Signal Processing 121 (2016) 1–16

Contents lists available at ScienceDirect

Signal Processing

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

Multi-scale image hashing using adaptive local feature extraction for robust tampering detection

Cai-Ping Yan, Chi-Man Pun*, Xiao-Chen Yuan

Department of Computer and Information Science, University of Macau, Macau SAR, China

ARTICLE INFO

Article history:
Received 29 April 2015
Received in revised form
21 October 2015
Accepted 27 October 2015
Available online 10 November 2015

Keywords:
Tampering detection
Multi-scale image hashing
Adaptive Feature Point Detection
Image authentication

ABSTRACT

The main problem addressed in this paper is the robust tampering detection of the image received in a transmission under various content-preserving attacks. To this aim the multi-scale image hashing method is proposed by using the location-context information of the features generated by adaptive and local feature extraction techniques. The generated hash is attached to the image before transmission and analyzed at destination to filter out the geometric transformations occurred in the received image by image restoration firstly. Based on the restored image, the image authentication using the global and color hash component is performed to determine whether the received image has the same contents as the trusted one or has been maliciously tampered, or just different. After regarding the received image as being tampered, the tampered regions will be localized through the multi-scale hash component. Lots of experiments are conducted to indicate that our tampering detection scheme outperforms the existing state-of-the-art methods and is very robust against the content-preserving attacks, including both common signal processing and geometric distortions.

1. Introduction

With the ease of digital image manipulation, ensuring credibility of the image contents has been becoming a common concern. The rapid development of image editing software dramatically increases the doctored photographs. If tampered images are extensively used in the official media, scientific discovery and forensic evidence, will undoubtedly reduce trustworthiness and produce serious impact on various aspects of the society. Tampering detection, a scheme that identifies the integrity and primitivism of the digital multimedia data, has been proposed in recent years [1]. Generally, there are two kinds of tampering, copy-move forgery [2] and splicing. The main problems existing in this area are the authentication of the image received in a communication and the location of the regions of the image which have been tampered. To address these problems, two main categories of tamper detection approaches have been introduced: the watermarking based approaches and the signature based approaches.

In the watermarking based tamper detection approaches, the watermark [3–5] was embedded into the host image without perceptual distortion and then was extracted to judge if there was a malicious manipulation on the received image. This category of approaches should ensure that the watermark can survive against the common attacks such as lossy compression and noise addition; meanwhile, the watermark should be sensitive to the distortions introduced in malicious manipulation. Unfortunately, in this way, the watermark should previously be encoded and which will distort the contents of the host image. Different from watermarking based approaches, the signature based approaches require no

* Corresponding author. Tel.: +853 88224369.
E-mail addresses: yb47428@umac.mo (C.-P. Yan), cmpun@umac.mo (C.-M. Pun), xiaochen_yuan@ieee.org (X.-C. Yuan).

http://dx.doi.org/10.1016/j.sigpro.2015.10.027
0165-1684/

embedding process. Instead, secured image hashing which maps an input image to a small and robust string is utilized to generate the hash/signature. The hash/signature is exclusively attached for each host image and it may slightly change when the content-preserving manipulations are applied. The general procedure in this aspect is: 1) a robust hash designed for content-based identification is attached to the host image; 2) the hash is analyzed at the destination to verify the reliability of the received image.

Different signature/hash based tamper detection approaches have been recently proposed. Venkatesan et al. [6] first introduced the image hashing concept. They used the non-reversible compression of wavelet coefficients as descriptors to generate the hash, which took the robustness against compression, geometric distortions, and some other attacks into consideration. Roy et al. [7] first developed the hashing method that can localize image tampering using a short signature. However, they only investigated the robustness against some limited attacks such as rotation, cropping, and compression. Motivated by Singular Value Decomposition (SVD) [8], a new dimension reduction method called Non-negative Matrix Factorization (NMF) [9,10] was introduced. The NMF significantly enhanced hashing robustness under a large class of perceptually insignificant attacks while allowing an acceptably small length of hashing, but suffered from brightness changes and large geometric transforms. In order to estimate the parameters of the geometric transforms (i.e., rotation and scaling) so that the tampered areas can be located, several image alignment techniques have been proposed [11–13]. In [11], the geometric transform estimation was completed by exploiting information extracted through Radon transform and scale space theory, which was necessary to implement further integrity check such as tampering localization. In [12], SIFT features were encoded into a compact visual words representation for geometric transform estimation, and a hybrid construction using both SIFT and block-based features were used to detect and localize image tampering. In [13], a more robust image alignment method by encoding spatial distribution of features to deal with highly textured and contrasted tampering patterns was proposed, and the geometric transformation was estimated based on a voting procedure in the parameter space of the model. A wavelet based image hashing method was developed in [14], which was robust to most content-preserving operations and can be used to detect tampered regions. Zhao et al. [15] used Zernike moments and local features to design image hash, yet which can only detect salient regions and can tolerate limited attacks. Lv et al. [16] proposed a novel shape-contexts-based image hashing approach using SIFT-Harris detector, which divided the image into rings and sectors. The method was robust to a wide range of geometric attacks and can be applied for image tampering detection. However, when the tampered area was located in the center of the image, the central orientation estimation of the tampered image based on radon transform would be error, which would directly lead to the failure of tampering localization.

Most of the above-mentioned methods have good performance in certain aspects, but they may not complete tampering detection comprehensively. For example, many of the existing tamper detection methods cannot detect the tampered area with arbitrary size and position; in addition, the locations of the tampered regions are difficult to detect accurately, especially when under various content-preserving attacks. Recently, a perceptual image hash method by combining image-block-based features and key-point-based features was proposed in [17]. Although this method can achieve the goal of localizing tampered regions accurately, its hash length is tens of thousands. In order to solve these problems, we propose a robust tampering detection scheme based on the multi-scale image hashing and adaptive local feature extraction in this paper. Our approach has several desirable contributions: First, an adaptive local feature extraction method is proposed based on the popular Scale Invariant Feature Transform (SIFT) [18] for more robust feature descriptors. Second, a multi-scale image hashing method and the location-context generation technique which encoding the geometric distribution and image content together are proposed. Third, an effective image authentication and tampering localization methods are proposed successively to accurately detect the tampering for different tampered images. Based on the above contributions, the constructed system for tampering detection is robust against various content-preserving attacks, including both common signal processing and geometric distortions. A comprehensive testing dataset is created with tampered images of various types in which the inserted/removed regions are of different sizes and located at different positions, and tampered images with various attacks, to verify the effectiveness and robustness of the proposed tampering detection scheme.

The remainder of the paper is organized as follows. Section 2 presents the proposed multi-scale image hashing method by using Adaptive Feature Point Detection and local feature generation. Section 3 explains the proposed tampering detection scheme in detail, including the image restoration, image authentication, and tampering localization. Section 4 demonstrates the experiments and analysis of the results. Finally, the conclusions and future work are drawn in Section 5.

2. Multi-scale image hashing using adaptive local feature extraction

An image hash [6–17] is a distinctive signature which represents the visual content of the image in a compact way. The image hash should be robust against common operations and meanwhile should be different from the one computed on a different/tampered image. Image hashing techniques are considered extremely useful to validate the authenticity of an image received through a communication channel. In this paper, we have proposed the novel multi-scale image hashing approach which can not only achieve the binary decision task related to image authentication, but also accomplish the request of localization of the tampered regions.

The framework of the proposed multi-scale image hashing approach is shown in Fig. 1. First, the Adaptive Feature Point Detection method is proposed by using the SIFT algorithm. Then the local features are generated by applying the Stationary Wavelet Transform (SWT) [19,20]