

Source Camera Identification Issues: Forensic Features Selection and Robustness

Yongjian Hu, South China University of Technology, China, and University of Warwick, UK

Chang-Tsun Li, University of Warwick, UK

Changhui Zhou, South China University of Technology, China

Xufeng Lin, South China University of Technology, China

ABSTRACT

Statistical image features play an important role in forensic identification. Current source camera identification schemes select image features mainly based on classification accuracy and computational efficiency. For forensic investigation purposes; however, these selection criteria are not enough. Consider most real-world photos may have undergone common image processing due to various reasons, source camera classifiers must have the capability to deal with those processed photos. In this work, the authors first build a sample camera classifier using a combination of popular image features, and then reveal its deficiency. Based on the experiments, suggestions for the design of robust camera classifiers are given.

Keywords: *Camera Identification, Digital Image Forensics, Image Feature Selection, Pattern Classification, Robust Camera Classifier*

INTRODUCTION

Influenced by classical steganalysis (Farid, 2002; Avcibas, 2003), the use of statistical image features becomes common for source imaging device (e.g., camera, scanner) identification. Source imaging device identification can be thought of as a process of steganalysis if device noise in images is regarded as a disturbance caused by externally embedded messages. As

a result, the statistics of the images captured by different cameras are believed to be different.

A variety of image features have been proposed and studied in prior arts of steganalysis. In Farid and Lyu (2002), they found that strong higher-order statistical regularities exist in the wavelet-like decomposition of a natural image, and the embedding of a message significantly alters these statistics and thus becomes detectable. Two sets of image features were studied. The mean, variance, skewness and kurtosis of the subband coefficients form the first feature

DOI: 10.4018/jdcf.2011100101

set while the second feature set is based on the errors in an optimal linear predictor of coefficient magnitude. A total of 216 features were extracted from the wavelet decomposed image to form the feature vector. Support vector machines (SVM) were employed to detect statistical deviations. Avcibas *et al.* (2003) proved that steganographic schemes leave statistical evidence that can be exploited for detection with the aid of image quality features and multivariate regression analysis. To detect the difference between cover and stego images, 19 image quality metrics (IQMs) were proposed as steganalysis tools.

Statistical image features were introduced for forensic image investigation as soon as this research field emerged. In one early camera identification scheme, Kharrazi *et al.* (2004) studied a set of features that designate the characteristics of a specific digital camera to classify test images as originating from a specific camera. CFA (color filter array) configuration, demosaicing algorithms and color processing/transformation were believed to have great impact on the output image of camera. Thus, three average values in RGB channels of an image, three correlations between different color bands, three neighbor distribution centers of mass in RGB channels as well as three energy ratios between different color bands were used for reflecting color features. Moreover, each color band of the image was performed with wavelet decomposition, and the mean of each subband was calculated, just as in Farid and Lyu (2002). In addition to color features, 13 IQMs were borrowed from Avcibas *et al.* (2003) to describe the characteristics of image quality. The average identification accuracy for their SVM classifier was 88.02%. This scheme was re-implemented on different camera brands and models in (Tsai *et al.*, 2006).

In one early scanner identification scheme, Gou *et al.* (2009) proposed a total of $30+18+12=60$ statistical noise features to reflect the characteristics of the scanner imaging pipeline and motion system. The mean and STD (standard deviation) features were extracted using 4 filters (i.e., averaging filter, Gaussian

filter, median filter, and Wiener adaptive filters with 3×3 and 5×5 neighborhood) in each of three color bands to form the first $2\times 5\times 3=30$ features. The STD and goodness of Gaussian fitting were extracted from the wavelet decomposed image of each color band in 3 orientations to form another $2\times 3\times 3=18$ wavelet features. Two neighborhood prediction errors were calculated from each color band at two brightness levels to form the last $2\times 3\times 2=12$ features. The outcome of their SVM classifier had the identification accuracy over 95%.

Another scanner identification scheme was proposed by Khanna *et al.* (2009). Unlike Gou *et al.* (2009) that used three types of features, only statistical properties of the sensor pattern noise (SPN) were used. The SPN was first proposed for correlation-based camera identification in Lukas *et al.* (2006). The major component of SPN is the photo response non-uniformity noise (PRNU). Due to the similarity between camera and scanner pipelines, the PRNU-based detection was extended for scanner identification. However, the camera fingerprint is a 2-D spread-spectrum signal while the scanner fingerprint is a 1-D signal. So Khanna *et al.* (2009) proposed a special way to calculate the statistical features of the linear PRNU. The mean, STD, skewness, and kurtosis of the row correlations and the column correlations form the first eight features on each color channel of the input image. The STD, skewness, and kurtosis of the average of all rows and the average of all columns form the next six features. The last feature for every color channel is representative of the relative difference in periodicity along the row and column directions of the sensor noise. The results from their SVM-based classifier were often better than those in Gou *et al.* (2009). The robustness of this PRNU features-based scanner classifier was evaluated when subject to JPEG compression, contrast stretching and sharpening.

Other image features-based schemes include Filler *et al.* (2008), Tsai *et al.* (2007), Tsai and Wang (2008). In Filler *et al.* (2008), four sets of image features related to PRNU were used. In Tsai *et al.* (2007), the impact of image content on camera identification rates

13 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/source-camera-identification-issues/62074

Related Content

Development of Secured Log Management System Over Blockchain Technology

Sagar Shankar Rajebhosale and Mohan Chandrabhan Nikam (2019). *International Journal of Cyber Research and Education* (pp. 38-42).

www.irma-international.org/article/development-of-secured-log-management-system-over-blockchain-technology/218896

Cyber Laws for Preventing Cyber Crimes Against Women in Canada

Debarati Halder and K. Jaishankar (2012). *Cyber Crime: Concepts, Methodologies, Tools and Applications* (pp. 765-776).

www.irma-international.org/chapter/cyber-laws-preventing-cyber-crimes/60980

The Unprecedented Rise in Cybercrime and the Role of the Human Vulnerability Factor

Nabie Y. Conteh and Malcolm D. Royer (2021). *Ethical Hacking Techniques and Countermeasures for Cybercrime Prevention* (pp. 32-43).

www.irma-international.org/chapter/the-unprecedented-rise-in-cybercrime-and-the-role-of-the-human-vulnerability-factor/282223

Source Camera Identification Based on Sensor Readout Noise

H. R. Chennamma and Lalitha Rangarajan (2010). *International Journal of Digital Crime and Forensics* (pp. 28-42).

www.irma-international.org/article/source-camera-identification-based-sensor/46045

Design a Wireless Covert Channel Based on Dither Analog Chaotic Code

Pengcheng Cao, Weiwei Liu, Guangjie Liu, Jiangtao Zhai, Xiao-Peng Ji, Yuewei Dai and Huiwen Bai (2021). *International Journal of Digital Crime and Forensics* (pp. 115-133).

www.irma-international.org/article/design-a-wireless-covert-channel-based-on-dither-analog-chaotic-code/272837