IMAGE QUALITY FEATURE BASED DETECTION ALGORITHM FOR FORGERY IN IMAGES

Shrishail Math[1] and R.C.Tripathi[2]

Indian Institute of Information Technology, Allahabad, India,211012
[1]ssm@iiita.ac.in [2]rctripathi@iiita.ac.in

ABSTRACT
The verifying of authenticity and integrity of images is a serious research issue. There are various types of techniques to create forged images for various intentions. In this paper, Attempt is made 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.

KEYWORDS
Image forgery, image quality, moments, Multiblock cosine transforms

1. INTRODUCTION

Once photographs are known for their authenticity and considered as a evidences. However today any one with basic knowledge of computer and image editing softwares like photoshop, GIMP etc maybe able to manipulate photographs easily.

The Advances in image processing and photo realistic softwares, higher capable digital camera, and other handheld portable image acquisition devices, high speed internet and social networking and image photo managing and sharing softwares like picasa Microsoft office manager etc provided easy platform for a image manipulations.

Images are manipulated for various reasons. Fun, entertainment, education, etc, however, recently image manipulations are used to misrepresent images, altering the meaning of pictures and contexts with malicious intention.

Figure 1-1: Original Picture of Joseph Stalin and Nikolai Yezhov

Figure 1-2: Manipulated image Nikolai Yezhov was erased

2. RELATED WORK

The recently researchers made efforts to detect the image forgery detection, the different methods are proposed for different types of forgeries [1,2,3,4,5]. Images forgery detection based on active methods such as digital watermarking[5], a digital signature[5] but those requires embedded of information or a data such a holograms either at image acquisition stage or image formation step .The detection method methods verifies the integrity of imbedded information, other method is blind or passive image forgery detection. This method doesn't require any pre imbedded information or a data.

The blind methods becoming popular since it don't require any extra hardware or softwares and its natural. The forgery detection based on near duplicate concepts are proposed[11], inconsistencies of light properties[15] ,noise features[16] and chromatic aberration[13],camera parameters[18], are reported The image forgery detection methods for, JPEG compression[17] and image splicing[] are also reported

3 PROPOSED METHOD

Our image forgery detection model based on image qualities and markov process based features the frame work of model is shown in fig 3-1

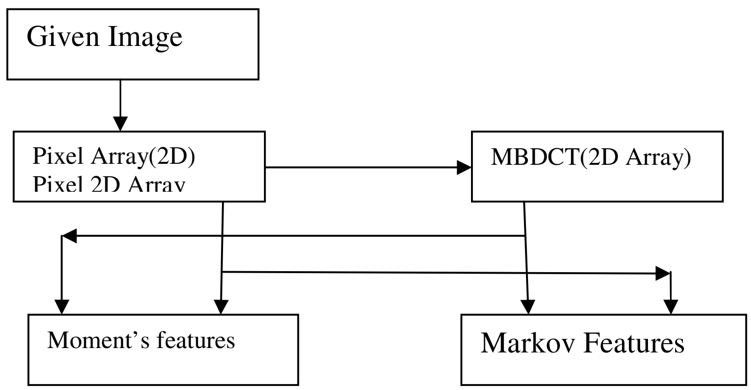

Fig 3-1: Image forgery detection model

3.1Image qualities*:* In computer vision research there is rich set of literature available on image qualities. We selected a image quality features based on study of Avcibas. In[19,20] ,Aviabas present a large set of image quality features, which are sensitive to discriminative to based few features of forgeries such as compression, watermarking, blurring and distortions. We selected such a eighteen features which are sensitive to image forgery operations .Those features are Mean Errors (D1-D4), Correlation (C1-C5), Spectral Errors (S1-S5), HSV Norms (H1-H2)

a.) Mean error features : Mean absolute error D1, mean square error D2, modified infinity norm D3, L*a*b perceptual error D4
b) Normalized cross-correlation C1, image fidelity C2,,Czenakowski correlation C3, mean angle similarity C4, mean angle-magnitude similarity C5.
c) Pratt edge measure E1, edge stability measure E2.
d) Spectral phase error S1, spectral phase-magnitude error S2, block spectral magnitude error S3, block spectral phase error S4, block spectral phase-magnitude error S5.
e). HVS absolute norm H1, HVS L2 norm H2.

3.*2Moment based Features*: The forgery operation assumed to be disturbs the continuity, smoothness, regularity pattern, smoothness, consistency and periodicity of pixel correlations

Our moment based feature extraction procedure is shown in fig 3-2.

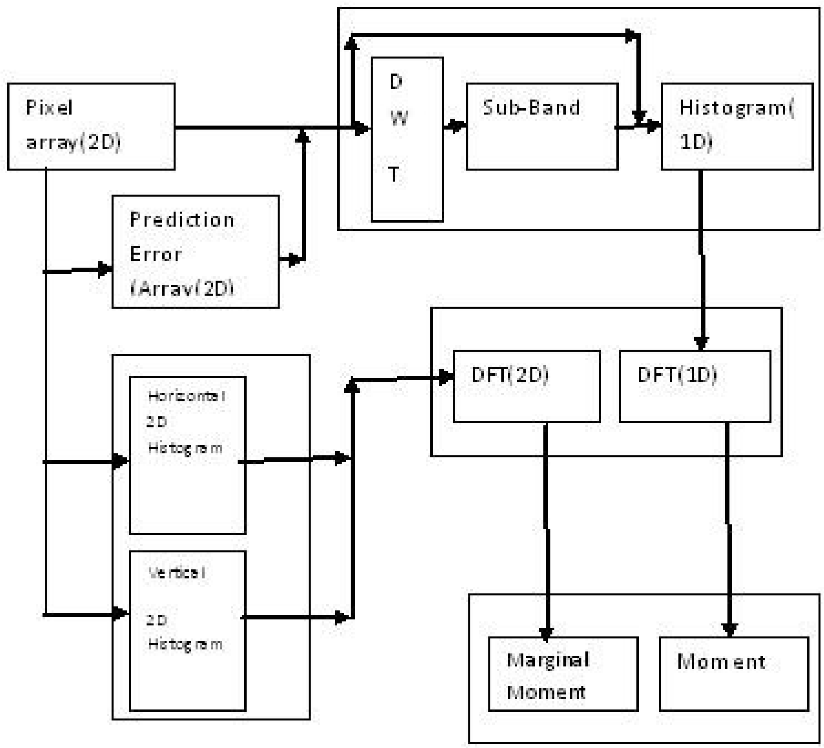

Figure 3-2: Moment extraction Procedure

Multi- block discrete cosine transforms (MBDCT): the block discrete cosine transforms co-efficient are able to reflect the disturbances (changes) in the local frequency distributions. We use multibolck discrete cosine transform to pick up local frequency disturbances effectively.

The 2D block DCT coefficients are represented by

$$F(s,t) = \frac{2}{n}\sum_{x=0}^{n-1}\sum_{y=0}^{n-1} V(x)V(y)\cos\frac{\pi(2x+1)s}{2n}\cos\frac{\pi(2y+1)t}{2n} f(x,y) \qquad \text{----1}$$

Where f(x,y),x,y=0,1 denotes a nXn image

3.3 ***Prediction Error 2D Array*****:** This is used for dimension reduction purpose. It also serves the additional purpose of enhancing the statistical artifacts introduced by forgeries. The prediction context is shown in Fig 3-3

x	a
b	c

Figure 3-3: Prediction Context

We predict the pixel value x using the neighbouring pixel a,b and c, the prediction 2D array is represented as

$$\hat{x} = sign(x) \cdot \{|a| + |b| + |c|\} \quad \text{------------------------2}$$

The prediction error 2D array can be expressed by

$$\Delta x = x - \hat{x} = x - sign(x) \cdot \{|a| + |b| + |c|\} \quad \text{-------------------------3}$$

Discrete wavelet transforms

The wavelet transforms are suitable to pick up transient and localised changes in spatial and frequency domain.

Moments and Marginal moments

The 1D Characteristic function (CF) is the DFT of the first order histogram of each wavelet sub band.

The absolute moments of 1D CF are defined by

$$M_l = \frac{\sum_{i=1}^{K/2} x_i^l |H(x_i)|}{\sum_{i=1}^{K/2} |H(x_i)|} \quad \text{---------------------------4}$$

Where $H(x_i)$ is the CF component t frequency x_i ,

Here K= total number of different values assumed by all of coefficients in the sub-band under consideration, and L= order of moment, which is a integer value

The 2D characteristic function is the 2D DFT of the second order histogram of the image and MBDCT coefficient 2D array.

The second order histogram is defined as

$$h_d(j_1, j_2, \zeta, \theta) = \frac{N(j_1, j_2, \zeta, \theta)}{N_T(\zeta, \theta)} \quad \text{-------------------------------5}$$

Where

ζ @ the distance between two pixel,

θ @ Angle of line linking these two pixels with respect to the horizontal axis

$N(j_1, j_2, \zeta, \theta)$ @ Number of pixel pairs for which the first pixel value is J1 while second is J2

$N_T(\zeta, \theta)$ -Total number of pixel pair in the image with separation (ζ, θ,).

Two Marginal moments of the 2D CF are given by

$$M_{u,j} = \frac{\sum_{j=1}^{k/2} \sum_{i=1}^{k/2} u_i^{l} \left| H(u_i, v_j) \right|}{\sum_{j=i}^{k/2} \sum_{i=1}^{k/2} \left| H(u_i, v_j) \right|} \quad \text{----------------------6}$$

$$M_{v,j} = \frac{\sum_{j=1}^{k/2} \sum_{i=1}^{k/2} v_j^{l} \left| H(u_i, v_j) \right|}{\sum_{j} \sum_{i=1}^{k/2} \left| H(u_i, v_j) \right|} \quad \text{------------------------7}$$

Where $H(u_i,vj)$- 2D CF component at DFT frequency (u_i,vj), l- order of moment, integer

4 THE EXPERIMENT AND RESULTS

4.1 Algorithm:

1.) Extract Image Quality Metrics (IQMs).

a. Divide test image into 4 regions.

b. Extract features from every region.

2) Extract moment based features.

a. Apply wavelet transform to this image and obtain all the sub-bands including the test image itself.

b. Obtain histogram for each sub-band.

c. Apply DFT to the histogram of each sub-band to obtain its characteristic function.

d. Apply Eqn -- (4) to calculate moments.

e. Apply Eqn-- (3) to obtain prediction-error 2-D array.

f. Repeat a. to d. To obtain prediction-error 2-D array.

g. Obtain 2-D histograms for the test image.

h. Apply 2-D DFT to each 2-D histograms to obtain the 2-D characteristic function.

i. Apply Eqn--(6) and Eqn-- (7) to calculate marginal moments.

3) Apply 2×2, 4×4, 8×8, …n×n BDCT to the given image. Round those BDCT coefficients to nearest integers, and repeat step 2).

4) Repeat step 1) to 3) to obtain features of all images.

5) Obtain the best parameter of C and g which will be used in training.

6) Train a part of images using SVM and obtain SVM model.

7) Predict the remaining images using SVM model.

4.2 Experiment: We used the image dataset[22] of Columbia image splicing detection and evaluation data.Other images are collected from internet:933 authentic images and 912 forged(spliced images) and 55 forged as well as same number of authentic images were collected from various resources from internet. SVM classifier and matlab code [21] is used for randomly selected 65%,75% and 85% images from the above databases for training purposes and renaming are used for testing purposes. The results are shown in table1below.

Data set	Training ratio	True Positive %	True negative %	Accuracy %
Columbia data set	65%	82.37%	79.41%	80.89%
	75%	83.56%	80.73%	82.14%
	85%	85.41%	82.11%	83.76%
Collected data set(Internetresources	65%	80.81%	77.15%	78.98%
	75%	81.32%	78.38%	79.85%
	85%	83.12%	80.19%	81.65%

Table 1: Results

5. CONCLUSIONS AND DISCUSSIONS

Sensitive quality features [19, 20] and markov process model was utilised to detect forgery in image data using histogram moments. Experiments were conducted on famous databases providing and authentic and forged images to find true positive and true negative results duly these training of data for various ratios like %65,75% and 85% respectively .Results obtained are superior to so far used techniques.

REFERENCES

[1] Hany farid" Image forgery Detection a survey",IEEE signal processing magazine ,march 2009, pp 16-25

[2] H.T.sencar and N.memon ,"Overview of state of the art in digital Image forensics",WSPC proceedings sept 2007

[3].. Luo weigi,Qu Zhenhua, et. Al,"A survey of passive technology for digital Image forensics", Front. Computer Science china 2007

[4]. N.Krawetz" A picture's worth… Digital Image Analysis and Forensics",black hat briefings USA2007.

[5]. .Babak Mahdian and stnislav siac,"Blind methods for detecting Image fakery",ICCST2008.

[6] Zhen zhang,yuan ren ,et al.," a survey on passive blind image forgery by doctored method detection",seventh ICML&C,kunming ,july2008.

[7] .Tran van lanh kai-sen chong et .al ," a survey on digital image forensic methods",ICME2007

[8] .Kusam ,pawanesh abrol,devanand,"Digital tampering detection techniques:A review",IJIT2009.

[9]. Yu-Feng Hsu and Shih-Fu Chang"Statistical fusion of multiple cues for image tampering detection"Asilomar Conference on Signals, Systems, and Computers 2008

[10] Tian-Tsong Ng and Shih-Fu Chang and Jessie Hsu and Lexing Xie and Mao-Pei Tsui"Physics-motivated features for distinguishing photographic images and computer graphics", MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia, November 6--11, 2005, Hilton, Singapore

[11]. Alin C. Popescu and Hany farid"Exposing digital forgeries by detecting duplicated image regions" 2004

[12]. Alin C. Popescu and Hany farid"Blind removal of lens distortion" Journal of the Optical Society of America 2001

[13]. Micah K. Johnson and Hany Farid," Exposing digital forgeries through chromatic aberration" ACM Multimedia and Security Workshop 2006

[14] Micah K. Johnson and Hany Farid,""Detecting photographic composites of people" 6th InternationalWorkshop on Digital Watermarking, Guangzhou, China

[15]. Micah K. Johnson and Hany Farid"Exposing digital forgeries by detecting inconsistencies in lighting" ACM Multimedia and Security Workshop 2005

[16]. H. Gou, A. Swaminathan, and M. Wu, "Noise features for image tamperingdetection and steganalysis" in Proc. IEEE Int. Conf. Image Processing, San Antonio,TX, 2007, vol. 6, pp. 97–100.

[17]. J. He, Z. Lin, L. Wang, and X. Tang, "Detecting doctored JPEG images viaDCT coefficient analysis", in Proc. European Conf. Computer Vision, Graz, Austria,2006, pp. 423–435.

[18]. Z. Lin, R. Wang, X. Tang, and H.-Y. Shum, "Detecting doctored imagesusing camera response normality and consistency",in IEEE Computer So-ciety Conference on Computer Vision and Pattern Recognition, vol. 1, June2005, pp. 1087

[19]. Avcibas I,B.Sankur,K. Sayood,"Statistical Evaluation of Image Quality Measure", Journal of Electronic Imaging,11,206-223,2002

[20] .Avcibas I,N.menon,B.Sankur,"Steganalysis Using Image Quality Metrics",IEEE Transactions on Image processing,12,221-229,2003

[21]. C. C. Chang and C. J. Lin, LIBSVM: A Library for Support Machines, http://www.csie.ntu.edu.tw/~cjlin/libsvm/index.html.

[22]. Columbia DVMM Research Lab, Columbia Image Splicing DetectionEvaluation Dataset, http://www.ee.columbia.edu/ln/dvmm/downloads/AuthSplicedDataSet/AuthSpliced DataSet.htm.

Authors

Shrishail Math received his B.E(Electronics&Communication enginerring) ,M.Tech(Computer science &Engineering from University of Mysore and Manipal university,respectively in year 1996 and 2001.Currently Doctoral student at Indian Institue of information Technology,Allahabad, his research interest are information assurance & security, and multimedia forensics.

R.C.Triapthi is a Dean (R&D) Research and Development, head of IPR division, worked as senior Director Ministry of Communication &Information Technology (MCIT), Govt .of India, published three books and several research articles and papers in international and national journals. He worked as a co-chairman of 1st International conference onIntelligent Interactive Multimedia (IITM) 2010 sponsored by ACM