COPY MOVE FORGERY DETECTION USING GLCM BASED STATISTICAL FEATURES

Gulivindala Suresh[1] and Chanamallu Srinivasa Rao[2]

[1]Department of ECE, JNTU-K UCE, Kakinada, AP, INDIA
[2]Department of ECE, JNTU-K UCE, Vizianagaram, AP, INDIA.

ABSTRACT

The features Gray Level Co-occurrence Matrix (GLCM) are mostly explored in Face Recognition and CBIR. GLCM technique is explored here for Copy-Move Forgery Detection. GLCMs are extracted from all the images in the database and statistics such as contrast, correlation, homogeneity and energy are derived. These statistics form the feature vector. Support Vector Machine (SVM) is trained on all these features and the authenticity of the image is decided by SVM classifier. The proposed work is evaluated on CoMoFoD database, on a whole 1200 forged and processed images are tested. The performance analysis of the present work is evaluated with the recent methods.

KEYWORDS

GLCM, CMFD, SVM Classifier, Detection rate

1. INTRODUCTION

Digital images have a significant role in conveying the information. Digital Image manipulation became very easy with the availability of advanced photo editing tools. But, due to the manipulation the trustworthiness of digital images is lost. Hence, detection of image forgery is important and is achieved in passive mode without embedding any signature in the original image. Passive image forgery detection works on the discrepancies in the statistical features of the forged image. Copy-Move tampering is a very common method of tampering digital image where in some portion of an original image is copied and pasted at some other location in the same original image. In general, this is done with intent to conceal a region in the image. The copied portions are within the image, so the changes in texture, variations in intensity or any statistical property may match with the remaining portion of the original image. Hence, it is challenging for detecting the forged portion based on HVS [1]. An exhaustive search can be used to identify the significant features of copied and pasted portions on the tampered image. This mechanism needs more time for detection and is computationally complex [2]. Therefore, similarity measure can be used on the identical image regions for detecting the forgery successfully [2]. Figure 1(a) and 1(b) illustrates Copy move forgery.

DOI: 10.5121/ijci.2016.5419

a.Original Image b. Copy-Move Forged Image

Figure 1. Illustration of copy-move forgery

A comprehensive report on passive methods for forgery detection in images is available in [3]. Here, the works based on textural features are reviewed. Shikha Dubey et al. [4] used local descriptors for textural features and block matching is performed using clustering technique. In [5], the Gabor magnitude of the image is computed and a histogram is formed as a feature vector. Gabor Wavelets and Local Phase Quantization [6] are used to extract texture features for image forgery detection. In [7], features are extracted based on GLCM and Histogram of Oriented Gradient (HOG) and KNN classifier is used for image forgery detection.

2. Methods

2.1. GLCM

GLCM is the key process of this work. The Gray Level Co-occurrence Matrix (GLCM) provides information on the occurrence of various combinations of pixel intensities in a gray image. It is a statistical approach [8] of exploring the spatial relationship among pixels. GLCM computes in what a way a pixel with intensity i occur horizontally, vertically or diagonally to a pixel with intensity j.

GLCM exhibits certain properties regarding the spatial relationships of gray intensities in the image.

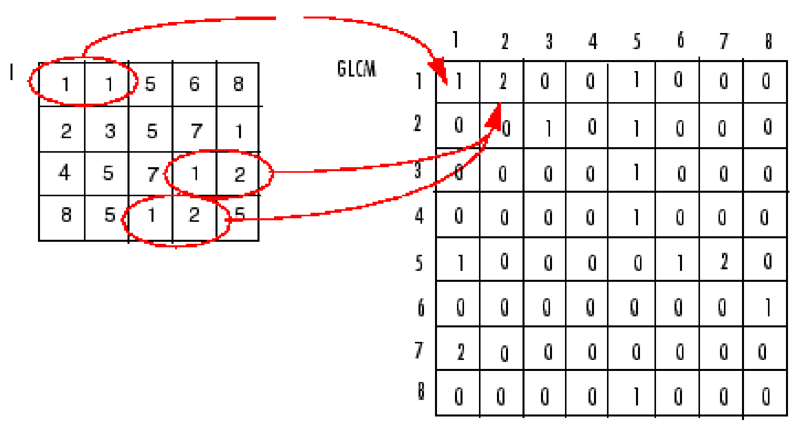

Figure 2. Formation of GLCM

The process involved in GLCM formation is shown in Figure 2. The statistical features that are computed from GLCMs are as follows:

$$Energy = \sum_{i,j} P(i,j)^2 \tag{1}$$

$$Entropy = -\sum_{i,j} P(i,j)\log P(i,j) \tag{2}$$

$$\text{Homogeneity} = \sum_{i,j} \frac{1}{1+(i-j)^2} P(i,j) \tag{3}$$

$$\text{Inertia} = \sum_{i,j} (i-j)^2 P(i,j) \tag{4}$$

$$\text{Correlation} = -\sum_{i,j} \frac{(i-\mu)(j-\mu)}{\sigma^2} P(i,j) \tag{5}$$

$$Shade = \sum_{i,j} (i+j-2\mu)^3 P(i,j) \tag{6}$$

$$\text{Prominence} = \sum_{i,j} (i+j-2\mu)^4 P(i,j) \tag{7}$$

$$\text{Variance} = \sum_{i,j} (i-\mu)^2 P(i,j) \tag{8}$$

$$where\ \mu = \mu_x = \mu_y = \sum_i i \sum_j P(i,j) = \sum_j j \sum_i P(i,j)$$

$$\text{and}\ \sigma = \sum_i (i-\mu_x)^2 \sum_j P(i,j) = \sum_j (j-\mu_y)^2 \sum_i P(i,j)$$

$$\text{Contrast} = \sum_i \sum_j (i-j)^2 P(i,j) \tag{9}$$

$$\text{Angular Second Moment} = \sum_i \sum_j \{P(i,j)\}^2 \tag{10}$$

$$\text{Inverse Difference Moment} = \sum_i \sum_j \frac{1}{1+(i-j)^2} \{P(i,j)\} \tag{11}$$

$$\text{Autocorrelation} = \sum_i \sum_j (ij) P(i,j) \tag{12}$$

$$\text{Dissimilarity} = \sum_i \sum_j |i-j| P(i,j) \tag{13}$$

$$\text{Maximum Probability} = \underset{i,j}{MAX}\ p(i,j) \tag{14}$$

$$\text{Sum Entropy} = -\sum_{i=2}^{2N_g} P_{x+y}(i) \log\{P_{x+y}(i)\} \tag{15}$$

$$\text{Difference Variance} = Variance\ of\ p_{x-y} \tag{16}$$

$$\text{Difference Entropy} = -\sum_{i=0}^{N_{g-1}} P_{x-y}(i) \log\{P_{x-y}(i)\} \tag{17}$$

$$\text{Information Measures of Correlation} = \frac{HXY - HXY1}{\max\{HX, HY\}} \tag{18}$$

$$= (1-\exp[-2.0(HXY2 - HXY)])^{1/2} \tag{19}$$

$$\text{Inverse Difference} = \sum_i \sum_j \frac{1}{1+|i-j|} \{P(i,j)\} \tag{20}$$

2.2. Support Vector Machine

Vapnik proposed SVM [9], basically a statistical learning concept. The SVM works on the fundamental principle of inserting a hyperplane between the classes, and it will keep at highest distance from the nearest data points. Data points appear nearest to the hyperplane are defined as Support Vectors. Popular kernels are Linear kernel, Polynomial kernel of degree'd', Gaussian radial basis function (RBF), and Neural Nets (sigmoid). Here, in this work, the RBF kernel is used.

3. Proposed Method

A Copy-Move Forgery Detection (CMFD) method is proposed using GLCM and SVM. The proposed method is detailed below and is shown in Fig.3.

i. The standard database CoMoFoD consists of original, forged and processed images is considered in the performance analysis.
ii. The images in the database are converted to gray scale.
iii. The statistical features are computed on GLCMs developed from the gray scale images.
iv. The Support Vector Machine is trained with those 20 statistical features for every image in the database using RBF kernel.
v. Statistical features of the testing image are obtained in similar process using steps 2 and 3.
vi. The SVM classifier classifies the image either to be authentic or forged.

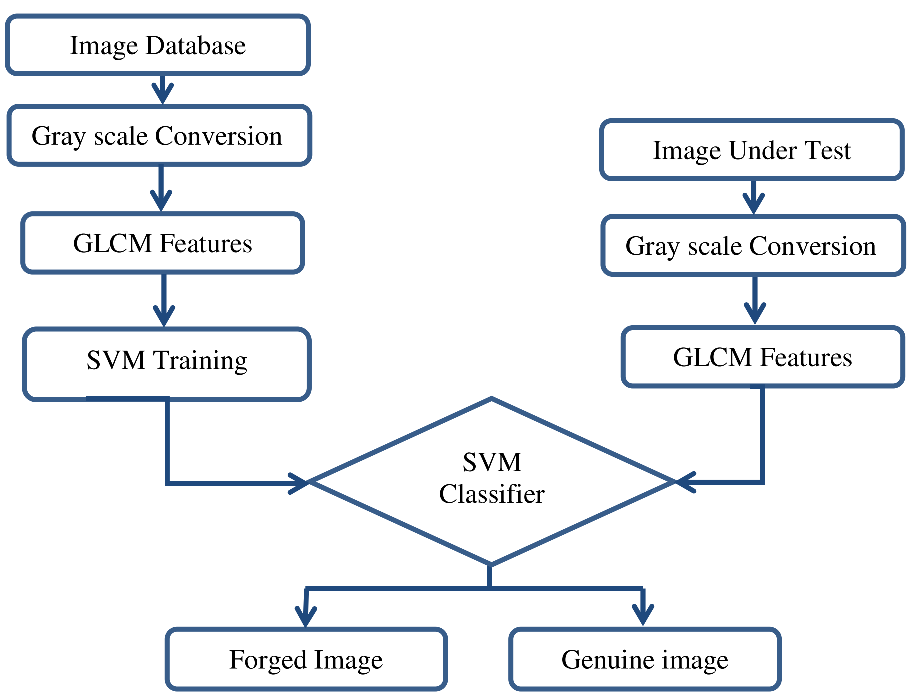

Figure 3. Process Flow of Proposed Method

4. EXPERIMENTATION AND RESULTS

The proposed method is evaluated on a standard database CoMoFoD [10] using the parameters TPR and FNR. This database contains original, forged and post-processed images after forgery.

True Positive Rate (TPR) = (Forged images declared Forged) / Forged Images
False Negative Rate (FNR) = (Forged images declared Genuine) / Forged Images

In the proposed method, 200 images of size 512x512 are considered. The operations such as scaling and rotation are performed before pasting the copied portion. It is evident from the Table 1 that the TPR value reduces if the copied portion is rotated much. As well, for small scaling factors the TPR is less and when the scaling factor is high TPR is high.

Table 1: TPR Values for Rotation and Scaling attacks

Rotation		Scaling	
Rotated angle	TPR	Scaled factor in %	TPR
3	95.31	40	75
5	75	70	84.37
40	68.75	95	89.5
90	62.50	105	96.87

The post-processed images with the below attacks are considered for evaluation.

i. "JC" - JPEG compression with quality factor ranging from 20 to 100,
ii. "IB" - Image Blurring with mean = 0, variance values of 0.009, 0.005 and 0.0005,
iii. "NA" - Noise Addition with averaging filter masks 3x3, 5x5, 7x7,
iv. "BC" - Brightness Change varies between 0.01- 0.95, 0.01- 0.9 and 0.01- 0.8,
v. "CR" - Color Reduction 32, 64, 128 levels per color component
vi. "CA" - Contrast Adjustments varies between 0.01- 0.95, 0.01- 0.9 and 0.01- 0.8.

The present method is appraised by considering 50 forged images in each post-processing attack category, so at the outset 1200 forged and processed images are tested.

Table 2: TPR and FNR of our proposed method for various post-processing attacks

Attack Description	TPR in %	FNR in %
No Attack	100	0
Brightness Change (0.01, 0.95)	92	8
Brightness Change (0.01, 0.9)	100	0
Brightness Change (0.01, 0.8)	100	0
Contrast Adjustment (0.01, 0.95)	66	34
Contrast Adjustment (0.01, 0.9)	68	32

Contrast Adjustment (0.01, 0.8)	76	24
Color Reduction 32	98	2
Color Reduction 64	94	6
Color Reduction 128	94	6
Image Blurring μ = 0, σ2 = 0.009	60	40
Image Blurring μ = 0, σ2 = 0.005	68	32
Image Blurring μ = 0, σ2 = 0.0005	88	12
Noise Adding 3x3	100	0
Noise Adding 5x5	96	4
Noise Adding 7x7	78	22
JPEG Compression QF=20	70	30
JPEG Compression QF=30	74	2
JPEG Compression QF=40	74	26
JPEG Compression QF=50	80	20
JPEG Compression QF=60	90	10
JPEG Compression QF=70	94	6
JPEG Compression QF=80	100	0
JPEG Compression QF=90	100	0
JPEG Compression QF=100	100	0

It is evident from Table 2 that the proposed method withstand attacks JPEG compression, Image blurring, Color reduction, brightness change and Noise addition in a better manner when compared to the attacks Contrast adjustment and Image blurring. It is evident from Table 3 that our method outperforms the other two methods [4, 6] in terms of TPR under no attack.

Table 3: Comparative Analysis of the proposed method

Method	Robust to Affine attacks	TPR %
Method in [4]	RST invariant	95.48
Method in [6]	No	99.83
Proposed Method	RST Invariant	100

5. CONCLUSIONS

In recent times, GLCM features are exploited to identify forgery related to Human faces in digital images. But, in our proposed method it is explored for all kinds of images such as buildings, plants, vehicles, people and textures. The simulation results indicate that our proposed method withstands all the post-processing attacks except Contrast Adjustment and Intensity Blurring. The proposed method outperforms the two methods [4, 6]. Proposed method is also invariant to rotation and scaling attacks to some extent. In future, the work can be extended to localize the tampered regions.

REFERENCES

[1] Shivakumar B L, Santhosh Baboo S. Detecting Copy-Move Forgery in Digital Images: A Survey and Analysis of Current Methods, Global Journal of Computer Science and Technology, 2010, 10 (7), pp. 61-65.

[2] Khan S, Kulakarni A, A reduced time complexity for detection of copy-move forgery detection using Discrete Wavelet Transform, International of Computer Applications, 2010, 6 (7), pp.31-36.

[3] Mahdian B,Stanislav S. . A bibliography on blind methods for identifying image forgery, Signal Processing: Image Communication, 2010, 25 (6), pp. 389-99.

[4] Shikha Dubey, A Sarawagi, Manish Srivastava, "Image Forgery Detection based on Local Descriptors and Block Matching using Clustering Technique" International Journal of Computer Applications, Vol.141, No.10, May 2016, pp.11-14.

[5] Jen Chun Lee, "Copy-Move image forgery detection based on Gabor magnitude", Journal of Visual Communication and Image representation, Vol.31, 2015, pp.320-334.

[6] Meera Mary Isaac, M Wilscy, "Image forgery detection based on Gabor Wavelets and Local Phase Quantization" Proceedia Computer Science, Vol. 58, 2015, pp.76-83.

[7] Liya Baby, Ann Jose, "Digital Image Forgery Detection Based on GLCM and HOG Features" International Journal of Advanced research in Electrical, Electronics and Instrumentation Engineering, Vol.3, Issue.5, Dec.2014, pp.426-430.

[8] Mryka Hall-Beyar,GLCM Tutorial,February 2007 [Online] Available: http://www.fp.ucalgary.ca/mhallbey/tutorial.htm (February 21, 2007).

[9] Vapnik, Vladimir. The nature of statistical learning theory. springer, 2000.

[10] http://www.vcl.fer.hr/comofod/comofod.html

AUTHORS

Gulivindala Suresh is a research scholar in the Department of ECE at JNTU-K University College of Engineering, AP, India. He obtained his M.Tech from Biju Patnaik University of Technology, Orissa, India. He obtained his B.Tech from JNTU, AP, India. His research interests are Digital Image Processing and VLSI. He is a member of IETE. Presently, he is working as Assistant Professor in the Department of ECE, GMR Institute of technology, Rajam, AP, INDIA.

Srinivasa Rao Ch is currently working as Professor in the Department of ECE, JNTUK University College of Engineering, Vizianagaram, AP, India. He obtained his PhD in Digital Image Processing area from University College of Engineering, JNTUK, Kakinada, AP, India. He received his M. Tech degree from the same institute. He published 40 research papers in international journals and conferences. His research interests are Digital Speech/Image and Video Processing, Communication Engineering and Evolutionary Algorithms. He is a Member of CSI. Dr Rao is a Fellow of IETE.