KIPS Tr. Comp. and Comm. Sys.
Vol.3, No.5 pp.163~172 pISSN: 2287-5891

http://dx.doi.org/10.3745/KTCCS.2014.3.5.163

SIFT 기반 카피-무브 위조 검출에 대한 타켓 카운터-포렌식 기법

Munkhbaatar Doyoddorj† · 이 경 현††

요 약

Scale Invariant Feature Transform (SIFT)은 높은 매칭 능력과 회전이나 스케일 조정 시 안정성으로 인해 이미지 특징 매칭을 위해 많은 응용에서 사용되어지고 있으며, 이러한 특성으로 인해 카피-무브 위조 검출을 위한 핵심 알고리즘으로 각광받고 있다. 하지만 SIFT 변환은 이미지 조작의 증거를 감출 수 있는 안티포렌식의 가능성이 높음에도 불구하고 이에 대한 연구는 거의 없으므로, 본 논문에서는 의미론적으로 허용될 수 있는 왜곡을 적용하여 SIFT 기반 카피-무브 위조 검출을 방해하기 위한 타켓 카운터-포렌식 기법을 제안한다. 제안 기법은 공격자가 유사성 매칭 절차를 속일 수 있는 동시에 SIFT 키포인트의 변형을 통한 추적을 방해하여 이미지 조작의 증거를 숨길 수 있는 방안을 제공한다. 또한 제안 기법은 의미론적 제약 하에서 가공된 이미지와 원본 이미지 간의 높은 충실도를 유지하는 특성을 가진다. 한편, 다양한 조건의 테스트 이미지에 대한 실험을 통해 제안 기법의 효율성을 확인하였다.

키워드 : 카운터-포렌식, SIFT, 카피-무브 위조 검출, 추적 숨김

A Targeted Counter-Forensics Method for SIFT-Based Copy-Move Forgery Detection

Munkhbaatar Doyoddorj† · Kyung-Hyune Rhee††

ABSTRACT

The Scale Invariant Feature Transform (SIFT) has been widely used in a lot of applications for image feature matching. Such a transform allows us to strong matching ability, stability in rotation, and scaling with the variety of different scales. Recently, it has been made one of the most successful algorithms in the research areas of copy-move forgery detections. Though this transform is capable of identifying copy-move forgery, it does not widely address the possibility that counter-forensics operations may be designed and used to hide the evidence of image tampering. In this paper, we propose a targeted counter-forensics method for impeding SIFT-based copy-move forgery detection by applying a semantically admissible distortion in the processing tool. The proposed method allows the attacker to delude a similarity matching process and conceal the traces left by a modification of SIFT keypoints, while maintaining a high fidelity between the processed images and original ones under the semantic constraints. The efficiency of the proposed method is supported by several experiments on the test images with various parameter settings.

Keywords : Counter-Forensics, SIFT, Copy-Move Forgery Detection, Hiding Traces

1. Introduction

Digital image has experienced tremendous growth in recent decades, and digital camera images have been used for a growing number of applications. With such increasing popularity and the availability of low-cost image editing software, the integrity of digital image contents can no longer be taken for granted. Therefore, the research on digital image forensics and tamper detection has gained a great concern. The basic idea of image forensics is that a number of traces are remained in the media when a processing tool is applied to digital contents. Several methods have been proposed to leverage on these traces

※ 이 논문은 부경대학교 자율창의학술연구비(2013년: 2013-0472)에 의하여 연구되었음.
† 비 회 원 : 부경대학교 정보보호학과 박사
†† 종신회원 : 부경대학교 IT융합응용공학과 교수
Manuscript Received : December 30, 2013
First Revision : February 12, 2014
Accepted : February 13, 2014
* Corresponding Author : Kyung-Hyune Rhee(khrhee@pknu.ac.kr)

and reach some conclusions on the past history of the object; there are techniques for integrity verification, source identification or classification, analysis of near-duplicates dependencies and many others[1].

However, every image forensics tool has assumed that the image forger has not taken any countering measure to remove its trace. In reality, likewise other information security fields, vulnerabilities in the existing forensic tools can be exploited, and the modified images can not only fool our eyes, but also pass safely through detection techniques. Thus there is an essential need to re-evaluate all existing forensics tools to take countering measure into account[2]. In the field of forensics sciences, countermeasures to the investigation activities are known under the name of counter-forensics or anti-forensics. The counter-forensics aims at concealing the traces introduced by processing tools when the user edits or tampers image contents. Harris[3] defined counter-forensics techniques as any attempt to compromise the availability or usefulness of evidence for the forensics process. Under this interpretation, the simple wiping-off of fingerprints on a crime scene can be considered as a counter-forensics act. In a similar way, multimedia counter-forensics involves all those means that allow covering traces of image manipulation, or, more precisely, to make manipulation invisible to the existing detection methods. Hence, the study of counter-forensics methods misled as forensics techniques by tamper hiding or concealing traces of manipulations, is becoming a hot research topic[4].

Most of the tamper hiding algorithms are came from steganalytic research [5]. Both of them try to achieve undetectability by preserving image properties as many as possible. Yet, steganography and tamper hiding differ in the amount and source of information to hide, and the extent to which an image can be altered. Most steganographic methods are designed to embed a given message by minimizing the number of changes to the cover (hence, keep its semantics) while tamper hiding merely conceals the information that larger parts of the original medium have been modified with the aim to change its semantics [6]. Nevertheless, counter-forensics techniques do not have the requirement to transmit a message, so the modification is more flexible.

The research on attacks against forensics techniques is important to investigate forensics detectors, as steganography for steganalysis and vice versa. Kirchner et al.[7] introduced the concept of fighting against image forensics. The distinction of this concept is between post-processing and integrated techniques, and between targeted and universal ones. A counter-forensics technique belongs to the post-processing class if it consists of two steps: first the attacker performs the tampering, thus obtaining a desired modified content, then he processes the content so to conceal or erase the detectable traces left during the first step. On the contrary, an integrated counter-forensics technique modifies the image so that it does not expose detectable traces. It is easy to guess that, developing integrated methods is much harder in most cases. The second distinction is related to the target of the counter-forensics method: if it aims at removing the trace searched for by a specific detector, then it belongs to the targeted family. A universal method, instead, attempts to maintain statistical properties as many as possible, so to make the processed image hard to be detected with tools unknown to the attacker.

One of the most common types of image forgeries is the copy-move forgery, where a region from one part of an image is copied and pasted onto another part in same image, thereby concealing the image content in the latter region. Copy-move is one of the easiest way to make a forged image, hence the attacker can actively use a copy-move technique for image forgery. The capability of SIFT to discover correspondences between similar visual content, in fact, allows the forensics analysis to detect even very accurate and realistic copy-move forgery[8, 10, 11].

The remainder of the paper is organized as follows: In section 2, we summarize previously published papers concerned with the topic of this paper. Impeding a SIFT-based copy-move forgery detection method is presented in Section 3. The experimental results are provided in Section 4. Conclusion is drawn in Section 5.

2. Related Work and Contribution

The literature on counter-forensics techniques is still very limited compared to the fast growth of publications on image forensics techniques. A SIFT is a powerful instrument to recognize and retrieve object, an analysis on SIFT security becomes very important also in the case of Content Based Image Retrieval (CBIR) systems in

order to assess if an attacker is able or not to succeed in deluding the image recognition process.

For the region duplication detection, several recent methods have explored the use of matched image keypoints to identify duplicated regions. In Huang et al. [8], keypoints and features based on the scale invariant feature transform [9] are used to account for illumination changes in the detection of copy-move region duplication. However, the robustness of SIFT keypoints and features to image distortions is not fully exploited, which prevents this method from being extended to detect affine transformed duplicated regions. In Pan et al. [10], they describe an SIFT-matching-based detection method that can locate duplicated regions with rotation or scaling. Another recent work Amerni et al. [11] uses SIFT keypoint matching to estimate the parameters of the affine transform and recover matched keypoints.

As a countermeasure to the aforementioned SIFT-based solutions, an intuitive method is to remove original feature points or insert fake feature points in an image while maintaining certain visual quality. A paper by Hsu et al. [12], in which first the impact of simple attacks is analyzed and then a method to strengthen SIFT features (keypoints) is proposed. Following this work, Do et al. [13-14] focused on a SIFT-based CBIR scenario and devised a number of interesting attacks. The aim of the previous works is to modify the SIFT feature descriptor of a keypoint but they are not interested in the complete removal of the keypoints. A pioneer work on this has been presented in Cadelli et al. [15] where an attack based on local warping techniques derived from image watermarking was proposed. All these studies have demonstrated that devising methods to attack SIFT feature is not a trivial task.

The main contribution of our work is to demonstrate a counter-forensics research against image forgery detection based on the counter-forensics techniques by concealing manipulation traces. The actual reliability of such methods can only be estimated by considering what an attacker can try to do to invalidate detection techniques. The key insight of our work is investigated in this paper by analyzing countermeasure method against SIFT algorithm to recreate keypoints in a keypoints removed image while still avoiding keypoint matching for a copy-move forgery detection. The keypoint creation sometimes inevitably accompanies the keypoint removal. Additionally, keypoint removal and insertion are harmful to scale-space image feature extraction. Also, keypoint removal or creation mechanism is not suitable for image counter-forensics, because the forensic analyst can easily identify the traces of manipulation. In order to solve this problem, we propose an attack which is successful in deluding a SIFT-based copy-move forgery detection, that can simultaneously remove and create the keypoints in the image to conceal the traces by keeping with the same keypoints removal and creation rate.

To provide an experimental validation, we need to choose a specific scenario. This consists of selecting a detector for the forensics analyst and a processing tool for the adversary. During the whole procedure, the adversary can exploit the knowledge of the detector used by the forensics analyst since we are aiming at a targeted counter-forensics method.

Basically, our attack aims at identifying the security weakness of the SIFT that employ scale-space keypoint detection mechanism and should not be interpreted as the conventional attacks (signal processing or geometric attacks) that are blind in destroying the keypoints.

3. Impeding SIFT-Based Copy-Move Forgery Detection

The research community has recently started to approach SIFT-based copy-move forgery detection from the perspective of the attacker, whose goal is to hide the features causing similar blocks or keypoints to match. In this section, we describe an attack scenario to impede a SIFT-based copy-move forgery detection methods. Our activity is countermeasure against the exact detection of feature points in digital image. In presence of a copy-move manipulation the extracted SIFT keypoints from the copied and the original regions have similar descriptor vectors. Therefore, matching among SIFT features adopted to detect if an image has been tampered with and, subsequently, localize such forgery. In this sense, the investigation of the attacker is considered on the keypoints extraction of tampered image.

Lowe [9] has presented a powerful framework to recognize or retrieve objects. The SIFT approach can be viewed as a texture descriptor composed by four major stages:

1. Scale-space extrema detection
2. Keypoint localization
3. Orientation assignment
4. Keypoint description

Our main intention is investigated in this paper by avoiding the local extrema in the scale-space extrema detection stage. At first, we introduce a scale-spaces for the extraction of SIFT descriptor, and then we present our proposed method. Our method modifies the selection of local extrema on DoG space by using semantically admissible distortion.

3.1 Scale-Spaces

1) Gaussian Scale-Space

The SIFT detector and descriptor [9] are constructed from the Gaussian scale-space of the source image $I(x,y)$, which is defined as a function $L(x,y,\sigma)$. This is produced from the convolution of $I(x,y)$ with a variable-scale Gaussian $G(x,y,\sigma)$:

$$L(x,y,\sigma) = G(x,y,\sigma) \bullet I(x,y) \tag{1}$$

where $\bullet$ is the convolution operation in x and y, and

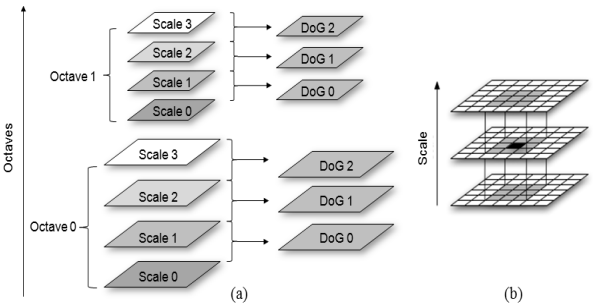

Fig. 1. Scale-space representation. (a) Gaussian scale-space, (b) Scale-space extrema detection

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

where $G(x,y,\sigma)$ is an isotropic Gaussian kernel of variance σ^2, x and y are the spatial coordinate and σ is the scale coordinate.

Since the scale-space $G(x,y,\sigma)$ represents the same information (the image $I(x,y)$) at different levels of scale, it is sampled in a particular way to reduce redundancy as shown in Fig. 1(a). The domain of the variable σ is discretized in logarithmic steps arranged in O octaves. Each octave is further subdivided in S sub-levels. The distinction between octave and sub-level is important because at each successive octave the data is spatially downsampled by half. Octaves and sub-levels are identified by a discrete octave index ϕ and sub-level index s, respectively. An example of Gaussian scale-space representation is illustrated in Fig. 1.

The octave index ϕ and the sub-level index s are mapped to the corresponding scale σ by the formula,

$$\begin{gathered}\sigma(\phi,s) = \sigma_0 2^{\phi+s/S}, \\ \phi \in \phi_{\min} + [0,\ldots,O-1], \quad s \in [0,\ldots,S-1]\end{gathered} \tag{3}$$

where $\sigma_0 \in R_+$ is the base scale level, $S \in N$ is the scale resolution. Note that it is possible to have octaves of negative index. The spatial coordinate x and y are sampled on a lattice with a resolution which is a function of the octave. We denote x_ϕ and y_ϕ the spatial index for octave ϕ; this index is mapped to the coordinate x and y by

$$\begin{gathered}x = 2^\phi x_\phi, \quad y = 2^\phi y_\phi \\ x_\phi \in [0,\ldots,M_\phi - 1], \quad y_\phi \in [0,\ldots,N_\phi - 1], \quad \phi \in Z\end{gathered} \tag{4}$$

where (N_ϕ, M_ϕ) is the spatial resolution of octave ϕ.

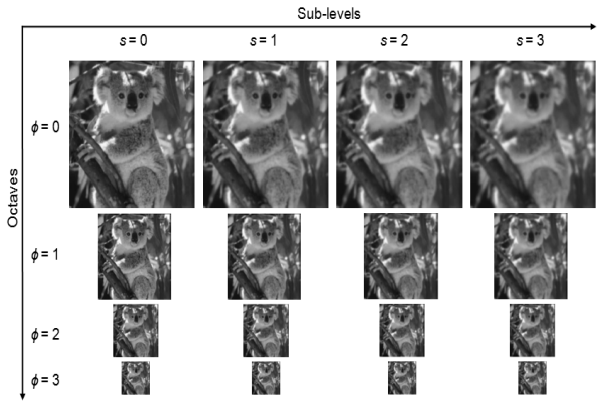

Fig. 2. An example of a Gaussian scale-space representation

If (N_0, M_0) is the resolution of the base octave $\phi = 0$, the resolution of the other octaves is obtained as

$$M_\phi = \left\lfloor \frac{M_0}{2^\phi} \right\rfloor, \quad N_\phi = \left\lfloor \frac{N_0}{2^\phi} \right\rfloor \tag{5}$$

It will be useful to store some scale levels twice, across different octaves. We do this by allowing the parameter s to be negative or greater than S. Formally, we denote the range of s as $[s_{\min}, s_{\max}]$. We also denote the range of the octave index ϕ as $[\phi_{\min}, \phi_{\min}+O-1]$, where $O \in N$ is the total number of octaves. Table 1 for a summary of these symbols used in this paper.

2) Difference of Gaussian Scale-Space

To efficiently detect stable keypoint locations in scale-space, the algorithm make use of another scale-space too, called difference of Gaussian (DoG), which is, coarsely speaking, the scale derivative of the Gaussian scale-space $G(x,y,\sigma)$ along the scale coordinate σ, as shown in Fig. 1(a). The difference of Gaussian pyramid is generated from a single input image.

Table 1. Scale-space parameters

Symbols	Descriptions
$G(x,y,\sigma)$	Gaussian scale-space
$D(x,y,\sigma)$	DoG scale-space
$*(\cdot,\sigma(\phi,\cdot))$	Octave data
σ_0	Base scale offset
$\sigma(\phi,s)=\sigma_0 2^{\phi+s/S}$	Scale coordinate formula
$\phi \in [\phi_{min}, \phi_{min}+O-1]$	Octave index and range
$s \in [s_{min}, s_{max}]$	Scale index and range
$M_0, \quad N_0$	Base spatial resolution (octave $=0$)
$M_\phi = [\frac{M_0}{2^\phi}], \quad N_\phi = [\frac{N_0}{2^\phi}]$	Octave lattice size formulas
$x = 2^\phi x_\phi, \quad y = 2^\phi y_\phi$	Spatial coordinate formula
$x_\phi \in [0,...,M_\phi-1], \quad y_\phi \in [0,...,N_\phi-1]$	Spatial indexes and ranges
$\pi_B(x,y)$	A random permutation of the indices belonging to the B-th block
$\Delta(x,y)$	A $i.i.d$ random variables uniformly distributed in the interval $[-\Delta_{max}, \Delta_{max}]$

The output is a pyramid of several images, each being a unique difference of Gaussians. To generate the pyramid, the input image is repeatedly blurred; the difference between consecutive blur amounts is then output as one octave of the pyramid. One of the blurred images is downsampled by a factor of two in each direction, and the process occurs again with output in a different size. It is given by

$$G(x,y,\sigma(s,\phi)) = (G(x,y,\sigma(s+1,\phi)) - G(x,y,\sigma(s,\phi))) \quad (6)$$
$$\bullet\ I(x,y) = L(x,y,\sigma(s+1,\phi)) - L(x,y,\sigma(s,\phi))$$

Lowe's [9] implementation uses the following parameters:

$$\sigma_n = 0.5, \quad \sigma_0 = 1.6 \bullet 2^{1/S}, \ \phi_{\min} = -1, \ S = 3$$

In order to compute the octave $\phi = -1$, the image is doubled by bilinear interpolation (for the enlarged image $\sigma_n = 1$. In order to detect extrema at all scales, the difference of Gaussian scale-space has $s \in [s_{\min}, s_{\max}] = [-1, S+1]$. Since the difference of Gaussian scale-space is obtained by differentiating the Gaussian scale-space, the latter has $s \in [s_{\min}, s_{\max}] = [-1, S+2]$. The parameter O is set to cover all octaves (i.e. as big as possible).

The feature points are chosen from the local maxima or minima in the DoG space. Each point in $D(x,y,\sigma(s,\phi))$ will be compared with its 26 neighboringpixels, of which 8 pixels located in current scale image and others located in the scale above and below. As shown in Fig. 1(b), the candidate pixel in black is compared with those other 26 pixels in white. The candidate pixel will be considered to be a feature point and its coordinate pixel value is larger than all those 26 pixels values or smaller than them.

In order to impede the detection of local maxima or minima, we applied semantically admissible distortion on the DoG space. As a result, the detected keypoints are found on totally different positions by effect of the keypoint localization and the orientation assignment processes.

3.2 Review of Semantically Admissible Distortions

The random pre-warping must be strong enough to avoid that registration techniques can undo the warping and, in the meantime, it must guarantee the invisibility of the distortion. For this reason gathering information about the subset of semantically admissible geometric distortions is a vital requirement. In a general case of a geometric distortion can be seen as a transformation of the position of the pixels in the image. It is possible to distinguish between global and local geometric distortions.

A global transformation, in fact, is defined by a mapping analytic function that relates the points in the input image to the corresponding points in the output image. It is defined by a set of operational parameters and performed over all the image pixels.

Local distortions, in fact, refer to transformations affecting in different ways the position of the pixels of the same image or affecting only part of the image. A general model which comes to mind to do this is a distortion according to which each pixel of the image is assigned a random displacement vector

$$\Delta(x,y) = (\Delta_h(x,y), \Delta_v(x,y)),$$

where $\triangle_h(x,y)$ and $\triangle_v(x,y)$ are i.i.d random variables uniformly distributed in the interval $[-\triangle_{\max}, \triangle_{\max}]$. The main problem in a so defined transformation is that it does not take into account the way the Human Visual System (HVS) perceives geometrical distortions. In the following models to treat geometric transformations are sketched. The models are analyzed by means of visual inspection under semantic constraint.

3.3 Attacks or Local Distortions on DoG Space

Our goal is to take into account the HVS to find a perceptually admissible subset of the possible distortions that can be applied to the DoG space.

As explained above, a generic local distortion can be described, for example, by a permutation of the position of pixels on DoG space. Of course this kind of distortion introduces an annoying degradation. A way to overcome this problem could be to fix a maximum displacement of the position of pixels, i.e. to perform a block-based local permutation.

1) Block-Based Local Permutation (B-LP)

This model consists in partitioned the $b \times b$ blocks on DoG space and obtaining the distorted DoG space by allowing random permutations within each block. Here, the size of the partitioned block should be smaller, which provides a semantic constraint. Each spatial coordinates x and y on DoG space (base $D(x,y,\sigma(s,\phi))$) is tiled by non-overlapping blocks a size of $b \times b, (b=3)$ pixels. Blocks are horizontally slid by b pixels rightwards starting with upper left corner and ending with the bottom right corner. The total number of non-overlapping blocks for each spatial coordinates of $M_\phi \times N_\phi$ pixels are $B_\phi = (M_\phi/b) \times (N_\phi/b),\ \ \phi \in \phi_{\min} + [0,...,O-1]$.

Let $D(x,y,\sigma(s,\phi))$ be a generic pixel of the distorted DoG space belonging to the B-th block in ϕ-th octave, then

$$D(x,y,\sigma(s,\phi)) \leftarrow D(x,y,\sigma(s,\phi)) \bullet \pi_B(x,y) \tag{7}$$

where $\pi_B(x,y)$ is a random permutation of the indices belonging to the B-th block. Increasing the size of image allows to consider a larger number of transformations but, at the same time, affects the image quality leading to increasingly annoying artifacts.

Hence we permuted the element of the base levels on DoG space, the detection of the local maxima or minima is chosen in the different locations. Thus, the block-based local permutation allows to impede the detection of local invariant features with eliminating or creating a local features under lower rate value. This property provides the hiding traces left in an image counter-forensics area.

2) Cancellation-Based Local Permutation (C-LP)

In this model, we add to the previous one the possibility of duplicating and canceling sample values so that it is also possible to model local expansions and shrinkings. Furthermore in this way we allow for a larger number of possible distortions. Let $D(x,y,\sigma(s,\phi))$ is a generic pixel of the distorted DoG space, we have

$$D(x,y,\sigma(s,\phi)) \leftarrow D(x+\triangle_h, y+\triangle_v, \sigma(s,\phi)) \tag{8}$$

where $\triangle_h$ and $\triangle_v$ are sequences of i.i.d integer random variables uniformly distributed in the interval $[-\triangle_{\max}, \triangle_{\max}]$.

Important property of invariant feature is measured by repeatability measure. The same feature can be found in several images despite geometric and photometric transformations. We test the B-LP and the C-LP distortions, respectively, in order to alter the detection of SIFT keypoints as shown in Fig. 3. As a results, the B-LP distortion can provide more stable and repeatability properties than the C-LP distortion during the increases of the size for divided blocks. From this results, we have chosen the B-LP distortion under semantic constraint to impede the detection of keypoints well.

4. Experimental Results

In this section, we extensively evaluate the proposed counter-forensics method in a realistic scenario, and show that it yields good results in hiding traces while retaining a high image quality for the attacked image.

We simulated our method under a PC with 3.2G Hz Core i5 CPU, 8G RAM, and Windows 8 platform. The simulation was carried out using Matlab version R2008a. We test our method on commonly used 8 gray-scale images of size 512×512 pixels for performance evaluation (Lena, Barbara,...,etc.) and Benchmark data for image copy-move detection dataset including 120 authentic and

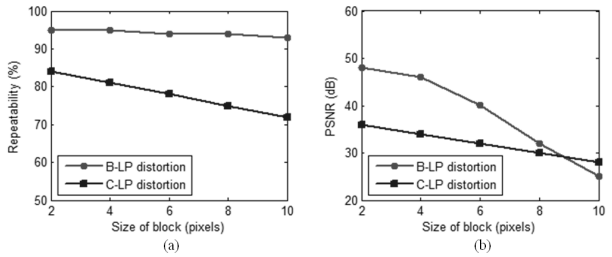

Fig 3. The repeatability (a) and the perceptual quality (b) measures for the block-and the cancellation-based local permutations

124 forged color images of size 3888×2592 pixels with different outdoor scenes for copy-move forgery, as shown in Fig. 4. For Benchmark data, the authentic images were taken by different digital cameras. All tampered images in this dataset are generated from the authentic images by crop-and-paste operation using Adobe Photoshop CS3 version 10.0.1 on Windows XP. The tampered regions are from the same authentic image. In the following tests, the keypoints have been computed by means of VLFeat, Vedaldi and Fulkerson's implementation of SIFT [16]. (DoG peak and edge thresholds set to 4 and 10, respectively). The threshold for keypoint matching is fixed to 0.6, as suggested by Lowe in [9].

Fig. 4. Examples of Benchmark data (3888×2592)

4.1 Efficiency for Tamper Hiding and Impeding Keypoint Matching

In this section, we present an analysis on the efficiency of the proposed procedure for impeding keypoint matching. The experimental tests carried out to check the keypoint detection of the proposed method.

In Fig. 5, the number of original keypoints (blue) are detected by VLFeat algorithm, and the detected keypoints (green) after the proposed processing tool (adversary) are described on the Lena image.

During the procedures, the processing tool can eliminate some keypoints by an effect of semantically admissible

Fig. 5. Results of keypoint detection. For Lena, PSNR=43.02dB. Size of block (3×3) on DoG space

distortion. However, similar keypoints are generated as such effect, also the removal percent is almost equal to the generation percent (around ±10%).

This means our method can provide tamper hiding scenario, that hiding traces left by processing tool. In other words, the forensics analist can detect forged image, that the processed keypoints by the keypoint removal or creation procedure. In our impeding method, the keypoints are eliminated while created, and rates of two procedures are almost equal to 1 (α=Removed_KP(%)/Created_KP(%) ≈ 1).

4.2 Analysis for Copy-Move Forgery Detection

In this section, we report some experimental results on images where a copy-move attack has been performed by taking into account the context.

1) Evaluation of the Detection Accuracy

In Fig. 6, detection results are pictured by presenting on the tampered images for (a, c) a SIFT-based copy-move forgery detection method [17] and (b, d) the corresponding one, where matched keypoints and clusters, attacked by our processing tool, are highlighted. As a result, an interesting situation concerns that our method can impede the (similarity) keypoint matching process and to make false matching results.

In order to quantify the accuracy of detection, the true positive ratio (TPR) and the false positive ratio (FPR) are

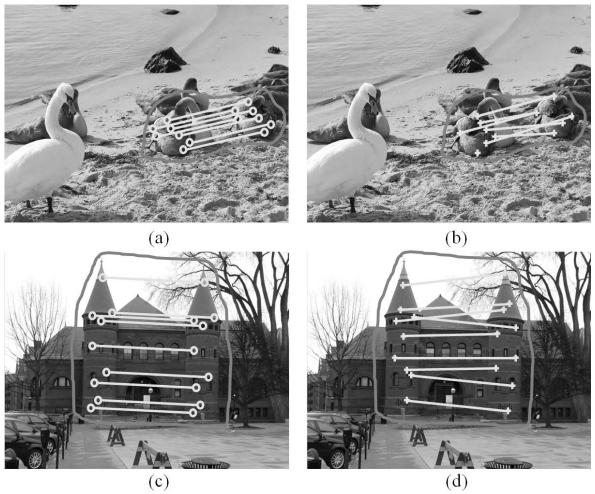

Fig. 6. An examples of a SIFT-based copy-move forgery detection method [17] is pictured in (a,c), and corresponding detection results of our attacking is reported in (b,d)

employed, as follows:

$$TPR = \frac{|\Omega_1 \cap \Omega_2| + |\overline{\Omega_1} \cap \overline{\Omega_2}|}{|\Omega_1| + |\Omega_2|}, \qquad (9)$$
$$FPR = \frac{|\Omega_1 \cup \Omega_2| + |\overline{\Omega_1} \cup \overline{\Omega_2}|}{|\Omega_1| + |\Omega_2|} - 1.$$

where Ω_1 and Ω_2 are the original copied region and the detected copied region, while $\overline{\Omega_1}$ and $\overline{\Omega_2}$ are the forged region and the detected forged region, respectively.

Our goal is to minimize the TPR while maintaining a higher the FPR. The horizontal axis corresponds to the false positive rate (incorrectly labeling an image as altered) and the vertical axis corresponds to the true positive rate (correctly labeling an image as altered). We applied two different processing tools in order to avoid the detection accuracy of similarity matching for the copy-move forgery, such as the B-LP and the C-LP distortions on the DoG space, respectively. Fig. 7 shows the small sized B-LP can achieve higher the image quality while maintaining degraded the detection accuracy compared with non-countered method [17]. For example, with a false positive rate of 0.1, we achieve a true positive rate of 0.34. But, in the increased size of B-LP, the detection accuracy is reduced significantly. For the C-LP distortion, the detection accuracy is lower than other two cases, because, the cancellation is strongly affected to change the value of point on the DoG space, while also can decrease the image quality PSNR=22.31dB. As a result of the detection accuracy, the small sized B-LP approach is efficient to impede the similarity matching methods without higher rate of changes for the image quality.

2) Efficiency for Tamper Hiding

We calculated the removal and creation rates, respectively, which was conducted on each test image of Benchmark dataset (N = 124 images), as shown in Fig. 8. Each histograms of the removal and creation rates are concentrated on around a value of ±10%. For this case, our method can also successfully provide better the tamper hiding scenario on the number of images for Benchmark dataset.

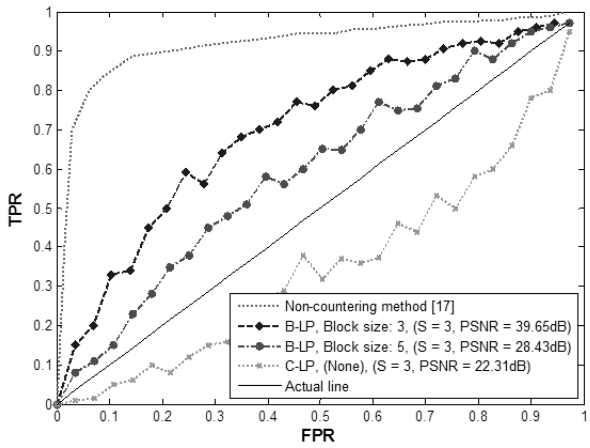

Fig. 7. Results of the detection accuracy for copy-move forgery

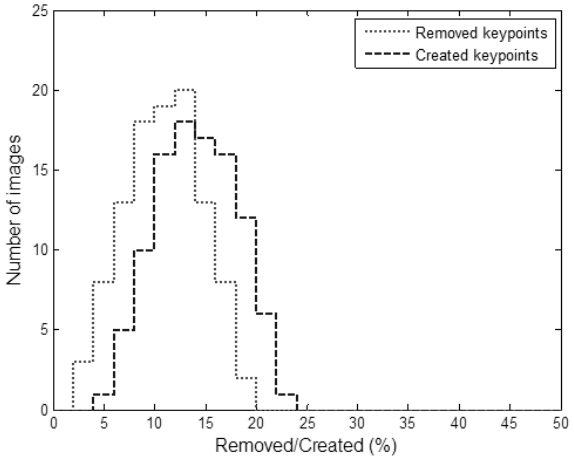

Fig. 8. Efficiency of our method respect to the tamper hiding. Curves correspond to the envelopes of removal and creation rate histogram, obtained by analyzing the manipulated Benchmark data

5. Conclusion

In this paper, we proposed a targeted counter-forensics method for SIFT-based copy-move forgery detection by applying semantically admissible distortions in processing tool. Our activity is a countermeasure against the exact detection of feature points in digital image. In the case of a copy-move manipulation, the extracted SIFT keypoints from the copied and the original regions have similar descriptor vectors. Therefore, the matching among SIFT features can be adopted to detect whether an image has been tampered or not and, subsequently, we can localize such a forgery. In this sense, the investigation of the attacker is considered as the keypoints extraction of the tampered image. Our proposed processing tool has considered on the DoG space of SIFT algorithm, where we applied the semantically admissible distortion in order to alter the detected keypoints under the semantic constraint. The proposed method allows the attacker to delude a similarity matching process and conceal the traces remained by the modification of SIFT keypoints, while maintaining a high fidelity between the processed and original images under the semantic constraint.

Reference

[1] J. Redi, W. Taktak, J.-L. Dugelay, "Digital image forensics: a booklet for beginners," Multimedia Tools and Applications, 51, pp.133-162, 2011.

[2] L. Chen, S. Wang, S. Li, J. Li,."Countering Universal Image Tampering Detection with Histogram Restoration," International Workshop on Digital Forensics and Watermarking, IWDW 2012, LNCS 7128, pp.282-289, 2012.

[3] R. Harris, "Arriving at an Anti-Forensics Consensus: Examining How to Define and Control the Anti-Forensics Problem," Digit. Investig. 3(Supplement 1), pp.44-49, 2006.

[4] R. Bohme, M. Kirchner, "Counter-Forensics: Attacking Image Forensics," In Digital Image Forensics, ed. by H.T. Sencar, N. Memon, Springer, New York, pp.327-366, 2013.

[5] Y.Q. Shi, C. Chen, G. Xuan, W. Su, "Steganalysis Versus Splicing Detection," In IWDW 2007, LNCS 5041, pp.158-172, 2008.

[6] M. Kirchner, R. Bohme, "Tamper hiding: Defeating image forensics," In Information Hiding, LNCS 4567, pp.326-341, 2007.

[7] M. Kirchner, R. Bohme, "Hiding Traces of Resampling in Digital Images," IEEE Transactions on Information Forensics and Security, Vol.3, No.4, pp.582-592, 2008.

[8] H. Huang, W. Guo, Y. Zhang, "Detection of copy-move forgery in digital images using SIFT algorithm," In Proc. IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, 2008.

[9] D. Lowe, "Distinctive image features from scale-invariant keypoints, "International Journal Computer. Vision, Vol.60, No.2, pp.91-110, 2004.

[10] X. Pan, S. Lyu, "Detecting image region duplication using SIFT features," In Proc. ICASSP, Dallas, TX, 2010.

[11] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, G. Serra. "Geometric tampering estimation by means of a SIFT-based forensic analysis," In Proc. ICASSP, Dallas, TX, 2010.

[12] C.Y. Hsu, C.-S. Lu, S.-C. Pei, "Secure and robust SIFT," In ACM Multimedia Conf., pp.637-640, 2009.

[13] T.-T. Do, E. Kijak, T. Furon, L. Amsaleg, "Understanding the Security and Robustness of SIFT," In Proc. of the International Conference on Multimedia, MM'10, pp. 1195-1198, 2010.

[14] T.-T. Do, E. Kijak, T. Furon, L. Amsaleg, "Deluding Image Recognition in SIFT-Based CBIR Systems," In Proc. of the 2nd ACM Workshop on Multimedia Forensics, Security and Intelligence, MiFor'10, pp.7-12, 2010.

[15] R. Caldelli, I. Amerini, L. Ballan, G. Serra, A. Costanzo, "On the Effectiveness of Local Warping Against SIFT-Based Copy-Move Detection," In Proc. of Int'l Symposium on Communications, Control and Signal Processing (ISCCSP), May, 2012.

[16] A. Vealdi, B. Fulkerson, VLFeat: An open and portable library of computer vision algorithms. (http://www.vlfeat.org/), 2008.

[17] I. Amerini, L. Ballan, R. Caldelli, A.D. Bimbo, G. Serra, "A SIFT-based forensic method for copy move attack and transformation recovery," IEEE Transactions on Information Forensics And Security, Vol.6, No.3, pp.1099-1110, 2011.

Munkhbaatar Doyoddorj

e-mail : mbtrdd@gmail.com

2003년 National University of Mongolia (학사)

2011년 부경대학교 정보보호학과(석사)

2014년 부경대학교 정보보호학과 박사

관심분야 : 스테가노그래피, 워터마킹, 이미지 포렌식

이 경 현

e-mail : khrhee@pknu.ac.kr

1982년 경북대학교 수학교육과(학사)

1985년 한국과학기술원 응용수학과(석사)

1992년 한국과학기술원 수학과(박사)

1993년~현 재 부경대학교 IT융합응용공학과 교수

관심분야 : 정보보호론, 공개키 암호, 신원기반 암호, 멀티미디어 정보보호, 그룹 키 관리