International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 04 Issue: 10 | Oct -2017 www.irjet.net p-ISSN: 2395-0072

Image Copy-Move Forgery Detection Using Block Matching Probabilities

Rituja D. Akojwar[1], Prof A. P. Khandait[2]

[1]*MTech, Dept. of Electronics Engineering, PCE, Maharashtra, India*
[2]*Professor, Dept. of Electronics Engineering, PCE, Maharashtra, India*

Abstract – *Digital image plays an important role in expressing and transmitting visual information. It is the data representing the two dimensional scene. But in today's world these images can be easily tampered and synthesized with the availability of powerful tools and software such as Adobe Photoshop, coral draw etc. Such image tampering is very dangerous as it can be used for misleading the general public, as false evidence in court, or to hide a very important data. This motivates the need to develop image tampering detection tools. Recovering people's confidence has become very much essential. In this paper we propose a detection technique to detect copy-move forgery also known as cloning where the part of image is copied and pasted into the same image. In the proposed method we firstly divide the image into overlapping blocks. Then use fusion methodology for feature extraction, and Support Vector Machine for feature matching whose output is the probability of each matched block and then compare it with a threshold value*

Key Words: **Digital image, Copy-Move forgery, Cloning, Fusion methodology, Support Vector Machine, Threshold.**

1. INTRODUCTION

Images are natural carrier of information. Nowadays they play an important role in our community as they are used in wide variety of application such as in military purpose, surveillance system, insurance processing, internet, TV, advertisement media, forensic investigation etc. But due to availability of powerful, low cost image editing tools these images can be easily tampered. Therefore, the authenticity of images has become questionable. There are mainly 3 methods in which image can be tampered. First splicing, second is copy-move forgery or cloning and third is image retouching.

1.1 IMAGE SPLICING

The splicing is the technique where the part of an image is copied and pasted into another image. This can alter the visual message of the image more aggressively than image retouching or the copy move forgery.

The fig 1. depicts the image slicing technique. It shows the two original image and the spliced image. First image is the original image of sunset and second is the original image of a man surfing. For the third, the second original image is copied and pasted into the first original image giving rise to spliced image.

Fig 1. Original and spliced image i) Original image of sunset ii) Original image of man surfing iii) Spliced image.

1.2 COPY-MOVE FORGERY

This is the most common kind of image tampering technique used, where one needs to cover a part of the image in order to add or remove information. In this tampering a part of an image is copied and pasted into the same image.

Fig 2. shows image copy-move technique. First is the original image and the part of that image is copied and pasted into the original image giving rise to copy-move forgery.

Fig 2. Example of copy-move forgery ;(left) original Image; (right) tampered image with two sun

1.3 IMAGE RETOUCHING

This type of forgery enhances the image by adjusting colors, contrast, noise, sharpness etc. It is less harmful than other two type of forgery. It is widely used by magazine photo editors.

Fig 3. shows the original image which contains spots on face whereas there is retouched image which vanishes all the spots.

Fig 3. Example of image retouching ;(left) original Image; (right) retouched image

2. METHODOLOGY

In this paper we have developed a detection technique to detect the copy-move forgery where part of the image is copied and pasted into the same image to hide or create false information. The Fig 4. shows the block diagram of methodology used in the paper

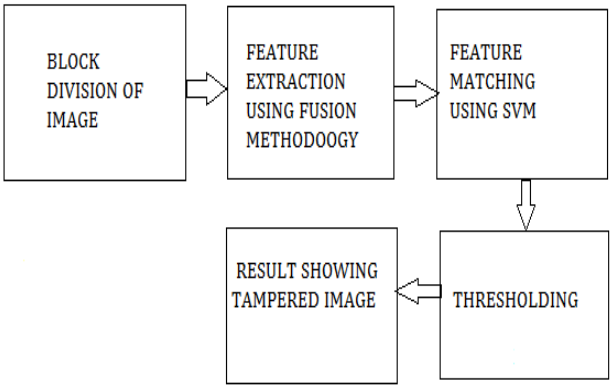

Fig 4. Block diagram

Consider the block diagram shown in Fig 4. Initially an image is divided into a number of overlapping blocks. Next step is feature extraction. There are mainly two feature extraction techniques, first is block based method and second is key-point based method. In this paper we have used three block based methods and one key point based method, that is four detectors are applied simultaneously to extract the features of an image. The detectors used for block based methods are 1) edge mapping 2) colour mapping 3) morphological features whereas for key-point based method used is SURF. Next the Support Vector Machine method (SVM) is used for feature matching. The output of this matching is the probability of the matched blocks. Then a threshold value is applied and blocks having probability greater than the threshold value are considered to be tampered. The brief description of each block is given below.

2.1 BLOCK DIVISION OF IMAGE

The image is divided into a number of overlapping blocks. Fig 5. Shows the image and its overlapped block division.

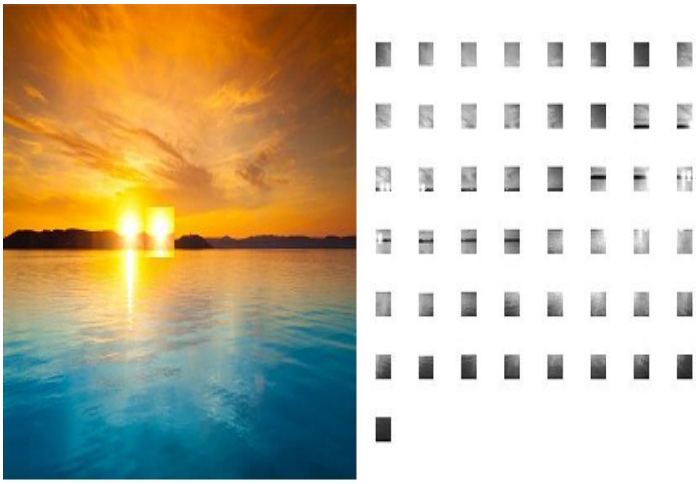

Fig 5. Example of block division of an image

2.2 FEATURE EXTRACTION

Every image contains information or features not visible to human eye. With the help of feature extraction techniques, we extract this important features. Described below are the feature extraction techniques used in the paper.

EDGE MAPPING: Edge detection is the name for a set of mathematical methods which target at classifying points in a image at which the image intensity varies sharply. There are two types of edge detection techniques they are laplacian and gradient. In this paper we have used the edge detection using the Extended Epanechnikov function

For extended Epanechnikov [1] functions a fuzzy set is described by an infinite number of membership functions at the same time a weakness and strength: uniqueness is sacrificed at the advantage of flexibility, thus making the "adjustment" of a fuzzy model possible. On the four-dimensional feature space it define the fuzzy set membership functions for the six classes as extended Epanechnikov functions by Equation's. for any input feature vector ***x***. The extended Epanechnikov functions are shown here with small diameters for clarity. In practice, they overlap so that each input feature vector falls into one or more of the fuzzy set membership functions. Such functions are dome shaped.

$$\mu_0(x) = max\left\{0,1 - \frac{\|x-c0\|}{w}\right\} for\ class\ 0$$

$$\mu_0(x) = max\left\{0,1 - \frac{\|x-c1\|}{w}\right\} for\ class\ 1$$

$$\mu_0(x) = max\left\{0,1 - \frac{\|x-c2\|}{w}\right\} for\ class\ 2$$

$$\mu_0(x) = max\left\{0,1 - \frac{\|x-c3\|}{w}\right\} for\ class\ 3$$

$$\mu_0(x) = max\left\{0,1 - \frac{\|x-c4\|}{w}\right\} for\ cass\ 4$$

$$\mu_0(x) = max\left\{0,1 - \frac{\|x-c5\|}{w}\right\} for\ cass\ 5$$

COLOR MAPPING: A colour mapping may be referred to as the algorithm that results in the mapping function or the algorithm that transforms the image colour. It gives the colour value of each pixel. The histogram is plotted between the number of pixel of that colour (Y- axis) to the gray levels (X- axis). Vector quantization (VQ) is one of the bases of palette-based images (color quantized images), which divides the color space into finitely quantized cells (by LBG or minimization of variance) as shown in Figure 1. The colors falling in those cells are represented by some colors (centroids). This is also called color quantization [2,3] as it truncates all colors of the 24-bit color image to a finite number of colors (equal to number of quantization cells) which is combined with spatial distribution provided by dithering based on error diffusion algorithm.

MORPHOLOGICAL FEATURES: Morphological image processing is a collection of non-linear operations related to the shape or morphology of features in an image. Morphological techniques probe an image with a small shape or template called a structuring element

Mean value:

$$\bar{x} = \frac{x_1 + x_2 \ldots\ldots + x_n}{n}$$

Standard deviation :

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$$

SURF: The SURF detector focuses its attention on blob-like structures in the image. These structures can be found at corners of objects, but also at locations where the reflection of light on specular surfaces is maximal (i.e. light speckles).

The SURF detector algorithm can thus be summarized by the following steps:

1. Form the scale-space response by convolving the source image using DoH filters with different σ
2. Search for local maxima across neighbouring pixels and adjacent scales within different octaves
3. Interpolate the location of each local maxima found
4. For each point of interest, return x, y, σ, the DoH magnitude, and the Laplacian's sign.

The SURF feature detector is based on the Hessian matrix because of its good performance in accuracy. The Hessian matrix is defined as H(x, σ) for a given point $\boldsymbol{x} = (x, y)$ in an image as follows

$$H(p,\sigma) = \begin{bmatrix} L_{xx}(p,\sigma) & L_{xy}(p,\sigma) \\ L_{yx}(p,\sigma) & L_{yy}(p,\sigma) \end{bmatrix}$$

where Lxx (x,σ) is the convolution of the Gaussian second order derivative g(with the image I in point x and similarly for Lxy (x,σ) and Lyy (x,σ). These derivatives are called as Laplacian of Gaussians.

The SURF descriptor is extracted from an image in two steps : the first step is assigning an orientation based on the information of a circular region around the detected interest points. The orientation is computed using Haar wavelet responses in both x and y direction. Once the Haar-wavelet responses are computed, they are weighted with a Gaussian with σ = 2.5s centered at the interest points. In a next step the dominant orientation is estimated by summing the horizontal and vertical wavelet responses within a rotating wedge which covering an angle of π/3 in the wavelet response space. The resulting maximum is then chosen to describe the orientation of the interest point descriptor.

In a second step, the region is split up regularly into smaller square sub-regions and a few simple features at regularly spaced sample points are computed for each sub-region. The horizontal and vertical wavelet responses are summed up over each sub-region to form a first set of entries to the feature vector. The responses of the Haar wavelets are weighted with a Gaussian centred at the interest point in order to increase robustness to geometric deformations and the wavelet responses in horizontal dx and vertical Directions dy are summed up over each sub-region. Furthermore, the absolute values |dy| and |dx |are summed in order to obtain information about the polarity of the image intensity changes. Therefore each sub-region has a four-dimensional descriptor vector

2.3 FEATURE MATCHING

The features obtained with the feature extraction methods are matched in this step. That is each block is compared with other blocks and blocks having similar features are considered to be tampered. In this project we have used Support Vector Machine algorithm for feature matching.

Support Vector Machine (SVM) is primarily a classifier that performs classification tasks by constructing hyper-planes in a multidimensional space separating cases of different class labels. According to SVM the decision boundary should be as far away from the data of both classes as possible. Let us consider that we have data points belonging to two classes, A+ and A– . Each point in the dataset comes with a class label y, +1 or 1, indicating one of two classes A+ and A– .[4, 5]

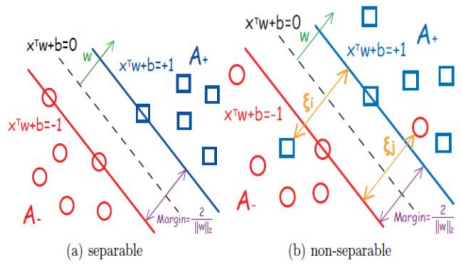

Fig 6: Linear SVM for separable and non separable data

2.4 THRESHOLDING

In this step a threshold value is selected which ranges between 0 to 1. The blocks having probability less than the threshold will be ignored and the blocks having probability above the threshold will be termed as tampered image

3. OBSERVATION AND RESULT

The experiment is performed in MATLAB . The image is first compressed to size 512*512 then image is divided into overlapping blocks of size 128*128. The result showed the following parameters.

PARAMETERS	VALUES
TPR	100%
FPR	42%
ACCURACY	79%
PRECISION	70.4%
F-MEASURE	82.3%
RUNNING TIME	900s

Table-1: Parameter table

It was observed that our algorithm has improved the True Positive Rate factor which give rise to high F-measure and improved accuracy. But due to False Positive Rate being high the precision is affected which is high compared to the existing methods.

4. CONCLUSION

Image tampering detection is a hard problem to solve as it involves different methodologies and abilities. This way, it is impossible that just one image tampering detection approach reveals perfectly an image manipulation. Also, any given image manipulation detector might be deceived by anti-forensic operations created by a forger. To address this problem, we explored approaches to combine methods that take the best of two worlds in the copy-move detection problem: block-based and points of interest detection methods. In this sense, the combination of different detectors is promising and paramount, as it can explore complementary properties from the combined detectors. The proposed approaches have shown to perform better than existing ones for fusion and for individual detectors

REFERENCES

[1] C.G. Looney, "A Fuzzy Classifier Network with Ellipsoidal Epanechnikovs", Technical Report, Computer Science Department, University of Nevada, Reno, NV, 2001.

[2] P. Heckbert, (1982),"Color image quantization for frame buffer display", Computer Graphics, 16, 3, 297-307.

[3] . M. T. Orchard, C. A. Bouman, (1991)," Color Quantization of Images", IEEE Trans. Signal Processing, 39, 12, 2677-2690.

[4] Yuh-Jye Lee, Yi-Ren Yeh, and Hsing-Kuo Pao, "An Introduction to Support Vector Machines", National Taiwan University of Science and Technology, Taipei, Taiwan

[5] Chih-Chung Chang, Chih-Wei Hsu, and Chih-Jen Lin," The analysis of decomposition methods for support vector machines". IEEE Transactions on Neural Networks, 11(4):10031008, 2000