JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

IMAGE FORGERY DETECTION USING ADAPTIVE OVER SEGMENTATION AND FEATURE POINT MATCHING

[1] ADARI VENKATA DURGA GAYATHRI
[1] M.Tech Student with CST, Department of Computer Science and Systems Engineering
Andhra University College of Engineering, Visakhapatnam, AP, India

Abstract

In this paper we propose a scheme to detect the copy-move forgery in an image, mainly by extracting the key points for comparison. The main difference to the traditional methods is that the proposed scheme first segments the test image into semantically independent patches prior to key point extraction. As a result, the copy-move regions can be detected by matching between these patches. The matching process consists of two stages. In the first stage, we find the suspicious pairs of patches that may contain copy-move forgery regions, and we roughly estimate an affine transform matrix. In the second stage, an EM-based algorithm is designed to refine the estimated matrix and to confirm the existence of copy-move forgery. Experimental results prove the good performance of the proposed scheme via comparing it with the state-of-the-art schemes on the public databases.

Keywords: Copy-Move Forgery Detection, Adaptive Over-Segmentation, Local Color Feature, Forgery Region Extraction

1. INTRODUCTION

Image forgery has been an issue since the advent of traditional photography in the 19th century, however it is a much more prevalent problem in the digital age. The primary issue is that photographs are often used as concrete evidence of an event, and are generally seen by the public as truthful and trustworthy. Images that are forged, therefore abusing this trust, can have many wide-reaching social impacts. For example, CCTV images are often used in a court of law in order to provide solid evidence either by the defence or by the prosecution. If the confidence in these images are put in to doubt and the jury is unable to put their utmost trust in them, then the trial is put in to repute. Detecting manipulation and forgery within these images is therefore of the utmost importance. Similarly, forged images are extensively used within the media, either deliberately or accidentally. Tabloid newspapers, magazines and marketing campaigns routinely modify images of models or famous figures in order to make them look more aesthetically pleasing to the viewer. This can be a simple case of adding a filter or modifying the contrast of the image, but it is often much more extreme; improved muscle definition, more toned body parts and wrinkle removal are examples of commonly achieved results.

The issue has become so prevalent and well known that the verb "photoshopped", referring to the popular image editing application Adobe Photoshop, has become a neologism for manipulating and modifying digital images. One of the most pressing issues is that there are many different ways of modifying an image, and due to a digital images' complex nature it's impossible to have an algorithm that detects every type of image forgery. Because of this, image forgery detection isn't widely used in the professional world. The underlying concept would be highly useful in the majority of professional fields that deal with images on a day to day basis, where the reliability and credibility of these images is crucial. In addition, with the large increase in the use of social media, individuals would also benefit greatly from being able to detect forgeries within images. Convincingly manipulated images are widely circulated on social media platforms [17], and are able to be spread rapidly within communities who believe them to be true. In order to detect these image forgeries, it is required that we understand some typical methods used in order to manipulate images.

• **Copy-paste Cloning** - This is where existing areas within an image are cloned, allowing regions to be covered or objects to be duplicated. This is a commonly used method as the forgeries have the potential to look very convincing, due to the fact that they have come from the source image to begin with. • Image Splicing - Whereby objects from another image are spliced together with the source image, adding objects that weren't present in the original image.
Various blending techniques exist, such blurring edges, reducing the contrast and utilizing cloning to help disguise the new object in with the surrounding area.
• **Modification of existing regions** - This is similar to copy-paste cloning, but instead of being an exact duplication, existing regions are modified in order to suit the needs of the forgery. This can include simply resizing the object, mirroring or skewing it, or splicing two existing objects together. In all of these cases however, the duplicated region has been resampled, meaning that it has been modified enough not to be recognized by any clone detection algorithm. Whereas existing projects have worked on the comparison of image forgery detection methods, these are often limited in scope and only compare variants of the same algorithm on images that are specifically created for that type of method. For example, JPEG Analysis and Edge Detection have been compared [1], however no reason is given as to why these specific implementations were chosen over others, as both tend to detect similar kind of forgeries. In addition, no detail of the images that were used in the research is provided; for example, it is unknown if they are standard library

images or images tailored for this kind of forgery detection algorithm. In addition, pre-existing image forgery detection applications are often of an academic nature (proof of concept or of prototype quality), or very simple. Searching for forgery detection mainly brings up academic papers on the subject, however the most downloaded results on the popular open-source site Source Forge return fairly trivial applications that only detect metadata tags embedded within images [2][3][4]. Whilst this is a useful measure, and something which will also be tested in addition to the main algorithms within our implementation, metadata tags are easily removed or manipulated and it is therefore not an accurate measure of whether an image has been forged or not. Although the implementation within this project is mainly a proof of concept and used purely for research purposes, it is a starting point that could be developed in to a fully-fledged application. Creating a polished, user-friendly interface for the chosen algorithm is then fairly trivial, bringing that type of forgery detection to the mass market.

A technique that works by 1st applying principal element analysis to little mounted - size image blocks to yield a reduced dimension illustration was planned by Alin C Popescu et al. (2004). Whereas performing arts the on top of technique we are able to realize some duplicate pictures (noises). Then the duplicate regions are detected by lexicographically (the follow of aggregation dictionaries).Sorting the whole image blocks. This can be terribly wonderful and actual appropriate technique to yield a reduced dimension illustration. It's sensitive to jpeg lossy compression and additionally it's additive to noise.

A methodology to discover copy- move forgery by dividing the image into overlapping blocks of equal size, extracting feature for every block and representing it as a vector and typing all the extracted feature vectors victimization the base sort, was planned by Hwei-jen sculpture et.al (2009). Base type dramatically reduces the time complexness and also the adopted options enhance the aptitude of resisting of varied attacks like JPEG compression and mathematician noise. Each potency and high detection rates are incontestable.

Sevinc Bayram et al.(2009) projected to use Fourier- Mellin Transform (FMT) options that square measure invariant to scaling and translation. A replacement detection scheme that creates use of investigation bloom filters is additionally introduced by them. It detects copy move forgery terribly accurately albeit the cast image is turned, scaled or extremely compressed. This detection scheme improves the potency. However the hardiness of the tactic is reduced.

B.L.Shivakumar et al. (2011) proposed a method to detect duplication regions. Because one of the common image forgery methods is copy move forgery (CMF). Identification of the CMF can be detected by the duplication regions using Speeded Up Robust Features (SURF) keypoints. These SURF keypoints are extracted from images. The duplication region can be detected with different sizes. The result shows that CMF with minimum false match for images with high resolution. A few small copied regions were not successfully detected.

Irene Amerini et al. (2011) proposed a method to support image forgery detection based on SIFT algorithm. Thus, the algorithm is used to detect the regions which are duplicated and determine the geometric transformation applied to perform such tampering. But, the main drawbacks of this technique, it is unable to detect the image with uniform texture and salient keypoints.

Pravin Kakar et al. (2012) has proposed a method based on transforming-invariant features. These got y utilizing the features from MPEG-7 image signature devices.This method achieved good results, accuracy and extremely low false positives. Thus, these features are invariant to common image processing operations. This method cannot detect regions which have undergone affine transformations and/or multiply copied.

"A. J. Fridrich, B. D. Soukal, and A. J. Lukáš, "Detection of copy-move forgery in digital images," in in Proceedings of Digital Forensic Research Workshop, 2003."
We describe an efficient technique that automatically detects duplicated regions in a digital image. This technique works by first applying a principal component analysis to small fixedsize image blocks to yield a reduced dimension representation. This representation is robust to minor variations in the image due to additive noise or lossy compression.

"A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2004-515, 2004."
Fridrich et al. [1] proposed a forgery detection method in which the input image was divided into over-lapping rectangular blocks, from which the quantized Discrete Cosine Transform (DCT) coefficients of the blocks were matched to find the tampered regions. Popescu and Farid [2] applied Principal Component Analysis (PCA) to reduce the feature dimensions. Luo et al. [3] used the RGB color components and direction information as block features. Li et al. [4] used Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD) to extract the image features. Mahdian and Saic [5] calculated the 24 Blur-invariant moments as features. Kang and Wei [6] calculated the singular values of a reduced-rank approximation in each block. Bayram et al. [7] used the Fourier-Mellin Transform (FMT) to obtain features. Wang et al. [8, 9] used the mean intensities of circles with different radii around the block center to represent the block features. Lin et al. [10] used the gray average results of each block and its sub-blocks as the block features. Ryu et al. [11, 12] used Zernike moments as block features. Bravo-Solorio and Nandi [13] used information entropy as block features.

As an alternative to the block-based methods, key point-based forgery detection methods were proposed, where image key points are extracted and matched over the whole image to resist some image transformations while identifying duplicated regions. In [14-16, 18], the Scale-Invariant Feature Transform (SIFT) [20] was applied to the host images to extract feature points, which were then matched to one another. When the value of the shift vector exceeded the threshold, the sets of corresponding SIFT feature points were defined as the forgery region. In [17, 19], the Speeded up Robust Features (SURF) [21] were applied to extract features instead of SIFT. However, although these methods can locate the matched key points, most of them cannot locate the forgery regions very well; therefore, they cannot achieve satisfactory detection results and, at the same time, a sustained high recall rate [22].

2. TYPES OF DIGITAL IMAGE FORGERY

Now day's fake images have become more in society. Tampering images are common for making controversies. For example, it can be used for sensational news, spread political news and rumors. As the quality of pictures suffers, it's necessary to plot techniques so

as to verifytheir genuineness and trait of pictures.

Picture sterilization is characterized as ”adding, changing, or deleting some vital options from a picture while not exploit any obvious trace. There are totally different techniques used for formation a picture. Taking under consideration the ways wont to create cast pictures, digital image forgery are often isolated into 3 primary classifications: Copy-Move forgery, Image splice, and Image resampling.

A. Copy-Move Forgery

In copy move forgery, in original image some part of the image with any size is copies and pasted in that only in some area of image it can be show in figure1. As the copiedpart originated from the same image, its essential properties such as noise, color and texture don't change and make the recognition process troublesome.

B. Image Forgery using Splicing

Image splicing uses cut-and-paste systems from one or alot of pictures to make another pretend image. once conjunction is performed exactly, the borders between the spliced regions will visually be unbearable. Splicing, however, disturbs the high order Fourier statistics. These insights will so be utilized as a locality of identifying phony. Figure 2, demonstrates an honest sample of image conjunction during which images the photographs the images of the shark and also the eggbeater ar unified into one picture.

Fig 1. (a) Original Image (i); (b) copy Image (ii); (c) forged image

C. Image Resampling

To make associate astounding fake image, some elite regions got to undergo geometric transformations like rotation, scaling, stretching, skewing, flipping and then forth. The interpolation step plays a vital role within the resampling method and introduces non-negligible applied mathematics changes. Resampling introduces specific periodic correlations into the image. These correlations are often utilized to acknowledge phony caused by resampling. In Figure three, the image on the left is that theoriginal image whereas the one on the proper is that the cast image obtained by rotation and scaling it.

Fig 2. (a) The real image (b) final result of imageretouching

3. EXISTING SYSTEM

The goal in copy-move forgery detection is detecting duplicated image regions, even if they are slightly different from each other. One direct solution to this problem would be an exhaustive search, which involves comparison of the image to every cyclic-shifted version of itself. However this approach would be computationally very expensive and would take (MN) 2 steps for an image of size MxN. Also, this type of search might not work in the case where the copied area has undergone some modifications. A second and more efficient approach, which was proposed by Fridrich et al. in [6], is the use of autocorrelation properties. Nevertheless, this approach is shown to be effective only when the duplicated regions were a large portion of the image. Another approach, which is the main interest of this paper, is block-matching procedure. In this approach, the image is segmented into overlapping blocks first. The task here is to detect connected image blocks that are copied and moved, instead of detecting the whole duplicated region. Note that the copied region would consist of many overlapping blocks and since each block would be moved with the same amount of shift, the distance between each duplicated block pair would be the same, as well. Therefore, the forgery decision can be made only if there are more than a certain number of similar image blocks within the same distance and these blocks are connected to each other so that they form two regions of the same shape.

Adaptive Over-Segmentation Algorithm

The Adaptive Over-Segmentation algorithm, which is similar to when the size of the host images increases, the matching computation of the overlapping blocks will be much more expensive. To address these problems, we proposed the Adaptive Over-segmentation method, which can segment the host image into non-overlapping regions of irregular shape as image blocks afterward, the forgery regions can be detected by matching those non-overlapping and irregular regions. Segmentation method, the non-overlapping segmentation can decrease the computational expenses compared with the overlapping blocking; furthermore, in most cases, the irregular and meaningful regions can represent the forgery region better than the regular blocks. However, the initial size of the super pixels in SLIC is difficult to decide. In practical applications of copy-move forgery detection, the host images and the copy-move regions are of different sizes and have different content, and in our forgery detection method, different initial sizes of the super pixels can produce different forgery detection results; consequently, different host images should be blocked into super pixels of different initial sizes, which is highly related to the forgery detection results.

4. PROPOSED SYSTEM

I. OVERVIEW OF THE PROPOSED CMFD SYSTEM AND IMAGE SEGMENTATION

In this section, via revisiting the important issues involved in CMFD we first give the framework of our proposed scheme, and then we explain the reason for using image segmentation

A. CMFD Revisiting and the Framework of the Proposed Scheme

In order to obtain a convincing detection, result we would always like to acquire as much forensic information as possible from the test image. So, the mission of CMFD is not only to determine if an image has some regions containing identical contents, but also to locate these tampered regions. To this end, we can describe the image with a set of local patches, like the blocks or key points in traditional CMFD schemes, and transfer CMFD into a problem of comparison among these local patches. The comparison process may be time-consuming if the number of the patches is too large. For example, the block-based methods usually need a huge amount of time to detect an image. So it is important to decrease the number of patches for comparing. In this regard, the keypoint based methods are faster and more favourable than the block-based ones, because the number of the image keypoints is smaller than that of the divided blocks. However, on the other hand, keypoint-based method also has the following two problems. Firstly, the keypoints lying spatially close to each other should not be compared because they may be naturally similar. The determination of the shortest distance between two comparable keypoints is tricky. Most prior arts empirically select this threshold but neglect its relationship with the image size and content. Secondly, it is uneasy to accurately localize and distinguish the copying source region and the pasting target region, because, unlike the overlapping blocks, the keypoints are often not concentrated together. To deal with this problem Amerini et al. proposed a method based on clustering the matched keypoints, which was also adopted by the CMFD evaluation framework [1]. This method was further improved, where the clustering object became a vector associated to the candidate transform estimation. It is shown that the new clustering-based CMFD scheme significantly raise the accuracy of localization of CMF regions. We know that an image is seldom forged aimlessly. Hence the copy- move regions should have a certain meaning. Then the CMFD can be performed by matching these patches, as long as the pasting target and copying source regions are not in the same patch. We note that this is not the first CMFD system that employs image segmentation technique. Farid proposed to detect the duplication in science images by grouping the pixels with similar properties. However, being designed for the science images such as gel and micrograph, Farid's method is not efficient and robust enough for normal images that are content rich and contain many different textures. Recently, Liu et al. also proposed a forgery detection method using JPEG features and local noises discrepancies , where segmentation is proved to be useful to splicing detection. In our proposed CMFD scheme, after segmenting the image, we perform the first stage of affine estimation. During this stage we first extract the keypoints from the whole image and construct a k-d tree. Then the KNN (k-nearest neighbour) search is performed in each region for each keypoint to find a possible correspondence.

B. Image Segmentation

In order to separate the copying source region from the pasting target region, the image should be segmented into small patches, each of which is semantically independent to the others. This job is best done by an expert with much experience of digital forensics. In our implementation, however, we only consider the automatic approach and leave the expert interfering method for future work. After testing four famous image segmentation methods, it is observed that the segmentation method does not greatly influence the CMFD's efficiency. Among them the methods are more favourable owing to their comparatively lower complexity. In most cases, one image sized 800×600 can be segmented in 15 seconds using a personal computer (3.3GHz CPU, 4G RAM). Figure 3 gives an example of image segmentation. One may concern the scenario that segmentation cannot help us to separate the CMF regions into different patches. As mentioned above, in order that two CMF regions do not exist in the same patch, we should not coarsely segment the image. In our implementation, each image is empirically segmented into no less than 100 patches (refer to Section V for a further explanation), and thus, a CMF region may be in two or more patches (refer to Figure 3). In consequence the useful information for CMFD is reduced in each patch. However, to obtain a convincing detection result we need not a large number of keypoints (sometimes four is enough).

To get the relationship between the frequency distribution of host images and the initial size of the superpixels a large number of experiments are performed. By using the haar wavelet 4-level DWT technique is applied to the host image then the image I divided into the low-frequency energy ELF and high-frequency energy EHF can be calculated using (1) and (2), respectively. The percentage of the low-frequency distribution in (3) can be calculated by using the low-frequency ELF and high-frequency energy EHF. The initial size of the superpixels can be defined in (4)

$$E_{LF} = \sum |CA_4| \quad (1)$$

$$E_{HF} = \sum_i \left(\sum |CD_i| + \sum |CH_i| + \sum |CV_i| \right), i = 1,2,\dots,4 \quad (2)$$

$$P_{LF} = \frac{E_{LF}}{E_{LF} + E_{HF}} . 100\% \quad (3)$$

$$S = \begin{cases} \sqrt{0.02 \times M \times N} & P_{LF} > 50\% \\ \sqrt{0.01 \times M \times N} & P_{LF} \leq 50\% \end{cases} \quad (4)$$

Where S means the initial size of the superpixels; M N indicates the size of the host image; and P LF means the percentage of the low-frequency distribution

II. FIRST STAGE OF MATCHING

In this section we will introduce the first stage of the matching process of our proposed CMFD system. The three steps involved in this stage will be detailed in the following three subsections.

A. Keypoint Extraction and Description

In our implementation, we employ vlFeat3 [35] software to help us to detect and describe the keypoints. There are many kinds of keypoint detection and description methods. The common co-variant keypoint detection and description algorithms, such as difference of Gaussian (DoG), Harrisaffine and Hessian-affine, can provide similar detection performance. In our implementation we just employ the default setting of vlFeat for keypoints detection and description, namely SIFT [16]. Although the methods of keypoint detection and description are not rather important, note that the number of the keypoints should be larger than 2000 for good performance.

B. Matching between Patches

Next we look for the suspicious pairs of patches that have many similar keypoints. This process is performed by comparing each patch with the rest. Refer to Figure 4, assume that patch A is considered at this time. Define the distance between two keypoints by the L-2 norm of the difference between their descriptors. In patch A for each keypoint we search its K nearest neighbours that are located in the other patches. Considering there are usually more than one couple of copy-move regions in the image, we set K = 10 in our implementation. We should not take all the K searched keypoints into consideration, but only if the difference is smaller than a threshold (0.04 in our implementation), the two keypoints are considered to be matched. In other words, each keypoint in patch A is corresponding to no more than K keypoints in the remaining patches. We know that the target and source regions should have a large proportion of matched keypoints. If a large proportion of the matched correspondences of A are located in another certain patch, say B in Figure 4, A and B are considered to be a suspicious pair of patches where we may find CMF regions. So a threshold ϕ is defined to find the matched patches. In our implementation, ϕ is empirically set as 10 times the average number of keypoints per patch.

III. SECOND STAGE OF MATCHING

In the first stage of matching process, we have found the suspicious pairs of patches as well as the transform matrix between them. Although RANSAC [20] can provide us with a robust estimation of transform matrix, it is still not accurate enough. Furthermore, some of these detected patches may be just false alarm containing not any CMF regions. In this section, we will introduce our second stage of matching process where the estimation of the transform matrix is refined via an EM-based algorithm. And the false alarm patches might also be eliminated in this stage.

A. CMF Determination Based on Probability

In the first stage of matching process, we made use of the detected keypoints in the copying source region and pasting target region to estimate a transform matrix H. This process follows the traditional way of computer vision [21]. In particular, the pixels not around the keypoints are abandoned. It is mainly because computer vision usually focuses on the research of transform estimation of two distinct images, in which case we are able to obtain a comparatively larger number of matched keypoints. However, in the CMFD case the forgery regions are sometimes so small that only a limited number of keypoints can be detected there. As a result, the detection result of the first stage is not convincing because we do not have enough keypoints. So in the second stage we propose to exploit all the pixels in the matched patches to find out a more accurate estimation $\bar{H}$. Meanwhile, the pixels belonging to the CMF regions would be more clearly distinguished from the background. Since the really matched pixels in the copying source region and pasting target region should be close to each other, Where S means the initial size of the super pixels; M N indicates the size of the host image; and P LF means the percentage of the low-frequency distribution.

B. Obtaining the New Correspondences of the Pixels

Denote the transform matrix we estimated in the first stage by H0 for differentiation here. As H0 is not accurate enough, the ~x0 obtained by (2) may not be the real correspondence of ~x. So we search a new correspondence of ~x in the pasting target region, such that the pixel located at the new correspondence position is more similar to the pixel at ~x than the old correspondence in terms of their local feature descriptions.

C. Iterative Re-estimation of the Transform Matrix

Using the newly matched pixel pairs we wish to estimate a more convincing matrix $\bar{H}$. Please note that some of these pixel pairs are outliers that are located outside the CMF region. Furthermore, some correspondences are not accurate enough because they may be at the smooth image regions. One natural solution is RANSAC as it is rather good at handling outliers. However, there usually are a large number of pixel pairs and hence RANSAC is too time-consuming. We have two classes of pixels in each segmented patch. One is the CMF region, the other is the background. Distinguishing the CMF region from the background is the same problem as classifying these two kinds of pixels. We propose to employ the EM algorithm [27] to this end. The EM algorithm is a useful method for statistical parameter estimation of the samples with underlying distributions. The algorithm repeats proceed.

5. IMPLEMENTATION

1) JPEG compression: the JPEG compressed images are the forgery images. The compression can be with a quality factor varying from 100 to 20, in steps of -10. So here we have to test the total of 489=432 images.

2) Rotation: the regions which are copied are rotated by the rotated angle varying from 2º to 10º, in steps of 2º, and the rotation angles are about 20º, 60º and 180º as well. So here we have to test the total of 488=384 images.

3)Scaling or Noise: The regions which are copied are scaled by using the scale factor varying from 91% to 109% in steps of 2%, and the scale factor is about 50%, 80%, 120%, and 200%. As well. So here we have to test the total of 4814=672 images.

4) Median filter: Total 48 forged host images are present in the dataset. These images are scaled down from 90% to 10% in steps of 20%. So here we have to test the total of 485=240 images.

EDGE-BASED SEGMENTATION

The focus of this section is on the segmentation methods based on detection in sharp, local changes in intensity. The three types of image feature in which we are interested are isolated points, lines , and edges. Edge pixels are pixels at which the intensity of an image function changes abruptly. We know local changes in intensity can be detected using derivatives. For reasons that will become evident, first- and second-order derivatives are particularly well suited for this purpose

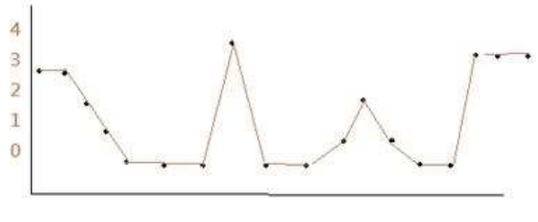


Image strip	3	3	2	1	0	0	4	0	0	1	2	1	0	0	4	4	4
1st. derivative	0	-1	-1	-1	0	4	-4	0	1	1	-1	-1	0	4	0	0	0
2nd. derivative	0	-1	-0	0	1	4	-8	4	1	0	-2	0	1	4	-4	0	0

The intensity histogram of image

We have following conclusions from Figure . The intensity histogram of imageFigure First-order derivatives generally produce thicker edges in an image.

1. Second-order derivative have a stronger response to fine detail, such as thin lines, isolated points, and noise.
2. Second-derivatives produce a double-edge response at ramp and step transitions in intensity.
3. The sign of the second derivative can be used to determine whether a transition into an edge is from light to dark or dark to light.

ISOLATED POINTS

It is based on the conclusions reached in the preceding section. We know that point detection should be based on the second derivative, so we expect Laplacian mask.

1	1	1
1	-8	1
1	1	1

The mask of isolation point

We can use the mask to scan the all point of image , and count the response of every point, if the response of point is greater than T(the threshold we set), we can define the point to 1(light), if not, set to 0(dark).

$$G(x,y)=\begin{cases}1 & \text{if } |R(x,y)|\geq T\\ 0 & \text{otherwise}\end{cases}$$

LINE DETECTION

As discussion in section2.1,we know the second order derivative have stronger response and to produce thinner lines than 1st derivative. We can get four different direction of mask.

-1	-1	-1
2	2	2
-1	-1	-1

Horizontal

2	-1	-1
-1	2	-1
-1	-1	2

+45

-1	2	-1
-1	2	-1
-1	2	-1

Vertical

-1	-1	2
-1	2	-1
2	-1	-1

-45

Line detection masks.

Let talk about how to use the four masks to decide which direction of mask is better than others.

Let R_1, R_2, R_3 and R_4 denote the response of the masks in Figure 2.3. If at a point in the image , $|R_k|>|R_j|$,for all j ≠ k. $|R_1|>|R_j|$ for j=2,3,4,that point is said to be more likely associated with a line in the direction of mask k.

EDGE DETECTION

(a) Two region of constant intensity separated by an ideal vertical ramp edge.(b)Detail near the edge, showing a horizontal intensity profile.

We conclude from the observation that the magnitude of 1st derivative can be used to detect the presence of an edge at a point in an image. The 2nd derivative have two properties : (1) it produces two values for every edge in an image.(2)its zero crossings can be used for locating the center of thick edges.

Basic Edge Detection (gradient)

The image gradient is to find edge strength and direction at location (x,y) of image, and defines as the vector.

The magnitude (length) of vector ∇f , denoted as M(x,y)

$$\text{mag}(\nabla f) - \sqrt{g_x + g_y} \qquad \alpha(x,y) = \tan^{-1}\left[\frac{g_y}{g_x}\right] \qquad \nabla f \equiv \text{grad}(f) \equiv \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$$

The direction of the gradient vector is given by the angle. The direction of an edge at an arbitrary point (x,y) is orthogonal to the direction. We are dealing with digital quantities, so a digital approximation of the partial derivatives over a neighborhood about a point is required.

SOBEL OPERATOR

-1	-2	-1
0	0	0
1	2	1

(e)

-1	0	1
-2	0	2
-1	0	1

(f)

0	1	2
-1	0	1
-2	-1	0

(f)

-2	-1	0
-1	0	1
0	1	2

(g)

(a)~(g) are region of an image and various masks used to compute the gradient at the point labe z_5

$$g_x = \frac{\partial f}{\partial x} = (z_7 + 2z_8 + z_9) - (z_1 + 2z_2 + z_3)$$

$$g_y = \frac{\partial f}{\partial y} = (z_3 + 2z_6 + z_9) - (z_1 + 2z_4 + z_7)$$

$$M(x,y) \approx |g_x| + |g_y|$$

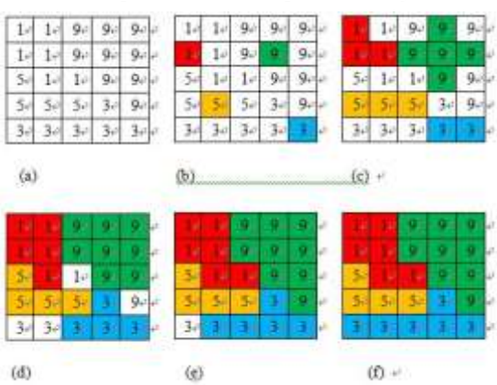
The Sobel mask uses 2 in the center location for image smoothing. The Prewitt masks are simpler to implement than Sobel masks, but the Sobel masks have better noise-suppression (smoothing) characteristics makes them preferable.

In the previous discussion, we just discuss to obtain the g_x and g_y. However, this implementation is not always desirable, so an approach used frequently is to approximately the magnitude of the gradient by absolute values:

Region-Based Segmentation

Region growing segmentation is an approach to examine the neighboring pixels of the initial “seed points” and determine if the pixels are added to the seed point or not.

Step1. Selecting a set of one or more starting point (seed) often can be based on the nature of the problem. Step2. The region are grown from these seed points to adjacent point depending on a threshold or criteria (8-connected) we make.
Step3. Region growth should stop when no more pixels satisfy the criteria for inclusion in that region



(a)Original image (b)Use step 1 to find seed based on the nature problem.(c) Use Step 2(4-connected here) to growing the region and finding the similar point. (d)(e) repeat Step 2. Until no more pixels satisfy the criteria. (f) The final image.

Then we can conclude several important issues about region growing :

1. The suitable selection of seed points is important. The selection of seed points is depending on the users.

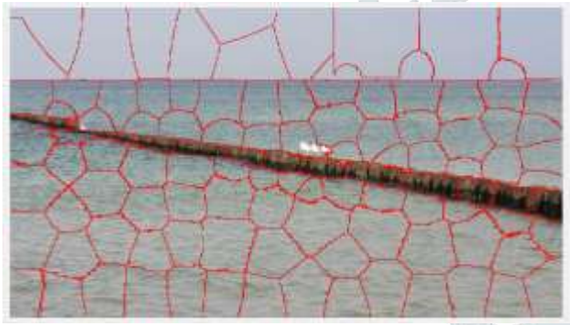
2. More information of the image is better. Obviously, the connectivity or pixel adjacent information is helpful for us to determine the threshold and seed points.
3. The value, “minimum area threshold”. No region in region growing method result will be smaller than this threshold in the segmented image.
4. The value, “Similarity threshold value“. If the difference of pixel-value or the difference value of average gray level of a set of pixels less than “Similarity threshold value”, the regions will be considered as a same region.
5. The result of an image after region growing still have point’s gray-level higher than the threshold but not connected with the object in image.

6. RESULTS AND ANALYSIS

In this section, a series of experiments square measure conducted to evaluate the effectiveness and lustiness of the planned image forgery detection theme exploitation adaptive over-segmentation and have purpose matching. Within the following experiments, the image dataset in [22] is employed to check the proposed methodology. This dataset is made supported forty eight high-resolution uncompressed PNG true color pictures, and the average size of the pictures is 1500*1500. within the dataset, the copied regions square measure from the classes of living, nature, man-made and mixed, and that they vary from to a fault swish to highly textured; the copy-move forgeries square measure created by copying, scaling and rotating semantically substantive image regions.

Original Image
We select an original image from a folder

Forged Image
Based on the original image, the forged image that is named just like original image name, followed by extension '_copy' is selected



Segments of an Image
Here, the image is segmented into separate sections.

Forgery Detection
Here, the forged section pixels are displayed.

.

Forgery Region Detection
Here, the total forged area is displayed.

Image Detection at Forged Region
Here, the image at forged part is displayed.

CONCLUSION

In this paper we presented a CMFD scheme based on image segmentation. Although the CMF regions are detected mainly by comparing the key points extracted in the image, we cannot simply classify the proposed scheme as a key point-based one. It can be seen as a combination of both existing schemes because in the two stages of matching process both key points and pixel features are employed. Our main contributions can be concluded to the following two aspects.

1) Considering the CMF regions usually have certain meaning, we propose to segment the image into semantically independent patches, such that the CMFD problem can be solved by partial matching among these segmented patches.

2) The matching process between segmented patches consists of two stages. In the second stage, an accurate estimation of transform matrix can be obtained by an EM-based algorithm. One may concern the computational complexity of the proposed scheme. Compared with the keypoint-based schemes, the proposed scheme mainly needs two more steps, namely the image segmentation and the transform estimation refinement. If using some efficient methods, we are able to segment an image in several seconds. The re-estimation of transform matrix is more complex because it needs an iterative procedure (refer to Section IV-C). However, owing to the threshold set in (1), only a few patches (about one tenth) need the second stage of matching for transform matrix re-estimation. In our future work, we will try to improve the detection speed of the proposed scheme by means of parallel programming.

REFERENCES

[1] A. J. Fridrich, B. D. Soukal, and A. J. Lukáš, "Detection of copy-move forgery in digital images," in in Proceedings of Digital Forensic Research Workshop, 2003.

[2] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," Dept. Comput. Sci., Dartmouth College, Tech. Rep. TR2004-515, 2004.

[3] W. Luo, J. Huang, and G. Qiu, "Robust detection of region-duplication forgery in digital image," in Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, 2006, pp. 746-749.

[4] G. Li, Q. Wu, D. Tu, and S. Sun, "A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD," in Multimedia and Expo, 2007 IEEE International Conference on, 2007, pp. 1750-1753.

[5] B. Mahdian and S. Saic, "Detection of copy–move forgery using a method based on blur moment invariants," Forensic science international, vol. 171, pp. 180-189, 2007.

[6] X. Kang and S. Wei, "Identifying tampered regions using singular value decomposition in digital image forensics," in Computer Science and Software Engineering, 2008 International Conference on, 2008, pp. 926-930.

[7] S. Bayram, H. T. Sencar, and N. Memon, "An efficient and robust method for detecting copy-move forgery," in Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, 2009, pp. 1053-1056.

[8] J. Wang, G. Liu, H. Li, Y. Dai, and Z. Wang, "Detection of image region duplication forgery using model with circle block," in Multimedia Information Networking and Security, 2009. MINES'09. International Conference on, 2009,

[9] J. Wang, G. Liu, Z. Zhang, Y. Dai, and Z. Wang, "Fast and robust forensics for image region-duplication forgery," Acta Automatica Sinica, vol. 35, pp. 1488-1495, 2009.

[10] H. Lin, C. Wang, and Y. Kao, "Fast copy-move forgery detection," WSEAS Transactions on Signal Processing, vol. 5, pp. 188-197, 2009.

[11] S. Ryu, M. Lee, and H. Lee, "Detection of copy-rotate-move forgery using Zernike moments," in Information Hiding, 2010, pp. 51-65.

[12] S. J. Ryu, M. Kirchner, M. J. Lee, and H. K. Lee, "Rotation Invariant Localization of Duplicated Image Regions Based on Zernike Moments," Ieee Transactions on Information Forensics and Security, vol. 8, pp. 1355-1370, Aug 2013.

[13] S. Bravo-Solorio and A. K. Nandi, "Exposing duplicated regions affected by reflection, rotation and scaling," in Acoustics, Speech and Signal Processing (ICASSP), 2011 IEEE International Conference on, 2011, pp. 1880-1883.

[14] H. Huang, W. Guo, and Y. Zhang, "Detection of copy-move forgery in digital images using SIFT algorithm," in Computational Intelligence and Industrial Application, 2008. PACIIA'08. Pacific-Asia Workshop on, 2008, pp. 272-276.

[15] X. Y. Pan and S. Lyu, "Region Duplication Detection Using Image Feature Matching," Ieee Transactions on Information Forensics and Security, vol. 5, pp. 857-867, Dec 2010.

[16] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A sift-based forensic method for copy–move attack detection and transformation recovery," Information Forensics and Security, IEEE Transactions on, vol. 6, pp. 1099-1110, 2011.

[17] X. Bo, W. Junwen, L. Guangjie, and D. Yuewei, "Image copy-move forgery detection based on SURF," in Multimedia Information Networking and Security (MINES), 2010 International Conference on, 2010, pp. 889-892.

[18] P. Kakar and N. Sudha, "Exposing Postprocessed Copy–Paste Forgeries Through Transform-Invariant Features," Information Forensics and Security, IEEE Transactions on, vol. 7, pp. 1018-1028, 2012.

[19] B. Shivakumar and L. D. S. S. Baboo, "Detection of region duplication forgery in digital images using SURF," IJCSI International Journal of Computer Science Issues, vol. 8, 2011.

[20] D. G. Lowe, "Object recognition from local scale-invariant features," in Computer vision, 1999. The proceedings of the seventh IEEE international conference on, 1999, pp. 1150-1157.

[21] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in Computer Vision–ECCV 2006, ed: Springer, 2006, pp. 404-417.

[22] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An Evaluation of Popular Copy-Move Forgery Detection Approaches," Ieee Transactions on Information Forensics and Security, vol. 7, pp. 1841-1854, Dec 2012