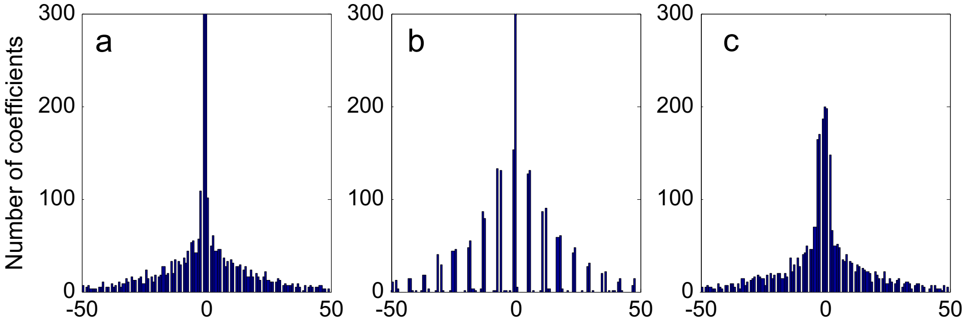


Fig. 1. Histogram of DCT coefficients of the (2, 2) subband. (a) From an original image, (b) from the same image after JPEG compression, QF=75, (c) from the anti-forensically modified image [3].

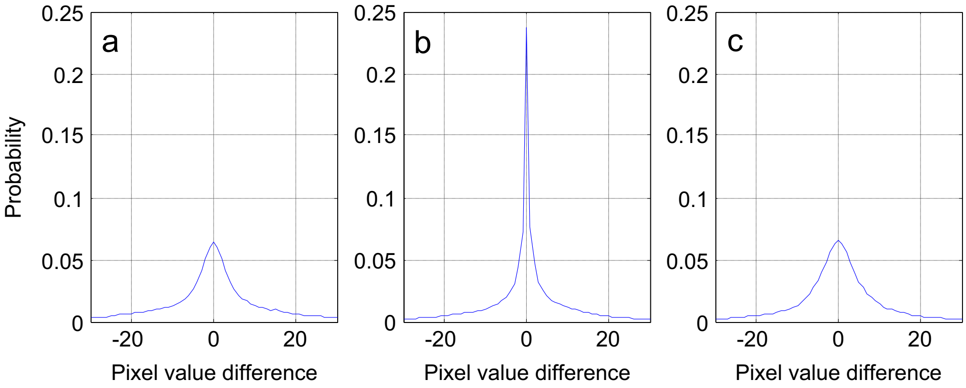


Fig. 2. Histogram of pixel value difference. (a) From an original image, (b) From the same image after median filtering and (c) From the anti-forensically modified image [5].