Passive Approach for Copy-Move Forgery Detection for Digital Image

V. Rathod[1], J. Gavade[2]

[1]PG Scholar, Department of Electronics, Textile and Engineering Institute Ichalkaranji, Maharashtra, India
[2]Assistant Professor in Electronics, Textile and Engineering Institute Ichalkaranji, Maharashtra, India
{vishwajarathod@gmail.com, jayashree2k2@gmail.com}

Abstract: In the present era of the digital world, digital images and videos are the main carriers of information. However, these sources of information can be easily manipulated by using readily available photo editing softwares such as Photoshop etc. Nowadays, it is possible to add or remove some important objects or features from the image without leaving any traces of tampering. Such type of forgery is called as copy move forgery in which objects from the image are copied and pasted within same image. To detect the copy move forgery in digital images have become the most hot research area nowadays. To solve the above problem, this paper proposes a hybrid method which is a combination of Dyadic Wavelet Transform (DyWT) and Scale Invariant Feature Transform (SIFT). In this method, DyWT is used to decompose image into four parts LL, LH, HL, and HH. Since LL part contains most of the information SIFT is applied on LL part to extract the features. Finally sorting algorithm RANSAC is used to match the features and to take the decision about forgery. The performance of proposed algorithm is tested on various images and the results show that algorithm works efficiently.

Keywords: ***DyWT (Dyadic Wavelet Transform), SIFT (Scale Invariant Feature Transform), Copy move forgery, RANSAC etc.***

1 Introduction

In today's digital era digital images have become one of the main medium of communication. Everyday millions of images and videos are uploaded on social media sites. The images that are shared on such sites gains more attention and also affect the societal opinions regarding particular incident because its natural tendency of human being to accept the things as it is which are seen by eyes. A lot of image editing softwares are easily available with the help of which one can change or modify the contents of image. This is called as Image forgery or Tampering. This tampering can be done as a fun also and sometimes intentionally for specific use, for example in order to hide some contents or objects in an image.

When the contents of the image are modified intentionally then this issue becomes serious as it may lead to some danger situations or controversies. Hence it has become very important to check the trustworthiness of any image. Tampering of image can be done in various ways but the most common method of tampering is Copy Move Forgery in which a part of image is copied, modified and pasted within the same image. Since the forgery is performed within single image the forged area has almost same properties as that of original image.

Hence it becomes very much difficult to identify such type of forgery.While tampering the digital image the copied portion of image is not pasted as it is in the image rather some post processing operations such as compression, rotation, blurring are performed on copied area first and then this modified area is pasted somewhere else in the same image. This done so that nobody can identify the forged area and forgery detection will become more difficult [1].Fig. 1 shows the example of copy move forgery. The first image is original while the second is forged in which the bird is copied and pasted in same is image. To detect the image forgery basically two approaches are used as follow:

B. Iyer, S. Nalbalwar and R.Pawade (Eds.)
ICCASP/ICMMD-2016. Advances in Intelligent Systems Research.
Vol. 137, Pp. 466-473.

1.1 Active Approach

In this approach some digital information in the form of digital signature or watermarking embedded in the original image when it is generated, and thus it has a limited scope.

1.2 Passive Approach

This approach does not require any embedded information for forgery detection. Here the detection is done based on the statistical features of the images which has undergone forgery. Hence this approach is difficult. The rest of the paper is arranged as follow: Section 2 takes the brief review of previous work. Section 3 comments on the proposed method. Section 4 shows algorithm of proposed method. Section 5 discusses the results and performance of proposed system. Section 6 gives evaluation measures of proposed method. Section 7 discusses the conclusion and future scope.

a. Original Image b. Forged Image

Fig.1. Copy move Forgery example

2 Related Work

Until now a lot of work has been done to address this problem. In the following paragraph we took review of the work done previously along with its advantages and disadvantages. In [4] Gupta, Saxena, Vasistha have proposed an efficient method to detect copy-move forgery objects in a single image. In this method first input image is divided into overlapping blocks. After that Discrete Cosine Transform (DCT) is applied to extract the features and then lexicographic sort is applied on these features to detect the doctored image blocks. But the limitations of this method are DCT is less efficient, less accurate and data loss is also occurred. In [5] a similar detection method is proposed by Popescu and Hany Farid. In their method Principal Component Analysis (PCA) is used to reduce dimensionality of image blocks. PCA has statical approach. It uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called Principal Components.

The number of Principal Components are less than or equal to the number of original variables. The limitations of this method are blocks are directly extracted from the image resulting in large number of blocks, thus affecting the efficiency of detection algorithm. And PCA does not work efficiently to detect forgeries in JPEG image. Li et al [6] developed method which reduces the overall computational time. Here the first Discrete Wavelet Transform (DWT) is applied to decompose the given image into four different sub-bands LL, LH, HL and HH. Since most of the information is present in the low frequency band, to extract the features LL band is divided into overlapping blocks. By doing this, numbers of blocks have been reduced and overall process has speed up. On these blocks, they applied Singular Value Decomposition (SVD).

In linear algebra the SVD is factorization of real or complex matrix. It has many useful applications in signal processing and statics. But the DWT and SVD suffers from the drawback that computation of SVD takes lot of time and it is computationally complex.

To reduce the time required for the computation, Khan and Kulkarni [7] has proposed wavelet based method. They have used Discrete Wavelet Transform (DWT).DWT is generally used as a compression technique to compresses the image. To find the similarities between overlapping blocks phase correlation is applied. In the first step as image compression is done, the time required for the computation is reduced. The drawback of this

method is that the rotated and scaled version of forged object cannot be detected as DWT is not shift invariant and data loss also occurs. Muhammad, Hussain [8] proposed method in which Dyadic Wavelet Transform (DyWT) is used. DyWT is not shift variant and captures the structural information in better way than DWT.

In this paper first the image is segmented and then decomposition using DyWT takes place. The stastical measure is used to detect similarities between two objects. But this algorithm is suitable for simple background. From the summary of Literature Review it is clear that present methods that are used for the image copy move forgery detection have various limitations. Still there are few gaps that are to be addressed to increase the accuracy and efficiency of forgery detection techniques such as:

- Detection accuracy should be more.
- Able to detect smallest forged area of an image.
- It should require computationally less time.
- If the image is forged by image processing operations such as scaling, rotation, blurring, filtering etc., this also can be detected by algorithm.

3 Proposed Method

From the literature review we can see that previous methods have some drawbacks. To overcome these drawbacks, this paper suggests new passive blind forgery detection technique based on combination of Dyadic Wavelet Transform and SIFT features. Dyadic Wavelet Transform is used to decompose the image and SIFT is used to extract the dominant features from the decomposed image. The basics of DyWT and SIFT are as given below:

3.1 Dyadic Wavelet Transform (DyWT)

Generally Wavelet Transform is used for data compression as well as image compression. Many previous methods used DWT for copy move forgery detection to decompose image into 4 parts i.e. LL, LH, HL, HH. Decomposition is used in order to reduce further computational processing time. In DWT when decomposition is performed, every second coefficient is selected. It is referred as shift variant. For object and pattern recognition the data must be shift invariant. As the data is shifted its descriptors also get shifted. Therefore the place of the copied and pasted object is not same. Due to this forgery is not detected. To overcome this problem related with DWT, Mallet and Zhong introduces Dyadic Wavelet Transform which is shift invariant. As the DyWT does not involve down sampling so that number of coefficients does not reduce between scales like DWT. So DyWT is most prominent than DWT. Let I be the image to be decomposed, and h[k] and g[k] be the scaling (low pass) and wavelet (high pass) filters. Start at scale j = 0, and take I^0= I, and compute the scaling and wavelet coefficients at scales j = 1, 2…j using Equations. (1) and (2):

$$c^{j+1}[n] = \sum_k h[k]c^j[n + 2^j k] \quad \text{……..(1)}$$

$$d^{j+1}[n] = \sum_k g[k]c^j[n + 2^j k]. \quad \text{……...(2)}$$

Let $h^j[k]$ and $g^j[k]$ be the filterscoefficients obtained by inserting 2^j-1 zeros between the terms of h[k] and g[k]. Then we can perform DyWT using filtering as follows:

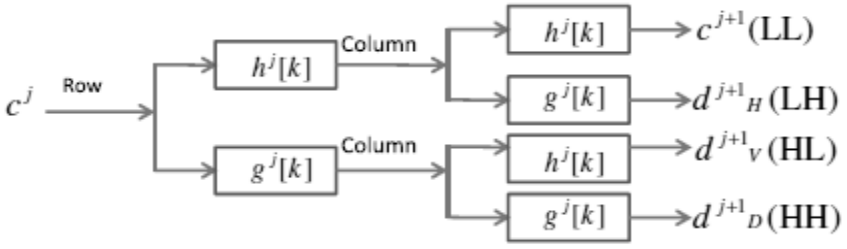


Fig.2. One level decomposition of DyWT of 2D image

As mentioned, there is no down sampling involved in DyWT. In the wavelet transform, $h^j[k]$ are low pass filter coefficients, $g^j[k]$ are high pass filter coefficients and c^j are the image coefficients. First image coefficients are row wise given to the $h^j[k]$ and $g^j[k]$. Then image coefficients are converted into low pass sub band and high pass sub band. After that these two sub bands are individually again given column wise to the $h^j[k]$ and $g^j[k]$ and finally four sub bands are formed i.e. LL, LH, HL, HH as shown in Fig. 2.

3.2 Scale-Invariant Features Transform (SIFT)

SIFT is an algorithm which is used to detect and describe local features of an image. SIFT features are invariant to any transformation such as scaling and rotation in image domain. SIFT contains four steps that are as follows:

3.2.1 Scale-space extrema detection:

The first stage of computation searches over all scales and image locations, which is used to find the local extremas in scale-space. To detect scale space extrema scale space filtering is used. In which Laplacian of Guassian (LoG) is used. But LoG is computationally expensive so that Difference of Guassian (DoG) is used which is approximation of LoG. DoG of an image is obtained as difference of Guassian by blurring an image with different values of σ, where σ is the scaling parameter. After getting the DoG, searches for overall scale and image locations by comparing one pixel with its 8 neighbours, 9 pixels in next scale and 9 pixels in previous scales. If its value is minimum or maximum than all pixels then point is an extrema.

3.2.2 Key point localization:

The next step is key point localization. In this step, more accurate key points are located. Taylor series expansion of scale space is used to get more accurate location of extrema, and if the intensity at this extrema is less than a threshold value, then it is rejected.

3.2.3 Orientation assignment:

Orientation to each key point is assigned to achieve invariance to rotation. To assign the orientation to key point histogram is used.

3.2.4 Key point descriptor:

The local image gradients are measured at selected scale in the region around each key point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

3.3 Sorting Algorithm

RANSAC is used as sorting algorithm. The RANdom SAmple Consensus (RANSAC) algorithm is a general parameter estimation approach designed to cope up with large proportion in input data. RANSAC is resampling technique that generates candidate solutions by using minimum number of observations (data points) required to estimate the underlying model parameters. Unlike the conventional sampling techniques that use as much as data as possible to obtain an initial solution and then proceed to prune, RANSAC uses the smallest set possible and proceeds to enlarge this set with consistent data points. The input to the RANSAC algorithm is a set of observed data values, a way of fitting some kind of model to the observations, and some confidence parameters. RANSAC achieves its goal by repeating the following steps:

3.3.1 Select a random subset of the original data. Call this subset as the hypothetical inliers.
3.3.2 A model is fitted to the set of hypothetical inliers.
3.3.3 All other data are then tested against the fitted model. Those points that fit the estimated model well, according to some model-specific loss function, are considered as part of the consensus set.
3.3.4 The estimated model is reasonably good if sufficiently many points have been classified as part of the consensus set.

3.3.5 Afterwards, the model may be improved by re-estimating it using all members of the consensus set.

4 Algorithm of Proposed method

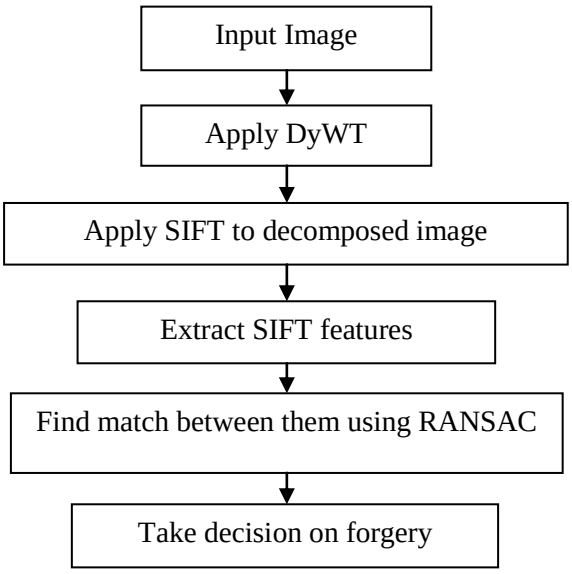


Fig.3. Flowchart of Proposed System

First input image is taken. DyWT is applied to the given image to decompose it into four parts (LL, LH, HL, and HH). As LL part contains most of the information, SIFT is applied to LL part to extract the features. Then RANSAC algorithm is used to match the features. According to matching the decision is taken whether image is forged or not.

5 Simulation Results

In order to test the performance of the proposed algorithm, algorithm is applied on standard database MICC-F220 which is combination non-tampered images and tampered images. Tampered images are forged using various attacks such as scaling, rotation, blurring etc. The simulation is performed on MATLAB2009b software. Following Figures make comments on the simulation results. Fig. 4 shows the result of tampered image while Fig. 5 shows the result of non-tampered image. Firstly the tampered image is taken and DyWT is applied to decompose it. Fig. 4.a shows the original image. Fig.4.b shows the LL part of the original image. Then the SIFT is applied to obtain the key points and to extract the features. Fig.4.c shows the SIFT descriptors of the image. Fig. 4.d shows the detection of forged area. To test the accuracy of the algorithm, the non-tampered image is given as input image. As shown in Fig.5.d this is detected as not tampered.

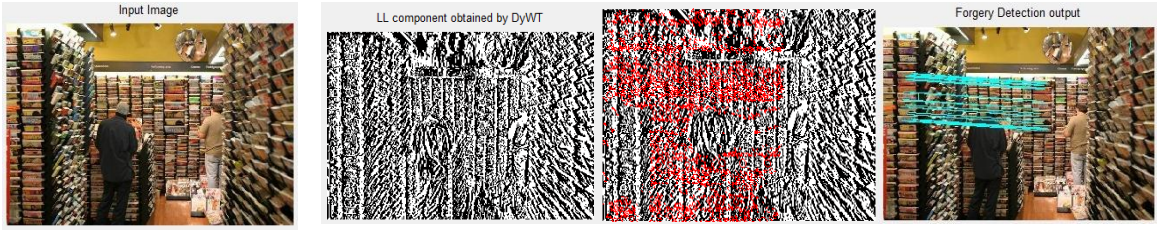
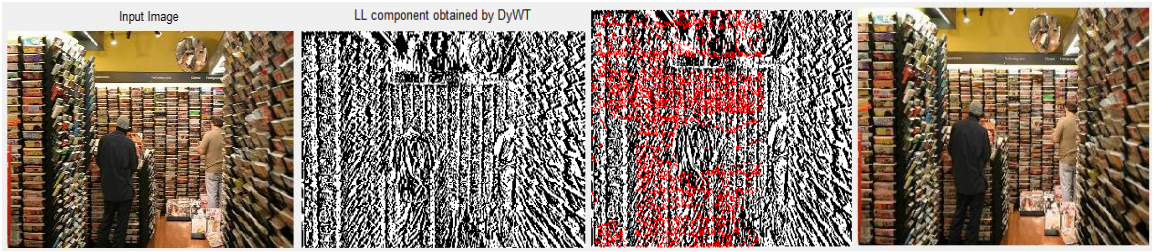


Fig.4. (a) Input Image, (b) LL part obtained by DyWT (c) SIFT descriptors (d) Forgery Detection Output



(a) (b) (c) (d)

Fig.5. (a) Input Image (b) LL part obtained by DyWT (c) SIFT descriptors (d) Forgery Detection Output

6 Evaluation Measures for Proposed Method

In order to verify the performance of the algorithm, we have considered four performance metrics given as follow:

- True Positive (TP): Authentic is detected as Authentic.
- False Positive (FP): Authentic is detested as forged.
- False Negative (FN): Forged is detected as Authentic.
- True negative (TN): Forged is detected as forged.
- Precision (p): It represents the probability of truly detecting a forgery.
- Precision Rate:= TP/(TP+FP)……(5)
- Recall (r): It represents probability that a forged image have been detected; it may be either true or falsely forged.
- Recall Rate= TP/(TP+FN)……(6)
- Detection Accuracy= (TP+TN)/ (TP+TN+FP+FN)…….. (7)

We have tested our algorithm on 15 images which was combination of forged image and Original Image. Out of this 15, 10 images are original images while remaining 5 images are forged images. Also we have compared the performance of this algorithm with only SIFT detection algorithm in terms of above parameters. The following table summarizes the results.

Table1: Performance Measure of Algorithms

Parameters	TP	FP	FN	TN	Precision	Recall	Detection Accuracy
DyWT + SIFT	10	0	0	5	100%	100%	100%
SIFT	9	1	1	4	90	90	86%

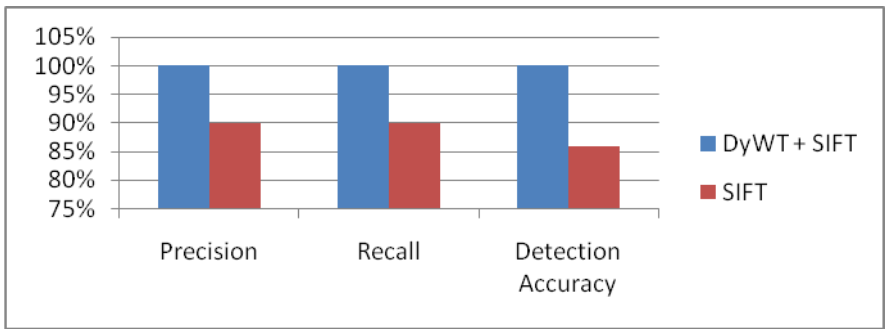


Fig.6: Bar Chart Representation for comparison of Performance of Both Methods

7 Conclusion and Future Scope

From the literature review we can see that there are different types of techniques proposed to detect copy move forgery in digital image. But these methods are not 100% efficient to detect copy move forgery. Each method has its own limitations. This paper promotes a hybrid algorithm for detecting the most common type image forgery i.e. copy move forgery. We proposed a method which is the combination of DyWT and SIFT. DyWT is shift-invariant so the image size is intact. As the low frequency component contains most of the image information, SIFT is applied on low frequency component to extract the features and then matching is obtained between feature descriptors to conclude that given image is forged or not. Also we have compared proposed algorithm with only SIFT based detection algorithm. From the simulation result it is clear that performance of proposed algorithm is good as compared to the SIFT algorithm.
But computational time required for the proposed algorithm is higher than SIFT. While testing the performance of said method, we have not considered the various signal processing attacks that can be done on forged area. In future we will consider various attacks such as scaling, rotation and compression to check the performance of said algorithm and according will make modifications in algorithm if required for better performance. We have tried the proposed algorithm only on single copy pasted area. In future, we will test the algorithm for detecting multiple copy pasted areas in same image. Also we will try to detect accurately the smallest forged area from the forged image.

References

[1] Tu Huynh-Kha, Thuong Tien, Synh Ha-Viet, Khoa Huynh, Marie Luong, "A Robust Algorithm of Forgery Detection in Copy Move and Spliced Images", International Journal of Advanced Computer Science and Applications, Vol. 7, No. 3, 2016
[2] Raj deep Kaur, Amandeep Kaur, "A Review of Copy-Move Forgery Detection Techniques", IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555 Vol.6, No.2, Mar-April 2016
[3] Swapnil H. Kudke, A. D. Gawande, "Copy-Move Attack Forgery Detection by using SIFTS", International Journal of Innovation Technology and Exploring Engineering (IJITEE). ISSN: 2278-3075, Volume-2. Issue-5, April 2013.10
[4] A.Gupta, N. Saxena, S. K. Vasistha, "Detecting copy move forgery using DCT", International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013.4
[5] C. Popescu and Hany Farid, "Exposing digital forgeries by detecting duplicated image regions", Technical Report , TR2004-515, Department of Computer Science, Dartmouth College, pp. 758-767,2006. 5
[6] G.Li, Q.Wu, D.Tu and Shaojie Sun, "A sorted neighborhoods approach for detecting duplicated regions in image forgeries based on DWT and SVD", IEEE International Conference in Multimedia & Expo, 2007.6
[7] S. Khan and A. Kulkarni, "Reduced time complexity for detection of copy-move forgery using Discrete Wavelet Transform", International Journal of Computer Applications, Volume 6-No. 7, September 2010, and Doi: 10.5120. **7**
[8] Ghulam Muhammad, Muhammad Hussain, George Bebis, "Passive copy move forgery detection using undecimated dyadic wavelet transform", Digital Investigation 9 (2012) 49-57. Doi:10.1016/j.diin.2012.04.004.
[9] Irene Amerini, Lamberto Ballen, Roberto Caldelli, Alberto Del Bimbo and Giuseppe Serra, "A SIFT-Based Forensic Method for Copy-Move Attack Detection and Transformation Recovery", IEEE Transactions on Information Forensics and Security, VOL. 6 No. 3, September 2011. Doi 10.1109/TIFS.2011.2129512.
[10] P. Mishra, N. Mishra, S. Sharma and R. Patel, "Region duplication forgery detection techniques based on SURF and HAC", The Science World Journal Volume 2013, Article 267691, 8 pages, Doi:10.1155/2013/267691
[11] N. Muhammad, M. Hussein, G. Muhammad and G. Bebis, "Copy-move forgery detection using Dyadic Wavelet Transform", 2011 Eight International Conference Computer Graphics, Imaging and Visualization Doi:10.1109/CGIV.2011.29. **8**
[12] M. F. Hashmi, V. Anand, A. G. Keskar, "Copy –move image forgery detection using an efficient and robust method combining un-decimated wavelet transform and scale invariant feature transform", 2014 AASRI Conference on Circuit and Signal Processing (CSP 2014),Doi:10.10.16/j.aasri.2104.09.015.

[13] M. Patil, B.Iyer and R.Arya, ""Performance analysis of PCA and ICA algorithms for facial expression analysis" ASIC(Springer), vol. 439, pp. 965-976, Mar. 2016.
[14] Kumar Sunil, Desai Jagan, Mukherjee Shaktidev," DCT-PCA Based Method for Copy-Move Forgery Detection" ICT and Critical Infrastructure: Proceedings of the 48th Annual Convention of Computer Society of India- Vol II Volume 249 of the series Advances in Intelligent Systems and Computing pp 577-583
[15] P. Deshpande, S. Sharma, S. Peddoju and A. Abraham, "Efficient data storage in cloud environment", Informatica: The journal of computer science (ACM), vol. 36, no. 2, pp. 501−510, 2015.