An Adaptive Hybrid Algorithm For Digital Image Copy-Move Forgery Detection

Dissertation submitted in partial fulfillment of requirement

for the award of degree of

Master of Engineering

in

Electronics and Communication Engineering

Submitted by

Amanjot Kaur Lamba

Roll No: 801461002

Under the guidance of

Dr. Neeru Jindal

(Assistant Professor, ECED)

Dr. Sanjay Sharma

(Professor & Head, ECED)

ELECTRONICS AND COMMUNICATION ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA – 147004 (PUNJAB)

JULY 2016

DECLARATION

I, Amanjot Kaur Lamba, hereby declare that the dissertation entitled **"An Adaptive Hybrid Algorithm For Digital Image Copy-Move Forgery Detection"**, is an authentic work carried out by me towards the partial fulfillment for the award of degree of Master of Engineering (M.E.) in Electronics and Communication (ECE), accomplished under the esteemed guidance of **Dr. Sanjay Sharma** (Professor & Head) and **Dr. Neeru Jindal** (Assistant Professor), Electronics and Communication Department (ECED), Thapar University, Patiala.

The content presented in this dissertation has not been submitted in any form in other University/Institute for the award of any other degree.

Date: 14 July 2016

Amanjot Kaur

Amanjot Kaur Lamba
Roll no.: 801461002

It is certified that the above statement made by the student is correct to the best of my knowledge and belief.

Dr. Neeru Jindal
Assistant Professor, ECED
Thapar University, Patiala
Date: _____

Dr. Sanjay Sharma
Professor & Head, ECED
Thapar University, Patiala
Date: _____

Countersigned by:

Dr. Sanjay Sharma
Professor & Head, ECED
Thapar University, Patiala
Date: _____

Dr. S. S. Bhatia
Dean of Academic Affairs
Thapar University, Patiala
Date: _____

ACKNOWLEDGEMENT

First and foremost, I am very grateful to Almighty GOD for the wisdom he bestowed upon me, the peace of my mind and the good health throughout the project duration.

I am highly indebted to **Dr. Sanjay Sharma**, Professor and Head, and **Dr. Neeru Jindal**, Assistant Professor, Electronics and Communication Engineering Department, Thapar University, Patiala who always as mentors encouraged me throughout the work. I am blessed to have the opportunity to work under their valuable guidance.

I convey my sincere thanks to **Dr. Amit Kumar Kohli**, Associate Professor, P.G. Coordinator and **Dr. Amit Mishra**, Assistant Professor, Programme coordinator (ECE), Electronics and Communication Engineering Department, Thapar University, Patiala who have provided me learning atmosphere and infrastructure for the completion of work. I am also thankful to the staff and faculty members of Electronics and Communication Engineering Department for their encouragements.

I thank all my friends for their support and motivation that strengthened me to complete the work with full dedication.

I would like to extend my appreciations to my parents and family for their immense love, support, encouragement and selfless help in the complete tenure of my work. They stood behind me with their blessings in every hardship of my work.

Amanjot Kaur

Amanjot Kaur Lamba

ABSTRACT

Due to the development of sophisticated cameras and image editing tools, digital image tampering techniques are frequently used without leaving visual cues behind. Digital image copy-move forgery is said to be an image manipulation which involves copying and pasting of certain section (or sections) within the same digital image. Generally, this is done with intention of hiding important information or providing false information in an image. This motivates a need for forgery detection systems that are transparent to such manipulations and can disclose whether a given image has been morphed just by investigating the dummy image.

Several methods have been presented for the copy-move forgery detection in recent years. Nearly all of the existing block-based methods are computationally expensive and robust to noise addition, JPEG compression, but are susceptible to geometrical attacks like rotation, translation and scaling. On the other hand, keypoint-based detection techniques are computationally efficient as well as perform better under geometrical attacks in comparison with block-based methods but suffer from low recall rate. The proposed technique is a hybrid one which incorporates both block-based and keypoint-based schemes in order to deal with their drawbacks. Focus of the proposed thesis work is on achieving 100% precision and recall at image level copy move forgery detection using adaptive algorithm. Firstly, adaptive image segmentation is performed on the test image resulting in image patches followed by detection and extraction of features of these patches. These features are matched patch-wise to obtain suspected keypoint pairs. An adaptive keypoint matching algorithm is used to extract matched keypoint pairs from the suspected keypoint pairs. Finally, an adaptive forgery region extraction is used to locate similar areas in the test image.

The evaluation results demonstrate that the proposed hybrid scheme is more robust under plain as well as various challenging situations such as down-sampling, up-scaling, down-scaling and JPEG compression than the prior state-of-the-art techniques. The proposed scheme achieved improved results with 100% precision, 100% recall and 100% F1 score at image level, while 95.01% precision, 87.18% recall and 90.92% F1 score at pixel level under plain copy-move attack. The proposed adaptive scheme can be extended to videos in future.

TABLE OF CONTENTS

ABBREVIATIONS AND ACRONYMS

AWGN	Additive White Gaussian Noise
BMP	Bitmap
CMF	Copy-Move Forgery
CMFD	Copy-Move Forgery Detection
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
DYWT	Dyadic Wavelet Transform
FT	Fourier Transform
IP	Image Patch
JPEG	Joint Photographic Experts Group
KPCA	Kernel Principal Component Analysis
MIFT	Mirror Reflection Invariant Feature Transform
MKP	Matched Keypoint Pair
PCA	Principal Component Analysis
PF	Patch Features
SIFT	Scale Invariant Feature Transform
SKP	Suspected Keypoint Pair
SLIC	Simple Linear Iterative Clustering
SURF	Speeded Up Robust Feature
SVD	Singular Value Decomposition
WPD	Wavelet Packet Decomposition

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

This chapter familiarizes with copy-move forgery (CMF), its classification and detection techniques.

1.1 PREAMBLE

Digital images can be modified without leaving distinct cues due to the technological advancements and immense growth in image processing tools [1]. It becomes difficult for naked eye to visually examine whether a given image is genuine or not. Hence, integrity and credibility of digital images have become indispensable. So, there is a need to develop precise methods that could verify their authenticity, taking into account the fact that the images are substantially treated as proofs in law courts or as newspaper commodities [2]. In this manner, digital image forgery detection has become one of the major objectives of image forensics.

The classification of digital image tampering detection schemes is done into two categories, namely active techniques and passive techniques. In active schemes, prior information about an image is absolutely necessary for authentication which limits their application. It involves the pre-processing of images like signature generation and watermark embedding. On the contrary, there is no need of such pre-processing methods or prior information in case of the passive methods like CMF and image splicing. Of the existing forgery detection techniques, the CMF is a popular and common forgery owing to its efficacy.

1.2 COPY-MOVE FORGERY

Digital image CMF is a kind of image manipulation which involves copying and pasting of certain section (or sections) within the same digital image [3]. This may be performed with some false intentions to hide an object or evidence in that image by pasting the copied patch over it. Before doing so, the replicated segments can be made to undergo one or a mixture of some geometrical transformations such as rotation, scaling, etc. in order to make them fit for the locations on which they are being moved to [4].

Textured areas, like grass, gravel, foliage, or any fabric having inconsistent framework, are ideal for such purposes as the similitude areas are more likely to mingle with the image's background and the humans are incapable in discerning any dubious artifacts easily. Those portions altered by CMF are often almost imperceptible and indistinguishable by a human eye; thus, detection of evidence of such actions is a primary concern in image forensics.

Because the copied patches are of same image, their color palette, dynamic range, noise component, and other key characteristics will be absolutely compatible with it. For making the manipulations more harder for detection, forgers often use the retouch tool or/and feathered crop for further masking any cues of the copied-and-moved patches. Hence, detecting any CMF in an image requires extensive exploration of its local patterns and region matches.

In Fig. 1.1, one can see less obvious CMF in which a part of foliage is used to hide a truck (compare both original and forged image). Still, it is not difficult to comprehend the imitated areas visually as both the copied-and-moved parts of the foliage bear a questionable similarity.

Fig. 1.1 An original image “Truck” (left) and its forged version (right). [5]

1.3 COPY-MOVE FORGERY DETECTION

CMFD methods are basically passive methods which rely on the assumption that tampering is expected to change the elementary statistics, even when it may not leave visual cues behind [6]. Such inconsistencies are used in CMFD algorithms.

The aim of CMFD is not only to detect CMF but also to localize the similar patches in forged images. Fig. 1.2 shows a forged image (above) and detection map (below) . The

similar regions in the forged image are represented by the white regions in the detection map. The location of the white regions in detection map are with respect to that of the duplicate regions in the forged image.

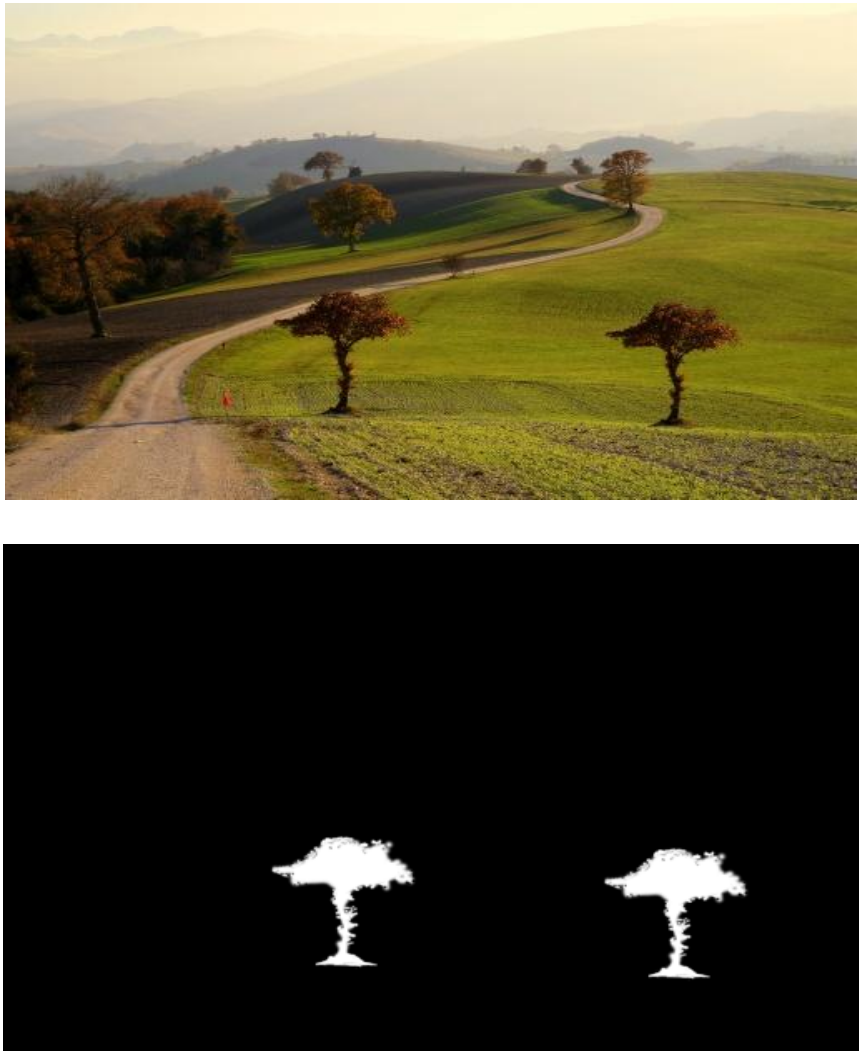


Fig. 1.2 A forged image (above) and its detection map (below). [4]

There is indeed an extremely low possibility that digital images used in forensics like criminals' photographs, fingerprint images, crime scene images, and many more would be unassailable by digital watermarks as well as digital signatures. Hence, Blind/ Passive algorithm is the most suitable approach to discover tampering in given images used by the department of forensics because a passive technique has no requirement of any kind of prior knowledge about the test image [7]. For the detection of forgery traces, blind techniques often employ image functions. Also, they consider the fact that forgeries often introduce some specific changes in images that are detectable.

All the CMFD techniques are forgery dependent techniques. These are designed specifically for this cloning detection. Since, in CMF, the copied and/or moved regions are the parts of the exact image, an extreme correlation exists among these portions which can be further employed as a criteria for CMFD. But the key challenge in CMFD is the high computational burden resulted from the extensive search done for formulating correlated/matched segments.

Also, these tampered images are usually instilled with some local noise for concealing tampering done to those images. It is a common practice among forgers to add random noise locally to the copied-and-moved parts of images [8]. Noise degradation significantly affects the results of almost all of the tamper detection methods.

1.3.1 Generalized Framework For CMFD

A generalized framework for CMFD approach has been discussed in Fig. 1.3. Regardless of the ample scope of algorithms that have been proposed for CMFD, most of them adhere to this common framework.

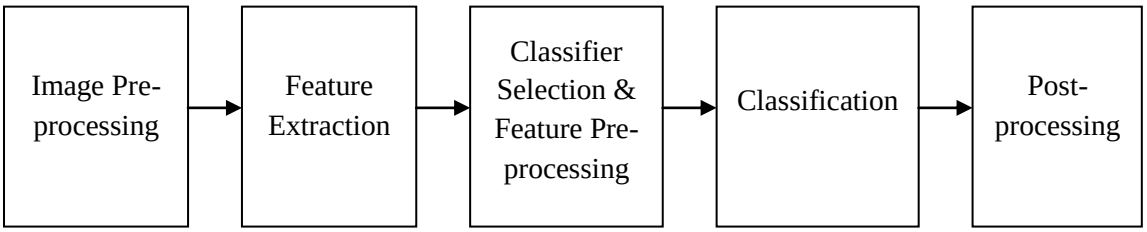


Fig. 1.3 Generalized framework for CMFD.

(a) Image pre-processing:

Before any feature extraction procedure, some of the processes are performed over a test image, which include resizing, cropping, gray-scale conversion from RGB color space etc. for the classification performance improvement. This is often done to make these images more suitable for the succeeding stages.

A forged image having size $L \times M$ is converted into grayscale format (excluding the methods which need RGB channels). In most of the block-based identification methods, the test image is firstly split into respective overlapping blocks each of $s \times s$ dimensions resulting in total number of $N = (L - s + 1) \times (M - s + 1)$ overlapping blocks. With the upsurge in the size of the trail image, there would be proliferate in computational load.

Segmentation is becoming popular among authors to get non-overlapping and semantically independent image patches (IP), especially in case of hybrid detection techniques.

(b) Feature extraction:

Features set is extracted for every class which is helpful in distinguishing it from all the other classes, meanwhile being essentially invariant to all the differences in attributes within a class from the host tampered data.

The overlapping blocks that are obtained in block-based techniques are often arranged in a row vector f_j each of dimensions $1 \times P$, where $P = s^2$. Features of each of the overlapping blocks are extracted using any appropriate technique (depending upon the algorithm) and is stored in the form of row vectors. The number of features to be extracted solely depend upon the authors.

In Keypoint-based techniques, a test image is scanned for keypoint detection, and for every keypoint, the feature vector f_j of size $1 \times P$ is calculated. Positions (x_j, y_j) of each keypoint are also stored in f_j, making its size $1 \times (P + 2)$. For K number of keypoints, a feature matrix of size $K \times (P + 2)$ is obtained.

In particular, significant and informative features need to be extracted and features that are selected must be responsive to image manipulations. The desirable attribute of chosen features and formed feature vectors is that they should have low dimensions, as this will lower the computational complications of training as well as classifications.

(c) Classifier selection and feature pre-processing:

On the basis of the extracted features, a suitable classifier needs to be chosen or developed. A large set of digital images for the classifier's training is selected and some salient parameters of the chosen classifier are obtained, that can be exploited for this classification.

Pre-processing of features is useful in reducing the features dimensionality excluding deterioration of performance of classification based on machine-learning. For dimensionality diminution, Principal Component Analysis (PCA) and Singular Value Decomposition (SVD) are generally used.

(d) Classification:

A classifier distinguishes any given image and categorizes it into two classes: genuine and tampered digital image. On the basis of extracted features set, a suitable classifier is further chosen or designed. A humongous priming set is more likely to give a better classifier in terms of performance.

Phase correlation and lexicographical sort are frequently used classifiers in case of block-based techniques; which result in bringing alike blocks closer. A threshold is used in searching matching feature vectors among the closest neighbors. Let F_i be a matched pair that consists of features f_u, f_v where $u!=v$ represents feature vector indices. Now, the shift vector, s_{uv} in these two matching blocks is determined as:

$$s_{uv}(dx, dy) = (x_u - x_v, y_u - y_v) \tag{1.1}$$

Here, (x_u, y_u) and (x_v, y_v) denotes the spatial location of the features f_m, f_n respectively, and dx, dy represents the difference between these spatial locations.

A counter of shift vectors, $C_s(i)$ is obtained (and increased by one) for each and every alike blocks pair having the equal shift using (1.2):

$$C_s(dx, dy) = C_s(dx, dy) + 1 \tag{1.2}$$

Grouping of the matched block pairs having the alike shift is done. Those groups of blocks which have $C_s(.)$ below a threshold T_p are discarded. T_p keeps in check the size of the tiniest detectable copy-move patch. This technique is named as thresholding. It reduces false alarms and yields forged regions. The value T_p is carefully chosen using the training sets. This is usually done in most of the existing block-based algorithms.

(e) Post-processing:

In post-processing operation, localization of replicated regions is done as investigated in Fridrich *et al*. [5], Muhammad *et al*. [9] and Ghorbani *et al*. [10].

This final step involves morphological operations which are performed with the intent to lower false positive rate. To discriminate various copy-move patches, matching patches belonging to equal shift vectors are marked by the same color, usually white, to visually locate the duplicated areas.

1.3.2 Classification Of CMFD Techniques

The existing CMFD schemes are, broadly, divided into two categories: block-based CMFD techniques [5], [9]-[25] and keypoint-based CMFD techniques [26]-[34]. The point of block-based CMFD techniques is dividing a given image into either overlapping or non-overlapping tiles followed by the application of any transform on each tile. Similar blocks are computed on the basis of some similarity criterion. However, the keypoint-based CMFD methods involves the extraction of interest points in an image. They make use of local features of the interest points for the identification of duplicated regions.

The existing block-based schemes, in fact, deal with a large number of blocks along with their feature vectors. This results in high computational burden. Although such techniques are robust to the addition of noise, JPEG compression and filtration, yet are vulnerable to geometrical transformations like translation, scaling and rotation. Due to the regular shape of blocks, these techniques yield very low recall rate. On the contrary, keypoint-based detection schemes are computationally efficacious and give better performance under geometrical transformations as compared to the block-based techniques. But the existing keypoint-based algorithms too result in low recall rate.

However, there is another class of CMFD schemes which integrates both block-based and keypoint-based schemes in order to cope with the limitations of the traditional techniques. In hybrid CMFD algorithms, host image is segmented into image blocks before keypoint extraction. In [35]-[37], authors employed hybrid method for forgery detection. These techniques are quite robust to the various geometrical attacks, in addition to the less computational burden.

1.4 CMF ATTACKS

A number of image handling operations that could be employed in empirical CMF are shown in Fig. 1.4. These operations are, here, grouped broadly into following two classes: intermediate and post-processing; depending upon when these are performed in the whole procedure of CMF.

Intermediate processes are mainly used in providing homogeneity and spatial synchronization in the replicated regions and their neighbors. These intermediate

operations include rotation, mirroring, scaling, chrominance modifying, or illumination modifying.

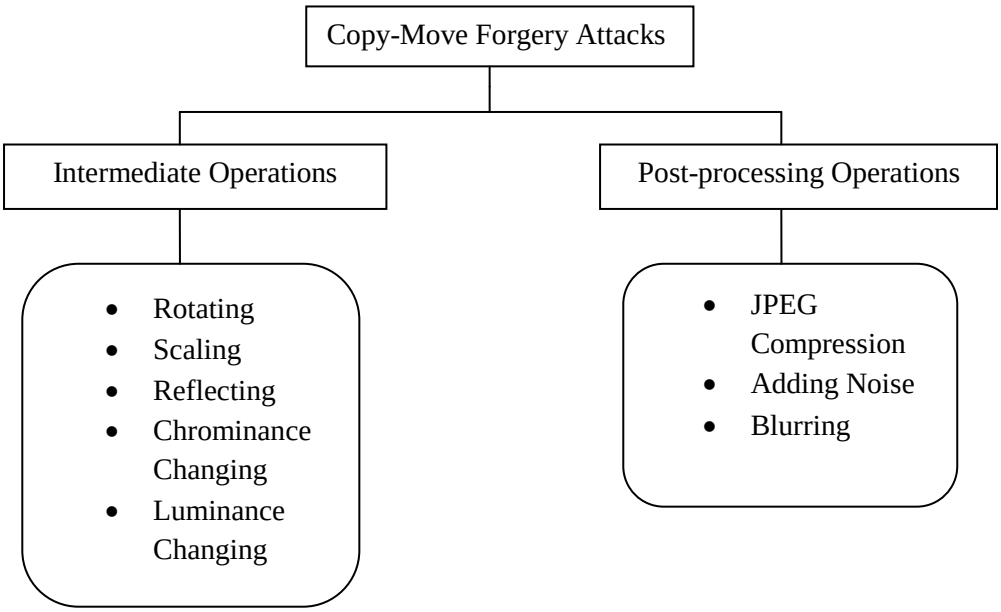


Fig. 1.4 Image processing operations related to CMF. [3]

Practically, a combination of two or more operations are performed by forgers in intermediate processing. A scaling transformation changes the size of the replicated patch, which often creates illusion. Alignment angle of the replicated patch is changed by rotational transformation. In some cases, luminance or/and chrominance are changed in order to reduce the similarities between the copied-and-moved regions.

The post-processing operations include compression, blurring and additive noise. Such operations usually efface any traceable artifacts of this cloning operation, for instance serrated edges. JPEG is one of the most prevailing and widely used compression standard. To comprehend if an image which is in Bitmap (BMP) format was earlier Joint Photographic Experts Group (JPEG) file format compressed is a chief concern for some of the image processing applications including image tampering detection.

1.5 ORGANIZATION OF THESIS

The work consists of following chapters:

CHAPTER 1: INTRODUCTION, briefs about CMF, types of CMFD schemes and different CMF attacks.

CHAPTER 2: LITERATURE SURVEY, discusses about developments in different types of CMFD techniques.

CHAPTER 3: DIGITAL FORENSIC TOOLS, briefly describes different types of digital image forgery detection tools.

CHAPTER 4: RESEARCH METHODOLOGY, explains the proposed hybrid CMFD technique in detail.

CHAPTER 5: EXPERIMENTAL RESULTS, briefs about the performance of the proposed adaptive scheme under plain CMF as well as CMF under various attacks, like down-sampling, up-scaling, down-scaling and JPEG compression and it's comparison with existing ones.

CHAPTER 6: CONCLUSION AND FUTURE SCOPE, summarizes the results obtained from the proposed method and discusses the possibility of future work.

CHAPTER 2

LITERATURE SURVEY

The literature survey is presented with many sections on the basis of detection techniques, namely block-based CMFD techniques, keypoint-based CMFD techniques and hybrid CMFD techniques.

2.1 BLOCK-BASED CMFD TECHNIQUES

The first and the foremost attempt in the determination of tampered regions using a block-based detection scheme was made by Fridrich *et al.* [5]. The authors divided the host image into overlapping tiles followed by the application of Discrete Cosine Transform (DCT) to every image tile to extract their features. Lexicographical representation was done to reduce the computational complications. Adjacent similar blocks pairs, then, were considered as the potential replicated regions. For the refining of results, a histogram was determined which counted the matching blocks that were equal distance apart. At last, a pre-defined threshold value is used in order to discard false alarms and finally determine the duplicated patches. This method achieved the best balance between complexity and performance but at the same time, this technique was incapable in detecting small replicated regions.

Popescu *et al.* [11] made use of PCA to represent overlapping image segments of the test image. The use of PCA resulted in decrease in feature dimensions which further reduced the computational load. This method achieved average detection accuracies of 50% with block size of 32×32 and 100% with block size of 160×160, both with JPEG quality = 95. Low JPEG qualities and small block sizes degraded the detection accuracy.

Luo *et al.* [12] presented a block-based method in which RGB components as well as direction information were used as block features. This algorithm exhibited robustness opposed to various attacks, like blurring, lossy compression and noise contamination. It achieved accuracy of 96.31% along with false negative rate of 9.66% in case of multiple attacks.

In [13], Wang *et al.* proposed a block-based CMFD technique which utilized Gaussian pyramid to reduce the image dimensions followed by the division of image into

overlapping blocks. Hu moment was employed by authors to each of the image block. The obtained eigen vectors were sorted lexicographically and then, their matching was done to find similar eigen vectors. This method not only exhibited good performance when the tampered images were maligned by the noise addition, lossy compression and Gaussian blurring, but also reduced the computational burden by narrowing the block-matching search space. However, this technique suffered from low detection rate.

Wang *et al.* [14] firstly, reduced the test image with Gaussian pyramid and extracted four features from each of the circlular block. The obtained feature vector was subjected to lexicographical sort followed by their matching using a threshold value. The proposed area threshold value was helpful in eliminating false negatives. This technique narrowed the block-matching searching space by reducing the total number of blocks, thereby, increasing the efficiency of the scheme. Moreover, this method proved to be robust to noise addition, blurring, rotation and JPEG compression with 65 quality factor.

Bashar *et al*. [15] proposed Kernel Principal Component Analysis (KPCA) and wavelet transform for key features extraction from the small blocks obtained from an image which were then subjected to lexicographical sort to indicate any resemblance in corresponding small blocks. The paper proposed algorithms that catch forged segments with translation, flip along with rotation. The cases of additive noise as well as lossy JPEG compression were also considered. KPCA is the best suitable for compressed and noisy data, also rotation as compared to PCA and wavelet-based. Also, wavelet-based features performed well in artifacts free environment. However, scaling and shearing were not considered.

Muhammad *et al.* [8] introduced a block-based technique based upon undecimated Dyadic Wavelet Transform (DYWT). The authors used DYWT for the decomposition of the digital image. They used both HH1 and LL1 rather than using LL1 sub-bands only. For each of the sub-band, wavelet coefficients calculated of each patch were considered as its feature vector. Euclidean distance, calculated between vector pairs, were arranged in increasing sequence for LL1 and in decreasing sequence for HH1. This generated two lists, that were reduced based upon the two thresholds. If any vector pair, with respect to their distance, seemed at the alike position in the both lists, then the corresponding block pair was determined as replicated. The accuracy claimed by this algorithm is 95.9% with false positives of 4.54%.

Ghorbani *et al*. [9] illustrated an algorithm based upon Discrete Wavelet Transform (DWT) which reduced the dimensions of the trial image. Decomposition of DCT-quantization coefficients was done to lower the dimensions of feature vector for CMFD. The sorting of feature matrix was done lexicographically and the replicated patches were detected by an extensive matching process. This proposed scheme is certainly capable of detecting such image manipulations accurately provided the copied patch is not rotated or/and scaled.

In [16], Bravo-Solorio *et al*. proposed an efficient method to find replicated regions transformed by reflection, scaling and/or rotation. Authors mapped overlapping image blocks to color-dependent features in addition to one-dimensional descriptor which is reflection and rotation invariant. These features were, then, lexicographically sorted in order to reduce the number of comparisons required in block-matching stage by bringing similar image closer to each other. The matched blocks are refined to get the potential replicates. This scheme exhibited excellent performance in terms of localisation and detection along with the less false alarm rate. However, the technique lacked in case of images having very little textural information over the large regions.

Cao *et al*. [17] presented a robust CMFD algorithm in 2012. The authors employed DCT coefficients' mean for the optimization of size of the feature vectors. In this technique, RGB to gray-scale conversion was performed and then, digital image was sliced up into N overlapping blocks. For every block B, DCT was implemented, and then, a circular block was utilised for the coefficients representation. The circular block was split up into four portions. Feature vector V was attained by the calculation of mean of coefficients of each part. Every single feature vector extracted was, then, lexicographical sorted. Calculation of the Euclidean distance between every consecutive vectors pair was done. Two blocks were said to be duplicate when their corresponding Euclidean distance was less than the pre-defined threshold. Along with the reduction in the feature vectors, this algorithm depicted quite good results as it could hold on to multiple CMF, and additive noise and blurring procedures too. Nonetheless, it was affected by scaling and rotation.

Nguyen *et al.* [18] proposed Radon transformation for the extraction of the block features as well as used phase correlation as classifier to discern the matching vector pairs. The proposed technique was performed quite well for the tampered images in which the

duplicated region's rotation angle was below 4°, had Gaussian noise addition and signal-to-noise ratio more than 35dB along with the block size of 4×4 pixels.

Wang *et al.* [19] introduced an approach based upon wavelets in order to identify area replication tampering. The image was divided into definite and equal sized overlapping blocks. The multiple level two-dimensional DWT was then applied to those blocks. The discriminative features were extracted from those wavelet coefficients for each patch. The feature vector of all the patches were lexicographical sorted. The patch-matching step was implemented to determine the duplicated regions. The output image was then generated to label all the replicated regions. This proposed method performed quite efficiently, especially whenever the forged images were heavily noise distorted.

Lynch *et al.* [20] developed an expanding block method for copied-and-moved region identification. The authors divided a test image into overlapping tiles and for each tile, grey value was computed being its supreme feature. A connection matrix was created based on this dominant factor. If this connection matrix had complete row of zeros, then that tile related to the row was disconnected to rest of the blocks in the bucket. In this manner, duplicated areas were found. This technique was good at localizing and determining the structure of the tampered patches, and furthermore, direct tile comparison could be completed without sacrificing the execution time. It was also exhibited that this expanding block method was capable in finding some specific forgeries, like JPEG compression, Gaussian blurring, or whenever the cloned regions were made somewhat lighter or darker. This method performed quite well in case of heavily noisy images.

CMFD method developed by Sekeh *et al.* [21] offered improved time complexity with the help of sequential block clustering. A cloned area identification model by applying block clustering followed by two-layer block-matching technique was proposed. This technique needed two kinds of block features: low-accurate, and high-accurate. The low-accurate feature was utilised in the clustering of blocks in their first matching, while the high-accurate feature was implemented in localised block-matching. Clustering resulted in the reduction of searching space for block-matching and improved timing complexity because it eliminated various block-comparing procedures. Whenever the count of clusters was higher than the threshold value, localised block-comparing was much reliable and efficient than lexicographical sorting algorithm.

A robust digital image CMFD method based upon DCT as well as SVD was presented by Zhao *et al.* [22]. The host image was sliced into some fixed-size overlapping tiles and to each of them, two-dimensional DCT was applied. The DCT coefficients were quantized in order to attain a much more robust block presentation succeeded by splitting these quantized tiles in non-overlapping sub-tiles. SVD was implemented on each of these sub-tiles, then features were extracted, so that, it reduced the dimensions of the blocks using their highest singular value. Sets of feature vector were lexicographical sorted, and replicated blocks were discriminated by some pre-calculated shift-frequency threshold. The proposed method could effectively comprehend multiple CMF and precisely locate those similitude regions as well, even when a digital image was falsified by AWGN, Gaussian blur, JPEG compression and also their compound operations.

Kashyap *et al.* [23] proposed a computationally productive algorithm for CMFD of an image using Wavelet Packet Decomposition (WPD) method. The authors performed the wavelet decomposition of the input test image using WPD and tiled it with overlapping grid or block, and for each block, Blur moments invariant representation was done. Then, PCA was used for the lowering of dimensions. The proposed method achieved 86.67 % accuracy in a small processing time. The experimental outcomes showed the high abilities of this CMF method to detect CMF in an image, also when the image was noisy, blurred in the copied areas of that image.

S. Ketenci *et al.* [24] applied two-dimensional Fourier Transform (FT) for feature extraction from each of the sliced blocks of an image. Fourier coefficients of predetermined number signify some information of the blocks. In the last step, similitude search among the nearest feature vectors was done to determine the tampering. The proposed method was capable in detecting the cloned sections with quite high accuracy, also when the image was duped with blurring and/or was JPEG compressed with different quality factors. This method ensured lower feature vectors with high accuracy rates. It decreased the load of computations.

Fattah *et al.* [25] introduced a new scheme based upon a block-matching algorithm. where two-dimensional DWT was performed upon the altered image. Only approximation DWT coefficients were utilized by the authors. To diminish the computational burden on the scheme, some of the candidate blocks were selected on the basis of similarity measure, from the non-overlapping blocks. Then, all of the overlapping

blocks were compared with that candidate blocks. Finally, to identify the tampered patches, a similarity criterion was introduced. Extensive experiments were carried out on various tampered images and it was concluded that this method can reliably and efficiently detect CMF.

2.2 KEYPOINT-BASED CMFD TECHNIQUES

Amerini *et al.* [26] proposed a novel methodology that utilized Scale Invariant Feature Transform (SIFT). It's foremost step consisted of SIFT features extraction along with keypoints matching. It's following step was devoted to the clustering of keypoints and tampering detection, and the next step estimated the occurrence of geometrical attacks, if cloning was detected. This method yielded fine accuracy on various kinds of post-processing operations including JPEG compression, noise, scaling, rotation and exhibited robustness against the compound attacks. Given any suspected image, it could definitively identify if a some patches have been replicated and, additionally, deduce the type of geometric transformation used for performing this forgery. It showed effectiveness with reference to some diverse operative scenarios such as multiple cloning as well as composite processing.

Pan *et al.* [27] developed an area replication method which involved the estimation of transformation matrix between matched SIFT key features which was robust against the distortions depending upon image feature-matching, such as illumination and geometrical distortions. Then all the pixels in the replicated sections were found after reforming the estimated transform matrices. The algorithm resulted in 99.08% average detection accuracy but was limited by the smaller region duplication as it was hard to detect because of fewer keypoints.

In [28], Shivakumar *et al.* proposed a region duplication technique based on Speeded up Robust Feature (SURF). The authors, firstly, extracted keypoints from the test image using SURF followed by the keypoint-matching step. The obtained keypoint pairs were filtered out to obtain matching pairs with a common pattern. This method detected CMF with minimum false positive for high resolution images. The method, also, performed well in the presence of Gaussian noise and geometrical attacks, like scaling and rotation. However, a few small duplicated regions got away undetected.

In [29], Bo *et al.* employed Hessian matrix in keypoint detection and Haar wavelets in orientation assignment. After the estimation of dominant orientation, orientation of the keypoint descriptor was described. The extraction of square regions was done around there keypoints. SURF descriptors in alignment with the dominant orientation were constructed and then, used for matching. This method exhibited robustness to the post-processing operations like scaling, blurring and rotation. However, it failed in detecting exact boundaries of the manipulated regions.

In [30], Hashmi *et al.* introduced a definitively detection method using both DWT as well as SIFT. DWT was utilized for dimension reduction, that further increased the detection results' accuracy. Firstly, DWT was employed on a digital image in order for its decomposition in four sub-bands, namely, LL, LH, HL, and HH. As the LL part holds maximum information, SIFT was employed on this LL part only for extraction of the key attributes and also for finding descriptor vector for each of these key attributes. Resemblance between various descriptors were found to sum up that if the given image was altered. This method allowed authors to detect whether CMF had occurred or not and also localized the cloned regions, that is, it denoted the location of cloned areas for visualization. This method reduced computational burden and improved the timing complexity as compared to other block-based algorithms.

To discern copy-move regions more efficiently, Jaberi *et al.* [31] adopted a powerful keypoint-based feature set, called Mirror Reflection Invariant Feature Transform (MIFT) which share almost all the properties of SIFT key features but MIFT features are invariant to mirror reflection transformations. For diminishing the false positive rate as well as negative rate, the authors used dense MIFT features to find similitude regions in images., rather than standard pixel correlation, in addition to morphological operations and hysteresis thresholding. The proposed method could detect replicated sections in CMF more precisely, especially with the small size of the replicated patches. It showed robustness to transformations that are geometrical in nature. This approach would not work well in the case of flat surface duplicated regions where no interest points could be detected.

Pandey *et al.* [32] presented a keypoint-based detection method using both SURF and SIFT. SURF and SIFT are invariant in response to illumination and geometrical transform. The proposed technique made it very fast as well as robust in detecting copy-

and-moved regions. It was insensitive to geometrical attacks like scaling and rotating. Experimental results demonstrated commendable performance of the technique in image CMFD as compared to other keypoint based techniques.

In [33], Chihaoui *et al.* introduced a method which detected replicated areas in the very same image automatically. Similar regions detection was carried out by discerning the local points of interest in the images using SIFT and by matching of identical features with the help of SVD method. Simulation results showed that this keypoint-based technique was robust to geometrical attacks and was capaable in detecting replicated patches with high performance.

In [34], Hashmi *et al.* proposed a unique algorithm for CMFD which could sustain pre-processing attacks by employing a DYWT and SIFT combination. DYWT usually does not assimilate down-sampling so the image dimensions remained unaffected. In this, firstly, DYWT was performed on an image for decomposing it into four sections, namely, LL, LH, HL, and HH. As the LL section contained maximum information, the authors employed SIFT on that LL section for extracting the key attributes, finding a descriptor vector for these attributes and then finding likeliness between these descriptor vectors to decide if there had been CMF done to that image. By employing DYWT along with SIFT, authors were able to get more key-attributes in number that were matched and hence, identified CMF more efficiently. This algorithm had higher matching rate.

2.3 HYBRID CMFD TECHNIQUES

In [35], Pun *et al.* introduced a hybrid CMFD scheme which integrated both block-based as well as keypoint-based detection algorithms. First, authors applied adaptive over-segmentation algorithm to segment the test image into irregular and non-overlapping patches in an adaptive manner. SIFT was utilized in extracting feature points from each patch and then, were matched to determine labelled key feature points. A copied region extraction method was proposed by the authors too, to determine and locate forgery regions. This technique exhibited better detection precision and recall as compared to the other existing detection methods.

In [36], Li *et al.* segmented a test image into non-overlapping independent patches. SIFT was then used for keypoint detection and extraction on each patch. K-d tree as constructed and K-nearest neighbor search was performed for each extracted keypoint.

Suspected regions were filtered out using a threshold value. This step was known as Patch matching. Affine transform matrix was estimated for these suspected regions only. This estimated affine transform matrix was then refined using Expectation-Maximization based algorithm to confirm CMF in the image. The scheme faced a drawback of high false alarms rate.

Ardizzone *et al.* [37] compared triangles instead of blocks, or keypoints. Keypoints were extracted from an image. Objects were modelled as a group of adjoined triangles built onto these interest points. With respect to the inner angles, color information and local feature vectors, triangle matching was done to detect similar regions. This scheme exhibited better performance in comparison to the state-of-the-art block-based and keypoint-based detection schemes even in the presence of geometrical transformations like scaling and rotation. However, the performance was affected in case of images having complex scenes due to the detection of large number of triangles.

Table 2.1 gives comparison of various CMFD algorithms.

Table 2.1: Survey of CMFD Methods

S.No.	Paper	Technique Used	Observations
1	Fridrich *et al.* [5]	DCT	Copy-move region is detected. Does not work well for noisy images.
2	Muhammad *et al.* [9]	DYWT	High accuracy and efficiency.
3	Ghorbani *et al.* [10]	DCT-DWT	Does not perform well in case of highly compressed images.
4	Bashar *et al.* [15]	DWT-KPCA	Remarkable performance under additive noise and JPEG compression.
5	Cao *et al.* [17]	DCT	Exact copy-move region detected. Works well if the image is noisy or blurred.

6	Nguyen *et al.* [18]	Radon Transformation	Performs well when forged region is rotated less than 4^{o}.
7	Wang *et al*. [19]	DWT	Works efficiently in case of heavily distorted images.
8	Sekeh *et al*. [21]	Sequential Block Clustering	Less time complexity.
9	Zhao *et al.* [22]	DCT-SVD	Effectively detects and locates duplicated regions. Robust to AWGN, JPEG compression, Gaussian blurring and their compound operations.
10	Kashyap *et al.* [23]	WPD	Accuracy of 86.67% within a small processing time. Less computational complexity.
11	Ketenci *et al.* [24]	Two-dimensional FT	Work detects multiple CMF.
12	Fattah *et al.* [25]	DWT	Efficiently detect CMF.
13	Amerini *et al*. [26]	SIFT	Robust to compound image processing.
14	Pan *et al*. [27]	SIFT	Smaller region duplication is hard to detect.
15	Hashmi *et al.* [30]	DWT-SIFT	Visually locates where the CMF has occurred.
16	Jaberi *et al.* [31]	MIFT	Works accurately for the small size of the replicated regions. Fails when the cloned regions are flat in nature where no keypoints can be detected.

17	Pandey *et al.* [32]	SURF-SIFT	Technique is invariant to geometrical and illumination transforms. Quick and robust in detecting duplicated regions.
18	Chihaoui *et al.* [33]	SIFT-SVD	Method is robust to geometrical transformations.
19	Hashmi *et al.* [34]	DYWT-SIFT	Robust against many pre-processing techniques
20	Pun *et al.* [35]	Over-segmentation and SIFT	Excellent performance against geometrical attacks: rotation, scaling, down-sampling and JPEG compression
21	Li *et al.* [36]	Expectation-Maximization-based algorithm	Performance deteriorates in case of highly textured images.

2.4 MOTIVATION

In recent years, the need for authentication of digital images has risen due to advances in image editing tools. Falsifiers manipulate images with wrong intentions of providing maligned information, or hiding some useful particulars. The simplest form of tampering is CMF. Many authors have proposed different techniques to detect CMF. Most of the existing block-based algorithms, in fact, deal with a large number of blocks and their feature vectors, which leads to high computational burden. Although these methods are invariant to noise addition, JPEG compressing and filtering (depending upon the selected features), yet they are susceptible to geometrical attacks like rotation, scaling and translation. Due to the regularly shaped blocks, these techniques give very low recall rate. On the other hand, keypoint-based methods are computationally efficient as well as perform well under geometrical attacks as compared to block-based techniques. But, the existing keypoint-based schemes too suffered from low recall rate. The need to overcome the shortcomings of these traditional methods has motivated in developing an adaptive

hybrid technique which is not only computationally efficient but can identify tampered images with high recall rate even under different challenging conditions like down-sampling, scaling and JPEG compression.

2.5 OBJECTIVES OF THESIS

The research achieves following objectives:

- To prepare an efficient adaptive CMFD algorithm and check its performance with parameters precision, recall and F1 score.
- To evaluate the proposed algorithm in the presence of JPEG compression and geometrical distortions like down-sampling, up-scaling and down-scaling.
- Comparison of proposed algorithm with existing ones and minimization of high false positives.

CHAPTER 3

DIGITAL IMAGE FORENSICS TOOLS

3.1 INTRODUCTION

As the advanced technologies are used in image tampering, the determination of integrity and credibility pose a real challenge to both the naked human eyes and machines. Hence, it is the need of the hour to develop robust tampering detection algorithms for the validating authenticity of digital images and identification of tampering operations. The existing digital image forensics techniques can be categorized into active and passive (blind) schemes.

3.2 ACTIVE AUTHENTICATION

In active algorithms, a digital signature or a digital watermark is embedded into an original digital image using some sort of embedding process. The embedded content can be extracted and further, used to examine the authenticity of the image. But these techniques suffer from a strong limitation that the watermark embedding process must be carried out by an authorized person or by the acquisition device (camera), which is impractical considering the fact that most image acquisition devices lack watermarking capabilities [38]. Active techniques work only in the presence of some prior information about the image. Therefore, such methods are inappropriate when we need to investigate images from some unknown or unreliable sources. Moreover, watermarking processes degrade the image quality considerably.

3.3 PASSIVE AUTHENTICATION

In passive approaches, also known as blind techniques, the source image is unavailable. There is no such need of any priori information of the image as in active techniques. These techniques are developed by assuming that the tampering processes inculcate artifacts in the forged images due to the alteration in their basic statistical properties, irrespective of the fact that the manipulation is visually indiscernible. Such inconsistencies are exploited to investigate forgery.

During tampering, an image can be subjected to different forms of attacks. The simplest of all the attacks is CMF, which involves the duplication of a certain region (or regions) within the image. Replication of image areas are, sometimes, done from other digital images as in image splicing. The tampered regions or the image itself can undergo various transformations to conceal the forgery. The digital forensics tools emphasize on identifying the forgery by utilizing different characteristics of image forgery operations.

The different passive image forgery detection algorithms can be grouped majorly under five classes i.e., pixel-based, camera-based, compression-based, geometric-based and physics-based schemes. Fig. 3.1 shows the different techniques used under these five classes.

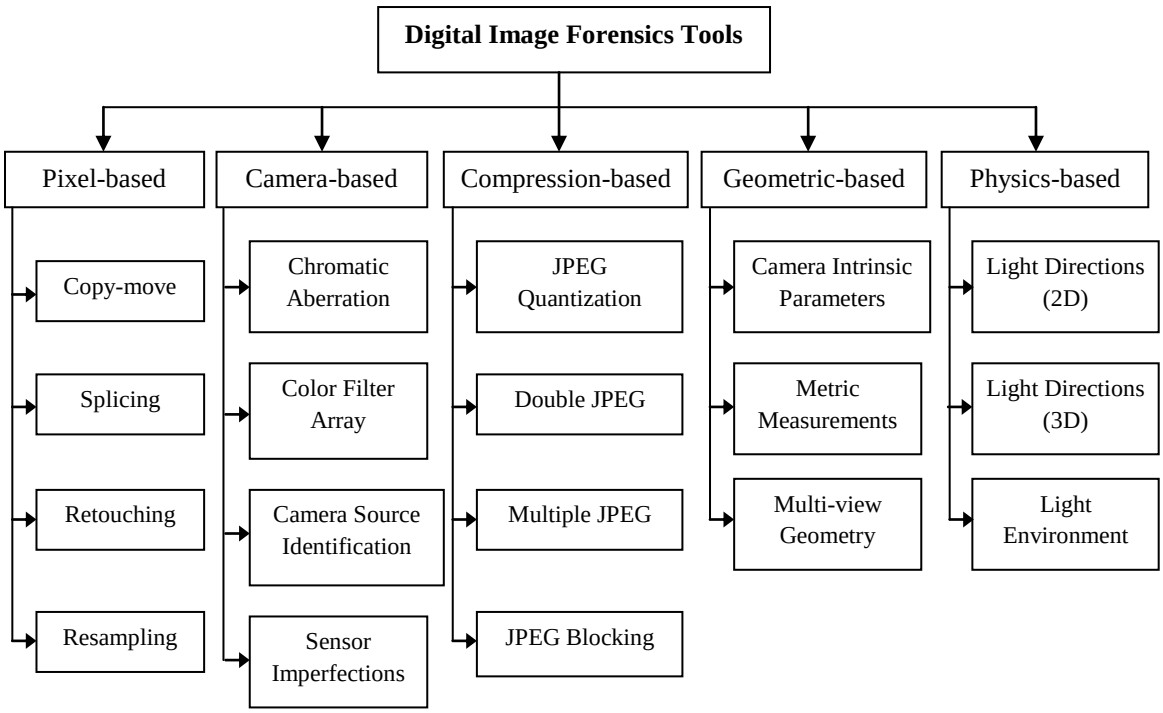


Fig. 3.1 Types of digital image forensics tools. [39]

3.3.1 Pixel-Based Forgery Detection Techniques

These schemes exploit the statistical deformities that occur in image pixels due to tampering for detection. Such schemes analyze inter-pixel correlations that develop due to a certain type of manipulation in either spatial or transformed domain. The most common and popular pixel-based algorithms are copy-move, image splicing, retouching, and resampling.

(a) CMFD techniques

Since, one or more sections is/are copied and pasted within the image in CMF (as shown in Fig. 3.2), a strong correlation exists in the similar regions which forms the basis of most of the existing CMFD algorithms. But the real challenge lies in finding effective features and matching algorithms to detect correlated sections.

As discussed in literature survey, most of the existing methods extract features either by dividing an image into overlapping tiles or by locating interest points in the entire image. Appropriate feature matching algorithm is, then, applied to find blocks (or keypoints) with similar features. Forgery regions are localized for visual inspection by representing matched regions corresponding to matched blocks (or keypoints) in different colors.

Fig 3.2 An example of CMF: original image (left) and forged image (right). [4]

(b) Image splicing detection techniques

Image splicing involves composition of two or more digital images and altering the authentic image remarkably resulting in a tampered one as shown in Fig. 3.3. If images having similar backgrounds are merged, then, it becomes difficult to identify their borders. Image splicing detection becomes quite challenging in such cases. However, the presence of abrupt and irregular variations between different combined sections along with their backgrounds, provide important cues for splicing detection in a given image.

A typical splicing scheme begins with a pre-processing stage which usually involves a conversion from RGB to gray-scale. Different features are obtained from both original and forged images in feature extraction stage. The obtained features are employed in training a classifier and then, this trained model is utilized in classification of the original and forged images. The localization of detected tampered regions is done in the post-processing stage.

(a) (b) (c)

Fig. 3.3 An example of image splicing: authentic images (a, b) and tampered image (c). [40]

(c) Image retouching detection techniques

Image retouching is one of the most commonly used image forgery tool for aesthetic and commercial uses. This process is mostly carried out to reduce or enhance image features as in Fig. 3.4. Since last few years, image retouching tools such as Adobe Photoshop, etc., are employed so that the images appear more natural and attractive. Images are investigated for retouching by finding the enhancements, blurring, illumination changes and color changes, if any. Retouching detection becomes easy with the availability of the original image, however, passive detection is a complex task.

(a) (b)

Fig. 3.4 An example of retouching: (a) original image and (b) retouched image. [41]

(d) Image resampling detection techniques

To create convincing tampered images, the targeted image portions are subjected to geometrical attacks like stretching, scaling, rotation, flip-ping, and skewing which require

resampling operation. Interpolation is the core step in resampling that results in statistical alterations in the forged image. The cues left by this step are exploited in forgery detection.

3.3.2 Camera-Based Forgery Detection Techniques

When an image is captured by any digital camera, it is subjected to a series of operations, which include quantization, correlation of color, gamma correction, filtering, white balancing, and JPEG compression, while moving from the camera sensor to the memory. These processing steps may vary from camera to camera. Different artifacts are introduced by the different stages of the image creation step. Chromatic aberration [42], color filter array [43], source camera identification [44], and sensor noise imperfections [45] are employed to evaluate camera artifacts. The camera-based forgery detection can be done by taking into account the inconsistencies in these artifacts [46].

3.3.3 Compression-Based Forgery Detection Techniques

Sometimes, forged images are compressed to make tampering detection a very complex task. JPEG is the most popular and commonly used compression standard used in variety of applications. To investigate whether BMP images have been JPEG compressed previously or not is a matter of concern for some image processing applications. In image forensics analysis, some JPEG compression properties are utilised for detection of the cues left by tampering. Such schemes can be classified into JPEG quantization based [47], double JPEG compression based [48], multiple JPEG compression based [49], and JPEG blocking based [50].

3.3.4 Geometric-Based Forgery Detection Techniques

These detection algorithms utilize geometrical constraints from perspective outlooks. Such schemes can be classified on the basis of camera's basic parameters (like principal point, focal length, skew, and aspect ratio) [51], metric measurement [52], and multiple view geometry [53]. For instance, in original image, the intersection of optical axis with image plane i.e., principal point is situated near the image's centre. Whenever a small section is translated or moved within the image, or two or more images are merged, it becomes challenge for the falsifier to keep the principal point of the tampered image in its

correct perspective. Therefore, employing projective geometry principles can prove helpful in developing efficient and effective tampering detection schemes.

3.3.5 Physics-Based Forgery Detection Techniques

Different photos are captured under varying lighting conditions. Whenever two or more images are combined, there is a huge possibility of having their lighting conditions unmatched. In physics-based forgery detection schemes, the lighting inconsistencies in light sources can be used to identify tampering [54].

3.4 SUMMARY

Digital image forgery detection algorithms are classified into active and passive schemes. Further, passive algorithms have been sub-categorized into five groups. The crux of detection techniques for different forgeries is the corresponding artifacts induced by them. However, the pixel-based forgery detection schemes are the simplest and most commonly used techniques.

CHAPTER 4

RESEARCH METHODOLOGY

4.1 INTRODUCTION

The proposed method integrates both block-based as well as keypoint-based detection techniques. This technique exploits the ability of adaptive segmentation and MIFT to detect similar regions in a tampered image. Framework of the proposed scheme is illustrated in Fig. 4.1. It consists of five stages: adaptive image-segmentation, feature extraction and description, patch feature matching, adaptive keypoint matching and adaptive forgery region extraction. The following sections present detailed description of each stage.

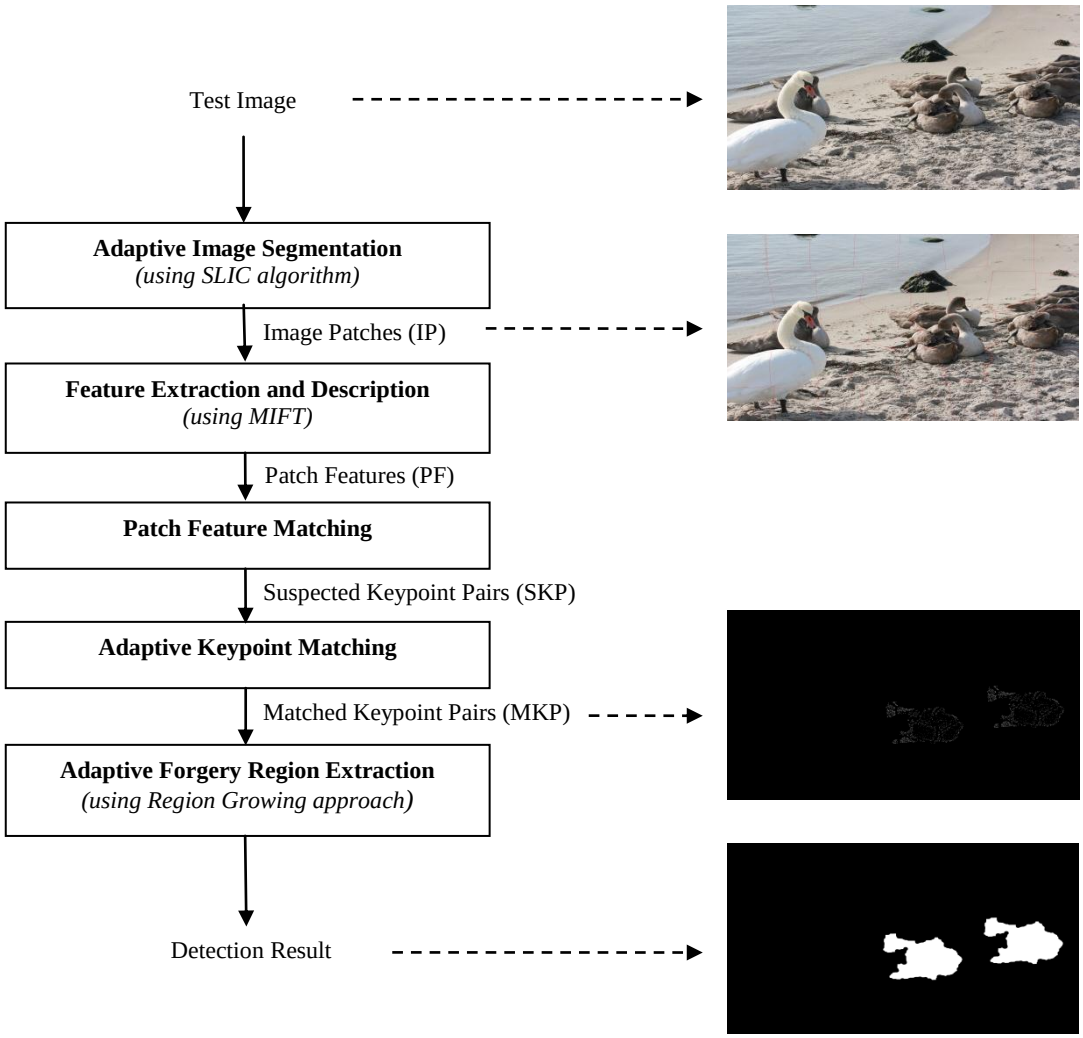


Fig. 4.1 Framework of the proposed CMFD algorithm.

4.2 ADAPTIVE IMAGE SEGMENTATION

The first and the foremost step of the proposed CMFD is to segment the test image in order to get semantically independent and irregular shaped patches. This is performed with the assumption that the duplicated regions will be segmented in a similar manner.

In the proposed algorithm, the Simple Linear Iterative Clustering (SLIC) algorithm [55] has been utilised to achieve digital image segmentation as it is fast, memory efficient and adheres to boundaries very well. This technique employed a k-means clustering approach to generate meaningful clusters of pixels, known as superpixels. Since these superpixels are irregular in shape, they can depict replicated regions better than the fixed-shape block representation.

SLIC algorithm selects cluster centres from the sampled uniform grid spaced S_r as given in (4.1).

$$S_r = \sqrt{N_p / N} \tag{4.1}$$

where N_p represents total image pixels and N signifies the desired number of the superpixels.

SLIC algorithm finds the similar pixels from each cluster centre to pixels for clustering within a $2S_r \times 2S_r$ neighborhood instead of in the entire image to generate superpixels quickly. SLIC integrates color proximity with spatial distances in a single distance measure D_m given in (4.2).

$$D_m = \sqrt{d_c^{\,2} + \left(\frac{d_s}{S_r}\right)^2 m_w^{\,2}} \tag{4.2}$$

where m_w is the weighting factor between color proximity, d_c and spatial proximity, d_s. d_c and d_s are defined in (4.3) and (4.4), respectively.

$$d_c = \sqrt{\left(l_u - l_v\right)^2 + \left(a_u - a_v\right)^2 + \left(b_u - b_v\right)^2} \tag{4.3}$$

$$d_s = \sqrt{\left(x_u - x_v\right)^2 + \left(y_u - y_v\right)^2} \tag{4.4}$$

Here, $[l, a, b]^T$ is the color representation of a pixel in the CIELAB color space, while $[x, y]^T$ represents the position of a pixel. The challenge faced in using SLIC algorithm is the setting initial size, S_r of the superpixels. Choosing a proper value for S_r is crucial as the forgery detection results are affected by it. A large initial size of superpixels will lead to inaccuracy in detection results accompanied with reduced recall rate, while a small IP size will increase computational burden. Hence, it is desired to choose an optimum value of S_r to get good detection results with less computational load.

Extensive experiments have been carried out to seek a relation between the initial size of the IP and the dimensions of the test images in order to get good CMFD results. For the initial patch size computation, firstly, approximate superpixels size, S_r' , has been calculated using (4.5).

$$S_r' = p_1 X^3 + p_2 X^2 + p_3 X + p_4 \tag{4.5}$$

Here, X is the total pixels in the host image while p_1, p_2, p_3 and p_4 are the coefficients. For $X \geq 22000$, the coefficient values are $p_1 = 3.6545 \times 10^{-19}$, $p_2 = -8.9288 \times 10^{-12}$, $p_3 = 9.0126 \times 10^{-5}$ and $p_4 = 74.4880$. If $X < 22000$, then the coefficients are set as $p_1 = p_2 = 0$, $p_3 = 0.0020$ and $p_4 = 7.8657 \times 10^{-16}$. The initial patch size, S_r is taken as the rounded value of S_r'. Fig. 4.2 gives the framework of Adaptive Image Segmentation algorithm.

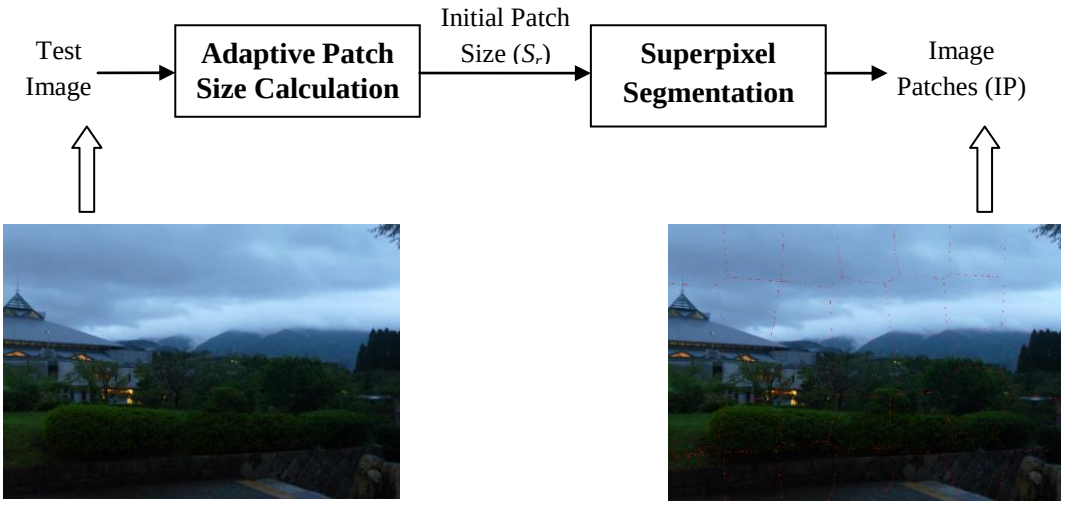


Fig. 4.2 Framework of Adaptive Image Segmentation method.

4.3 FEATURE EXTRACTION AND DESCRIPTION

The objective of feature extraction methods is to detect interest points, also called keypoints, along with their distinctive descriptors. These descriptors are desired to be invariant to common image distortions such as scale, translation, and JPEG compression. Widely used keypoint extraction methods include SIFT [56] and SURF [57] as they are robust to common image handling operations scaling, rotation, compression and blurring. SIFT has outstanding stability, robustness and distinctiveness. However, MIFT [58] has been employed which is an improvement to SIFT. MIFT features have almost all properties of SIFT features but are invariant to mirror reflection transformations.

Keypoint detection involves identification of locations along with scales that can be attributed under varying views of a patch repeatedly [56]. The scale space of an IP is described as a function, $F(a,b,\sigma)$. $F(a,b,\sigma)$ is the convolution of $G_s(a,b,\sigma)$, a variable-scale Gaussian and $I_p(a,b)$, the IP, as shown in (4.6).

$$F(a,b,\sigma) = G_s(a,b,\sigma) \otimes I_p(a,b) \tag{4.6}$$

where $\otimes$ is the convolution operation in *a* and *b*, which represent pixel coordinates, σ is the scale parameter and

$$G_s(a,b,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(a^2+b^2)/2\sigma^2}.$$

The difference-of-Gaussian function convolved with IP, $D_g(a,b,\sigma)$ is used for keypoint location detection. It is calculated using the differences in two scales separated by a constant multiplicative factor *m* as in (4.7).

$$\begin{aligned} D_g(a,b,\sigma) &= (G_s(a,b,m\sigma) - G_s(a,b,\sigma)) \otimes I_p(a,b) \\ &= F(a,b,m\sigma) - F(a,b,\sigma) \end{aligned} \tag{4.7}$$

MIFT rearranges the order of the cells and reorganises the orientation bins in each cell as depicted in Fig. 4.3. MIFT descriptors have been used in this work to determine replicated regions without mirror reflections as MIFT performs comparably well as SIFT in such cases too [58]. MIFT is applied on each IP obtained after the application of Adaptive Image Segmentation to extract features points and their descriptors, known as Patch Features (PF).

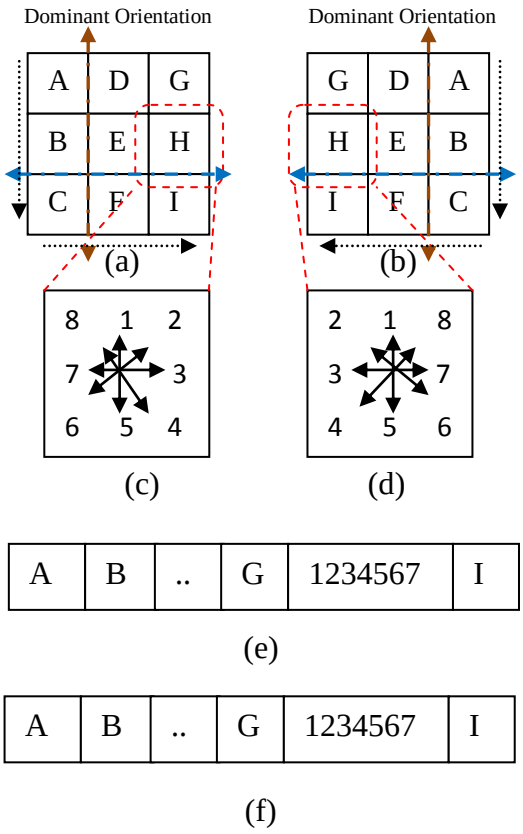


Fig. 4.3 Example of the MIFT descriptor organization with and without mirror reflection: (a) Interest region of a keypoint without mirror reflection; (b) horizontal reflection of (a); (c) 8 orientations distribution in the H cell of (a); (d) 8 orientations distribution in the H cell of (b); (e) MIFT descriptor of (a); (f) MIFT descriptor of (b).

4.4 PATCH FEATURE MATCHING

Next, the pairs of suspicious patches with many similar keypoint descriptors are checked. This process involves the comparison features of each patch with the rest. For P number of IP, $L = P(P-1)/2$ comparisons are required. For each keypoint in every patch, we look for K nearest neighbors in other IP. For a keypoint, complexity of finding K nearest neighbors is decreased to $O(nlogn)$ from $O(n^2)$, by using a k-d tree [36]. Here, $K = 3$ has been set taking into account that there can be multiple pairs of similar regions in the image. The two keypoints are considered to be Suspected Keypoint Pair (SKP) only if the L-2 norm of the difference between their descriptors is smaller than or equal to a threshold (0.15 in the implementation). A pair of patches is said to be suspected patch pair if it contains at least one SKP.

4.5 ADAPTIVE KEYPOINT MATCHING

To detect matched keypoint pairs (MKP) from SKP, an Adaptive Keypoint Matching method has been used. A patch keypoint matching threshold, P_K is calculated adaptively as in (4.9). Using this threshold value, undesired keypoint pairs are eliminated from SKP in order to yield MKP.

$$P_K = ceil(2 \times \frac{T_{SP}}{T_P}) \tag{4.9}$$

where T_{SP} is total number of SKP and T_P is total number of suspected patch pairs.

Using the location distance between SKP, their clusters are formed. SKP with similar rounded location distance values are grouped together. If the size of a cluster is less than or equal to P_K, then that cluster is eliminated. Now, a suspected patch pair with SKP less than or equal to P_K is removed. These two steps are repeated till no change in the number of clusters is observed. Those keypoint pairs which belong to these clusters are considered to be MKP and their respective suspected patch pairs are termed as matched patch pairs.

4.6 ADAPTIVE FORGERY REGION EXTRACTION

After the extraction of MKP, there is need to locate the duplicated regions, if any, in the test image. To attain this, an Adaptive Forgery Region Extraction algorithm has been proposed which is based on region growing. Fig. 4.4 demonstrates the flow diagram of Adaptive Forgery Region Extraction method.

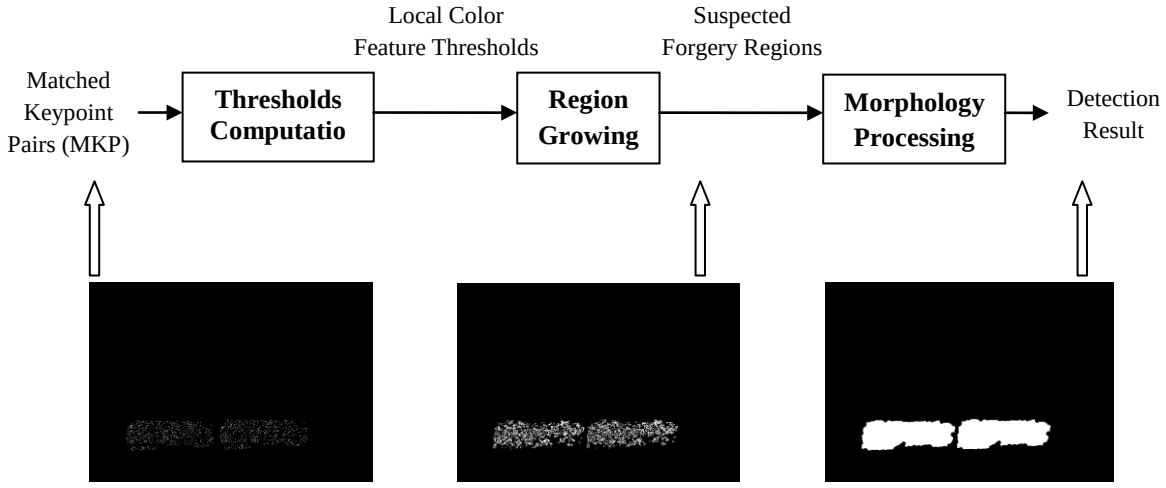


Fig. 4.4 Flow diagram of Adaptive Forgery Region Extraction method.

This approach to segmentation involves the examination of neighborhood pixels of initial seed points and decides whether these neighboring pixels should be added to the region. The criteria for adding neighboring pixels is defined using the local color feature of the neighboring pixels of 8-connectivity and the thresholds used in this case are calculated adaptively.

For each matched patch pair, the mean color intensities of test image pixels corresponding to MKP are calculated and are termed as *TR*, *TG* and *TB* corresponding to red, green and blue, respectively. Then region growing approach is applied by using MKP locations as seed points. Start by choosing any MKP as initial seed pixels. Compare the seed points with neighboring pixels. This region growing algorithm is subjected to constraints given in equation (4.10).

$$\begin{aligned} 0.45 \times TR \le T_{NR} \le 1.40 \times TR \\ 0.45 \times TG \le T_{GR} \le 1.40 \times TG \\ 0.45 \times TB \le T_{BR} \le 1.40 \times TB \\ 0.55 \times TM \le T_{NM} \le 1.30 \times TM \end{aligned} \tag{4.10}$$

Here, T_{NR}, T_{GR}, and T_{BR} are the red, green and blue intensities of neighboring pixels, respectively. *TM* represents the mean intensity of *TR*, *TG* and *TB* while T_{MN} signifies the mean of T_{NR}, T_{GR} and T_{BR}. From the seed pixel, region is grown by adding neighboring pixels to it that satisfy these constraints. Choose another MKP, which does not yet belong to any region, as seed pixels when the growth of one region stops and start again. Regions obtained at the end of region growing algorithm are referred as suspected forgery regions.

Finally, morphology processing is done. Close operation is applied to suspected forgery regions in order to close these regions while keeping their shapes intact. A circle of radius related to the test image has been used as the structuring element for close operation. After closing the regions, the holes are filled using fill morphological operation to get the final output.

4.7 SUMMARY

A hybrid algorithm to find CMFD based on adaptive method has been proposed. The Adaptive Image Segmentation algorithm segments the test image into irregularly shaped and non-overlapping IP. The initial size of these patches is determined adaptively, according to the test image, which increases the detection accuracy of the forgery, along

with the reduction in computational load. Feature are extracted as PF, followed by the extraction of SKP using the Patch Feature Matching. MKP are obtained using the Adaptive Keypoint Matching algorithm. Forgery regions are extracted using the proposed Adaptive Forgery region algorithm based on the region growing approach.

CHAPTER 5

EXPERIMENTAL RESULTS

This chapter is divided into four sections, namely test image database, performance characteristics, simulation results under plain copy-move attack and simulation results under different attacks.

5.1 IMAGE DATABASE

The dataset [4] is used to examine the effectiveness of the proposed method through series of experiments. It has 48 base images and 87 duplicated snippets which are used to make the forgeries. The original images vary from 800×533 to 2613×3900 in size. The duplicated regions are from various classes of man-made, nature, living, and mixed, with wide range of intensities varying from extremely smooth to overly textured. Therefore, this dataset has been chosen to objectively evaluate our scheme. We have tested 1488 images in total under different conditions for the performance evaluation.

5.2 PERFORMANCE CHARACTERISTICS

The performance and efficiency of the proposed adaptive CMFD algorithm has been evaluated at image level as well as pixel level. The performance characteristics used for the same are precision, *P* and recall, *R*.

At image level, *P* is defined as the ratio of rightly identified tampered images to the total number of images identified as tampered as in (5.1). In other words, it signifies the probability of detecting a true forgery. *R* is defined as the ratio of rightly identified tampered images to the total number of ground-truth tampered images as in (5.2). It signifies the probability of detecting a forged image. At pixel level, *P* is the ratio of total number of rightly detected replicated pixels to the sum total of detected replicated pixels while *R* is the ratio of total number of rightly detected replicated pixels to the sum total of replicated pixels in the ground-truth tampered images.

$$P = \frac{\left|D_f \cap G_f\right|}{\left|D_f\right|} \tag{5.1}$$

$$R = \frac{|D_f \cap G_f|}{|G_f|} \tag{5.2}$$

Here D_f and G_f are detected forgery pixels using the proposed scheme and ground-truth forgery pixels from the dataset, respectively. Average P and R are computed in order to lessen the impact of the randomness of the samples.

The pixel level metrics are used for examining the localization performance of the method with respect to the available ground truth data. However, the image level metrics signify the automated detection of manipulated images.

Another performance criterion F_1 score has been computed that combines both precision and recall as given in (5.3).

$$F_1 = 2 \times \frac{P \times R}{P + R} \tag{5.3}$$

In general, higher value of P, R and F_1 score signifies superior performance.

5.3 SIMULATION RESULTS UNDER PLAIN CMF

Several prior arts of block-based [11]-[12], [16], keypoint-based [26]-[29] and hybrid [35] forgery detection schemes have been chosen for comparison with our proposed method. All these existing techniques are discussed in Chapter 2. Firstly, the effciency of the proposed CMFD scheme has been examined by detecting plain CMF, that is without any attack or transformation. For this experiment, total 96 images have been taken of which 48 images are original while the rest of the 48 are plain copy-move forged images.

Tables I and II depict the forgery detection results under plain CMF at image level and pixel level, respectively. Table I clearly shows that our proposed adaptive scheme outperforms prior arts at image level with *P=100%* and *R=100%;* thus F_1*=100%*. At pixel level, *P=95.01%* and *R=87.18%;* thus F_1 *= 90.92%* is achieved by our scheme which is better than the prior arts as shown in Table II.

One may concern the scenario when the two replicated regions, with size smaller than initial patch size, S_r (as defined in Section 4.1), are quite close or connected. However, the possibility of the occurrence of this situation is too low that it can be ignored.

Moreover, there are no two copied regions that occur within a single irregular region in any image from the dataset used.

Table 5.1: CMFD Results Under Plain CMF At Image Level

Methods	**Precision, *P* (%)**	**Recall, *R* (%)**	**F_1 (%)**
Wang [11,12]	92.31	100	96.00
Bravo [16]	87.27	100	93.20
SIFT [26,27]	88.37	79.17	83.52
SURF [28,29]	91.49	89.58	90.53
Pun [35]	96	100	97.96
Proposed Scheme	**100**	**100**	**100**

Table 5.2: CMFD Results Under Plain CMF At Pixel Level

Methods	**Precision, *P* (%)**	**Recall, *R* (%)**	**F_1 (%)**
Wang [11,12]	98.69	85.44	90.92
Bravo [16]	98.81	82.98	89.34
SIFT [26,27]	60.80	71.48	63.10
SURF [28,29]	68.13	76.43	69.54
Pun [35]	97.22	83.73	89.97
Proposed Scheme	**95.01**	**87.18**	**90.92**

(a) Cattle

(b) Red Tower

(c) No Beach

(d) Christmas Hedge

(e) Bricks

(f) Statue

Fig. 5.1 Different original images.

(a) Cattle

(b) Red Tower

(c) No Beach

(d) Christmas Hedge

(e) Bricks

(f) Statue

Fig. 5.2 Different tampered images for CMFD.

(a) Segmented image of Cattle

(b) Segmented image of Red Tower

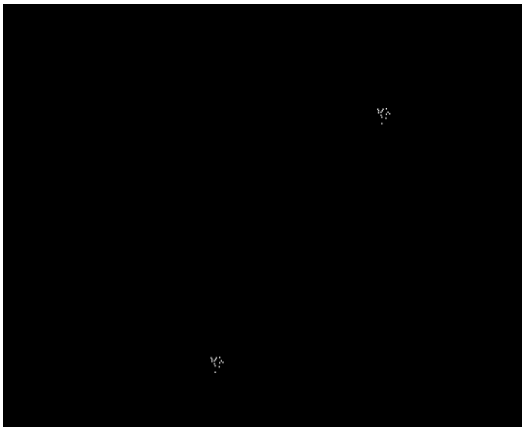
(c) Segmented image of No Beach

(d) Segmented image of Christmas Hedge

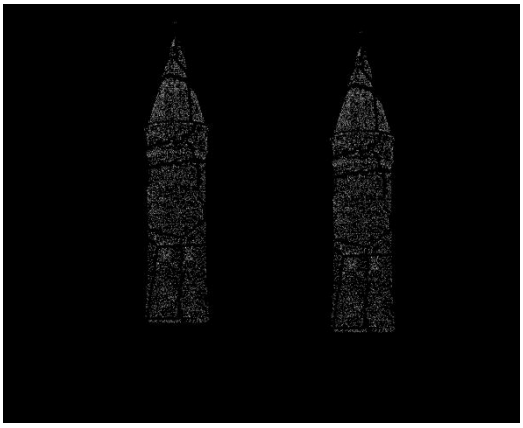
(e) Segmented image of Bricks

(f) Segmented image of Statue

Fig. 5.3 Different segmented images for CMFD.

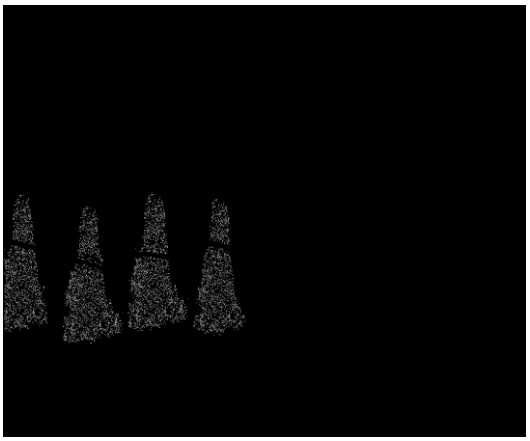


(a) Matched keypoints pairs of Cattle

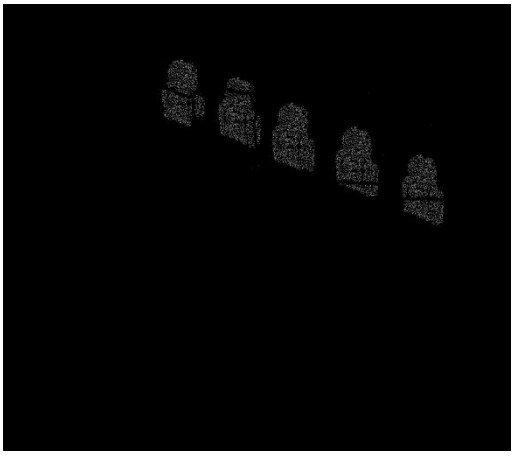


(b) Matched keypoints pairs of Red Tower

(c) Matched keypoints pairs of No Beach



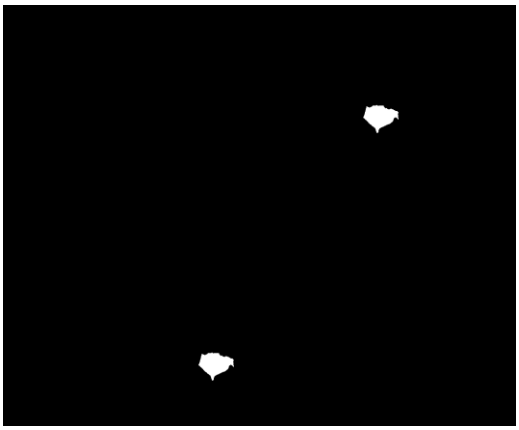
(d) Matched keypoints pairs of Christmas Hedge



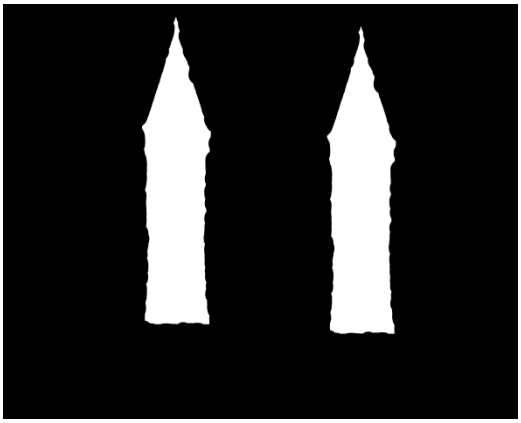
(e) Matched keypoints pairs of Bricks

(f) Matched keypoints pairs of Statue

Fig. 5.4 Matched keypoint pairs of different tampered images for CMFD.



(a) Detected forgery regions of Cattle



(b) Detected forgery regions of Red Tower

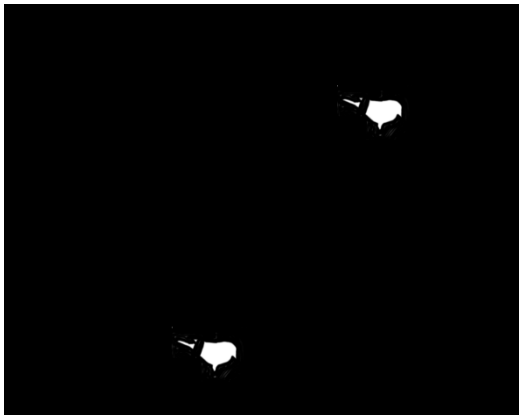
(c) Detected forgery regions of No Beach

(d) Detected forgery regions of Christmas Hedge

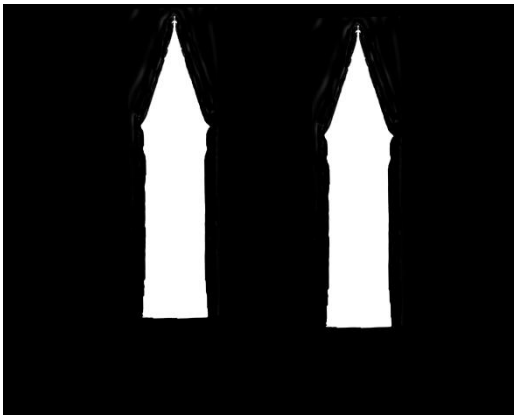
(e) Detected forgery regions of Bricks

(f) Detected forgery regions of Statue

Fig. 5.5 CMFD results of different tampered images.

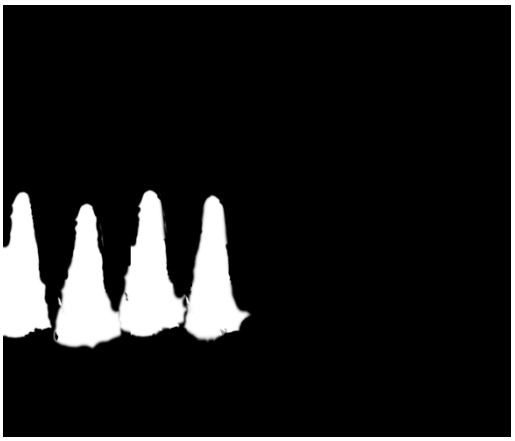


(a) Ground truth image of Cattle



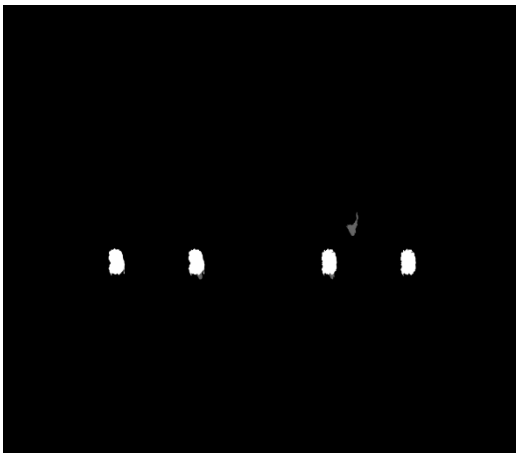
(b) Ground truth image of Red Tower

(c) Ground truth image of No Beach



(d) Ground truth image of Christmas Hedge

(e) Ground truth image of Bricks



(f) Ground truth image of Statue

Fig. 5.6 Different ground truth images for the comparison with detected CMF results

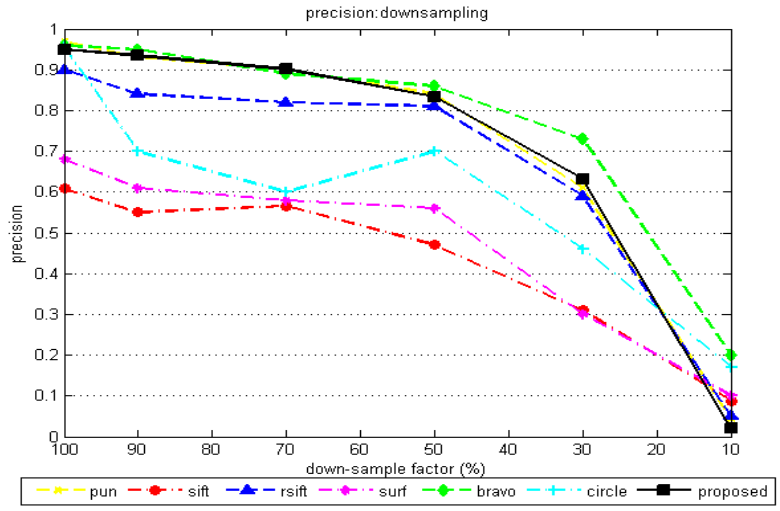
Fig. 5.2 shows the different copy-move forged images corresponding to the original images represented in Fig. 5.1. Fig. 5.3 represents the results of the Adaptive Image Segmentation step of the proposed hybrid technique. Matched Keypoints Pairs (MKP) are indicated in Fig. 5.4. These MKP are used to extract forgery regions which are represented in Fig. 5.5. For the evaluation of detected forgery results, ground truth images in Fig. 5.6 are used.

5.4 SIMULATION RESULTS UNDER DIFFERENT ATTACKS

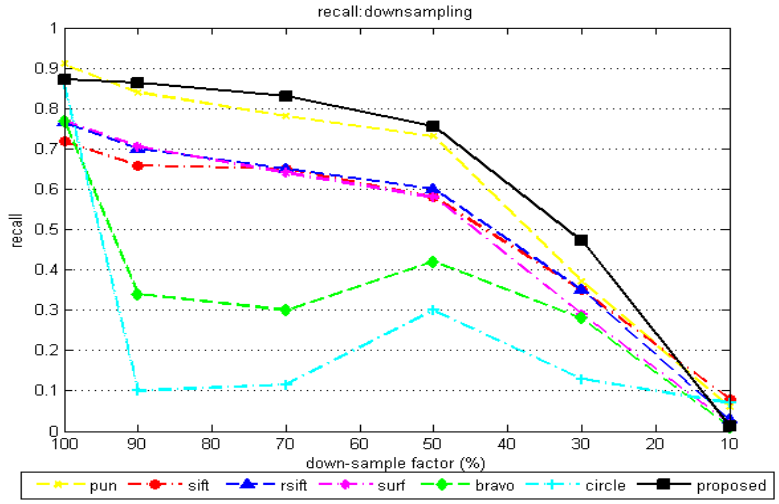
To evaluate the robustness of the technique, the proposed detection algorithm has been evaluated when the duplicated regions are maligned by different attacks in addition to the plain CMF. For this experiment, the tampered images are subjected to geometric transformations that include up-scaling, down-scaling and post-processing operations such as down-sampling and JPEG compression.

In case of down-sampling, 48 of the dataset's tampered host images are scaled down in steps of 20% from 90% to 10% . For up-scaling, the scale factor of duplicated regions is varied in steps of 2%, from 91% to 99%, whereas, for down-scaling, the scale factor ranging from 101% to 109%. For the case of JPEG compression, forged images are compressed using lossy JPEG format with quality factor varied in steps of −10, varying from 100 to 20.

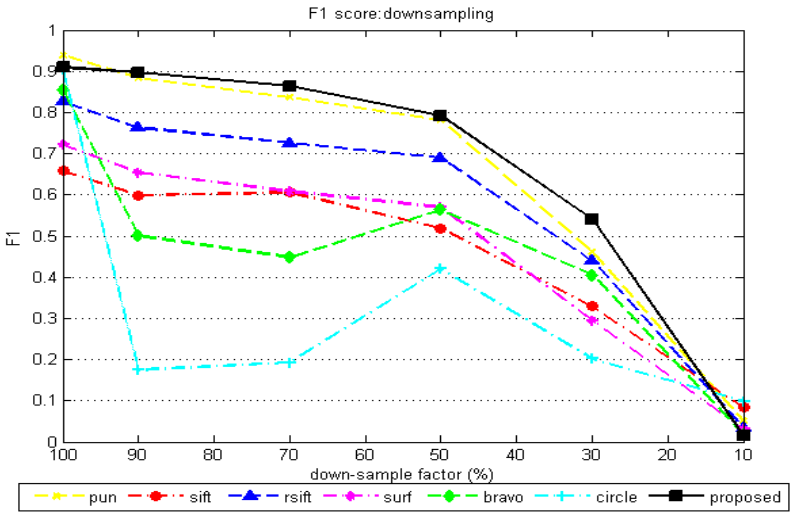
Figs. 5.7, 5.8, 5.9 and 5.10 represent the pixel level CMFD results under various distortions, down-sampling, up-scaling, down-scaling, and JPEG compression, respectively. The results depicted in black and marked ‘proposed’ represent the results obtained by the proposed scheme. The results shown in yellow and marked ‘Pun’ and the results indicated in blue and marked 'RSIFT' depict the results of the hybrid algorithm in [35] with adaptive-size blocking and fixed-size blocking, respectively. The results indicated in red and pink and marked as ‘SIFT’ and ‘SURF’ show the results of the keypoint-based schemes based on SIFT [26], [27] and SURF [28], [29], respectively, while the results in green and sky-blue and marked as ‘Bravo’ and ‘Circle’ depict the results obtained by the block-based schemes proposed by Wang *et al.* [11], [12] and by Bravo-Solorio *et al.* [16], respectively.



(a)

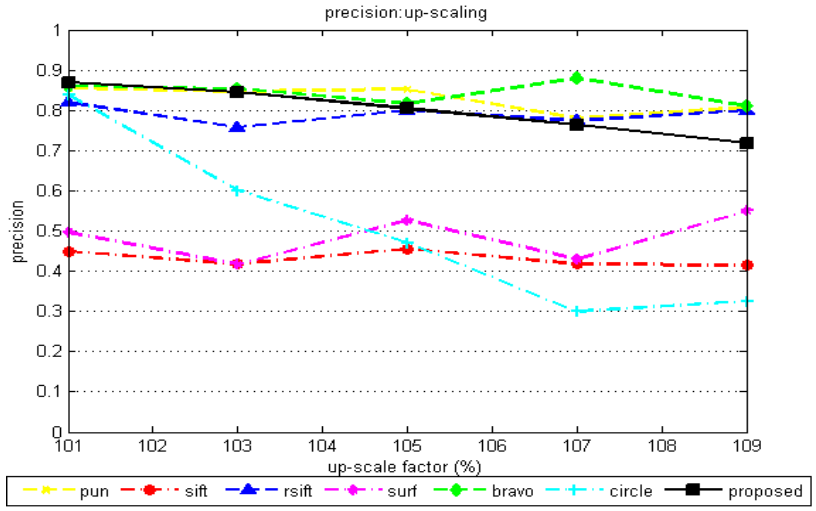


(b)

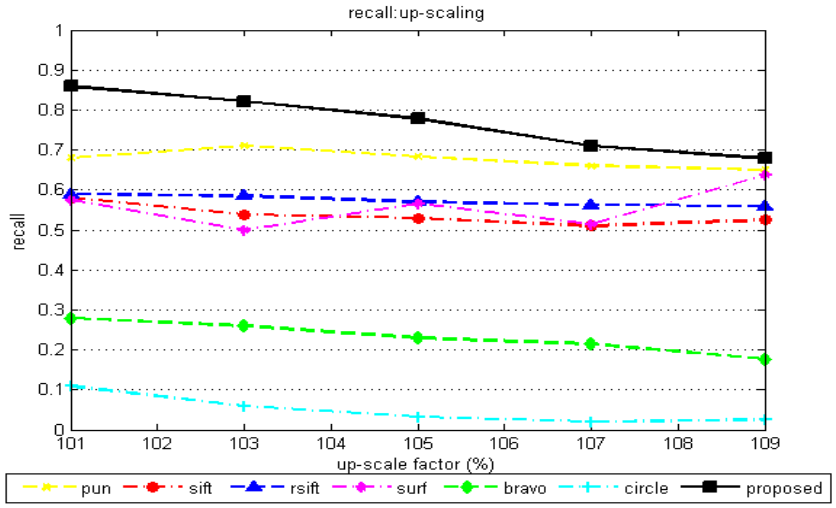


(c)

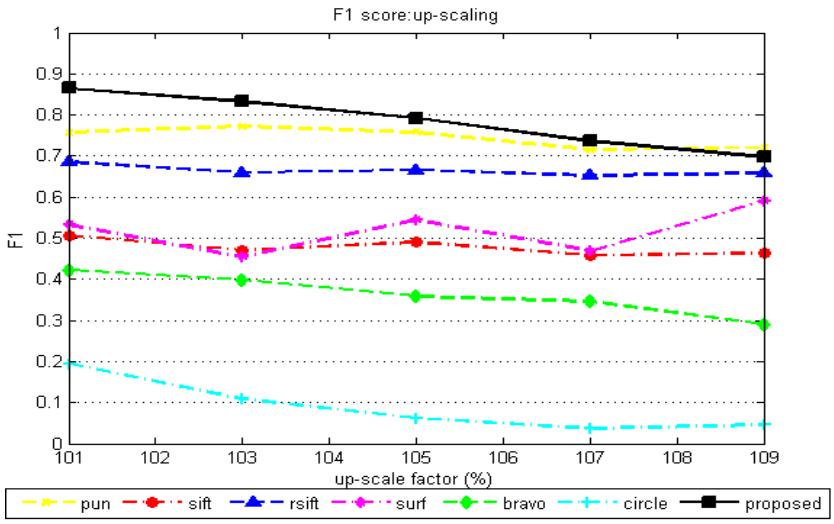
Fig. 5.7 CMFD results under down-sampling: (a) Precision, (b) Recall and (c) F1.



(a)

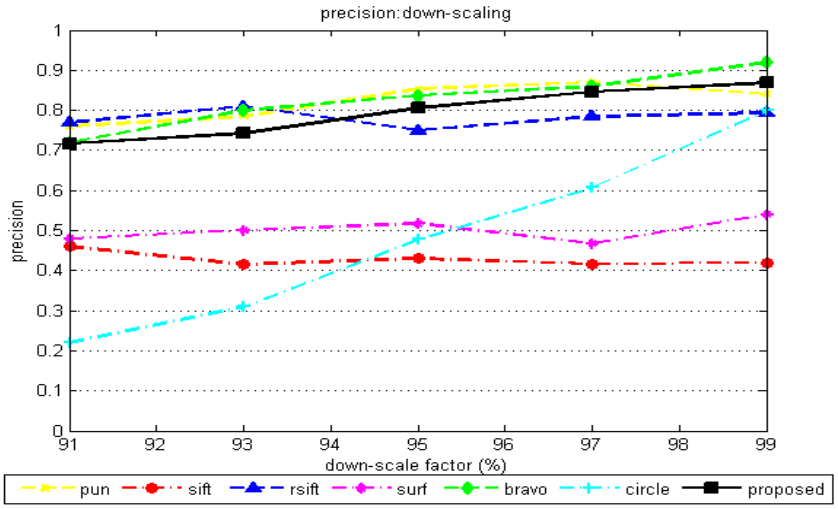


(b)

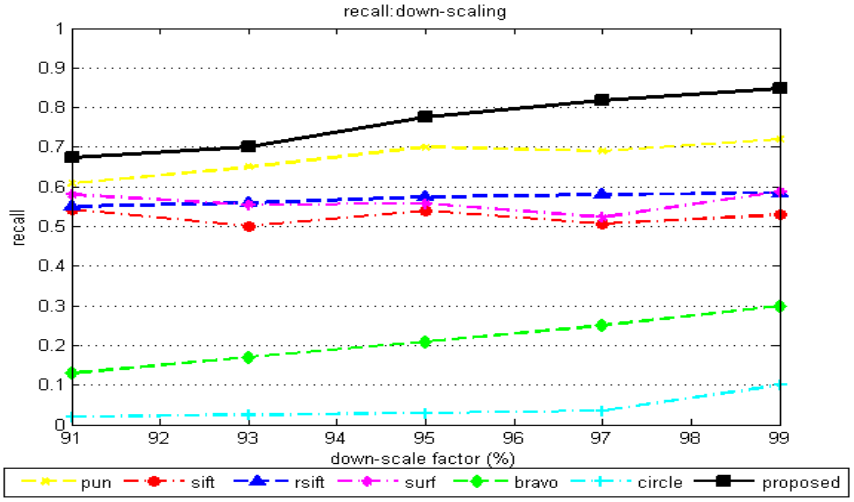


(c)

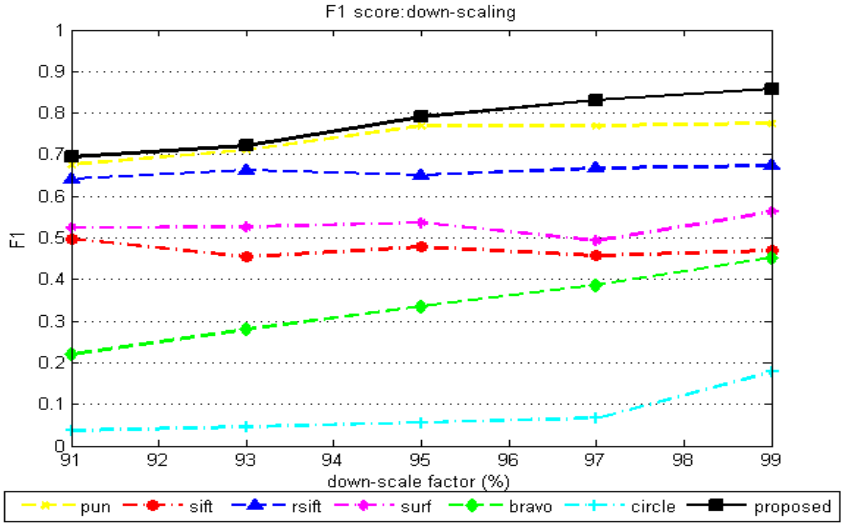
Fig. 5.8 CMFD results under up-scaling: (a) Precision, (b) Recall and (c) F1.



(a)

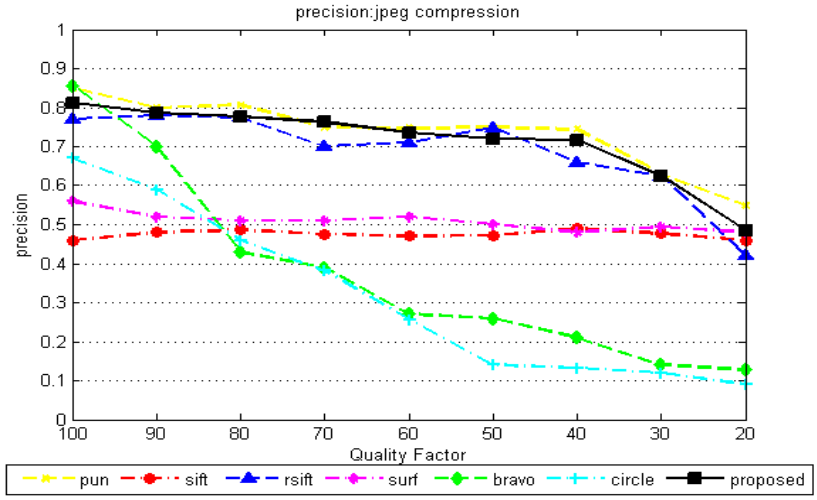


(b)

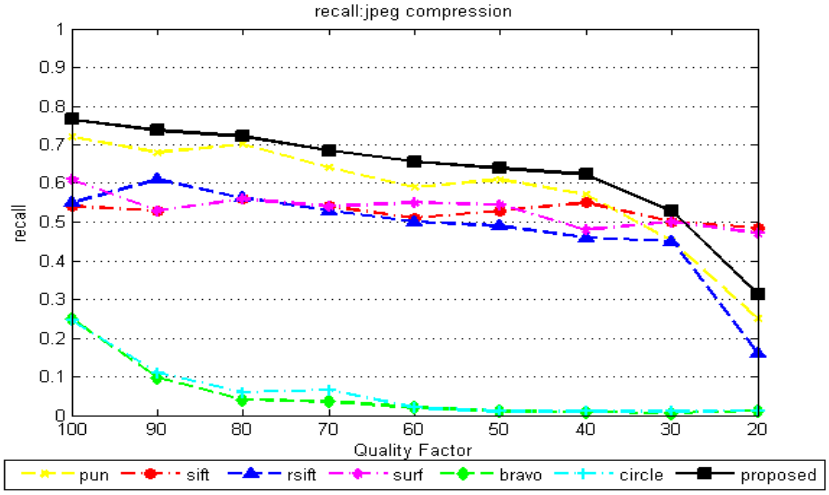


(c)

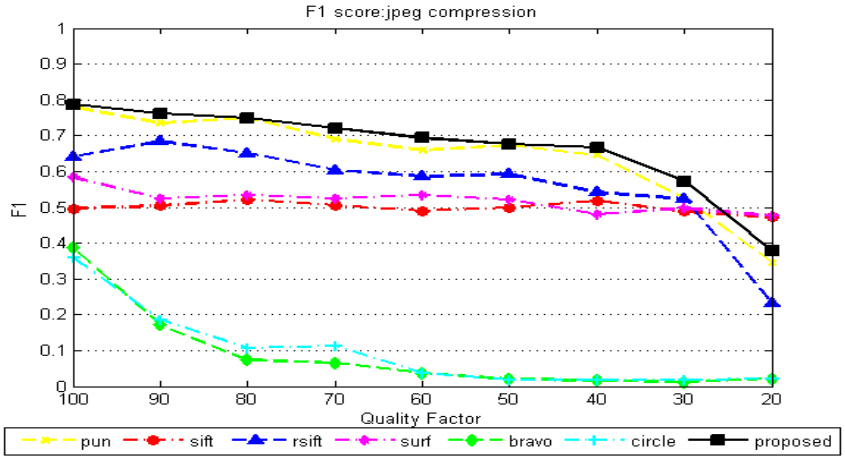
Fig. 5.9 CMFD results under down-scaling: (a) Precision, (b) Recall and (c) F1.



(a)



(b)



(c)

Fig. 5.10 CMFD results under JPEG compression: (a) Precision, (b) Recall and (c) F1.

5.5 SUMMARY

The robustness and efficacy of the proposed CMFD scheme has been demonstrated with various simulation results. In case of plain copy-move attack, the proposed technique performs better than existing techniques with 100% precision, 100% recall and 100% F1 score at image level, while 95.01% precision, 87.18% recall and 90.92% F1 score at pixel level. Under different challenging conditions, which include down-sampling, up-scaling, down-scaling and JPEG Compression, the simulation results indicate that precision of the proposed scheme is comparable to that obtained by existing state-of-the-art techniques but it gives the best performance in terms of recall and F1 score.

CHAPTER 6

CONCLUSION AND FUTURE SCOPE

6.1 CONCLUSION

This dissertation presents an adaptive technique to detect CMF which integrates both block-based and keypoint-based techniques. The experimental results state that under plain CMF, the proposed technique gives improved results in comparison to existing arts with P=100% and R=100%*;* thus F_1=100% at image level and P=95.01% and R=87.18%; thus F_1 =90.92% at pixel level.

The efficiency of the proposed algorithm is also tested under various attacks including down-sampling, up-scaling, down-scaling and JPEG compression. The summarized results under these attacks indicate that the precision of the proposed technique is somewhat similar to that obtained by block-based and hybrid methods but better than the keypoint-based techniques. The proposed technique exhibits the best recall results among all the schemes indicating that its capability in finding the largest number of similar regions. Better recall results are the outcome of Adaptive Forgery Region Extraction method which is used to extract the replicated regions from the tampered images as it helps greatly in reducing the probability of the tampering being undetected. It has been observed that the proposed hybrid CMFD scheme outdo the existing arts in terms of F1 score too. This measure clearly depicts that the CMFD results obtained by the proposed technique are better than that of the prior arts under the different attacks.

6.2 FUTURE SCOPE

This technique can be improved in future for the detection of similar regions under the influence of rotational geometrical attack and Gaussian blurring. Another aspect that will require improvement is the further reduction in computational burden due to feature matching process. However, the proposed hybrid scheme can be implemented on other forgery types, like image-splicing or on other media types, like video, in future.

REFERENCES

[1] J. A. Redi, W. Taktak, and J.-L. Dugelay, "Digital image forensics: A booklet for beginners," *Multimedia Tools and Applications*, vol. 51, no. 1, pp. 133–162, 2011.

[2] B. Mahdian and S. Saic, "Blind methods for detecting image fakery," *IEEE Aerospace and Electronic Systems Magazine*, vol. 25, pp. 18–24, 2010.

[3] S. Mushta and A. H. Mir, "Digital image forgeries and passive image authentication techniques: A survey," *International Journal of Advanced Science and Technology*, vol.73, pp.15–32, 2014.

[4] 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, no. 6, pp. 1841–1854, 2012.

[5] J. Fridrich, D. Soukal, and J. Lukas, "Detection of copy-move forgery in digital images," in *Proc. of the Digital Forensic Research Workshop*, Cleveland, OH, pp. 55–61, 2003.

[6] A. T. S. Ho and S. Li, "Handbook of digital forensics of multimedia data and devices," John Wiley & Sons, West Sussex, UK, 2015.

[7] O. M. Al-Qershi and B. E. Khoo, "Passive detection of copy-move forgery in digital images: State-of-the-art," *Forensic Science International*, vol. 231, no. 1, pp. 284–295, 2013.

[8] R. E. J. Granty, T. S. Aditya, and S. Madhu, "Survey on passive methods of image tampering detection," in *Proc. of the International Conference on Communication and Computational Intelligence*, pp. 431–436, 2010.

[9] G. Muhammad, M. Hussain, K. Khawaji, and G. Bebis, "Blind copy move image forgery detection using dyadic undecimated wavelet transform," in *Proc. of the 17th International Conference on Digital Signal Processing*, pp. 1–6, 2011.

[10] M. Ghorbani , M. Firouzmand, and A. Faraahi, "DWT-DCT (QCD) based copy-move image forgery detection," in *Proc. of the 18th International Conference on Systems, Signals and Image Processing*, pp. 1–4, 2011.

[11] A. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," Computer Science Department, Dartmouth College, Technical Report TR2004-515, 2004.

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

[13] J. W. Wang, G. J. Liu, Z. Zhang, Y. W. Dai, and Z. Q. Wang, "Fast and robust forensics for image region-duplication forgery," *Acta Automatica Sinica*, vol. 35, no. 12, pp. 1488–1495, 2009.

[14] J. Wang, G. Liu, H. Li, Y. Dai, and Z. Wang, "Detection of image region duplication forgery using model with circle block," in *Proc. of the International Conference on Multimedia Information Networking and Security*, pp. 25–29, 2009.

[15] M. Bashar, K. Noda, N. Ohnishi, and K. Mori, "Exploring duplicated regions in natural images," *IEEE Transactions on Image Processing*, 2010.

[16] S. Bravo-Solorio and A. K. Nandi, "Exposing duplicated regions affected by reflection, rotation and scaling," in *Proc. of the International Conference on Acoustics, Speech and Signal Processing*, pp. 1880–1883, 2011.

[17] Y. Cao, T. Gao, and Q. Yang, "A robust detection algorithm for copy-move forgery in digital images," *Forensic Science International*, vol. 214, no. 1–3, pp. 33–43, 2012.

[18] H. C. Nguyen, and S. Katzenbeisser, "Detection of copy-move forgery in digital images using radon transformation and phase correlation," in *Proc. of the 8th International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp. 134–137, 2012.

[19] Y. Wang, K. Gurule, J. Wise, and J. Zheng, "Wavelet based region duplication forgery detection," in *Proc. of the 9th International Conference on Information Technology*, pp. 30–35, 2012.

[20] G. Lynch, F. Y. Shih, and H.-Y. Mark Liao, "An efficient expanding block algorithm for image copy-move forgery detection," *Information Sciences*, vol. 239, pp. 253–265, 2013.

[21] M. A. Sekeh, M. A. Maarof, M. F. Rohani, and B. Mahdian, "Efficient image duplicated region detection model using sequential block clustering," *Digital Investigation*, vol. 10, no. 1, pp. 73–84, 2013.

[22] J. Zhao and J. Guo, "Passive forensics for copy-move image forgery using a method based on DCT and SVD," *Forensic Science International*, vol. 233, no. 1–3, pp 158–166, 2013.

[23] A. Kashyap and S. D. Joshi, "Detection of copy-move forgery using wavelet decomposition," in *Proc. of the International Conference on Signal Processing and Communication*, pp. 396–400, 2013.

[24] S. Ketenci and G. Ulutas, "Copy-move forgery detection in images via 2D-fourier transform," in *Proc. of the 36th International Conference on Telecommunications and Signal Processing*, pp. 813–816, 2013.

[25] S. A. Fattah, M. M. I. Ullah, M. Ahmed, and I. Ahmmed, "A scheme for copy-move forgery detection in digital images based on 2D-DWT," in *Proc. of the 57th International Midwest Symposium on Circuits and Systems*, pp. 801–804, 2014.

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

[27] X. Pan and S. Lyu, "Region duplication detection using image feature matching," *IEEE Transactions on Information Forensics Security*, vol. 5, no. 4, pp. 857–867, 2010.

[28] B. L. Shivakumar and S. S. Baboo, "Detection of region duplication forgery in digital images using SURF," *International Journal of Computer Science Issues*, vol. 8, no. 4, pp. 199–205, 2011.

[29] X. Bo, W. Junwen, L. Guangjie, and D. Yuewei, "Image copy–move forgery detection based on SURF," in *Proc. of the International Conference on Multimedia Information Networking and Security*, pp. 889–892, 2010.

[30] M. F. Hashmi, V. Anand, and A. G. Keskar, "Copy move forgery detection using DWT and SIFT features," in *Proc. of the 13th International Conference on Intelligent Systems Design and Applications*, pp. 188–193, 2013.

[31] M. Jaberi, G. Bebis, M. Hussain, and G. Muhammad, "Improving the detection and localization of duplicated regions in copy-move image forgery," in *Proc. of the 18th International Conference on Digital Signal Processing*, pp.1–6, 2013.

[32] R. C. Pandey, S. K. Singh, K. K. Shukla, and R. Agrawal, "Fast and robust passive copy-move forgery detection using SURF and SIFT image features," in *Proc. of the 9th International Conference on Industrial and Information Systems*, pp. 1–6, 2014.

[33] T. Chihaoui, S. Bourouis, and K. Hamrouni, "Copy-move image forgery detection based on SIFT descriptors and SVD-matching," in *Proc. of the 1st International*

Conference on Advanced Technologies for Signal and Image Processing, pp. 125–12, 2014.

[34] M. F. Hashmi, V. Anand, and A. G. Keshkar, "Copy-move image forgery detection using an efficient and robust method combining un-decimated wavelet transform and scale invariant feature transform," in *Proc. of the International Conference on Circuit and Signal Processing*, pp. 84–91, 2014.

[35] C. Pun, X. Yuan, and X. Bi, "Image forgery detection using adaptive oversegmentation and feature point matching," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 8, pp. 1705–1716, 2015.

[36] J. Li, X. Li, B. Yang, and X. Sun, "Segmentation-based image copy-move forgery detection scheme," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 3, pp. 507–518, 2015.

[37] E. Ardizzone, A. Bruno, and G. Mazzola, "Copy-move forgery detection by matching triangles of keypoints," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 10, pp. 1556–6013, 2015.

[38] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, "Digital Watermarking and Steganography," 2nd edition, Morgan Kaufmann, Francisco, CA, USA, 2008.

[39] M. A. Qureshi and M. Deriche, "A bibliography of pixel-based blind image forgery detection techniques," *Signal Processing: Image Communication*, vol. 39, pp. 46-74, 2015.

[40] CASIA Image Tampering Detection Evaluation Database V2.0, National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Science. Available: forensics.idealtest.org.

[41] 40 Amazing Before and After Photo Retouching Photos. Available: http://10steps.sg/inspirations/artworks/40-cool-before-and-after-photo-retouching-photos/.

[42] M. K. Johnson and H. Farid, "Exposing digital forgeries through chromatic aberration," in *Proc. of the 8th ACM Multimedia and Security Workshop*, New York, USA, pp. 48–55, 2006.

[43] J. Lukas, J. Fridrich, and M. Goljan, "Digital camera identification from sensor pattern noise," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 205–214, 2006.

[44] A. E. Dirik and N. D. Memon, “Image tamper detection based on demosaicing artifacts,” in *Proc. of the 16th International Conference on Image Processing*, pp. 1497–1500, 2009.

[45] M. Chen, J. Fridrich, M. Goljan, and J. Lukas, “Determining image origin and integrity using sensor noise,” *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 74–90, 2008.

[46] S. Bayram, H. Sencar, N. Memon, and I. Avcibas, “Source camera identification based on CFA interpolation,” in *Proc. of the International Conference on Image Processing*, vol. 3, pp. 69–72, 2005.

[47] G. Bhatia and A. S. Jalal, “Image forgery detection using feature based clustering in JPEG images,” in *Proc. of the 9th International Conference on Industrial and Information Systems*, pp. 1–5, 2014.

[48] B. Li, Y. Q. Shi, and J. Huang, “Detecting doubly compressed JPEG images by using mode based first digit features,” in *Proc. of the* 10th *Workshop on Multimedia Signal Processing*, Cairns, QLD, Australia, pp. 730–735, 2008.

[49] C. Pasquini, G. Boato, and F. Perez-Gonzalez, “Multiple JPEG compression detection by means of Benford–Fourier coefficients,” in *Proc. of the International Workshop on Information Forensics and Security*, Atlanta, GA, USA, pp. 113–118, 2014.

[50] Z. Fan and R. L. De Queiroz, “Identification of bitmap compression his-tory: JPEG detection and quantizer estimation,” *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 230–235, 2003.

[51] M. K. Johnson and H. Farid, “Detecting photographic composites of people,” in *Proc. of the 6th International Workshop on Digital Watermarking*, Guangzhou, China, ,pp. 19–33, 2007.

[52] M. K. Johnson and H. Farid, “Metric Measurements on a Plane from a Single Image,” Department of Computer Science, Dartmouth College, Technical Report TR2006-579, 2006.

[53] W. Zhang, X. Cao, Z. Feng, J. Zhang, and P. Wang, “Detecting photographic composites using two-view geometrical constraints,” in *Proc. of the International Conference on Multimedia and Expo*, pp. 1078–1081, 2009.

[54] E. Kee, and H. Farid, “Exposing digital forgeries from 3-D lighting environments,” in *Proc. of the International Workshop on Information Forensics and Security*, Seattle, WA, USA, pp. 1–6, 2010.

[55] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to the state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[56] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[57] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.

[58] X. Guo, X. Cao, J. Zhang, and X. Li., "Mift: A mirror reflection invariant feature descriptor," in *Proc. of the 9th Asian Conference on Computer Vision*, pp. 536–545, 2009.

LIST OF PUBLICATIONS

1. Lamba A. K., Jindal N. and Sharma S., An Adaptive Method for Copy-Move Image Forgery Detection, IEEE Transactions on Information Forensics and Security- *Communicated*, 2016.

Amanjot_Thesis

by Amanjot Lamba

Amanjot_Thesis

ORIGINALITY REPORT

15%	10%	11%	8%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Thapar University, Patiala Student Paper	1%
2	Qureshi, Muhammad Ali, and Mohamed Deriche. "A bibliography of pixel-based blind image forgery detection techniques", Signal Processing Image Communication, 2015. Publication	1%
3	Pun, Chi-Man, Xiao-Chen Yuan, and Xiu-Li Bi. "Image Forgery Detection Using Adaptive Oversegmentation and Feature Point Matching", IEEE Transactions on Information Forensics and Security, 2015. Publication	1%
4	www.sersc.org Internet Source	1%
5	dspace.thapar.edu:8080 Internet Source	1%
6	Bi, Xiuli, Chi-Man Pun, and Xiao-Chen Yuan. "Multi-Level Dense Descriptor and Hierarchical Feature Matching for Copy–Move Forgery Detection", Information Sciences, 2016.	1%