***I.J. Engineering and Manufacturing*, 2016, 6, 1-14**
Published Online November 2016 in MECS (http://www.mecs-press.net)
DOI: 10.5815/ijem.2016.06.01

Available online at http://www.mecs-press.net/ijem

A Discriminative Statistical Model for Digital Image Forgery Detection

Amira Baumy [a], Naglaa. F Soiliman [b,c], Mahmoud Abdalla [b], Fathi Abd El-Samie [d]

[a] *Obour Institute of Engineering and Technology, Obour, Egypt.*
[b] *Faculty of Engineering, Zagazig University, Zagazig, Egypt.*
[c] *Faculty of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia.*
[d] *Faculty of Electronic Engineering, Menoufia University,Menouf, Egypt.*

Abstract

The headway of modern technology and facility to use processing software leads to tamper and implicate of digital images. This tampering is being performed without leaving any a clear effect noted with the naked eye. The discrimination between different authentic and forged images can be based on its Probability Density Functions (PDFs). This paper introduces a new model for digital image forgery detection. This framework has two main phases; training and testing. In the training phase, the peak is calculated for the derivatives histogram of the illumination components by using homomorphic filter to separate the illumination components on each image. Firstly, the derivative of illumination histogram for authentic and forged images is calculated then the PDFs are estimated for authentic and forged images, finally the threshold is determined. In the testing phase, the determined threshold is tested with realistic dataset followed by using the selected bins for feature calculation in the prediction process. In the final prediction step, a detection and decision process is performed to obtain performance of the new model. This new model is provided a very effective performance. Different color image contrast systems RGB and HIS are studied and utilized for testing our model and compare between each channel for two systems to estimate performance and obtain more sensitive channel.

Index Terms: Forgery detection, Homomorphic filter, image histogram.

1. Introduction

Images have been considered the most reliability source that include and record the correct information. Due to the low cost and availability of digital camera, anyone can capture pictures and keep the real-time

* Corresponding author. Tel.: +201226521508; fax: +201140401292; fax: +201099388635; fax: +201009239012
E-mail address: mrmaryia@yahoo.com, nagla_soliman@yahoo.com, fathi_sayed@yahoo.com, mabdalla2010@gmail.com

information. The widespread and development of processing software technology can be used for hiding facts and evidences. Forgers can modify the information and create a new tampering one from merge of two or more images easily. Therefore the society can't trust the authenticity of digital images, when data become sensitive such as in forensics analysis, medical records, evidence in court, photos published in newspapers and magazines. The question that arises: “Is the image real or fake?” [1, 2, 3]. There are two forgery detection techniques. First, the passive techniques which no need information embedded into the image, it uses only received image without using signature or watermark [4, 5, 6]. Second, the active techniques which needed inserted information at the time of creation as watermarking, fingerprints or signature [7, 8, 9].

2. Related Work

Recently many researches are performed and concentrated on passive techniques to obtain and discriminate between original and forgery images. Forgery image may include one or more processing operations such as scaling, rotation and brightness. Especially if it contains another part from other images which has different backgrounds. Therefore processing operations are sign indicators to the tampering operation. All researches compute powerful statistical feature that can be used to detect the forgery and guarantee good statistical methods for discriminative between original and modified one. Z. Moghaddasi [10] used Singular Value Decomposition (SVD) to extract feature that used to identify forgery images. In which the Discrete Cosine Transform (DCT) has been applied to SVD-based features for image forgery detection. The accuracy of this detection algorithm is equal to 78.82% and the feature vector dimensional is equal to 50. The Support Vector Machine SVM has been used to classify and identify forgery images and original ones [11]. In [11] the dimensional of feature vector was reduced to 14. As the photo has been divided into different parts under perceptual grouping criterion, reducing the disassociation between sub segments and increasing combination within segments with normalized cut algorithm. Then features vector composed of mean and standard variance points, followed by feeding them into SVM classifier.

The idea of utilization rake-transform feature and edge statistics feature was applied in [12]. Feature vector dimensional is equals to 50 and SVM has been used as a classifier. Ghulam Muhammad [13] extracted chromatic channel from input images and wavelet transform was applied to the channel to separate lower sub band, then calculate Weber pattern (WP) from the sub band. The feature vector contains the WP histogram of the image and using SVM as a classifier. The main idea of Gajanan K. Birajdar [14] was estimating the rescaling factor and used zero-crossings characteristics of the second difference of the forgery image. Zhu Kaizhen [15] used the wavelet transform and calculated Image Quality Metrics (IQMs) and moments of characteristic functions for sub bands to composed features vector that contains 196-D features and using (SVM) as a classifier. In [16] the technique is used for identifying the location of copy-create and copy-move by using the JPEG Block and Direction Filter Techniques. J.Grim and Petr Somol [17] have been noted that the Gaussian mixture can be differentiate between forgery image and original one by statistical properties of the image. They get the samples of color pixel and calculate their probability distribution within the reference window. That paper was proposed a statistical scheme for image forgery detection. It depended on the estimation of the peak of differentiation illumination component histogram by utilizing Homomorphic Filter .U. A. Nnolim [18] described the analysis of the Homomorphic filtering algorithm, the equivalency between the frequency and spatial-domain methods and the implementation of low-pass and high-pass spatial domain Homomorphic filter in low power embedded devices. Radovan Jirık [19] medical ultrasonic images used homomorphic filter for two-dimensional deconvolution. Dileep MD [20] enhanced images by removing the illumination components which characterized by its low frequency of the images using homomorphic filter and retained the high frequency reflectance component. Homomorphic filtering break down some part of the image which has not effects on enhancement. In [21] the Homomorphic filtering is implemented by splitting illumination and reflectance parts and then weakens the low frequencies and reinforces the high frequencies. Exploitation of important information can be obtained from color image. Wei Wang [22] proposed method for forgery detection based on image chrome. G.LIANG [23] proposed accurate formulas for color transformation from RGB and HIS.

This paper presents a model for discrimination between different authentic and forged images, derivatives histogram of the illumination components has been used for the estimation PDFs to authentic and forged images. HIS and RGB has been investigated. The remainder of the paper is organized as follows: Section II presents related work. Section III presents the proposed detection algorithm. Section IV provides the simulation results followed by Section V concludes the paper.

3. The Proposed Detection Algorithm

The proposed method has been built a classifier to distinguish between the original and forgery images. Due to the fact that the image combined from two components according to its distinctive nature. It extracts the features from the histogram derivatives of illumination components and reflection components. The homomorphic filtering separates the low frequency illumination components and the high frequency reflection components from the image. The amount of light incident from source illumination which known as the illumination *I*(*x*, *y*) and the mount of illumination reflected by the object that known as the reflectance *R(x*, *y)* of the image intensity *F(x*, *y).* This relation can be estimated as follows:

$$F(x,y) = R(x,y)I(x,y) \tag{1}$$

$$logF(x,y) = logR(x,y) + logI(x,y) \tag{2}$$

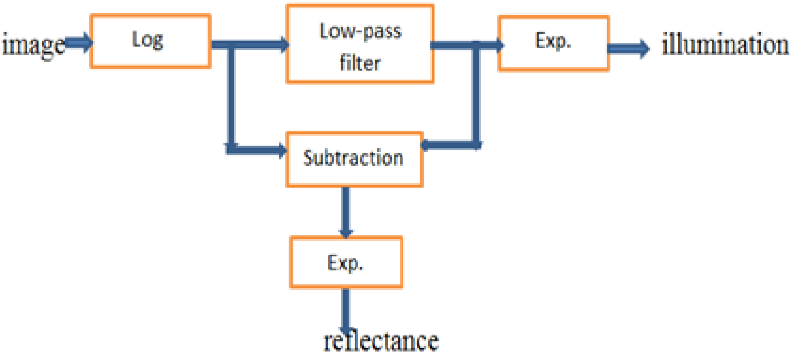


Fig.1. Homomorphic filtering.

The separation of the multiplied two components is hard therefore it preferably to utilize homomorphic filtering that explained at Fig.1. This filtering technique can be performed by applying a logarithm calculation on image to convert it from multiplying to summing relation. Such that the image intensity expressed as the sum of the logarithm form of the illumination and that of the reflectance components which is explained in equation (2).

The illumination components can be separated by using low-pass filter applied on the logarithm of the image intensity and the reflectance components can be produced by subtracting the logarithm of illumination components from the logarithm of the image intensity.

The estimation histogram graph has the same intensity value as that found in all image. It represents the statistical information for the reflection and illumination components. They represent the pixel values of the image and provided an implicit continuous distribution of colors values. It can be noted that the differentiation of the illumination component histogram of the forgery images had a large variation than that of the original images. Moreover the differentiation of the reflection component histogram of the forgery images had no marked change than that of the original images. The proposed method depends mainly on estimating the PDFs for each channel in two color system and selects the more sensitive channel. The proposed model has been included two phases;

1. Training phase in Fig.2.
2. Testing phase in Fig.3.

The proposed model can be implemented using the following steps which can be described in Fig.4;

1. Select the training dataset and select the channel for the color system.
2. Training dataset by defining two groups for the tampering and the original images.
3. For each image take the logarithm of the image intensity.
4. Use low-pass filter to extract the illumination components.
5. Obtain the histogram for the illumination components.
6. Obtain the peak of the histogram derivative.
7. For each group estimate the PDFs for the normalized peaks of the histogram derivative separately.
8. Determining the threshold.
9. Test this threshold with testing dataset to obtain performance of the new model.

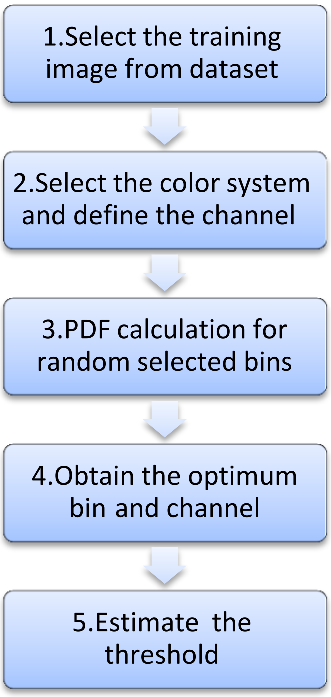

1.Select the training image from dataset
2.Select the color system and define the channel
3.PDF calculation for random selected bins
4.Obtain the optimum bin and channel
5.Estimate the threshold

Fig.2. Training phase

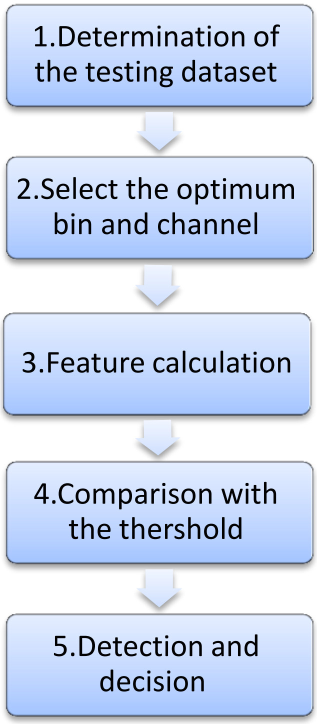

1.Determination of the testing dataset
2.Select the optimum bin and channel
3.Feature calculation
4.Comparison with the thershold
5.Detection and decision

Fig.3. Testing Phase

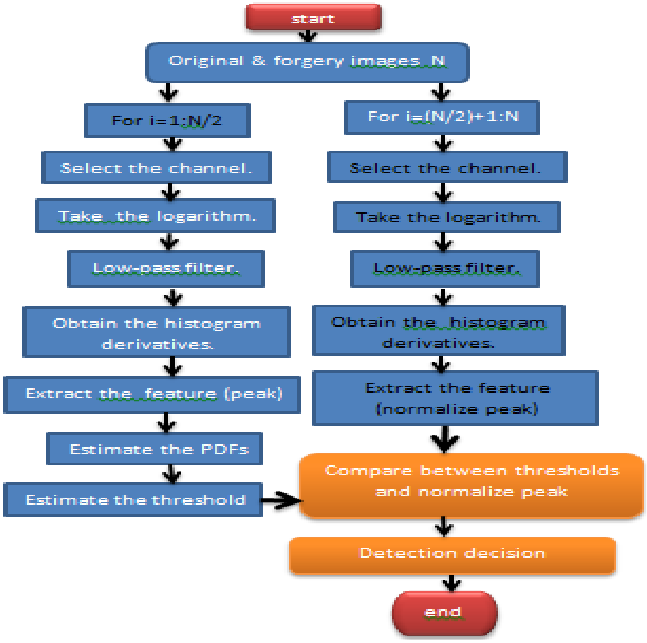

start
Original & forgery images N
For i=1:N/2
For i=(N/2)+1:N
Select the channel.
Select the channel.
Take the logarithm.
Take the logarithm.
Low-pass filter.
Low-pass filter.
Obtain the histogram derivatives.
Obtain the histogram derivatives.
Extract the feature (peak)
Extract the feature (normalize peak)
Estimate the PDFs
Estimate the threshold
Compare between thresholds and normalize peak
Detection decision
end

Fig.4. The Proposed Model Flowchart

4. Simulation Results

To measure the performance of the proposed technique and determine its quality, a realistic data base can be constructed. It is composed by 70 originals and 70 tampered images subdivided to 15 originals and 15 tampered images for training data and 55 originals and 55 tampered images for testing the new model. Photoshop program has been used to construct the tampering images. The tampering images have been composed by copy image part from one image and paste it to another image. The spliced part may be regular or irregular but its size not small, so it can be seen with the naked eyes. The proposed detection algorithm and the threshold have been examined with the dataset, each channels either RGB or HIS is tested. The performance of the proposed model has been tested by calculated True Positive Rate (TPR), False Positive Rate (FPR), True Negative Rate (TNR) and Accuracy where they are can be estimated as follows:

$$\mathrm{TPR} = \frac{\#\ \text{NO.of forged images detected as forged}}{\#\ \text{NO.of forged images}} \tag{3}$$

$$\mathrm{FPR} = \frac{\#\ \text{NO.of original images detected as forged}}{\#\ \text{NO.of original images}} \tag{4}$$

$$\mathrm{TNR} = \frac{\#\ \text{NO.of original images detected as original}}{\#\ \text{NO.of original images}} \tag{5}$$

$$Accuracy = \frac{\#\ \text{NO.of forged images detected as forged}+\#\ \text{NO.of original images detected as original}}{\#\ \text{NO.of forged images}+\#\ \text{NO.of original images}} \tag{6}$$

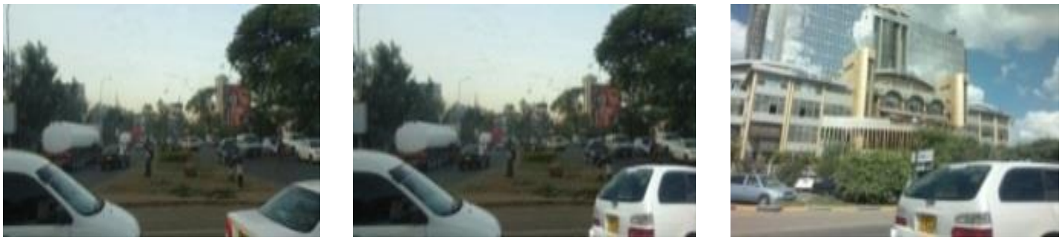
An example of forgery image can be explained in Fig. 5 (a), and the source images are depicted in Fig.5 (b and c). The forgery image has been combined by cutting and rescaled part from source (b) and pasted it in the source (c). The results of the proposed detection algorithm are shown in Fig. 6. The histogram of forgery image (Fig. 5-a) and the source image (Fig. 5-b) are shown in Fig. 6 (a, b) respectively. It can be noted from these figures that the histogram of tampered image has been changed more than the original one because it includes more processing operation. Fig. 7 (a, b) shows a comparison between the tampered illumination differentiation histogram and the original one. It is revealed that the tampered image has more variation than that of the original one.

A comparison of the PDFs between the selected bins (random variables) has been explained in Fig.8. The illumination differentiation histogram is achieved between the data set and the (2nd, 3rd, 5th, 10th, 20th, 50th) where they are explained in Fig. 8(a-F). All of these bins doesn't achieved any discriminative results. The best discriminative one is the peak of differentiation histogram shows in Fig. 8 (g). It can be noticed that the tampered illumination derivative histogram has a larger great peak than the original illumination derivative histogram. The PDF peaks of differentiation histogram of the reflection component have been depicted in Fig. 9. It can be noted that the differentiation histogram of reflection components don't achieved any noted difference for each channel because the reflectance component of an image varies suddenly unlike constancy of illumination component.

Figures 10, 11 and 12 show the PDFs for peak of differentiation histogram of illumination for blue, red and green channel, each channel gives accurate results for discrimination between original and tampered images. The intensity channel in HIS system is shown in Fig. 13, which achieves a good result because image feature become clear in intensity channel than hue and saturation channels which illustrated in Fig. 14 and Fig. 15. A comparison among red, green, blue channel and intensity channel is explained in Table 1.Red channel gives the best results because it has been achieved minimum threshold than each channel. The proposed threshold is selected to be equal to 0.0137 which achieves an accuracy of 90% compared with [10] which had a maximum accuracy of 78.82%. A comparison between the proposed algorithm and [10] is explained in Table [2]. It can be depicted in this table the best accuracy is performed by the red channel.

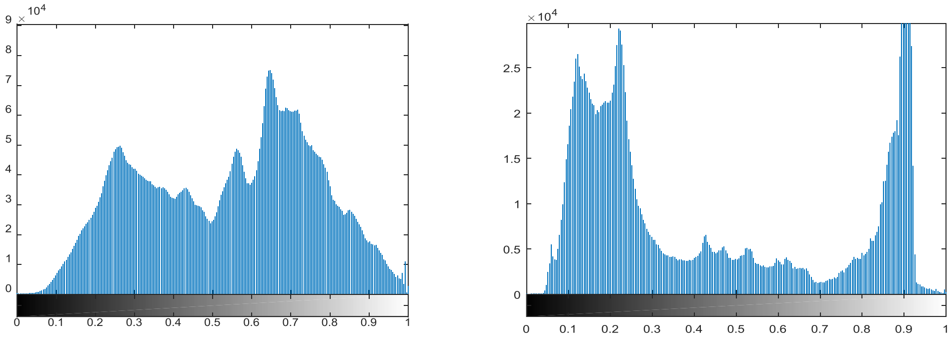
5. Conclusion

In this paper, a new and fast blind detection algorithm was proposed for the detection of forgery images. The illumination component of original image is changed when splicing occurs however there is not any effect on the reflection component. Different original and tampering image sets have been examined. The proposed model is tested using different channel and the threshold is determined for each. The red channel in RGB system gives the best results.



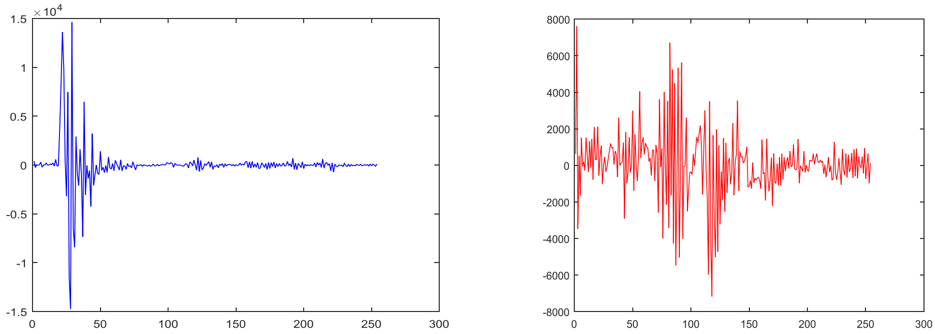
(a) Forgery image. (b) Source image. (c) Source image.

Fig.5. Example of Forgery Image.



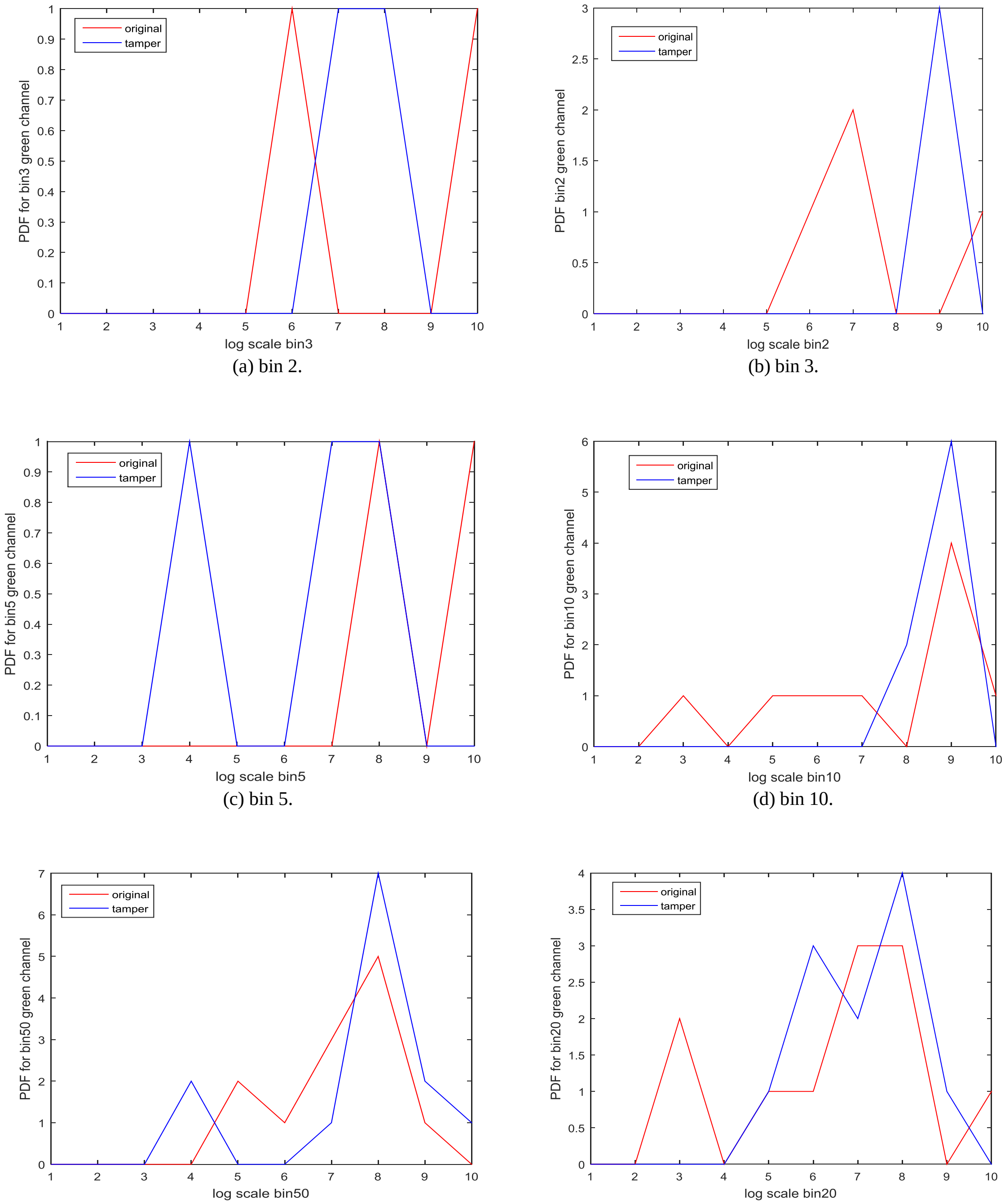
(a) The histogram of forgery (Fig. 5-a). (b) The histogram of source (Fig. 5-b).

Fig.6. A Comparison between the Histogram of Forgery Image and that of the Source Image.



(a) The tampered illumination derivative histogram. (b)The original illumination derivative histogram.

Fig.7. A Comparison between the Tampered Illumination Derivative Histogram and that of the Source Image.



(a) bin 2.

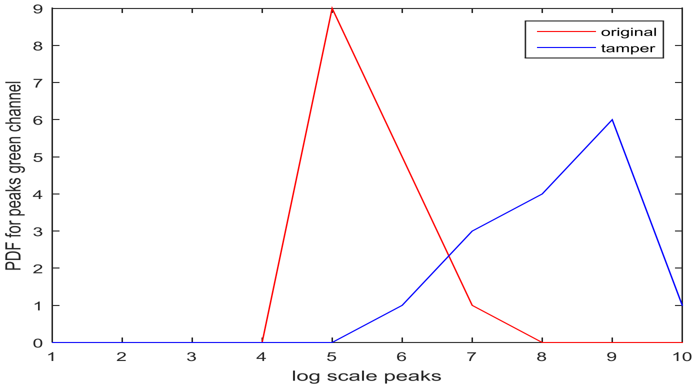
(b) bin 3.

(c) bin 5.

(d) bin 10.

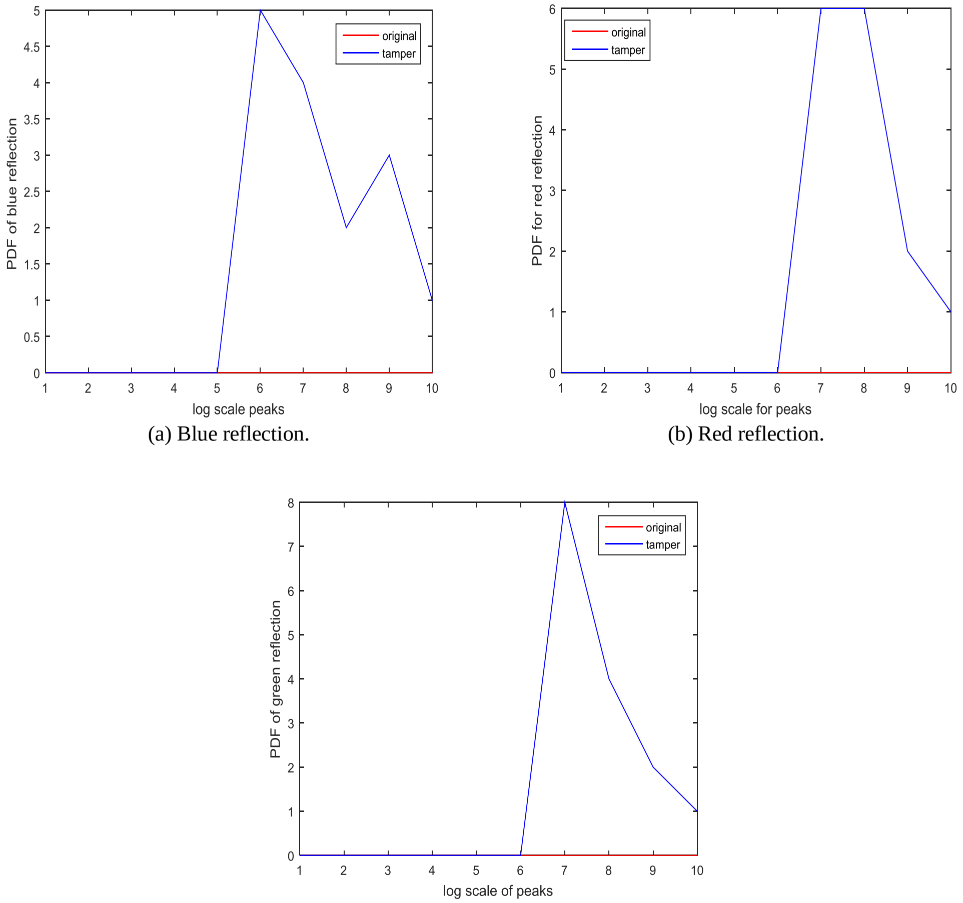
(e) bin 20.

(f) bin 50.



(g) peak of histogram.

Fig.8. Results of Random Selected Bins for Illumination Component.



(a) Blue reflection.

(b) Red reflection.

(c) Green reflection.

Fig.9. Results for Peaks of Derivative Reflection Component Histogram.

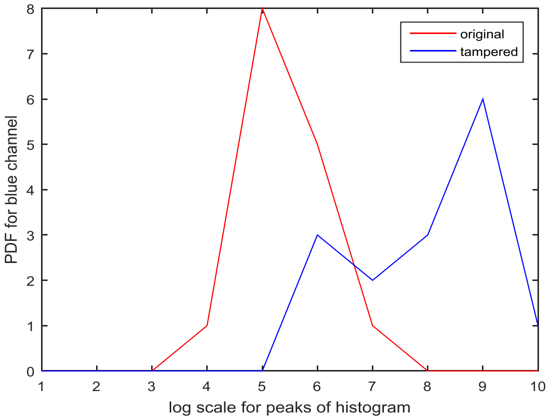


Fig.10. BLUE.

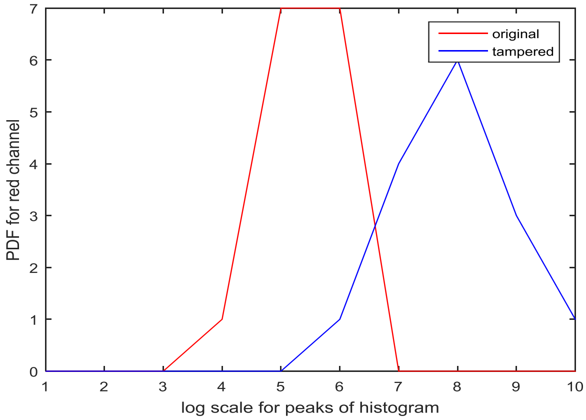


Fig.11. RED.

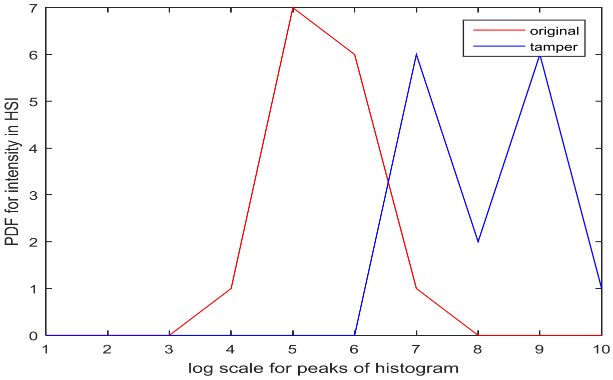


Fig.12. GREEN.

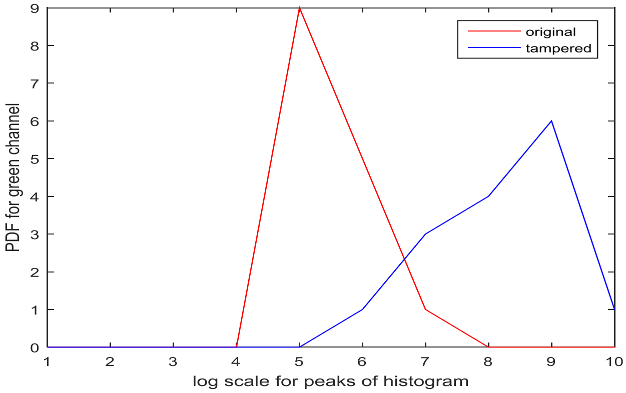


Fig.13. INTENSITY.

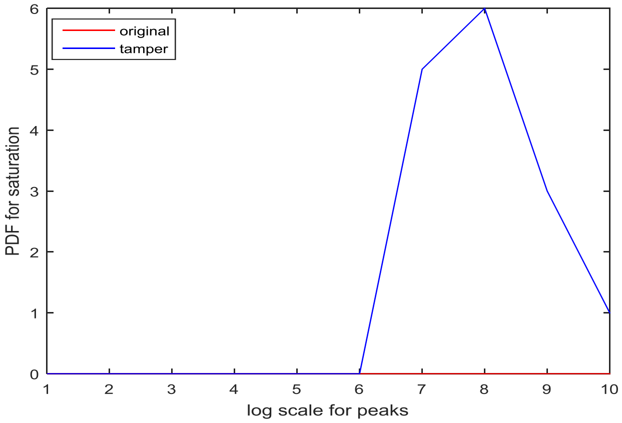


Fig.14. HUE.

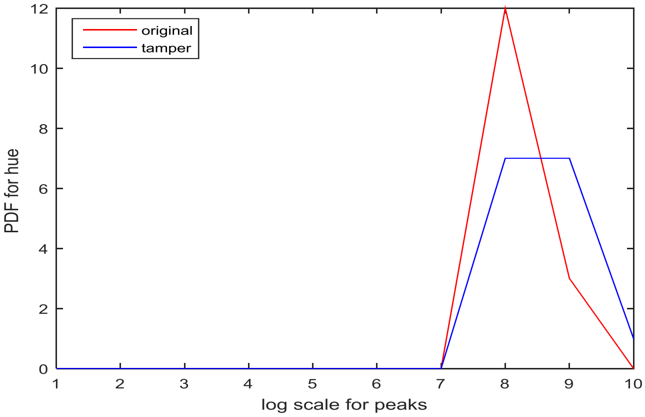


Fig.15. SATURATION.

Table 1. Comparison Between Red- Green-Blue Channel And Intensity Channel.

Channel	TPR	FPR	TNR	accuracy
RED	85.45%	5.4%	94.5%	90%
BLUE	50.91%	3.64%	96.36%	73.6%
GREEN	76.36%	5.4%	94.5%	85.45%
INTENSITY	70.91%	7.2%	92.72%	81.8%

Table 2. Comparison between Proposed Method and Z. Moghaddasi [10].

Methods	TPR	TNR	Accuracy
Z. Moghaddasi [10] SVD	76.92%	74.34%	75.63%
Z.Moghaddasi [10] SVD-DCT	77.56%	77.63%	77.6%
Z. Moghaddasi [10] SVD+ SVD-DCT	80.13%	77.63%	78.82%
Proposed model red channel	85.45%	94.5%	90%
Proposed model green channel	76.36%	94.5%	85.45%
Proposed model intensity channel	70.91%	92.72%	81.8%

References

[1] S. Mushtaq and A. H. Mir, “Digital Image Forgeries and Passive Image Authentication Techniques: A Survey ”, International Journal of Advanced Science and Technology vol.73, pp.15-32, 2014.

[2] M. D. Ansari, S. P. Ghreraa and V. Tragi, “Pixel-Based Image Forgery Detection: A Review”, IETE Journal of Education, 55:1, pp.40-46, 2014.

[3] D. Sharma, and P. Abrol, “ Digital Image Tampering – A Threat to Security Management”, International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 10, October, 2013.

[4] T. Qazi, K. Hayat1, S. U. Khan, and S. A. Madani, I. A. Khan1, “Survey on Blind Image Forgery Detection”, IET Image Process, vol. 7, Iss. 7, pp. 660–670, 2013.

[5] G. K. Birajdara, and V. H. Mankar,“ Digital Image Forgery Detection using Passive Techniques: A survey”, Digital Investigation 10, pp. 226–245, 2013.

[6] H. Farid, “A Survey of Image Forgery Detection”, IEEE Signal Proc Mag., vol. 2, no. 26, pp. 6–25, 2009.

[7] M. Sengupta and J. K. Mandal, “Authentication through Hough Transformation Generated Signature on G-Let D3 Domain (AHSG)”, International Conference on Computational Intelligence: Modeling Techniques and Applications, 2013.

[8] X. Wang, J. Xue, Z. Zheng, Z. Liu and N. Li, “Image Forensic Signature for Content Authenticity Analysis”, Vis. Commination. Image R., vol. 23, 2012.

[9] G. S. Spagnolo and M. DeSantis, “Holographic watermarking for authentication of cut images”, Optics and Lasers in Engineering, vol. 49, pp. 1447–1455, 2011.

[10] Z. Moghaddasi, H. A. Jalab, R. Md Noor, “SVD-based Image Splicing Detection”, International Conference on Information Technology and Multimedia (ICIMU) proceedings at IEEE, 2014.

[11] J. Hou, H. Shi, Y. Cheng, and Ran Li , “Forgery Image Splicing Detection by Abnormal Prediction Features”, International Conference of mechatronic and automation proceedings at IEEE, 2013.

[12] P. S. Yun Q.; S. W. Su; Tian-Tsong Ng, “Rake Transform and Edge Statistics For Image Forgery Detection”, IEEE Publication Year: 2010.

[13] G. Muhammad, “Multi-Scale Local Texture Descriptor For Image Forgery Detection”, IEEE Publication , 2013.

[14] G. K. Birajdar, “Blind Authentication of Resampled Images and Rescaling Factor Estimation” IEEE Publication , 2013.

[15] Z. Kaizhen, and Z. Zhang,“A Novel Algorithm of Image Splicing Detection” International Conference on Industrial Control and Electronics Engineering, IEEE Publication Year: 2012.

[16] S.Murali, B. Chittapur, “comparison and analysis of photo image forgery detection techniques” International Journal on Computational Sciences & Applications (IJCSA), I Publication Year: 2012.

[17] J. Grim, and P Somol, “Digital Image Forgery Detection by Local Statistical Models” IEEE Publication Year: 2010.

[18] U. A. Nnolim, “Implementation of Spatial Domain Homomorphic Filtering on Embedded Mobile Devices”, Nigerian journal of technology, 2015.

[19] R. Jirık, “High-Resolution Ultrasonic Imaging Using Fast Two-Dimensional Homomorphic Filtering”, IEEE Publication Year: 2006.

[20] D. MD, and A. S. Murthy, “A Comparison between Different Colour Image Contrast Enhancement Algorithms”, IEEE Publication Year: 2011.

[21] T. B. Adji, “Negative Content Filtering Based on Skin Texture Homomorphic Filter and Localizations”, International Conference of Electricial Engineering and Computer Science, IEEE Publication Year: 2014.

[22] Wei wang, “Effective Image Splicing Detection Based on Image Chrome”, IEEE Publication Year: 2009.

[23] G. LIANG, and D. CHANG , “Colour Image Enhancement with Exact HIS Colour Model”, International journal of innovative computing 2011.

Authors’ Profiles

Amira Baumy received the B.Sc.from the faculty of Engineering, Zagazig University, Egypt in 2009, She is currently teaching at the Department of Electronics and Communications Engineering, Obour-institute of Engineering and Technology. Her areas of interest are digital communications, signal processing, image processing.

Naglaa F. Soliman received the B.Sc., M.Sc., and Ph.D. degrees from the faculty of Engineering, Zagazig University, Egypt in 1999, 2004,and 2011, respectively. She was worked as a lecturer at the Department of Electronics and Communications Engineering, Faculty of Engineering,Zagazig University from 2012 up to 2015. She is currently at Faculty of Computer and Inforation Sciences, pricess Nourah Bint Abdulrahman University, Ray . Her areas of interest are digital communications,signal processing, image processing, and coding.

Mahmoud Abdalla received the B.Sc. and M.Sc., from the faculty of Engineering, Mansoura University, Egypt in 1979 and 1984 respectively, and Ph.D. degrees from the faculty of Engineering, Zagazig University, Egypt in 1989. He is currently work as a Professor of Electronic and Electrical Communication, Faculty of Engineering, University of Zagazig. His current research interests include digital Signal Processing ,Neural Networks image enhancement and digital communication.

Fathi E. Abd El-Samie received the B.Sc.(Hons.), M.Sc., and Ph.D.degrees from Menoufia University, Menouf, Egypt, in 1998, 2001, and2005 respectively. Since 2005, he has been a Teaching Staff Member with the Department of Electronics and Electrical Communications,Faculty of Electronic Engineering, Menoufia University. He is currently a researcher at KACST-TIC in Radio Frequency and Photonics for the e-Society (RFTONICs). He is a coauthor of about 200 papers in international conference proceedings and journals, and five textbooks.His current research interests include image enhancement, imagerestoration, image interpolation, super-resolution reconstruction of images, data hiding, multimedia communications, medical image processing, optical signal processing, and digital communications. Dr.Abd El-Samie was a recipient of the Most Cited Paper Award from the Digital Signal Processing journal in 2008.

How to cite this paper: Amira Baumy, Naglaa. F Soiliman, Mahmoud Abdalla, Fathi Abd El-Samie,"A Discriminative Statistical Model for Digital Image Forgery Detection", International Journal of Engineering and Manufacturing(IJEM), Vol.6, No.6, pp.1-14, 2016.DOI: 10.5815/ijem.2016.06.01