Information Sciences 345 (2016) 226–242

Contents lists available at ScienceDirect

Information Sciences

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

Multi-Level Dense Descriptor and Hierarchical Feature Matching for Copy–Move Forgery Detection

Xiuli Bi, Chi-Man Pun*, Xiao-Chen Yuan

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

ARTICLE INFO

Article history:
Received 17 September 2015
Revised 6 January 2016
Accepted 11 January 2016
Available online 4 February 2016

Keywords:
Copy–Move Forgery Detection (CMFD)
Multi-Level Dense Descriptor (MLDD)
Hierarchical Feature Matching
Color Texture Descriptor
Invariant Moment Descriptor

ABSTRACT

In this paper, a *Multi-Level Dense Descriptor* (*MLDD*) extraction method and a Hierarchical Feature Matching method are proposed to detect copy–move forgery in digital images. The *MLDD* extraction method extracts the dense feature descriptors using multiple levels, while the extracted dense descriptor consists of two parts: the *Color Texture Descriptor* and the *Invariant Moment Descriptor*. After calculating the *MLDD* for each pixel, the Hierarchical Feature Matching method subsequently detects forgery regions in the input image. First, the pixels that have similar color textures are grouped together into distinctive neighbor pixel sets. Next, each pixel is matched with pixels in its corresponding neighbor pixel set through its geometric invariant moments. Then, the redundant pixels from previously generated matched pixel pairs are filtered out by the proposed Adaptive Distance and Orientation Based Filtering method. Finally, some morphological operations are applied to generate the final detected forgery regions. Experimental results show that the proposed scheme can achieve much better detection results compared with the existing state-of-the-art CMFD methods, even under various challenging conditions such as geometric transforms, JPEG compression, noise addition and down-sampling.

1. Introduction

Recently, there has been much interest in Copy–Move Forgery Detection (CMFD). Copy–move forgery (CMF) is one of the most common types of image forgery. An image with CMF contains at least two regions with identical content where one or more distinct regions are copied and pasted into one or more destination locations in the same image to conceal important information. Sometimes, the copied content is modified using a pre-processing operation such as scaling, rotation, noise addition, or JPEG compression to make it match the surrounding regions in such a way that the tampering is not visible. Many CMFD schemes have been proposed for finding these tampered regions, and most of the detection schemes follow a common pipeline that includes three main steps [9]: (1) feature extraction, which extracts an appropriate feature from each of the blocks or interesting pixels; (2) feature matching, which obtains the best match of the blocks or interesting pixels based on corresponding extracted features; and (3) post-processing, which filters the offset field and the linking pixels with their nearest neighbors to improve the detection accuracy.

Generally speaking, the three-step pipeline can be applied to each pixel of the host image, in which case the field is dense. Alternatively, it can be applied to selected key points, in which case the field is sparse. Of the existing CMFD schemes, the keypoint-based methods [1,7,13,18,26,31,35] were proposed, which extract a relatively small set of pixels from the host

* Corresponding author. Tel.: +853 88224369.
E-mail addresses: yb47429@umac.mo (X. Bi), cmpun@umac.mo (C.-M. Pun), xiaochen_yuan@ieee.org (X.-C. Yuan).

http://dx.doi.org/10.1016/j.ins.2016.01.061
0020-0255/

images and then perform dense matching based on the feature descriptors of the keypoints. This approach is usually much faster than methods that are based entirely on dense matching. Pan and Lyu first proposed the keypoint-based algorithm in [26], where a Scale Invariant Feature Transform (SIFT) [21] was used to extract the keypoints and which can guarantee robustness against geometrical transformations. Similarly, SIFT was frequently used for feature extraction in CMFD [19,28]. Other well-known local descriptors such as Speeded Up Robust Features (SURF) [2] and DAISY [33] have also been considered for feature extraction and keypoint-based CMFD [1,13,31,35]. A benchmarking paper [6] evaluated the related approaches and clearly showed the performance gaps. Although the computational complexity of the sparse-filled methods is comparatively less because the number of keypoints represents only a relatively small set of all the pixels in the image, most of them are intrinsically less accurate than the dense-field methods, especially when copy–moves involve only smooth regions, which is typically the case with occlusive forgeries. In addition, the performance of the feature points-based forgery methods is highly related to the resolution of the host images [6]. When the host image is down-sampled, the performance of the feature points-based forgery methods drops considerably.

Considering those problems, we focus on the dense-field approach in this paper. Unfortunately, the complexity of the dense field approach is relatively high because all the pixels undergo the three-step pipeline. The solutions to this problem are to intrinsically simplify the feature extraction and speed up the matching phase itself as much as possible. In addition to considering detection accuracy and complexity, the robustness of the performance should be accounted for, which means that the performance should not be affected by common distortions such as JPEG compression, noise addition, or geometric distortions such as rotation and scaling. In the existing CMFD schemes, the dense field approaches are always known as block-based approaches because the host images are usually divided into overlapping blocks, each of which is considered to be an individual superpixel for calculating the corresponding pixel features. In the existing block-based CMFD methods [3–5,10,14–16,20,22,24,25,27,29,32,34,38], transforms such as the Discrete Cosine Transform (DCT) [5,10,14], Principal Component Analysis (PCA) [24,27], Wavelet [25], SVD [15,38], or Histogram Of Orientated Gradients (HOG) [17] are applied to extract the features, which improves the robustness. Fridrich et al. [10] calculated the quantized DCT coefficients of the overlapping blocks as feature descriptors. Popescu and Farid [27] used the PCA method to reduce the feature dimensions. Mahdian and Saic [24] employed the 24 blur-invariant moments as block features. However, these features are not particularly robust against geometric distortions such as rotation and scaling. In consequence, Ryu et al. [29] used Zernike moments as block features, and Li [20] applied the polar cosine transform to calculate the coefficients for the block features to achieve rotation invariance. Additionally, Wu et al. [34] proposed applying the Fourier–Mellin Transform to calculate the block features, thus achieving scale invariance. Lee et al. [17] applied HOG to each block and extracted statistical features to measure the similarity. Lynch et al. [23] used the average gray values of blocks as dominant descriptors. Recently, Cozzolino et al. [9] proposed an efficient dense field technique for CMFD and the fast approximate nearest-neighbor search algorithm, PatchMatch, was re-sorted. The experiments show that their dense-field technique is more reliable than the keypoint-based CMFD approaches.

The above description compels us to focus on the dense field approach. In this paper, we propose a novel *Multi-Level Dense Descriptor (MLDD)* extraction method, which includes the *Color Texture Descriptor* (*MLDD_CT*) and the *Invariant Moment Descriptor* (*MLDD_IM*) to extract the dense features instead of the existing sparse features. After obtaining the *MLDD* for each pixel, the pixels must be compared to each other to find the matched pairs. Then, to reduce the high computational complexity of the dense-field approach, to enhance the robustness against various attacks, and to make the result of the matching process be as accurate as possible, we propose a novel Hierarchical Feature Matching method that includes three main steps: 1. *Color Texture Based Filtering*, 2. *Geometrical Invariant Moments Based Matching*, and 3. *Adaptive Distance and Orientation Based Filtering*. Using the *Color Texture Based Filtering* technique, we sort the pixels of the host image according to their *MLDD_CT* values to generate the selected neighbor pixels set for each pixel. In this way, pixels with similar color texture will be grouped together into distinctive neighbor pixel sets. With the *Geometrical Invariant Moments Based Matching* technique, we match each pixel with its corresponding neighbor pixel set only through its *MLDD_IM*, which greatly reduces the computational complexity. Finally, the *Adaptive Distance and Orientation Based Filtering* technique can help to filter out redundant pixels and improve the detection accuracy.

In the following, we will provide the framework of the proposed CMFD algorithm using Multi-Level Dense Descriptors (MLDD) and Hierarchical Feature Matching in Section 2 and explain the proposed *Multi-Level Dense Descriptor* extraction and Hierarchical Feature Matching methods in detail. In Section 3, a large number of experiments will be conducted to demonstrate the effectiveness and robustness of the proposed method. Conclusions are drawn in Section 4.

2. Proposed Copy–Move Forgery Detection algorithm

The proposed CMFD method consists of two stages. The first stage is *Multi-Level Dense Descriptor Extraction*, in which the MLDD is generated as a pixel feature. Each MLDD contains two parts: a *Color Texture Descriptor* (*MLDD_CT*) and an *Invariant Moment Descriptor* (*MLDD_IM*). The second stage is *Hierarchical Feature Matching*, in which the pixels of the host image are first sorted according to their *MLDD_CT*, and a selected neighbor pixel set is generated for each pixel. In this way, pixels with similar color texture are grouped together into distinctive neighbor pixel sets. Then, each pixel is matched with its corresponding neighbor pixel set through the calculated *MLDD_IM*, and matched pixels are indicated as matched pixel pairs. Finally, the Adaptive Distance and Orientation based Matching method is proposed to filter out the redundant pixels from the previously generated matched pixel pairs, and the final Detected Forgery Regions can be generated from the remainder.