Mathematical Statistician and Engineering Applications
ISSN: 2094-0343
DOI: https://doi.org/10.17762/msea.v70i2.2445

Analysis of Keypoint-based Copy-Move Forgery Detection in Digital Images

Aditya Agnihotri
Department of Comp. Sc. & Info. Tech., Graphic Era Hill University,
Dehradun, Uttarakhand India 248002

Article Info
Page Number:1562 - 1568
Publication Issue:
Vol 70 No. 2 (2021)

Article History
Article Received*: 05 September 2021*
Revised*: 09 October 2021*
Accepted*: 22 November 2021*
Publication*: 26 December 2021*

Abstract
The images are used as the element that is primary of in multitudinous operations and the dupe- move phony is considerably used to induce the forged images by manipulating the content of digital images. With the advancement of various image editing software Photoshop that is including Draw, GIMP, etc make it's easier to manipulate the digital images. Thus, Copy- Move Forgery Detection is a task that is challenging. Some of the dupe- move phony discovery styles analogous as Symmetry predicated features that are original SIFT- Symmetry, Clustering with J-relation Algorithm, SIFT- KAZE crossbred Features, point Matching predicated on frippery- bin-First Algorithm, Binary Discriminant Featuresetc. are compared and analysed in this paper.

Keywords: Copy-move forgery, Digital images, Feature Matching, Clustering.

I. Introduction

The dupe- move phony is an image manipulation fashion that is applied to alter the information contained in it. It's an purposeful manipulation that aims to favor an existent's particular view. The dupe- move phony is the process of copying a region of an image and move or bury that region down, within the image that is same. It's mainly used to either turn a small group of goods to a large group or conceal the unwanted information in the image. An illustration of dupe- move phony discovery is shown inFig. 1. The image forensics aims to determine the authenticity and the source of the images that are digital. The dupe- move phony discovery scheme consists of detecting the authenticity of the image by checking whether the image is forged or not. The technique for detecting fake duplicate moves may be broadly categorised into two orders: the Block-based system and the Keypoint-based Method. It determines the relationship among the original region and its duplicated or copy-moved region by dividing the input image into imbrication blocks of pixels and using the Block- predicated Method. This system is also appertained to as thick- field Copy- Move Forgery Detection. The pixel rudiments in the image are matched to gain a more accurate result in the Keypoint- predicated Method(also called point- predicated system), the features or the keypoints. rather of blocks in block- predicated system, the original features analogous as SIFT(Scale Invariant Feature Transform) and suds(Speeded Up Robust Features) have been used in Keypoint predicated Method. firstly, the image is scanned to prize outside keypoints. also the point point and vector descriptor are calculated for every keypoints. When Compared with block-predicated styles, the keypoint- predicated styles are analyzed to be more robust to geometric changeovers analogous as paraphrase, rotation, andre- scaling that is large.

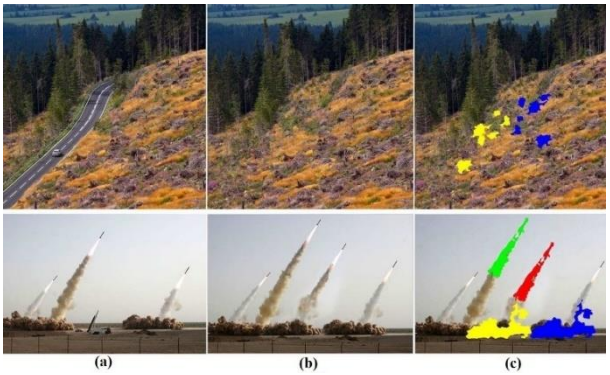

Fig. 1. (a) Two Original Images. (b) Forged Images created using the Original Digital Images. (c) The Duplicated Regions are Localized based on CMFD method.

Fig. 1 shows the localization and discovery of dupe- move forged regions using two samples. The illustration that is first of applying region duplication to hide unwanted regions analogous as road regions. In the alternate illustration, region duplication is applied to increase the count of succeeded dumdums to show the false print of the great achievement. In general, the dupe- move phony discovery fashion consistsof considered as features. Some CMFD styles use the combination of both the Block- predicated and Keypoint- predicated ways. The workflow that is typical of Keypoint- predicated CMFD system is shown in Fig. 2.

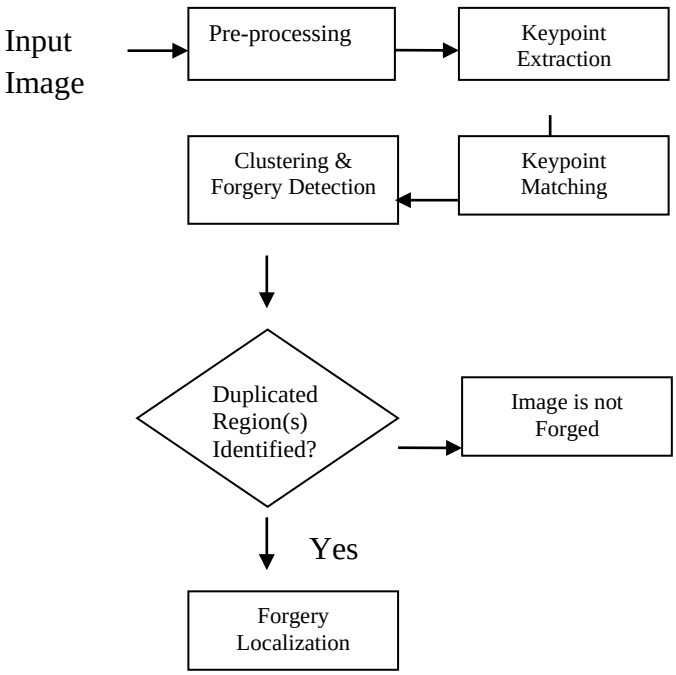

Fig. 2. Typical Workflow of Keypoint-based Copy-Move Forgery Detection.

This paper is organized as follows. Section II shows the commonly used feature that is keypoint-based techniques. In Section III, Literature survey of several papers obtained. Section IV describes the conclusion of the paper.

Keypoint-Based Feature Extraction Techniques

The feature extraction method of block-based method and keypoint-based method are variant as its nature of the detection procedures. Some of the commonly used feature extraction techniques are SIFT, SURF, ORB, KAZE, BRIEF.

SIFT

One of the most popular methods for extracting features from digital images is called the Scale Invariant Feature Transform (SIFT) Algorithm. This algorithm is used to identify features in digital pictures that serve as local keypoints. It locates and describes quantitative information of each keypoint, referred to as SIFT descriptor. The SIFT Technique has been verified to be invariant to rotation and scaling.

SURF

Speeded Up Robust Features (SURF) is partly inspired by SIFT Technique. The SIFT is several times faster than the SIFT. SURF uses an integer approximation of the determinant of Hessian blob detector to detect keypoints. The SURF descriptor has been used to locate and recognize objects, people or faces. It is also convenient to reconstruct scenes that are 3D to track objects and extract the keypoints.

A. ORB

The Efficient Keypoint matching algorithm deliver the best result. Oriented Fast and Rotated BRIEF (ORB) is used as an alternative to SIFT and SURF. It consists of the combination of FAST BRIEF and detector descriptor. To detect the keypoints, it uses FAST algorithm and ORB detects the keypoint at different scale. This is referred to as Multiscale Image Pyramid.

B. KAZE

The keypoint that is 2D method, KAZE features obtain keypoints in non-linear scale space through non-linear diffusion filtering. The feature extraction techniques such as SIFT and SURF are operated in Gaussian scale space. The KAZE features are invariant to rotation and scaling. The Accelerated KAZE (AKAZE) is an advanced version of KAZE.

C. BRIEF

The advanced technologies need to handle huge amount of data with limited resources, there is an expanding need for local descriptors that are fast to compute and match. The Binary Robust Independent Elementary Features which is very fast both to construct and to match and it use Binary Strings as an feature descriptor that is efficient. It outperforms other fast descriptors such as SIFT and SURF. The descriptor similarity can be evaluated using the Hamming Distance, which is identified as more efficient to compute, instead of applying commonly used approach, the distance that is euclidean.

Literature Survey

Moment's life style with the arising technologies also affects the credibility of the digital images and invite image phonies . The digital images are used as a kind of digital substantiation which plays a major part in numerous operations and occasions. There are colorful kinds of dupe- move phony discovery styles developed by experimenters to attack this challenge of different types of dupe- move phony attacks. Some of the recent Keypoint- grounded Copy- Move Forgery Detection Schemes are explained in the following

A. Feature Matching based on Best-Bin- First Algorithm

The region duplication discovery can be detected by using Image point Matching. To perform the point matching, keypoints in the image need to be collected. Keypoints are locales that carries meaningful information of the image content. In this system, SIFT fashion is used to prize keypoints and keypoints attained by this fashion is appertained to as SIFT Keypoints, which are

asleep to geometrical metamorphoses and illumination deformations.

The four way involved in SIFT Algorithm are (i) Scale- Space Extrema Detection, (ii) Keypoint Localization, (iii) exposure Assignment, (iv) Generation of 128- Dimensional Keypoint Vector from the histogram of original slants in its neighborhood, at each Keypoint. The dominant scale of the keypoint is used to establish the size of the surrounding area, and all important points are aligned with the dominant exposure of the point they belong to. Because of this, the point vector that was acquired is connected as being stable to gyration and scaling.

Also, the keypoints are matched grounded on their point vectors using the Best- caddy-First Algorithm. For a keypoint at position ‘ x ’ with point ‘ f ’, match it with a keypoint at another position ‘ x1 ’ whose point vector is ‘ f1 ’. The neighbor that's nearest of ‘ f1 ’ to ‘ f ’ is measured on the base of the Euclidean distance. The match that's stylish of the keypoint is might be falsehoods within close proximity. The Hunt operation is performed outside an 11 x 11 pixel-window centered at the keypoint to avoid the searching the nearest neighbors of a keypoint from the region that's same. Grounded on the point matching, the possible deformations of the regions that are duplicated estimated by modelling the deformation as affine metamorphosis of pixel equals. To pare out the mismatched keypoints, Random Sample Consensus Algorithm(RANSAC) is employed, which can estimate the model parameters with a high degree of delicacy.

The region correlation chart is attained and eventually, the duplicated regions can be located with the estimated affine transfigure. The SIFT keypoints aren't reflection steady. To handle the duplicated regions that are reflected, need to gain the imaged interpretation of the SIFT- keypoint, which is reflected around one image axis that's coordinate respects to the geometric center of the image aeroplane.

B. J-Linkage Clustering

Copy- Move phony Discovery system with SIFT Features and J- relation Clustering is proposed in this work. The first step consists of point birth and matching using SIFT fashion. The approach that is simplest to match keypoints is to set a threshold on the Euclidean Distance between point descriptors. But it obtains delicacy that is low to the high dimensionality of the point space. generality of Lowe’s matching fashion is used to deal with this issue and this fashion is also appertained to as g2NN Test. It uses the idea of Two Nearest Neighbor(2NN) Test, which finds the rate between neighbor that is closest and alternate closest neighbor and compares the rate with a predefined threshold. In g2NN, the 2NN test is repeated between the neighbor that is closest and alternate closest neighbor until this rate is lower than Threshold(set to0.5). It works well indeed the image contains multiple- copied regions. A clustering algorithm analogous as Hierarchical Agglomerative Clustering(HAC) can be used to descry the possible duplicated regions. In general, the clustering is performed by considering only the coordinates of the matched couples and not the constraints about the matching. So, it has two faults(i) also that ca n’t be separated if the duplicated regions are close to each other.(ii) When a patch contains keypoints with anon- invariant distribution that is spatial it's hard to identify the patch as single. To break these problems, the J-relation Clustering algorithm is introduced, which performs robust clustering. Finding the transitions among the original and the copied region can be accomplished with the help of the preference set vectors. In this approach, each preference set vector is placed in its own cluster. Additionally, the two clusters that have the least distance between them in the abstract space are mixed together. The preferences set vector of a cluster is considered to be the intersection of the preference sets of the couples who have a successful match, and the distance that exists between two clusters is considered to equal the Jaccard distance between the individual preference sets.

B. symmetry-based features that are local

In this system, the Original harmony Features are used to descry the dupe- move phony. First, the

features are pulled from the image by calculating the harmony that is original Value the keypoint descriptor for each keypoint is derived from this value of corresponding scale space using Log-Polar Transform(LPT). The calculation of the harmony value involves measuring the image's similarity along the reflection axis with an n-fold mirror. Harmony achieved by rotation. The point matching is carried out to determine which points of the forged areas fit up with one another. The Random Sampling Consensus (RANSAC) Algorithm is constantly applied with the matching points until carrying all of the matched points with separate copied and moved regions in order to tackle the numerous dupe-move phoney issues and false matches that are contained in it. During the replication process, clustering is used to distinguish between points that belong to copied regions and moved regions. Additionally, the transformation parameter is calculated for each cluster, and the localization of duplicate moved regions is carried out only if a phoney is found. The detection of keypoints in scale space is the method that can be used to accomplish scale invariance. In order to ensure blur invariance, Gaussian image conglomerations are created with a scale factor whose size is fixed throughout the process.

C. SIFT-KAZE Hybrid Features

A robust keypoint sensor KAZE is introduced and combined with SIFT is appertained to as a kind of mongrel point fashion. This system helps to prize further point points. The SIFT keypoints are represented with a 128- dimensional point vector and KAZE keypoints are described by using 64-dimensional point vectors. The KAZE point birth consists of calculating the regularized Hessian Matrix to descry original outside and the dominant exposure of the keypoints is determined in nonlinear scale space and which makes the point vectors gyration steady.

An bettered matching algorithm is used which can find the n-stylish matched features, which is same for both the SIFT and the KAZE keypoints. Also an effective filtering step grounded on image segmentation is executed to sludge out false matches. An replication strategy is developed to estimate metamorphosis matrices and determine the regions that are duplicated be located at pixel position. This system can precisely descry duplicated regions indeed after deformations similar as gyration, scaling, JPEG contraction and noise that's adding.

D. Binary Discriminant Features

Copy- Move phony Discovery using Binary Discriminant Features is the combination of both the Block- grounded Method and Keypoint- grounded Method. A double discriminational point descriptor(BDF) is used for point birth and matching to ameliorate the delicacy of phony discovery. The point descriptors can be astronomically classified into two orders(i) Floating- point Descriptors and(ii) double Descriptors. Image keypoints are considered as the features to identify the position of phony .

The point vector is reckoned on the base of Original grade Statistics, which is robust to deformations similar as illumination changes, standpoint shift and perspective distortion for each keypoint. Following the matching of the keypoints, an image blocking approach known as the Simple Linear Iterative Clustering (SLIC) algorithm is employed in order to partition the input image into Superpixels, which are irregular and do not overlap with one another. A number of pixels that share similar qualities can be combined to create what is known as a superpixel. It is generally agreed that the SLIC for Superpixel Segmentation is a memory system that works quickly and efficiently. A five-dimensional vector is used to define each individual pixel. Therefore, the clustering process takes place in the five-dimensional space that is defined by these vectors. The neighbouring blocks are also mixed in with the regions that are suspected based on colour similarity. This is done with the help of Colour Histogram Matching, and a final morphological close operation extracts the regions that have been found to be forged.

E. SIFT-Symmetry

The system that's SIFT- Symmetry on the CMFD with common geometric metamorphosis attacks

and in addition to that, reflection- grounded attacks. In this system, the SIFT- grounded CMFD system with harmony- grounded Matching, which is vindicated as high robust against dupe- move phony with reflection attack and any combinations of reflection attacks.

Three phases in this system are(i) Keypoint birth,(ii) Keypoint Matching and(iii) Keypoint Clustering and phony Discovery. The keypoint birth is reused in two way similar that first step is Keypoint birth with SIFT system to handle thenon-reflection attacks and the alternate step consists of Keypoint birth with glass- SIFT to handle the reflection- grounded attacks. The Keypoint Matching Phase also consists of two way. The g2NN(Generalized Two Nearest Neighbor) algorithm is used to match the SIFT keypoints. also proceeds directly with Clustering and phony Discovery Phase if the number of match points is lower than or equal to the threshold value(set as 5.

else, it performs harmony- grounded Matching using glass- SIFT Keypoints before the Clustering operation. It's delicate to localize the region that's dupe- moved the case that matching points attained are scattered. therefore, Hierarchical Agglomerative Clustering(HAC) fashion is applied. Eventually, the extensively used RANSAC algorithm is performed to sludge out the matches that are false in the result. This system is also steady to gyration and scaling.

F. FI-SIFT Algorithm

The SIFT algorithm isn't steady to the reflection attack. In this work, a Flip- Invariant SIFT(FI-SIFT) descriptor is proposed to ameliorate the robustness against reflection attacks. According to this FI- SIFT Algorithm, the SIFT- Symmetry is vindicated as hamstrung due to double matching.

The reflection is distributed into two types Horizontal Reflection and Vertical Reflection. The Vertical Reflection can be attained by rotating the vertical flipped interpretation of keypoints by 180 degrees. therefore, Vertical Reflection is considered in this system. The SIFT descriptor is reconstructed by a universal garbling fashion to induce FI- SIFT keypoint descriptor, which is steady to the reflection, gyration, scaling and affine metamorphosis. The g2NN Algorithm is espoused to deal with multiple keypoint matching. To pare out the dyads of mismatched keypoints, RANSAC Algorithm is employed, which robustly descry the affine parameters.

Conclusion

The Bulk of the Copy- Move Forgery Detection concentrates on thepost-processing attacks similar as contraction and noise addition and geometrical metamorphosis attacks similar as gyration, scaling and restatement. Now, the experimenters developing bettered styles to robust against the reflection attacks. The SIFT point Matching is extensively used for keypoint- grounded CMFD system. All these styles are invariance to gyration and scaling. According to this study, it's vindicated that the imaged interpretation of SIFTkeypoint descriptor can descry the reflection-grounded attacks. The g2NN Matching algorithm and RANSAC Algorithm are the most generally used styles in keypoint- grounded CMFD Schemes.

References

[1] Pan and S. Lyu, “Region Duplication Detection Using Image Feature Matching,” IEEE Transactions on Information Forensics and Security. (Dec 2010) 857-867.

[2] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, G. Serra, “A SIFT-based forensic method for Copy-Move Attack Detection and Transformation Recovery,” IEEE Transactions on Information Forensics and Security. 6 (2011) 1099–1110.

[3] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, L. Del Tongo, G. Serra, “Copy-move forgery detection and localization by means of robust clustering with J-linkage,” Signal Processing: Image Communication 28 (2013) 659–669,

[4] D. Vaishnavi, T.S. Subashini, “Application of Local Invariant Symmetry Features to Detect

and Localize Image Copy Move Forgeries.

[5] Fan Yanga, Jingwei Lia, Wei Lua, Jian Wengb, “Copy-move forgery detection based on hybrid features,” Engineering Applications of Artificial Intelligence.59 (2017) 73–83.

[6] You-Jian Yu, Guang-Fu Wang, and Jie Zhao, “FI-SIFT Algorithm for Exposing Image Copy-Move Forgery with Reflection Attacks,” International Journal of Network Security, First Online, Sept. 16, 2019.

[7] Mohsen Zandi, Ahmad Mahmoudi-Aznaveh, and Alireza Talebpour, “Iterative Copy-Move Forgery Detection Based on a New Interest Point Detector” IEEE Transactions on Information Forensics and Security.

[8] J.-M. Guo, Y.-F. Liu, Z.-J. Wu, "Duplication forgery detection using improved DAISY descriptor," Expert Syst. Appl. 40 (2013) 707-714.

[9] Yuanman Li, Student Member, IEEE and Jiantao Zhou, ”Fast and Effective Image Copy-Move Forgery Detection via Hierarchical Feature Point Matching,” IEEE Transactions on Information Forensics and Security. (2018) 1556-6013.

[10] Priya Mariam Rajua, Madhu S.Nair, “Copy-Move Forgery Detection using Binary Discriminant Features. Journal of King Saud University Computer and Information Science. (Nov 2018).

[11] B. Soni, P. K. Das, D. M. Thounaojam, “Keypoints based Enhanced Multiple Copy-Move Forgeries Detection System using Density-based Spatial Clustering of Application with Noise Clustering Algorithm. (Feb 2018), ISSN No: 1751-9659.

[12] N. B. A. Warif, A. W. A. Wahab, M. Y. I. Idris, R. Salleh,Fazidah Othman, “SIFT-Symmetry: A Robust

[13] Detection Method for Copy-Move Forgery with Reflection Attack,” Journal of Visual Communication & Image Representation, vol. 46, pp. 219-232, 2017.