# ABSTRACT

Digital images are easy to manipulate and edit due to availability of powerful image processing and editing software. Nowadays, it is possible to add or remove important features from an image without leaving any obvious traces of tampering. As digital cameras and video cameras replace their analog counterparts, the need for authenticating digital images, validating their content, and detecting forgeries will only increase. Detection of malicious manipulation with digital images (digital forgeries) is the topic of this paper. In particular, we focus on detection of a special type of digital forgery – the copy-move attack in which a part of the image is copied and pasted somewhere else in the image with the intent to cover an important image feature. In this paper, we investigate the problem of detecting the copy-move forgery and describe an efficient and reliable detection method. The method may successfully detect the forged part even when the copied area is enhanced/retouched to merge it with the background and when the forged image is saved in a lossy format, such as JPEG. The performance of the proposed method is demonstrated on several forged images.

# TABLE OF CONTENTS

# List of Figure

# List of Table

# 1. INTRODUCTION

The availability of powerful digital image processing programs, such as Photoshop, makes it relatively easy to create digital forgeries from one or multiple images. An example of a digital forgery is shown in Figure 1. As the newspaper cutout shows, three different photographs were used in creating the composite image: Image of the White House, Bill Clinton, and Saddam Hussein. The White House was rescaled and blurred to create an illusion of an out-of-focus background. Then, Bill Clinton and Saddam were cut off from two different images and pasted on the White House image. Care was taken to bring in the speaker stands with microphones while preserving the correct shadows and lighting. Figure 1 is, in fact, an example of a very realistic looking forgery. Another example of digital forgeries was given in the plenary talk by Dr. Tomaso A. Poggio at Electronic Imaging 2003 in Santa Clara. In his talk, Dr. Poggio showed how engineers can learn the lip movements of any person from a short video clip and then digitally manipulate the lips to arbitrarily alter the spoken content. In a nice example, a video segment showing a TV anchor announcing evening news was altered to make the anchor appear singing a popular song instead, while preserving the match between the sound and lip movement. The fact that one can use sophisticated tools to digitally manipulate images and video to create non-existing situations threatens to diminish the credibility and value of video tapes and images presented as evidence in court independently of the fact whether the video is in a digital or analog form. To tamper an analogue video, one can easily digitize the analog video stream, upload it into a computer, perform the forgeries, and then save the result in the NTSC format on an ordinary videotape. As one can expect, the situation will only get worse as the tools needed to perform the forgeries will move from research labs to commercial software.

Despite the fact that the need for detection of digital forgeries has been recognized by the research community, very few publications are currently available. Digital watermarks have been proposed as a means for fragile authentication, content authentication, detection of tampering, localization of changes, and recovery of original content [1]. While digital watermarks can provide useful information about the image integrity and its processing history, the watermark must be present in the image before the tampering occurs. This limits their application to controlled environments that include military systems or surveillance cameras. Unless all digital acquisition devices are equipped with a watermarking chip, it will be unlikely that a forgery-in the-wild will be detectable using a watermark. It might be possible, but very difficult, to use unintentional camera “fingerprints” related to sensor noise, its color gamut, and/or its dynamic range to discover tampered areas in images. Another possibility for blind forgery detection is to classify textures that occur in natural images using statistical measures and find discrepancies in those statistics between different portions of the image ([2], [3]). At this point, however, it appears that such approaches will produce a large number of missed detections as well as false positives. In the next section, we introduce one common type of digital forgeries – the copy-move forgery – and show a few examples. Possible approaches to designing a detector are discussed in Section 3. In Section 4, we describe the detection method based on approximate block matching. This approach proved to be by far the most reliable and efficient. The method is tested in the last Section 5 on a few forgeries. In the same section, we summarize the paper and outline future research directions


INTERNET

**Spies turn to high-tech info ops**

*PCs, Internet used for manipulating images, public opinion*

ALTERED IMAGE

1 2 3

**Altered Egos**
The top photo, of a fictitious meeting between President Clinton and Saddam Hussein, was generated by digitally altering and combining elements from three photos, shown below the altered image. Intelligence agencies plan to use similar techniques to create images and then disseminate them via the Internet in their efforts to influence events in foreign countries, such as Iraq.

**BY DANIEL VERTON**

Federal intelligence agencies are studying ways to use computers and the Internet, rather than just leaflets and radio broadcasts, to shape and disseminate information designed to sway public opinion in the world's hot spots.

As part of its so-called "perception management" program, the intelligence community has for decades created misinformation to trigger political change without direct political or military involvement in countries where the United States has vested interests, such as Iraq and North Korea.

Acting on congressional recommendations to bolster research and development in information technology, intelligence agencies are turning to PCs to develop more sophisticated means of manipulating and delivering digital photos, video clips and recorded sound to portray fictitious events in hopes of provoking desirable outcomes.

**Figure 1: Example of a digital forgery.**

## Copy-Move Forgery

Because of the extraordinary difficulty of the problem and its largely unexplored character, the authors believe that the research should start with categorizing forgeries by their mechanism, starting with the simple ones, and analyzing each forgery type separately. In doing so, one will build a diverse Forensic Tool Set (FTS). Even though each tool considered separately may not be reliable enough to provide sufficient evidence for a digital forgery, when the complete set of tools is used, a human expert can fuse the collective evidence and hopefully provide a decisive answer. In this paper, the first step towards building the FTS is taken by identifying one very common class of forgeries, the Copy-Move forgery, and developing efficient algorithms for its detection.

In a Copy-Move forgery, a part of the image itself is copied and pasted into another part of the same image. This is usually performed with the intention to make an object "disappear" from the image by covering it with a segment copied from another part of the image. Textured areas, such as grass, foliage, gravel, or fabric with irregular patterns, are ideal for this purpose because the copied areas will likely blend with the background and the human eye cannot easily discern any suspicious artifacts. Because the copied parts come from the same image, its noise component, color palette, dynamic range, and most other important properties will be compatible with the rest of the image and thus will not be detectable using methods that look for incompatibilities in statistical measures in different parts of the image. To make the forgery even harder to detect, one can use the feathered crop or the retouch tool to further mask any traces of the copied-and-moved segments.

In Figure 2, you can see a less obvious forgery in which a truck was covered with a portion of the foliage left of the truck (compare the forged image with its original). It is still not too difficult to identify the forged area visually because the original and copied parts of the foliage bear a suspicious similarity. Figure 4 shows another Copy-Move forgery that is much harder to identify visually. This image has been sent to the authors by a third party who did not disclose the nature or extent of the forgery. We used this image as a real-life test for evaluating our detection tools. A visual inspection of the image did not reveal the presence of anything suspicious.

**Figure 2: Forged test image "Jeep" (above) and its original version (below).**

# 2. OBJECTIVE OF THE STUDY

The object of this project is to identify a Copy Move Forgery Detection technique, with both feature extraction and feature matching technique, capable of obtaining better accuracy rate while maintaining the computational time seen with SURF. Also, the proposed CMFD method is compared with existing CMFD techniques in terms of performance.

# 3. LITERATURE REVIEW

There were several techniques proposed to detect image forgery in the literature of digital image forensics. Copy move forgery is one of the popular method to create the image forgery in which the part is copied and moved to the other place in the same Image. There are so many techniques to detect such type of forgeries. One approach to detect copy-move forgery detection, proposed by Fridrich et al. [3], basically performs a rigorous search by comparing the image to every cyclic-shifted versions of it. But the complexity of this approach is very high, it requires (mn) steps to execute for an image of size M×N so it is difficult to implement it practically. One of the distinguish property of copy move forgery detection is the feature extraction Process. Some methods are based on dimensionality reduction [4], [8], moments [9], [10], color properties [11], frequency domain transform [3]. Popescu et al [4] proposed a copy-move image forgery detection algorithm using block matching approach and Principal Component Analysis (PCA). In order to detect images through rotation, scaling and other operations quickly and efficiently, image tamper detection based on Radon and Fourier-Mellin transform is presented [5]. M.sridevi attempt to verify the authenticity of image using the image quality features like markov and moment based features. They are found to have their best results in case of forgery involving splicing [6]. There is a technique based on the Radon transform and phase correlation in order to improve the robustness in forgery detection. In this the proposed technique can detect forgeries even if the forged images were undergone some image processing operations such as rotation, scaling, Gaussian noise addition, etc. [7].

In the recent years many forgery detection methods and techniques are proposed. Digital watermarking is one effective tool for providing image realism and source details. Forgery detection is also offered by these digital watermarks. A number of watermarking techniques have been proposed. One uses a checksum on the image data which is embedded in the least significant bits of certain pixels [5]. Others identify the watermark by computing the spatial cross-correlation function of the sequence and the watermarked image and adding a maximal length linear shift register sequence to the pixel data [6]. Watermarks can be generated by modulating JPEG coefficients, or can be relied on image, using independent visual channels. Some watermarks are designed to be invisible, which can be blend in with camera or noise. Watermarks which are visible also do exist. In addition, a visually untraceable, strong watermarking scheme has been proposed which detects the change of a single pixel and can locate where the changes occur. The algorithms work for color images and can accommodate JPEG compression [5] [6]. The embedding of a watermark during the creation of the digital image limits its applications where the digital object generation mechanisms have built-in watermarking capabilities. Therefore, in the absence of widespread adoption of digital watermarking technology, it is necessary to resort to image forensic techniques [6]. As mentioned before, photo image forgery can include images which are tempered by copying one area in an image and pasting it onto another area. It is called as Copy-Move Forgery or Cloning [7]. It is done to add or remove some information. The second class of forgeries is copying and pasting areas from one or more images and pasting on to an image being forged this is formally referred to as image ―composition, defined as the digitally manipulated combination of at least two source images to produce an integrated result‖ by the image processing community. It is also known as Copy-Create Image Forgery [7]. Non-intrusive methods which help in dealing with copy-move forgery are commonly used and an overview on this method is given in [8]. Block matching is one of the most frequently used non-intrusive approaches for copy-move forgery detection. In block based methods, equal sized blocks are made for an image, and feature similarities between image blocks helps in detecting the tempered areas. The feature vectors are formed by extracting the features of the image. Sorting of the feature vectors are done so that related vectors are grouped together and neighboring information is analyzed. Similar feature vectors indicate that their corresponding image blocks are copies of each other [8]. In [9], the proposed algorithm shows a detection method based on matching the quantized lexicographically sorted Discrete Cosine Transform (DCT) coefficients of overlapping image blocks. After experimenting this algorithm, the results show unfailing decisions when the retouching operations are applied. However, the authors don't show robustness tests. Another method which is invariant to the presence of blur degradation, contrast changes and additive Gaussian noise is presented in [10]. Blur is used to represent the features of the image using moment invariants. The results of the experiment show that the algorithm gives an accurate result with the blurring filter and a lousy JPEG compression quality down to 70. In [11], another technique for copy-move forgery detection is discussed. This approach considers only shifting of copied regions. So, one more technique is discussed for fast-copy-move detection and a comparative study is made on both the algorithms. The computational complexity for Copy move algorithm is lower but the final result is not accurate while the Fast copy move approach is complex but produces an accurate result. But there is a disadvantage of using the second algorithm as it is not capable of detecting very small regions of the copied image. In [12] the algorithm presents the active and passive or blind techniques to detect the tempered images. Image forensics tools can be categorized into five categories: Pixel-based techniques, Format-based techniques, Camera-based techniques, Physics based techniques, Geometric-based techniques which is also mentioned above. The paper introduces Forgery detection using DCT followed by High Pass Filtering. In the proposed method the DCT is applied to the 8x8 image sub blocks which are derived after the division of the image. The paper mainly focuses on passive techniques to recognize image forgery detection. The method which is mentioned in [13] categorizes the image tampering based on different points of view, the most often performed operations in image tampering are:

- A region of the image is deleted or hided.
- A new object is added to image and information is misrepresented. Firstly, overlapping rectangular blocks are generated for the given image except the blocks where overlapping circular block are created. Secondly, some transformation technique is like DCT, PCA, DWT, SVD, LLE, etc. are used to reduce the area of search and to make the search unit as strong as possible to post processing like compression, Gaussian noise, scaling and rotation . Thirdly, after transformation the feature vectors are sorted lexicographically or using k-d tree. Comparison is made against the similarity parameters of the neighboring vectors to hint the duplication of region. The new techniques and methods that is currently available in the area of digital image forgery detection works on JPEG images only. At present, the majority of manipulation tools use JPEG standard for image compression.

# 4. REQUIREMENT ANALYSIS

## 4.1 Environmental specification:

The environmental specification specifies the hardware and software requirements for carrying out this project. The following are the hardware and the software requirements.

### 4.1.1 Hardware:

•1 GB RAM.

•320 GB HDD.

•Intel 2.4 GHz Processor core2duo

### 4.1.2 Software:-

•Windows XP Service Pack 2 / Windows 7

•Visual Studio 2008

•Windows Operating System

## 4.2 Feasibility study:

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. Three key considerations involved in the feasibility analysis are

•ECONOMICAL FEASIBILITY

•TECHNICAL FEASIBILITY

•SOCIAL FEASIBILITY

### 4.2.1 Economical Feasibility:-

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited .The expenditures must be justified. Thus the developed system as well within the budget and this was achieved because most of the technologies used are freely available.

### 4.2.2 Technical Feasibility:-

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

### 4.2.3 Social Feasibility:-

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently. The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised as he is the final user of the system.

# 5. SYSTEM DESIGN AND ARCHITECTURE

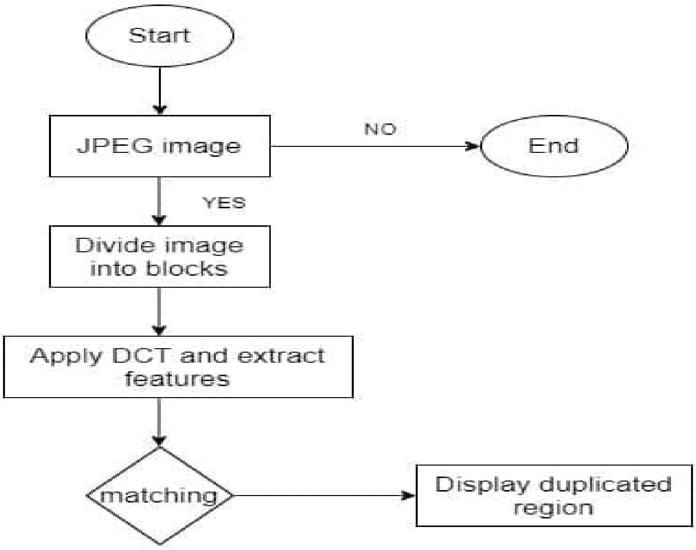


**Figure 3: Flow Chart**

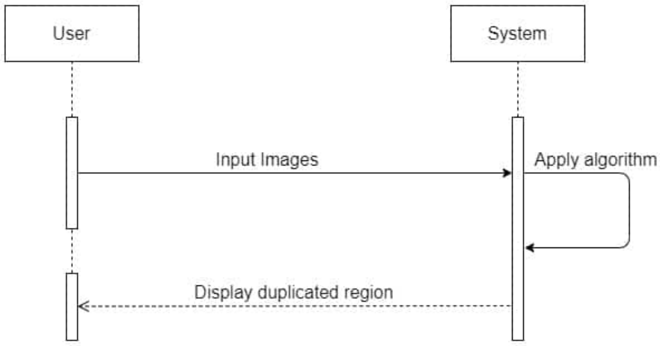


**Figure 4: Sequence Diagram**

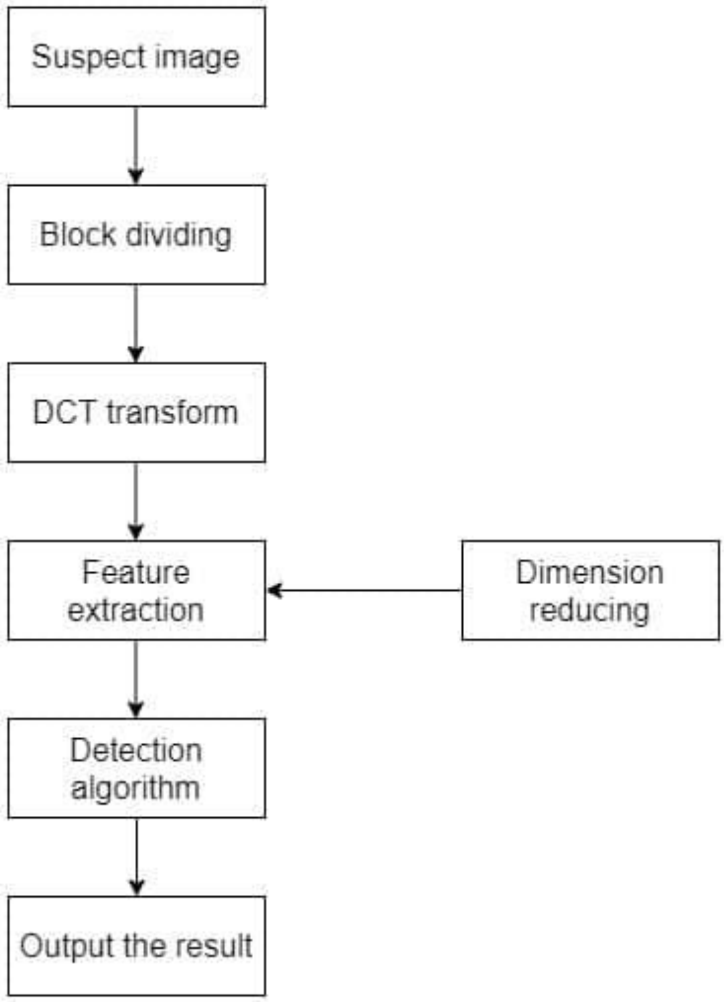


**Figure 5: System Design**

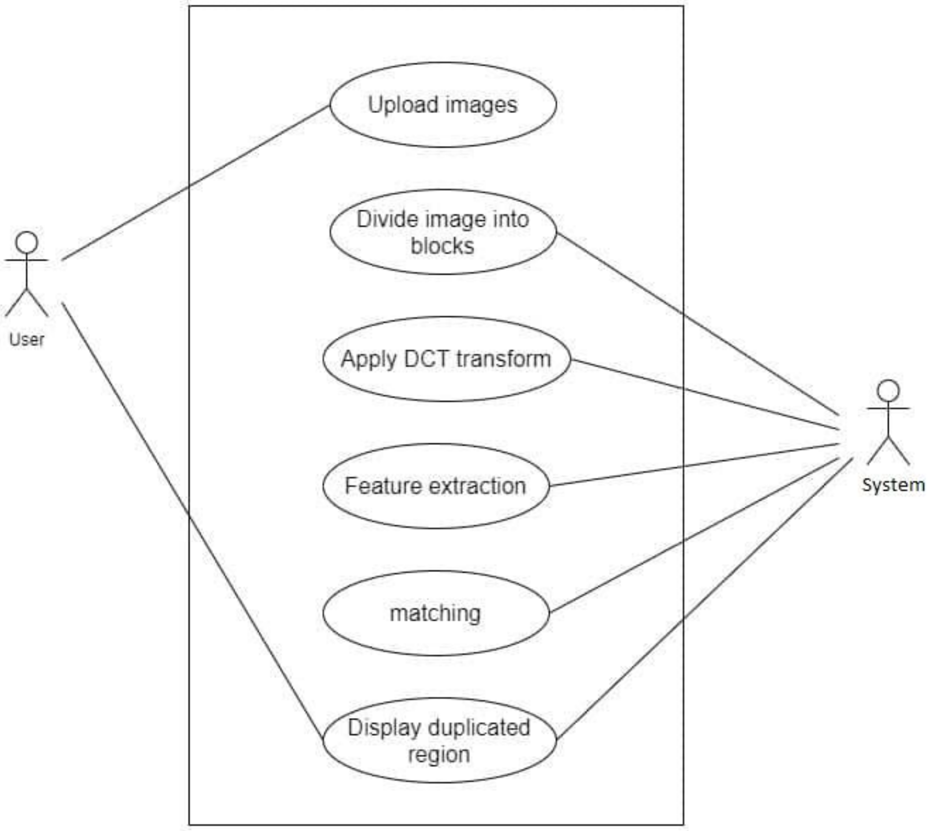


**Figure 6: Use Case Diagram**

# 6. METHODOLOGY

**Incremental Software Development Approach:**

The incremental build model is a method of software development where the model is designed, implemented and tested incrementally (a little more is added each time) until the product is finished. It involves both development and maintenance. The product is defined as finished when it satisfies all of its requirements.

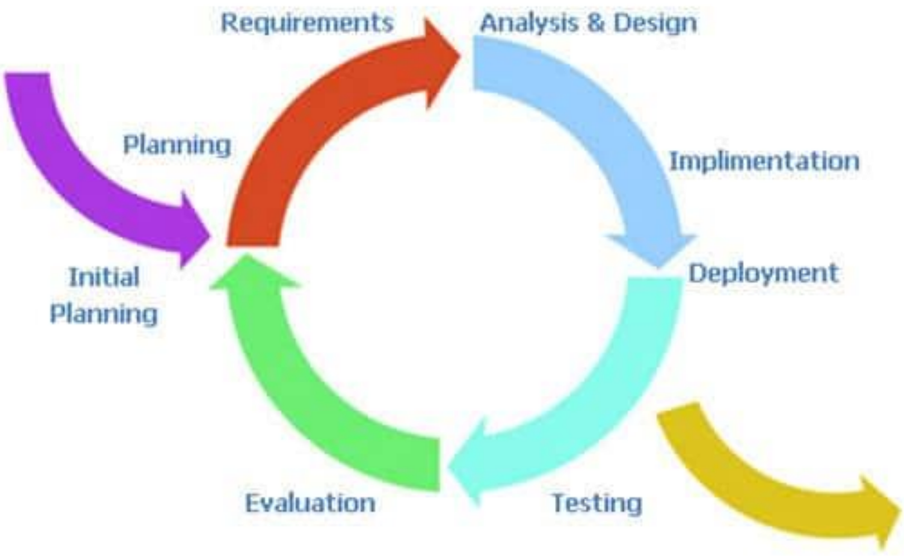


**Figure 7: Incremental Software Development Model**

**Algorithms:**

**Robust match**

The idea for the robust match detection is similar to the exact match except we do not order and match the pixel representation of the blocks but their robust representation that consists of quantized DCT coefficients. The quantization steps are calculated from a user-specified parameter Q. This parameter is equivalent to the quality factor in JPEG compression, i.e., the Q factor determines the quantization steps for DCT transform coefficients. Because higher values of the Q-factor lead to finer quantization, the blocks must match more closely in order to be identified as similar. Lower values of the Q-factor produce more matching blocks, possibly some false matches.

The detection begins in the same way as in the exact match case. The image is scanned from the upper left corner to the lower right corner while sliding a B×B block. For each block, the DCT transform is calculated, the DCT coefficients are quantized and stored as one row in the matrix A. The matrix will have (M– B+1) (N–B+1) rows and B×B columns as for the exact match case.

The rows of A are lexicographically sorted as before. The remainder of the procedure, however, is different. Because quantized values of DCT coefficients for each block are now being compared instead of the pixel representation, the algorithm might find too many matching blocks (false matches). Thus, the algorithm also looks at the mutual positions of each matching block pair and outputs a specific block pair only if there are many other matching pairs in the same mutual position (they have the same shift vector). Towards this goal, if two consecutive rows of the sorted matrix A are found, the algorithm stores the positions of the matching blocks in a separate list (for example, the coordinates of the upper left pixel of a block can be taken as its position) and increments a shift-vector counter C. Formally, let (i1, i2) and (j1, j2) be the positions of the two matching blocks. The shift vector s between the two matching blocks is calculated as

s = (s1, s2) = (i1 – j1, i2 – j2).

Because the shift vectors –s and s correspond to the same shift, the shift vectors s are normalized, if necessary, by multiplying by –1 so that s1 ≥ 0. For each matching pair of blocks, we increment the normalized shift vector counter C by one:

C (s1, s2) = C (s1, s2) + 1.

The shift vectors are calculated and the counter C incremented for each pair of consecutive matching rows in the sorted matrix A. The shift vector C is initialized to zero before the algorithm starts. At the end of the matching process, the counter C indicates the frequencies with which different normalized shift vectors occur. Then the algorithm finds all normalized shift vectors s (1), s (2)…..s (K), whose occurrence exceeds a user-specified threshold T: C(s(r)) > T for all r = 1… K. For all normalized shift vectors, the matching blocks that contributed to that specific shift vector are colored with the same color and thus identified as segments that might have been copied and moved.

The value of the threshold T is related to the size of the smallest segment that can be identified by the algorithm. Larger values may cause the algorithm to miss some not-so-closely matching blocks, while too small a value of T may introduce too many false matches. We repeat that the Q factor controls the sensitivity of the algorithm to the degree of matching between blocks, while the block size B and threshold T control the minimal size of the segment that can be detected.

For the robust match, we have decided to use a larger block size, B=16, to prevent too many false matches (larger blocks have larger variability in DCT coefficients). However, this larger block size means that a 16×16 quantization matrix must be used instead of simply using the standard quantization matrix of JPEG. We have found out from experiments that all AC DCT coefficients for 16×16 blocks are on average 2.5 times larger than for 8×8 blocks and the DC term is twice as big. Thus, the quantization matrix (for the Q-factor Q) that is used for quantizing the DCT coefficients in each 16×16 block has the following form

$$Q_{16} = \begin{pmatrix} Q'_8 & 2.5q_{18}I \\ 2.5q_{81}I & 2.5q_{88}I \end{pmatrix}, \text{ where } Q'_8 = \begin{pmatrix} 2q_{00} & 2.5q_{12} & \ldots & 2.5q_{18} \\ 2.5q_{21} & 2.5q_{22} & \ldots & 2.5q_{28} \\ \ldots & \ldots & \ldots & \ldots \\ 2.5q_{81} & 2.5q_{82} & \ldots & 2.5q_{88} \end{pmatrix}$$

and qij is the standard JPEG quantization matrix with quality factor Q and I is an 8×8 unit matrix (all elements equal to 1). We acknowledge that this form is rather ad hoc, but because the matrix gave very good performance in practical tests and because small changes to the matrix influence the results very little, we did not investigate the selection of the quantization matrix further.

Note regarding color images: In both Exact and Robust Match, if the analyzed image is a color image, it is first converted to a grayscale image using the standard formula I = 0.299 R + 0.587 G + 0.114 B, before proceeding with further analysis.

# 7. TIME SCHEDULE:

The Gantt chart shows planned and actual progress for a number of tasks displayed against a horizontal time scale. It is effective and easy-to-read method of indicating the actual current status for each of set of tasks compared to planned progress for each activity of the set. Gantt chart provide a clear picture of the current state of the project.

| S.N | Task | Month 1 | Month 2 | Month 3 | Month 4 | Month 5 | Month 6 | Month 7 |
|---|---|---|---|---|---|---|---|---|
| 1. | Planning | | | | | | | |
| 2. | Requirement Analysis | | | | | | | |
| 3. | Analysis and Design | | | | | | | |
| 4. | Coding | | | | | | | |
| 5. | Testing | | | | | | | |
| 6. | Evaluation | | | | | | | |
| 7. | Documentation | | | | | | | |

Table 1: Gantt chart

The Project is expected to complete in 7 month of total time including all major activities.

# 8. EXPECTED OUTPUT

In our project we have tampered several internet downloaded images by copying one Part of image and paste it on same image only. Our data set consist of 50 tampered images. This algorithm detects copy move forgery on these images correctly and efficiently.

# REFERENCES


[1]Tao Jing Xinghua li, Feifei Zhang, Image Tamper Detection Algorithm Based on Radon and fourier-Mellin Transform", pp 212-215 IEEE 2010.

[2]Sarah A. Summers, Sarah C.Wahl"Multimedia Security and Forensic Authentication of Digital images,"http://cs.uccs.edu/~cs525/studentproj/proj 52006/sasummer/doc/cs525projsummers Wahl.doc".

[3]J. Fridrich, D. Soukal, and J. Lukas, "Detection of Copy-Move Forgery in Digital Images", in Proceedings of Digital Forensic Research Workshop, August 2003.

[4]A. C. Popescu and H. Farad, "Exposing Digital Forgeries by Detecting Duplicated Image Regions," Technical Report, TR2004-515, Department of Computer Science, Dartmouth College, pp. 758-767, 2006.

[5]Guoqiang Shen, Lanchi Jiang, Guoxuan Zhang, "An Image Retrieval Algorithm Based on Color Segment and Shape Moment Invariants," Second International Symposium. Computational Intelligence and Design vol. 10, no.2, pp. 517-521, 2009.

[6]M .Sridevi, C.Mala and S.Sandeep "Copy – move image forgery detection", Computer Science & Information Technology (CS & IT) , Vol. 52 pp. 19–29, 2012.

[7]Hieu Cuong Nguyen and Stefan Katzenbeisser"Detection of copy-move forgery in digital images using Radon transformation and phase correlation" ,Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE, pp. 134-137,2012.

[8]X. Kang and S. Wei, "Identifying Tampered Regions Using Singular Value Decomposition in Digital Image Forensics," International Conference on Computer Science and Software Engineering, pp. 926-930, 2008.

[9]B. Mahdi an and S. Saic, "Detection of copy-move forgery using a method based on blur moment invariants.," Elsevier Forensic Science International, vol. 171, no. 2-3, pp. 180-189 Sep. 2007..

[10]S.-jin Ryu, M.-jeong Lee, and H.-kyu Lee, "Detection of Copy-Rotate- Move Forgery Using Zernike Moments," IH, LNCS 6387, vol. 1, pp. 51-65, 2010.

[11]W. Luo, J. Huang, and G. Qi, "Robust Detection of Region-Duplication Forgery in Digital Image," 18th International Conference on Pattern Recognition (ICPR'06), pp. 746-749, 2006.

[12]Amara Graps, "An Introduction to Wavelets", IEEE Computational Science and Engineering, 1992, 2(2):50-61.

[13]E. J. lentilucci, "Using the Singular Value Decomposition", Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, 29 May 2003.

[14]A.N. Myna, M.G. Venkateshmurthy, C.G. Patil, Detection of region duplication forgery in digital images using wavelets and log-polar mapping, in: ICCIMA '07: Proceedings of the International Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007), IEEE Computer Society, Washington, DC, USA, 2007, pp. 371–377.