International Journal of Pure and Applied Mathematics
Volume 117 No. 15 2017, 1077-1085
ISSN: 1311-8080 (printed version); ISSN: 1314-3395 (on-line version)
url: http://www.ijpam.eu
Special Issue

ijpam.eu

A SURVEY ON RECOGNITION OF IMAGE FORGERY USING GRAY LEVEL CO-OCCURRENCE MATRICES (GLCM) EXTRACTION

[1] *Clara Shanthi.G First- Research Scholar CSE Department,*
Dr.M.G.R.Educational And Research Institute University, Tamil Nadu, India
clarashanthi2@gmail.com
[2] *V.Cyril Raj, Second-Professor CSE & IT,*
Dr.M.G.R.Educational And Research Institute University, Tamil Nadu, India
cyrilraj@drmgrdu.ac.in

Abstract

We are certainly living in an era where we are vulnerable to an incredible array of visual imagery. Image forgery recognition is developing as one of the major research topic among researchers in the area of image forensics. This image forgery detection is addressed by two different types: (i) Active, (ii) Passive. Further consist of some different methods, such as Copy-Move, Image Splicing, and Retouching. In this paper, efficient forgery detection and classification technique is proposed by three different stages. At first stage, preprocessing is carried out using bilateral filtering to remove noise. At second stage, extract unique features from forged image by using efficient feature extraction technique namely Gray Level Co-occurrence Matrices (GLCM). Finally, forged image is detected by classifying the type of image forgery using Multi Class- Support Vector Machine (SVM). Also, the performance of the proposed method is analyzed using the following metrics: accuracy, sensitivity and specificity.

Keywords

Image Forgery Detection, Image splicing, Image Retouching, Copy-Move, Support Vector Machine, Gray Level Co-occurrence Matrices.

1. Introduction

From last decades, the image forgery detection has been emerged as a remarkable research in applications of computer vision, digital image processing, biomedical technology, criminal investigation, image forensics, etc. The Image Forgery Detection (IFD) techniques are classified into two types [1]. The classifications IFD technique is shown in following *figure1*

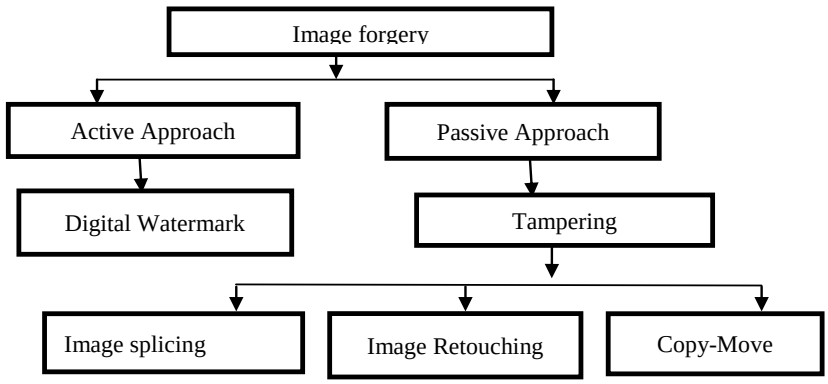

Figure: 1 Forgery detection techniques

The image forgery majorly classified in to active and passive approach. A digital watermark is a kind of marker covertly embedded in a noise tolerant signal

such as an audio, video or image data Image splicing, Due to the differences of tampering manipulations, the corresponding passive detection techniques are proposed to detect forgery images [3]. This paper focuses on algorithms and techniques for image Copy-Move, Retouching and image spliced detection techniques

1.1 Image Retouching

This technique is popular in magazine photo edition. This type of image forgery is present in almost all magazines cover that employed technique to improve certain features of image; also it is an attractive method [2].

Figure: 2 Image Retouching Attack on Images

There is lot of changes between first and second part of the image.

1.2 Copy –Move

In this technique, desired part of image is possible to add or remove information. In a copy-move attack, the intention is to hide something in the original image with some other part of the same image [3].

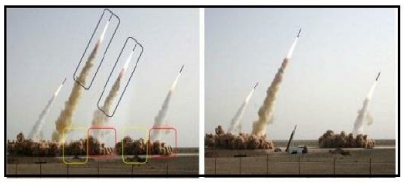

Figure: 3 Copy-Move Attacks on Images

1.3 Image splicing

This method is used to make a more aggressive forgery images. It is a simple process and can be crops and paste the desired regions from other sources.For example the following figure shows the image splicing based forgery image.

Figure: 4 Image splicing attack on Images

It can be used to copy a spliced portion from original image to desired image [4]. This paper presents an efficient forgery detection and classification

techniques are preprocessing, feature extract technique namely Gray Level Co-occurrence Matrices (GLCM). Lastly, forged image is detected by classifying the type of image forgery using Multi Class- Support Vector Machine (SVM).

2. Literature Review

Baby, L.,& Jose, A. (2014) explained about the forgery recognition based on texture and edge features. In this paper, a forensic technique that focus on inconsistencies in the illumination color of images. GLCM is a statistical method of by considering the spatial relationship of pixels. Then classifications of face pairs are performed using KNN classifier.

Dhevana, S., & Jayasri, C., (2015) explained to identify the forgeries in the image with the help of the GLCM. Usually the image handling procedures engaged on the digital images induces the interfering on the image, where the noise will also get joined with the image.

Although the researches Xuanjing Shen et al (2016) described Splicing Image Forgery Detection using Textural Features based Gray Level Co-occurrence Matrices(TF-GLCM). In TF-GLCM, the GLCM is measured by using Difference Block Discrete Cosine Transform (DBDCT) arrays that is used to capture the textural information and the spatial relationship among the image pixels sufficiently.

3. Forgery detection framework

The proposed functional arrangement of image forgery detection process is shown in following figure. Hence the process requires major stages such as Image Preprocessing, Feature Extraction, and Classifier. Image pre-processing is also an improvement stage of the image data, which can suppress the unwanted distortions and to enhance the important features of the remaining processing.

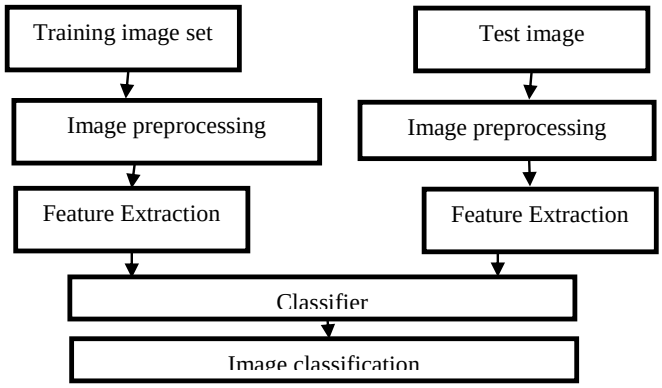

Figure: 5 block diagram for proposed Forgery detection

In this research, the process of image preprocessing includes two major steps: fast bilateral filtering and canny edge detection. Edge detection output is an input for

feature extraction; the input is given through image preprocessing.

3.1 Bilateral Filter

The bilateral filter is used to remove the noise and reduces the blurriness of color images. It is employed to filter the noise range of image which acts as a domain filter [6]. The range filtering select values based on the desired amount of low pass filtering. A low pass domain filter image is derived as

$$\mathrm{I(x):k_d\,(c)} = \iint_{\infty}^{\infty} \mathrm{c(x,y)dy} \quad (1)$$

Where kd(x) - normalized constant;
C(x, y) is used calculate the geometric closeness among neighborhood center x &y,
h(x) -output image;

$$\mathrm{h(x)} = \mathrm{k_d^{-1}} \iint_{-\infty}^{\infty} \mathrm{i(y)c(x,y)dy} \quad (2)$$

$$\mathrm{k_r\,(x)} = \iint_{-\infty}^{\infty} \mathrm{s(i(x)i(y)dy} \quad (3)$$

s(f(x),f(y)) calculated the photometric similarity of neighborhood pixel center x & y.

$$\mathrm{h(x)} = \mathrm{k_r^{-1}} \iint_{-\infty}^{\infty} \mathrm{i(y)} * \mathrm{s\big(i(x),i(y)\big)} * \mathrm{c(x,y)dy} \quad (4)$$

$$\mathrm{k_r\,(x)} = \iint_{-\infty}^{\infty} \mathrm{s(i(x),i(y)} * \mathrm{c(x,y)dy} \quad (5)$$

The output image is redefined by,

$$\mathrm{s(i(x),i(y)} = \mathrm{e}^{\frac{-(\mathrm{i(y)-i(x)})^2}{2\sigma_r^2}} \quad (6)$$

$$\mathrm{c(x,y)} = \mathrm{e}^{\frac{-(\mathrm{i(y)-i(x)})^2}{2\sigma_r^2}} \quad (7)$$

Pixel location (x,y) &output I(x,y) of bilateral filtermeasured as follows:

$$\mathrm{I'(X,Y)} = \in \mathrm{y}_{\in \mathrm{N(X)}} \mathrm{e}^{\frac{-(\mathrm{i(y)-i(x)})^2}{2\sigma_r^2}} \mathrm{e}^{\frac{-(\mathrm{i(y)-i(x)})^2}{2\sigma_D^2}} \mathrm{i(x,y)}$$

Where σd&σr - spatial & intensity of fall-off of weights control parameter,
N(x) - pixel of I(x), - geometric spread parameter, used to select the required low pass filter.

3.2 Edge Detection

Edge detection technique is used to reduce the image size and filter-out the information. Various edge detection techniques are used for efficient operation such as Canny, Sobel, Prewitt, Roberts, and so on. Among the various methods, 'Canny method' is mostly preferable for detect both strong and weak edges. Hence the 'σ' is used to control size of the Gaussian filter. Small value of 'σ' suggests a smaller Gaussian filter and that can limit the amount of blurring in image.

3.3 Feature Extraction

3.3.1 Gray level Co-occurrence Matrices (GLCM)

GLCM is one of the popular statistical methods used to measure the textural information of images. GLCM is also called as two-dimensional

dependent matrix which can reflect the broad information of direction and spatial relationship among image pixel & nearest neighbors [7]. Let f (p, q) - image source, and it could be analyzing the size of A×B; gray value of image pixel is quantized to B_g levels. Hence the GLCM function is derived in Equation (9). Consider the feature dimension & the proportional relation of image.

$$G = \begin{bmatrix} h(1,1) & \cdots & h(1,Bg) \\ \vdots & \ddots & \vdots \\ h(Bg,1) & \cdots & h(Bg,Bg) \end{bmatrix} \quad (9)$$

Where h (i, j) shows the relative frequencies; i and j- pixel couplevalues of image.h(i, j) can be derived as follow:

$$h(i,j) = \#\{(p_1,q_1),(p_2,q_2) \in A * B \setminus \ f(p_1,q_1) = i, f(p_2,q_2) = J\} \quad (10)$$

Where # - number of elements in a set.

The different combinations of distance (denoted by d) and angles (denoted by θ) between the two pixels can influence the way of calculating the number of the pixel couples in GLCM. In TF-GLCM, the distance is set at 1 and the angles are 0°, 45°, 90°, and 135°. After being normalized, for instance, the four GLCM extracted from the horizontal difference coefficient array can be given by

$$G_{d=1,\theta=0^*}(F_H(i,j)), \ldots . G_{d=1,\theta=135^*}(F_H(i,j)) \quad (11)$$

Although the researcher Haralick has extracted totally twenty four parameters from co-occurrence matrix, but selectively seven parameters are commonly used which has given as follows: energy, entropy, contrast, local homogeneity, correlation, cluster shade & prominence.

3.3.1.1 Energy: It is called as uniformity or angular second moment also it is used to measure the textural uniformity that is pixel pair repetitions.

$$energy(E1) = \sum i \sum j \, p_{ij}^2 \quad (12)$$

Detects the disorders in textures.

Energy reaches a maximum value equal to one. Energy (E1) is the sum of the square of the elements in the GLCM, which can measure textural uniformity.

3.3.1.2 Entropy: Entropy can be used to measure the disorder or complexity of desired images. The following entropy equation shows the strong formation and the complex textures tends to high entropy

$$entropy\ (ent) = -\sum i \sum j \, p_{ij} \, log_2 p_{ij} \quad (13)$$

3.3.1.3 Contrast: Contrast is used to measures the difference between the lowest and highest values of a contiguous set of pixels

$$constract(con) = \sum i \sum j \, (i-j)^2 \, pij \quad (14)$$

3.3.1.4 Homogeneity: Homogeneity is also called as inverse difference moment. Homogeneity decreases if contrast increases while energy is kept constant.

$$homogeneity\ (hom) = \sum i \sum j \frac{1}{1+(i-j)^2} p_{ij} \quad (15)$$

3.3.1.5 Correlation: $\text{correlation} = \sum_{I,J=0}^{N-1} P_{ij} \frac{(i-\mu)(j-\mu)}{\sigma^2}$ (16)

Correlation is used to measures the joint probability occurrence of pixel pairs and similarity of two images.

Where P_{ij} - texture image pixel value in a position of (i,j), N - Number of image gray levels,

$\mu = \sum_{i,j=0}^{N-1} i p_{ij}$; μ - texture image , $\sigma^2 = \sum_{i,j=0}^{N-1} p_{ij} (i-\mu)^2$; σ^2- variance of the texture image.

3.3.1.6 Variance

This static is a measure of heterogeneity and is strongly correlated to heterogeneity and is strongly correlated to first order statistical variable such as standard deviation.

Variance increases when the gray level values differ from their mean.

$$variance\ (var) = \sum i \sum j\ (i-j)^2\ p_{ij} \quad (17)$$

Where, μ -mean of pixel value in position (p_{ij})

Hence Difference variance = variance of P_{X-Y}

Difference of entropy = $\sum_{l=0}^{N-1} p_{x-y}(i) log\{p_{x-y}(i)\}$ (18)

3.3.1.7 Cluster shade

$$S_{CS} = \sum_{i,j=1}^{N} (i - M_x + j - M_y)^3\ p(i,j) \quad (19)$$

Skewness of the matrix is measured under cluster shade and the lack of symmetry is called as cluster shade feature.

3.3.1.8 Cluster prominence

$$S_{Cp} = \sum_{i,j=1}^{N} (i - M_x + j - M_y)^4\ p(i,j) \quad (20)$$

Where, $M_X = \sum_{i,j=1}^{N} i p(i,j)$; $M_y = \sum_{i,j=1}^{N} j p(i,j)$;

$$mean(m) = \sum_{i=0}^{l-1} z_i\ p(z_i) \quad (21)$$

m- Average intensity.

If the cluster prominence is low, peak co-occurrence matrix around the mean values.

3.4 Classifier

Multi Class-Support Vector Machine (MC-SVM) classifier is employed to classify the forgery detection. MC-SVM classifier is constructed by f 1, f 2, . . $f_{M,}$ each set is trained by each class. Muti-class maximum output is derived from g^i (x):

$$g^j(x) = \sum_{i=1}^{m} y_i \alpha_i^j\ k(x, x_i) + b^j \quad (22)$$

The detection of forged image is done using SVM classifier on systematic basis by designing a simple process consisting of two phases, namely training phase and testing phase.

4. Results and Discussion

The proposed set of experiments has established the better performance and effectiveness of image forgery detection using GLCM.

4.1 Accuracy: It is a ratio of number of correct assessment to the total number of assessments. It is measured as the term of (%)

$$\text{Accuracy } (\%) = \frac{((TN + TP))}{((TN + TP + FN + FP} \quad (23)$$

4.2 Sensitivity: It is defined as the ratio of the number of true positive assessments to the total number of true positive and false negative assessments.It is a degree of positive values that are correctly recognized.

$$\text{Sensitivity } (\%) = \frac{\text{TP}}{((\text{TP}+\text{FN}))} \quad (24)$$

4.3 Specificity: It is defined as the ratio of the number of true negative assessments to the total number of true negative and false positive assessments.

$$\text{Specificity } (\%) = \frac{\text{TN}}{((\text{TN}+\text{FP}))} (25)$$

5. Conclusion

This paper presents an efficient forgery detection and classification technique that can propose three major stages. At first stage, preprocessing is carried out using bilateral filtering to remove noise. At second stage, extract unique features from forged image by using efficient feature extraction technique namely Gray Level Co-occurrence Matrices (GLCM). Here, the GLCM improved the feature extraction accuracy. Lastly, forged image is detected by classifying the type of image forgery using Multi Class- Support Vector Machine (SVM). Also,the performance of the proposed method is analyzed using metrics such as accuracy, sensitivity and specificity.

6. References

1. Ardizzone, E., Bruno, A., &Mazzola, G. (2015). Copy–move forgery detection by matching triangles of key points. IEEE Transactions on Information Forensics and Security, 10(10), 2084-2094.
2. Giuseppe Cattaneo, Gianluca Roscigno, (2014). A Possible Pitfall in the Experimental Analysis of Tampering Detection Algorithms", the 17th International Conference on Network-Based Information Systems (NBiS), Salerno, Italy.
3. Baby, L., & Jose, A. (2014) Digital Image Forgery Detection Based on GLCM and HOG Features.
4. Pun, C. M., Yuan, X. C., & Bi, X. L. (2015). Image forgery detection using adaptive over segmentation and feature point matching. IEEE Transactions on Information Forensics and Security, 10(8), 1705-1716.
5. Hsu, C. M., Lee, J. C., & Chen, W. K. (2015, May). An efficient detection algorithm for copy-move forgery. In Information Security (AsiaJCIS), 10th Asia Joint Conference on (pp. 33-36).
6. Chaudhury, K. N., & Dabhade, S. D. (2016). Fast and provably accurate bilateral filtering. IEEE Transactions on Image Processing, 25(6), 2519-2528.
7. Shen, X., Shi, Z., & Chen, H. (2016). Splicing image forgery detection using textural features based on the grey level co-occurrence matrices. IET Image Processing, 11(1), 44-53.
8. Rajesh, M., and J. M. Gnanasekar. "Annoyed Realm Outlook Taxonomy Using Twin Transfer Learning." International Journal of Pure and Applied Mathematics 116 (2017): 547-558.
9. Rajesh, M. & Gnanasekar, J.M. Wireless Pers Commun (2017),https://doi.org/10.1007/s11277-017-4565-9

10. Rajesh, M., and J. M. Gnanasekar. "GCCover Heterogeneous Wireless Adhoc Networks." Journal of Chemical and Pharmaceutical Sciences (2015): 195-200.
11. Rajesh, M., and J. M. Gnanasekar. "CONGESTION CONTROL IN HETEROGENEOUS WANET USING FRCC." Journal of Chemical and Pharmaceutical Sciences ISSN 974: 2115.
12. Rajesh, M., and J. M. Gnanasekar. "GCCover Heterogeneous Wireless Ad hoc Networks." Journal of Chemical and Pharmaceutical Sciences (2015): 195-200.
13. Rajesh, M., and J. M. Gnanasekar. "CONGESTION CONTROL USING AODV PROTOCOL SCHEME FOR WIRELESS AD-HOC NETWORK." Advances in Computer Science and Engineering 16.1/2 (2016): 19.
14. RAJESH, M. "TRADITIONAL COURSES INTO ONLINE MOVING STRATEGY." The Online Journal of Distance Education and e-Learning 4.4 (2016).
15. Rajesh, M. "Object-Oriented Programming and Parallelism."
16. Rajesh, M., K. Balasubramaniaswamy, and S. Aravindh. "MEBCK from Web using NLP Techniques." Computer Engineering and Intelligent Systems 6.8: 24-26.