Chapter 1

INTRODUCTION

1.1 Introduction

As image is better than thousands of words, World Wide Web nowadays contains a large amount of digital images used for effective communication process sometimes also used as official documents [1]. It becomes very trivial for professionals or non-professionals to edit any pre-existing photographs by using freely available image editing tools, such as Photoshop. Now a days, it is an easy task to create digital image forgeries. Some of the tools found in the editor are: cropping, rotating, resizing an image, removing unwanted part of an image, merging pictures together. Therefore it's become a challenging task to verify the originality of an image.

Digital image forgery detection methods are classified into two approaches, which are active and passive approach. In active approaches, the digital images require signatures, watermarking etc. which are pre-processing. Passive approach is different from active approaches; passive approach does not need any watermark embedded in advance [2].

There are three main classes of image forgery, namely image retouching, copy-move forgery and image splicing. Image re-touching involves rotation, scaling, contrast enhancement and resizing. In copy-move forgery, an object or portion of the image is copied and pasted on to the other location in the same image so as to hide some information or alter the original content of the image. Here, the basic characteristics of the image remain same such as colour, texture. Thirdly, the image splicing is a forgery that is a combination of two or more than two images and results into a new image. Image splicing sources are heterogeneous as contents can be picked from anywhere.

1.2 Aims and Objectives

The existing method to detect Copy –Move Forgery lacks an efficient approach. In a Copy-Move forgery, a region of the image itself is copied and pasted into another region of the same image.

(a) (b)

Figure 1.1 Example of Copy – Move Forgery

In the above figure, the first image depicts an original image whereas the second one shows the forged image where in which truck object is covered with bushes from the same image.

The aim of this project is to develop an efficient method for detection of copy and move forgery.

1.3 Scope

The scope of the project is to detect copy move image forgery using hybrid wavelet transform to improve the efficiency. The combination of required block size and overlap to properly detect the forgery is found by evaluation of several test cases of the same image. The copy-moved image blocks, which have been scaled, rotated or transformed in any other way (except translation), will not be detected.

Chapter 2

REVIEW OF LITERATURE

2.1 Domain Explanation

In imaging science, image processing is processing of images using mathematical operations by using any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image.

2.2 Existing Solutions

Various methods have been proposed in literature to detect copy move forgery. The simplest method to detect this forgery is exhaustive search but this method is computationally expensive for medium-sized images. Huang et al. used DCT (discrete Cosine transform) as the discriminative features [3], but it cannot detect duplicated region with rotation of more than 5˚.

Popescu et al. used a principal component analysis (PCA) on image blocks as features [4]. However the above method fails for any type of geometric transformations of the block such as rotation, scaling etc. G. Lynch et al. used manual grouping for collecting similar blocks in a same class [5].

Jing Zhang used Discrete Wavelet Transform (DWT) instead of PCA for dimensionality reduction and located copy move regions by pixel matching [6]. To reduce the ration of false matching blocks Jie Hu developed a method that used the distance of Eigen vectors instead of DCT coefficients to fulfil the block [7].

Yang Wang, Kaitlyn Gurule, Jacqeline Wise, Jun Zheng proposed a wavelet based copy move forgery detection and applied multi-level 2D DWT, since the DWT approach gives significant results for perceiving the cloning forgery [8]. They detail about the application of the DWT on the input image to yield a diminished dimensional representation. Then, the compressed image is separated into blocks. These blocks are then arranged and fake blocks are recognized. Because of the DWT use, the location is initially completed on the least level image

representations. Thus, this copy move detection methodology builds accuracy of the recognition procedure.

Tanuja K. Sarode, Naveen Vaswani proposed a method in which a hybrid wavelet transform is generated from two different orthogonal transforms and is applied on forged image to detect forgery (E.g. Hybrid of DCT and Walsh) [9].

Chapter 3

ANALYSIS AND REQUIREMENTS

3.1 Functional Requirements

1. The system should search for images from the specified path.
2. The selected image which is to be detected should be classified into 8-bit (Grey Scale) or 24-bit (RGB) for proper processing.
3. The system should perform detection algorithm.
4. The system should slice 24-bit images into planes of 8-bits for further processing.
5. The image should be processed block by block in an overlapping manner and should save the values computed for further comparisons in order to detect Copy-Move Forgery.

3.2 Non-Functional requirements

1. **Easy access:** The system should be cost effective hence it can be implemented without much hassle.
2. **Accuracy:** Accuracy of the system is how close the output is to the required output.
3. **Usability:** Ease of use and learnability of the system. System should not be complicated, easy to use and user friendly.
4. **Robustness:** It is the ability of the system to cope with errors during execution.
5. **Reliability:** The probability that the system will perform the required function without failure under stated condition for a stated period of time is called reliability.

3.3 Software and Hardware Requirements

3.3.1 Software Requirement

- Operating System: Windows 7/8/10
- Programming Language: Matlab

3.3.2 Hardware Requirement

- Processor: Minimum 1.3GHz/Pentium 4 processor or higher.
- RAM: 2048 MB or More
- Memory: 100 MB

Chapter 4

DESIGN

4.1 Design Consideration

While designing the system, all the previous developments towards the forgery detection, and the potential flaws, have been taken into account. Following are the major considerations which were thought of while designing the solution:

1. Deciding upon the best comparison algorithm to match the blocks in shortest execution time

The system should find out the matching blocks by comparing the results of hybrid wavelet transform that is applied in an overlapping manner. Upon finding a match, the system should mark the two blocks of NxN pixels by changing their pixel intensities to pure white. All of the comparisons should be done in a manner that will lead to shortest execution time.

2. No false claim of forgery detection

The block size and overlapping combination provided to the system should be optimal enough to not falsely claim forgery detection in plain objects or backgrounds. For example, in a single coloured wall or plain white paper, which are basically small repeated blocks of same pixel intensities.

3. Slicing the given image into 8-bit planes

This is done in order to speed up the execution by computing and comparing only those blocks of second plane that have given a match in the first plane, and so on.

4. Deciding upon optimal Image Transformations Algorithms

Various combinations of image transformation algorithms should be tested for hybrid wavelet transform, such that the overall execution time can be reduced without compromising on the accuracy of the forgery detection.

Chapter 5

IMPLEMENTATION

5.1 Working of the System

The System uses a hybrid wavelet image transform, which is generated from two different orthogonal transforms. The whole image is processed using this hybrid wavelet to detect image forgery. The system works in four phases, which are Segmentation, Hybrid Wavelet Transformation, Feature Vector Generation and Marking, in that order.

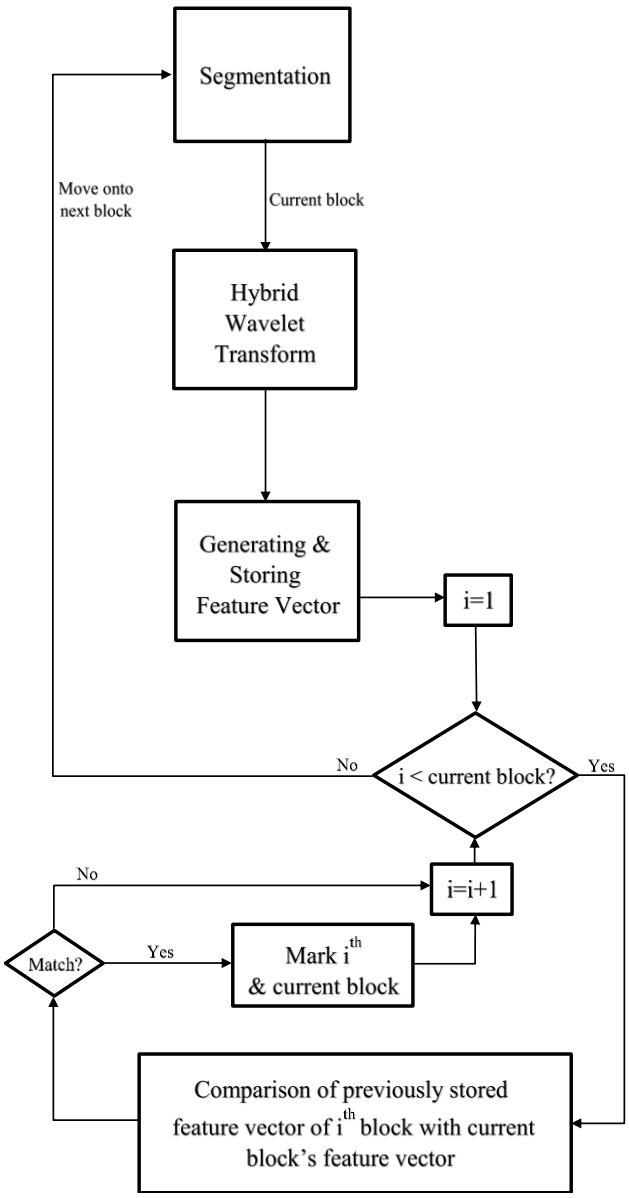

Figure 5.1 Flowchart showing phases of the proposed solution

5.2 Segmentation

This is the first stage of the system. If the image is of size N x N, then the image is divided into overlapping blocks of size K, where K < N/4. The block size varies from image to image.

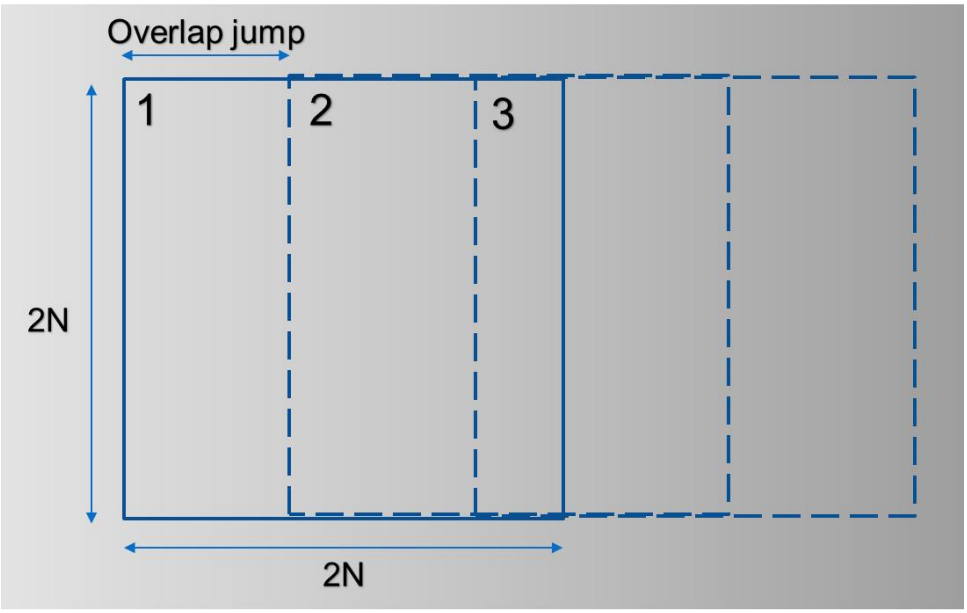

Figure 5.2 Segmentation with a block size of 2N x 2N

5.3 Hybrid Wavelet Transformation

This is the second stage of the system. Then the hybrid wavelet transform is applied on each block in order to extract feature vectors. The hybrid wavelet transform used in this system is that of Hadamard transform and Kekre wavelet transform.

5.3.1 Hadamard Transform

Hadamard transform is a Fourier transform, which performs orthogonal, linear, symmetric operation on an image of size 2^m. It is one of the simplest transform. It is generated recursively by taking kronecker product with Hadamard transform of size 2 x 2. It is entirely made up of 1 and -1. Fast Hadamard algorithm can be used to compute the Hadamard transform in n log n steps [10].

5.3.2 Kekre wavelet transform

Kekre Transform matrix can be of any size. It doesn't need to be an integer power of 2 like other orthogonal transforms. Hence it can be used for images of any size. In this matrix, all upper diagonal and diagonal elements are 1 whereas lower diagonal elements except the elements just below diagonal are zero. Kekre transform matrix can be represented using equation (1) [11, 12]. In equation (1), x and y represents the row-coordinate and column-coordinate of the transform matrix, respectively. Also, N represents the order of the square Kekre transform matrix i.e. NxN.

$$K_{xy} = \begin{cases} 1 & x \leq y \\ -N + (x-1) & x = y+1 \\ 0 & x > y+1 \end{cases} \tag{1}$$

5.3.3 Generating Hybrid Wavelet from orthogonal transform

The basic concept of wavelet transform is to select appropriate wavelet function called mother wavelet and then perform an analysis using shifted and dilated versions of mother wavelet. Wavelet transform gives time frequency analysis of a signal. Initially in study of wavelets Haar wavelet transform was emphasized. In recent study, wavelets of Walsh, Hartley, and Kekre have been proposed and experimented [13].

The hybrid wavelet of size N x N can be generated using two orthogonal transforms. Here, TAB is the hybrid wavelet. Matrix A and Matrix B are the orthogonal transforms matrices of size p x p and q x q respectively. While generating the hybrid wavelet, we get the first q rows by taking kronecker product of first row of transform A with complete transform B. The next row is generated by 2nd row of transform A and rest zeroes. Then this row is circularly right shifted to generate next q-1 rows. This process is repeated for rest of the rows of transform A. [14]

Figure 2 shows the generation of required hybrid matrix. Considering a block size of say N x N, a hybrid wavelet transform matrix of size N x N, is generated from the two transforms viz. Hadamard Transform and Kekre Wavelet Transform. Size of Hadamard matrix is kept as a constant 2 x 2, and that of Kekre Transform matrix is taken as N/2 x N/2. This combination yields every possible even sized detection block.

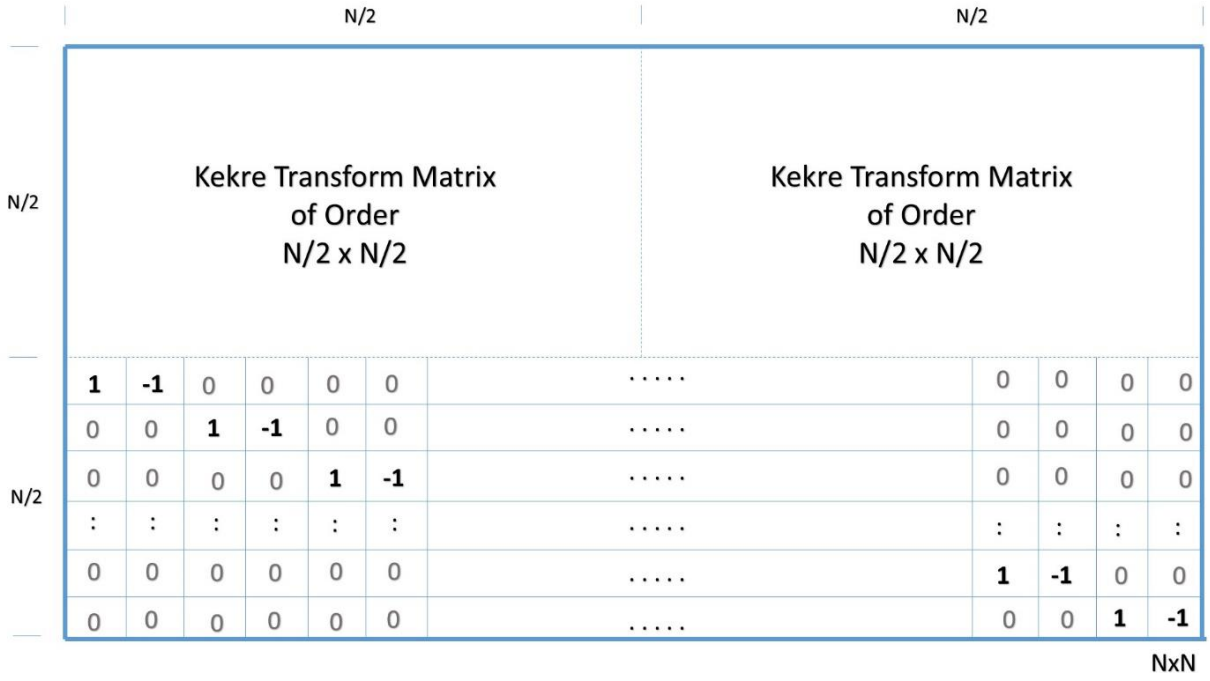

Figure 5.3 Required Hybrid Wavelet Transform Matrix TAB

5.3.4 Properties of Hybrid Wavelet Transforms

The hybrid wavelet transform matrix inherits orthogonal properties of two orthogonal transforms. The hybrid wavelet transform is orthogonal and non-involutional in nature [9].

5.4 Feature Vector Generation

This is the third stage of the system. Here, Energy signatures which are extracted from wavelet coefficients obtained after the application of hybrid wavelet transform are stored as the feature vectors i.e. mean, variance [8].

In equation (2) and (3) X_i represents coefficients obtained after applying the hybrid wavelet transforms, and n is the total number of coefficients obtained. $\bar{X}$ in equation (3) is the mean obtained in equation (2).

$$Mean = \frac{1}{n}\sum_{i=1}^{n} X_i \tag{2}$$

$$Variance = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2 \tag{3}$$

5.5 Marking

Each feature vector obtained in the previous stage corresponds to a block, and is compared with each of its following vector until a significant difference is found. If the value of this absolute difference of comparison is found to be within a certain threshold value, then these blocks are considered for matching. Two blocks are matched if all features of that block are same to a particular threshold value. The matched blocks are marked with white colour to indicate copy-move forgery.

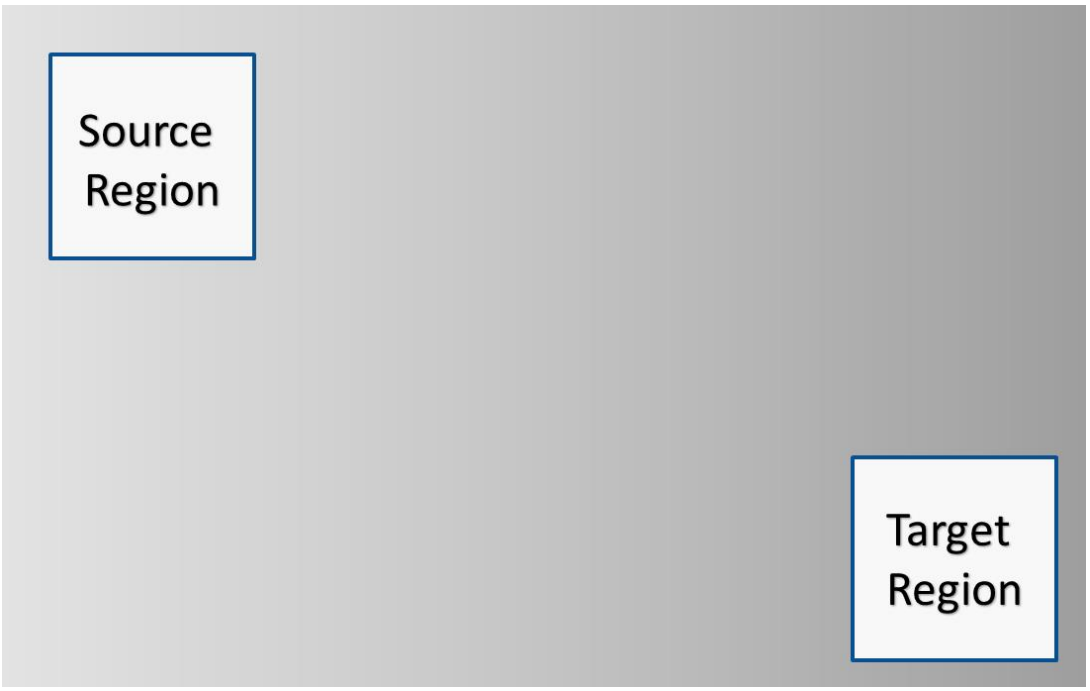

Figure 5.4 Marking of Copy – Move Forged Block in the image

Chapter 6

EXPERIMENTAL RESULTS

In order to check the accuracy of the proposed solution a set of images downloaded from internet are copy move forged and then tested. Also, Photoshop retouching tools like blur and brightness have been applied to few images post forgery. The images are in bitmap (.bmp) format.

For each image detection blocks of various sizes have been applied and the final result is the compilation of all the forged blocks detected for every block size applied. The detection blocks are squares having even pixel size.

Figure 6.1 (a) Original Image

Figure 6.1 (a) shows an image consisting of a single tree on green landscape. Figure 6.1 (b) is a copy move forged image, where the same tree has been copied and pasted on the same green landscape. Figure 6.1 (c) is the processed output, with white blocks showing the forgery [15].

Figure 6.1 (b) Copy – Move Forged Image

Figure 6.1 (c) Detection Result using Proposed Methodology

Figure 6.2 (a) shows an image of bus. Figure 6.2 (b) is a copy move forged image, where the top rectangular window has been copied and pasted in the same image, just few pixels right to the original. Also, the image has been blurred post forgery. Figure 6.2 (c) is the processed output, with white blocks showing the forgery [16].

Figure 6.2 (a) Original Image

Figure 6.2 (b) Copy – Move Forged Image

Figure 6.2 (c) Detection Result using Proposed Methodology

Figure 6.3 (a) shows an image consisting a dog having a look at the scenery. Figure 6.3 (b) is a copy move forged image, where the dog has been replicated besides the original. Also, the image's brightness has been increased post forgery. Figure 6.3 (c) is the processed output, with white blocks showing the forgery [17].

Figure 6.3 (a) Original Image

Figure 6.3 (b) Copy – Move Forged Image

Figure 6.3 (c) Detection Result using Proposed Methodology

Figure 6.4 (a) shows an image consisting a bush, upon which hangs a simple notice. Figure 6.4 (b) is a copy move forged image, where the notice has been hidden by copy moving the pixels of bush, from the same image. Figure 6.4 (c) is the processed output, with white blocks showing the forgery [18].

Figure 6.4 (a) Original Image

Figure 6.4 (b) Copy – Move Forged Image

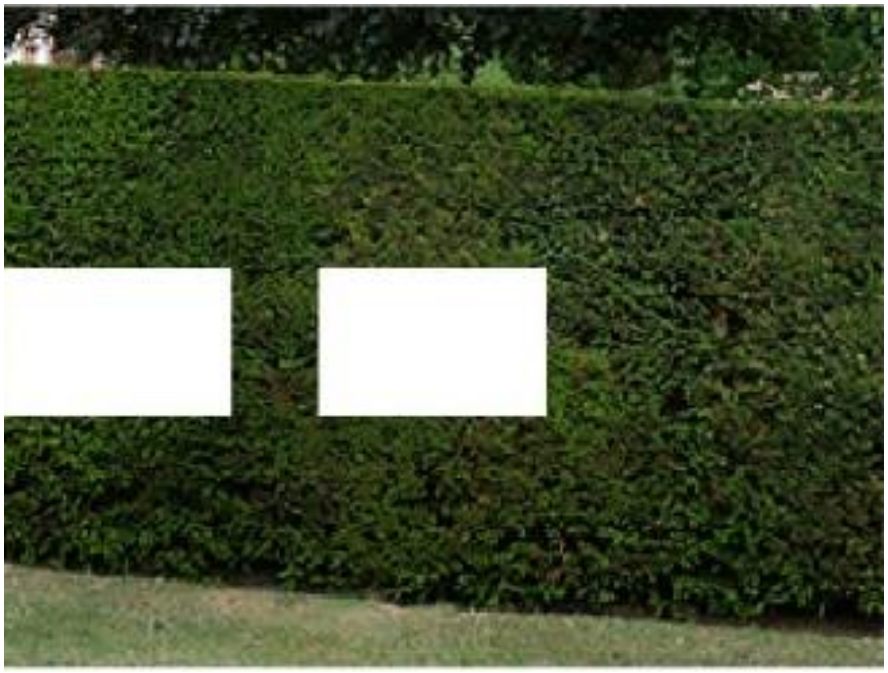

Figure 6.4 (c) Detection Result using Proposed Methodology

Figure 6.5 (a) shows an image of a keyboard. Figure 6.5 (b) is a copy move forged image, where certain keys have been replaced by copying the pixels of single key ‘T’, from the same image. Figure 6.5 (c) is the processed output, with white blocks showing the forgery [19].

Figure 6.5 (a) Original Image

Figure 6.5 (b) Copy – Move Forged Image

Figure 6.5 (c) Detection Result using Proposed Methodology

The forgery percentage and detection percentage has been calculated using equation (4) and equation (5).

$$Forgery\ Percentage = \frac{Total\ Forged\ Pixels}{Image\ Size} \times 100 \tag{4}$$

$$Detection\ Percentage = \frac{Total\ Detected\ Pixels}{Total\ Forged\ Pixels} \times 100 \tag{5}$$

The results for each figure have been shown in Table 1.

Table 6.1. Efficiency Table

Figure Number	**Forgery Percentage**	**Detection Percentage**
6.1	3.57%	57.72%
6.2	3.66 %	81.47%
6.3	1.05 %	100%
6.4	11.72 %	98.51%
6.5	11.596 %	100 %

For figures 6.2 – 6.5, the accuracy i.e. the detection percentage is pretty good, the minimum being 81.47%, and a maximum value of 100%. In these figures, the copy-move region was very similar to the detection blocks in shape. For figure 6.1, the detection percentage is 57.72%, which is because the copied and moved region isn't rectangular in nature i.e. it is not similar in shape with the detection blocks. Therefore, the inner regions of the copied tree in figure 6.1 have been properly detected, but the boundaries of tree have been not. Although, after looking at the image subjectively, it can easily concluded that the image is copy move forged.

Chapter 7

CONCLUSION

7.1 Conclusion

The Hybrid Wavelet Transform method, can be used to detect copy move forgery in digital images. The Kekre Wavelet Transform helps eliminating the limitation of most of the transforms related to size of detection blocks. Here, detection blocks of size 2N are obtained due to Kekre Wavelet Transform, whereas otherwise most of the other transforms facilitate having block sizes of order 2^N. For the generation of Hybrid Wavelet Transform, instead of Hadamard Transform, other orthogonal transforms can also be used along with Kekre Wavelet Transform. The accuracy of the proposed methodology, when tested on a dataset of images of varying resolutions, is observed to be 87.54 %. This method detects the forgery, even if the image has been entirely edited in characteristics of brightness, blurriness etc. post forgery. However, the method doesn't detects rotated or scaled copy move forged regions.

7.2 Future Work

Future work for this methodology includes implementing the methodology on a distributed environment, thereby reducing the execution time for the same. Also, research is to be done on finding the patterns of energies obtained from scaled or rotated images as compared with the original image, using the said hybrid wavelet transform. These patterns will prove to be helpful for implementing the same methodology for finding copy and move forged images where scaling or rotation has been also applied on the forged blocks.

REFERENCES

[1] Sondos M. Fadl, Noura A. Semary, and Mohiy M. Hadhoud: “Fan Search for Image Copy-Move Forgery Detection”, Second International Conference, AMLTA 2014, Cairo, Egypt, November 2014. pp 177-186.

[2] M. Sridevi, C. Mala, S. Sandeep: “Copy-Move Image Forgery Detection In a Parallel Environment”, CS & IT- CSCP 2012, pp. 19-29.

[3] Y. Huang, W. Lu, W. Sun and D. Long: "Improved DCT-based detection of copy-move forgery in images." Elsevier Forensic science international, vol. 206, pp. 178-184, no. 1, 2011.

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

[5] G. Lynch, F. Y.Shih and H.-Y. M. Liao, "An efficient expanding block algorithm for image copy-move forgery detection", Elsevier Science Inc. 2013, pp. 253-265.

[6] Jing Zhang, ZhanleiFeng, Yuting Su: “A new approach for detecting Copy- Move forgery in Digital images”, ICCS 2008, pp. 362-366.

[7] Jie Hu, Huaxiong Zhang, QiangGao, Hai Huang:“An improved lexicographical sort algorithm of Copy –Move Forgery Detection.

[8] Yang Wang, KaitlynGurule, Jacqeline Wise, Jun Zheng: “Wavelet Based Region Duplication Forgery Detection, 2012 Ninth International Conference on Information Technology –New Generations, pp. 30- 35.

[9] Tanuja K. Sarode, Naveen Vaswani: “Region Duplication Forgery Detection using Hybrid Wavelet Transforms”,International Journal of Computer Applications (0975 – 8887) Volume 90 – No 11, March 2014

[10] H. Huang, W. Guo, and Y. Zhang:"Detection of Copy-Move Forgery in Digital Images Using SIFT Algorithm," in Proceedings of IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, Vol. 2, pp. 272-276, 2008.

[11] H.B. Kekre, ArchanaAthawale, DipaliSadawarti: "Algorithm to Generate Wavelet Transform from an Orthogonal Transform", International Journal of Image Processing (IJIP), Vol.4, Issue 4, 2010.

[12] H.B. Kekre, Archana Athawale, Dipali Sadavarti: "Algorithm to Generate Kekre's Wavelet Transform from Kekre's Transform", International Journal of Engineering Science and Technology (IJEST), Vol. 2, No. 11, 2010, pp. 756-767.

[13] Dr. H.B. Kekre, Dr. Tanuja K. Sarode, SudeepThepade, Ms. SonalShroff: "Instigation of Orthogonal Wavelet Transforms using Walsh, Cosine, Hartley, Kekre Transforms and their use in Image Compression", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 9, No. 6, 2011, pp. 125-133.

[14] Dr. H.B. Kekre, Dr. Tanuja K. Sarode, SudeepThepade:
"Inception of Hybrid Wavelet using Two Orthogonal Transforms and its use in Image Compression", (IJCSIS) International Journal of Computer Science and Information Security, Vol. 9, No. 6, 2011, pp. 80-87.

[15] Image Manipulation Dataset, Unmodified/original images,https://www5.cs.fau.de/fileadmin/research/datasets/image_forensics_dataset/forensics_database/precomputed/orig_sd.zip, accessed February 2016

[16] Old Volkswagen Rusty Van,http://www.freepik.com/free-photo/old-volkwsagen-rusty-van_768803.htm, accessed February 2016

[17] Sitting on the pier, http://www.freepik.com/free-photo/sitting-on-the-pier_773436.htm, accessed February 2016

[18] Clone Detection – elsamuko, https://farm5.static.flickr.com/4094/4869570724_5439aef726_b.jpg, accessed February 2016

[19] Keyboard Detail, http://www.freepik.com/free-photo/keyboard-detail_766381.htm, accessed February 2016