International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2017 IJSRCSEIT | Volume 2 | Issue 6 | ISSN : 2456-3307

Feature Based Analysis of Copy-Paste Image Tampering Detection

Dr. Kusam Sharma[*1], Prof. Pawanesh Abrol[2], Prof. Devanand[3]

[*1]Department of Computer Science & IT, University of Jammu, Jammu, J&K, India
[2]Department of Computer Science & IT, University of Jammu, Jammu, J&K, India
[3]Department of Computer Science & IT, Central University of Jammu, Jammu, J&K, India

ABSTRACT

Authentication of a digital image is a challenging task. A tampered image is created by altering some of its contents using standard image processing tools. Copy-paste tampering is created by copying some part of an image and pasting it within the same image for covering unwanted information or an object, is the most used technique in digital image manipulation. The motive of copy-paste tampering detection technique is to locate regions that have been copied and pasted within the same image. A number of techniques are employed to detect copy-paste tampering; using image features / parameters is also one of them. In the present research work, a parametric non-overlapping block-based tampering detection model has been applied to ensure the presence of copy-paste tampering in a given digital image. The behaviour of different parameters has been analysed after their implementation onto a wide variety of digital images having different types, formats and dimensions. Statistical parameters of the input images of three different formats are computed, analysed and compared with those of their tampered images using specific threshold values. The model is tested for three different formats and for seven different selected block sizes. The results show that the proposed model identifies the tampered area for all the given images and works well with low to moderate copy-paste tampering. The results obtained can be used as the initial verification of the images for tampering and to enhance the tampering detection process by identifying most likely cases of possible image tampering. The proposed model is tested with larger domain of images having different types, formats and dimensions and for tampering within an image. However, the model has limitations with certain geometrical transformations.

Keywords : Copy-Paste Tampering, Block Based Tampering Detection Techniques, Overlapping Block Based Techniques, Non-Overlapping Block Based Techniques.

I. INTRODUCTION

With the advancement in technology, several pre-existing and latest image-processing tools are used to alter the material elements of an image thus resulting into a digital image tampering. Copy-paste tampering is the widely used and most common among the different types of digital image tampering and is broadly classified into copy-paste tampering in one image and digital splicing with different images. Copy-paste tampering in one image is a special type of tampering where a specific part of the image is copied and pasted somewhere else in the same image with a purpose to hide an object or information. As the copied part comes from the same image its colour palette, dynamic range and many other important features will almost be similar with the rest of the image [1]. Later on, some post processes like edge smoothing, blurring and noise addition are used to remove the visible clues. Several researchers are using different techniques for detection of copy-paste tampering. Some of the widely accepted and used copy-paste tampering detection techniques include SIFT [2], SURF [3], improved SVD [4], DWT-SVD [5], SVD [6], FMT [7], Lin's method [8], DCT-DWT [9], improved DCT [10, 11, 12], DCT-SVD [13], LUO's method [14], PCA [15], DWT [16], etc. In this research paper, block based copy-paste tampering detection techniques are considered. In block-based

techniques, the original image is further divided into either overlapping or non-overlapping blocks. Several transformations are applied on to these blocks to generate the feature vectors of the image features. Block based tampering detection techniques can be non-overlapping or overlapping. The block-based copy-paste tampering detection techniques are invariant to various other transformations like blurring, brightness changes and flipping.

Block based tampering detection techniques can be parametric or non-parametric. The parametric block based tampering detection technique analyses a wide variety of image parameters like statistical, geometrical and textural. The statistical parameters are very significant and manipulation in an image can be detected by analysing the behaviour of these statistical parameters. These techniques can be applied in various fields like image enhancement, image restoration, image denoising, digital image tampering detection and edge detection & eye gazing [17] etc. The analysis of these parameters of an image helps in determining and locating the tampered region within an image [18].

The present research work is carried out for the implementation of the proposed non-overlapping block based parametric tampering detection model for detecting copy-paste tampered regions with in an image and to see the behaviour of different parameters after their implementation onto a wide variety of digital images having different types, formats and dimensions. In this type of tampering, different features of an image are taken into consideration for further testing and tampering detection that includes image formats, dimensions, number and size of blocks. The model is tested and analyzed for different image databases (i.e. eyes, facial and topographical).

Understanding of the parameters in accordance with different image types, formats and dimensions may help in optimizing different image processing models especially in the field of tampering detection. There can be different statistical parameters like mean, median, mode, standard deviation, variance, covariance, skewness, kurtosis etc for detecting image tampering. In this research work, the behaviour of three statistical parameters i.e. variance, skewness and kurtosis are analysed for further tampering detection. Variance of an image is used where pixel variation of images belongs to particular class are same. Variance is normally used to find how each pixel varies from the neighbouring pixel or centre pixel and is used to classify them into different regions. The variance is used in identifying sharp details such as edges [19]. Skewness and kurtosis being the shape parameters characterizes the tails of a probability model rather than the central portion. Due to of which any two probability models with same skewness and kurtosis will have similar shapes [20] [21] [22]. Fig. 1 shows a Feature based copy-paste tampering wherein Fig. 1a) shows original image and Fig. 1b) shows tampered image obtained after feature based copy-paste tampering where a small segment of left top and bottom region is copied and pasted on the right region of the same image.

a. Original image **b.** Copy-Paste tampered image

Figure 1. Feature based copy-paste tampering.

The organization of the present research paper is as follows. Literature review and objectives are presented in the next two sections II and III. The proposed model and its experimental results and discussions are presented in section IV and V respectively. In section VI, inferences and conclusion is discussed and section VII presents limitations and future work.

II. METHODS AND MATERIAL

In order to perform feature based analysis for detecting copy-paste tampering, a parametric tampering detection model using non-overlapping block-based technique has been implemented. The schematic block diagram of the non-overlapping block based parametric l model for copy-paste tampering detection for images having different types, formats, dimensions and block sizes is presented in Figure 2.

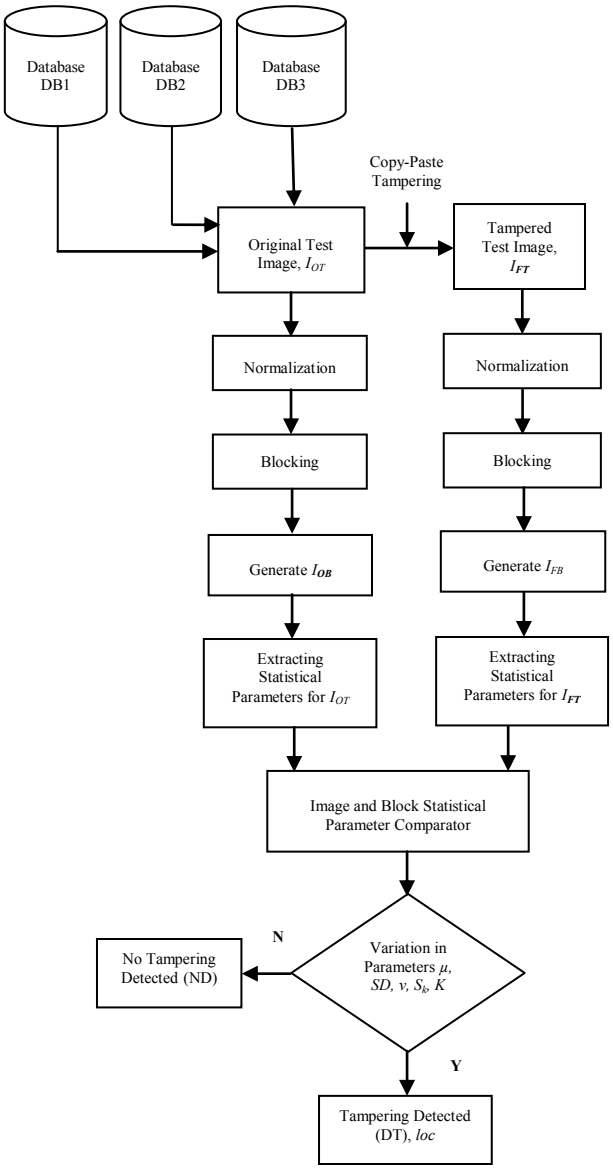

Figure 2. Schematic block diagram of proposed non-overlapping block-based statistical model for copy-paste tampering detection for images having different types, formats, dimensions and block sizes.

Initially an original test image (I_{OT}) of any type, format and dimension is taken from the image bank. After normalization, I_{OT} is divided into non-overlapping blocks of size [m x m] depending on the values of block row (B_r) and block column (B_c). A two or three-dimensional block matrix B is created depending on the number of colour bands (N_{CB}), resulting in the generation of blocked image (I_B) along with image residue (I_r) based on the size of image and block selected. After blocking of an image, the three parameters variance (v), skewness (S_k) and kurtosis (K) of the original image are computed. The entire process is repeated simultaneously for tampered image (I_{FT}), created after doing manual tampering in the original image. This copy-paste tampered image is again segmented into non-overlapping blocks and all parameters under study are computed. Each corresponding block of I_{OT} are analysed and compared with those of I_{FT} for further tampering detection. The statistical variation $\partial p = |P_{IO} - P_{IF}|$ is computed and analysed for all selected parameters for ascertaining the tampering and its extent.

The threshold t, defines the permissible proportion of variation in the parameter under study. The value l of threshold t is set after testing a wide range of image sets and analysing the behaviour of parameters. The tampering status DT (detected) or ND (not detected) is established on the basis of variations. Further analysis is done to find the specific location, i where the tampering is actually been done. The proposed interface for tampering detection is developed in MATLAB version 7.6.0.324 (R2008a) and 8.1.0.604 (R20013a) and is tested for more than 400 images taken from different image databases (i.e. eyes, facial and topographical) shown in Table I. The images taken from these databases are having different dimensions ranging between (10 x 10 to 1000 x 1000) and formats i.e. bmp, png and jpg and for different selected block sizes (i.e. standardized-2, non-standardized-5), S=5, 10, 20, 25, 32, 50 and 64. The experimental results thus obtained for S=25 are further analyzed and discussed in the next section.

Image domains	Some Selected images			
Facial				
Eyes				
Topography				

Table 1. Image databases used

III. RESULTS AND DISCUSSION

The experimental results are obtained after the implementation of the proposed parametric non-overlapping block-based model for copy-paste tampering detection for images taken from different image databases having different dimensions and three different formats. The cases shown in Table-II are selected cases out of 400 test results where there is significant variation of the different statistical parameters of original and tampered gray scale images in three formats and for block size S=25 out of selected

block sizes depicting copy-paste tampering detection based on threshold *t*. The statistical variation ∂ in the given above said parameters for original and tampered images and their blocks is computed for each parameter. The statistical variation in each parameter is given by ∂_v, ∂_{Sk} and ∂_K.

Selected block size, S = 25	Test image I_T	Image formats	Parametric variation (t>=*l*)			Tampering status
			t_1=3	t_2=0.6	t_3=3	
			∂_v	∂_{Sk}	∂_K	
	I_{T1}	BMP	2.66	0.06	0.28	DT
		PNG	2.66	0.06	0.28	
		JPEG	2.75	0.07	0.29	
	I_{T2}	BMP	1.01	0.06	0.19	DT
		PNG	1.01	0.06	0.19	
		JPEG	1.09	0.05	0.17	
	I_{T3}	BMP	5.60	0.31	0.91	DT
		PNG	3.78	0.26	0.79	
		JPEG	5.60	0.30	0.90	
	I_{T4}	BMP	7.78	0.66	2.05	DT
		PNG	7.78	0.66	2.05	
		JPEG	7.61	0.66	2.06	
	I_{T5}	BMP	28.80	2.27	8.96	DT
		PNG	24.56	4.95	13.05	
		JPEG	24.42	4.94	13.02	

Table 2. Parametric differences of original and tampered images having different formats depicting copy-paste tampering detection dased on threshold *t*, (*t*>=*l*)

The values *l* and *t* are set after testing a wide range of image sets and analysing the behaviour of statistical parameters. Based on the variation, the tampering status is established. Further analysis is done to find the specific location of the tampering in the image after dividing it into non-overlapping blocks. The value of threshold t_1 for ∂_v, t_2 for ∂_{Sk} and t_3 for ∂_K is set to 3, 0.6 and 3 respectively. All the images I_{T1} - I_{T5} are showing least variation for S_k and maximum variation for *v*. The statistical variation for all the three parameters is similar for images I_{T1}, I_{T2} and I_{T4} in bmp and png format. For test image I_{T3}, the statistical variation for bmp and jpeg is almost similar for all the parameters and for test image I_{T5}, the variation for png and jpeg is almost similar for all the selected parameters.

Moreover, the value of ∂_{Sk} for the images I_{T1} - I_{T5} is greater than 1.0 for all the three formats i.e. bmp, png and jpeg and for different dimensions, which signifies that the ∂_{Sk} is substantial and the distribution is far from symmetrical. These asymmetrical distributions will have long tail to the right and a positive skew. Also, it is deduced from the results generated above that the value of ∂_k for all the images in different formats is greater than 0. These positive kurtosis images would have a fairly uniform distribution of gray levels.

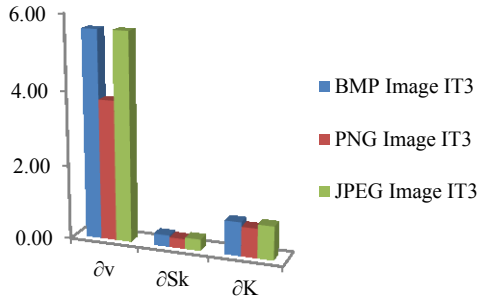

Figure 3. Parametric difference in ∂_v, ∂_{Sk} and ∂_K for BMP, PNG and JPEG original and tampered image I_{T3} for selected block size S=25

Figure 3 shows the parametric variation in variance, skewness and kurtosis for bmp, png and jpeg original and tampered image I_{T3} for selected block size S=25 thus depicting the presence of tampering. The variation ∂_v for image I_{T2} in bmp and jpeg format shows maximum variation than for png format.

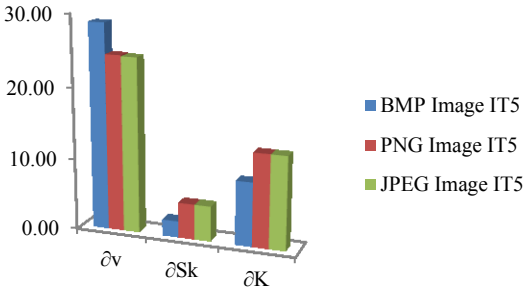

Figure 4. Parametric difference in ∂_v, ∂_{Sk} and ∂_K for BMP, PNG and JPEG original and tampered image I_{T5} for selected block size S=25

Variation in all the three parameters for bmp, png and jpeg original and tampered image I_{T3} can be seen in Fig. 4 for a selected block size S=25 thus showing the presence of tampering. The parametric variation ∂_v for image I_{T5} in bmp format shows maximum variation than for png and jpeg format. The variation ∂_v for image I_{T5} in bmp format shows maximum variation than for png and jpeg format.

The Table-III shows the parametric variation ∂_i along with corresponding threshold t_i for selected two different images I_{T4} and I_{T5} in three different formats i.e. bmp, png and jpeg and their corresponding blocks. These two images are selected from Table-II on the basis of significant variation in their parameters for different formats and their corresponding graphs. l for each parametric variation has been computed after testing a wide range of images. Any t_i less than corresponding l is considered as *ND* otherwise considered as *DT*. Further the table also shows the corresponding blocks of both the images in three different formats in which tampering has been observed.

Test image I_T	Image formats	Image blocks	Parametric variation (t>=*l*)			Tampering status & location, loc
			t_1=3	t_2=0.6	t_3=3	
			∂v	∂S_k	∂K	
I_{T5}	BMP	B_1	13.82	0.47	1.21	DT (B_1, B_2, B_3, B_4)
		B_2	31.38	8.86	7.04	
		B_3	45.39	1.49	0.69	
		B_4	17.35	0.35	1.04	
		I_{T5}	28.80	2.27	8.96	
	PNG	B_1	10.74	1.63	4.48	
		B_2	30.88	8.76	6.70	
		B_3	34.17	1.43	0.85	
		B_4	16.45	0.21	0.60	
		I_{T5}	24.56	4.95	13.05	
	JPEG	B_1	10.72	1.63	4.48	
		B_2	30.67	8.70	6.67	
		B_3	34.02	1.42	0.87	
		B_4	16.31	0.21	0.59	
		I_{T5}	24.42	4.94	13.02	

Table 3. Block-wisk parametric differences of original and tampered image having different formats depicting copy-paste tampering detection based on threshold t,(t>=l) for selected block size, s=25

Further in Table 3, blocks B_3, B_4 of test image I_{T4} and blocks B_1, B_2, B_3 and B_4 of test image I_{T5} for bmp, png and jpeg format shows parametric variation in all the three parameters thus confirming the existence of tampering. The blocks B_1 and B_2 of test image I_{T4} show no variation in any of its parameters thus depicting that no tampering has been done in these blocks.

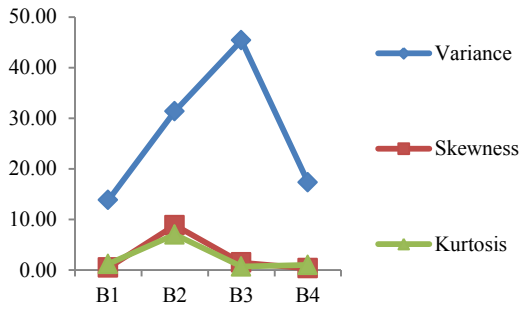

Figure 5. Block-wise parametric variation in *v*, S_k and *K* of original and tampered BMP image I_{T5} for selected block size S=25.

Figure 5 shows the relationship among three parameters variance, skewness and kurtosis for blocks of original and tampered image I_{T5} in bmp format for selected block size, S=25. The variation in these parameters shows tampering in all the blocks. ∂_{Sk} shows minimum variation whereas ∂_v shows maximum variation for all the blocks of image I_{T5} in bmp format.

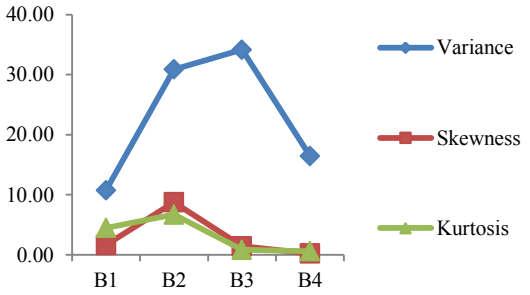

Figure 6. Block-wise parametric variation in *v*, S_k and *K* of original and tampered PNG image I_{T5} for selected block size S=25.

The block-wise parametric variation in *v*, S_k and *K* of original and tampered png image I_{T5} for selected block size S=25 is shown in Fig. 6. The variation in these parameters shows tampering in all the blocks. The block B_4 shows minimum variation for ∂_{Sk} whereas block B_3 shows maximum variation for ∂_v of image I_{T5} in png format.

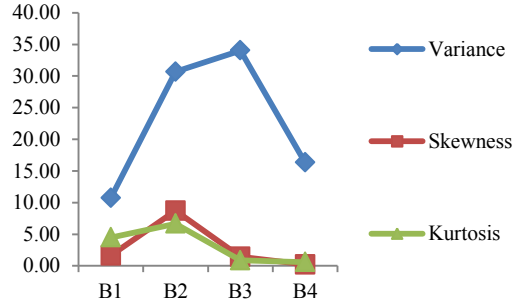

Figure 7. Block-wise parametric variation in v, S_k and K of original and tampered JPEG image I_{T5} for selected block size S=25.

Figure 7 shows the block-wise parametric variation in v, S_k and K of original and tampered jpeg image I_{T5} for selected block size S=25. The block B_4 shows minimum variation for ∂_{Sk} whereas block B_3 shows maximum variation for ∂_v of image I_{T5} in png format. ∂_{Sk} shows minimum variation whereas ∂_v shows maximum variation for all the blocks of image I_{T5} in jpeg format. The block B_4 shows minimum variation for ∂_{Sk} whereas block B_3 shows maximum variation for ∂_v of image I_{T5} in jpeg format.

During the process of detection of copy-paste tampering, it is very important to select and set the appropriate size of the smallest block of an image. For which several experiments have been conducted for block size and parameter selection. Larger size of image block increases computational complexity whereas smaller size causes too many false matches. Finally, seven block sizes S=5, 10, 20, 25, 32, 50 and 64 were selected and considered for further experimentation. The experimental results thus obtained for S=25 out of selected block sizes are further shown, analyzed and discussed in this research paper. This experimental model is tested on different images having different dimensions and formats. The experimental analysis is done for three very common and widely used image formats i.e. bmp, png and jpeg are considered.

Experimental results obtained deduce that the parameters of an original image and its blocks vary from the parameters of a tampered image and its blocks thus ensuring the existence of tampering in original images. From Table-II and Table-III, it is observed that the value of each parameter for each block of an original and tampered png and jpeg image and image itself are close to each other than the value of parameters of blocks of original and tampered bmp image. The value of variance, skewness and kurtosis for original png image is more close to those of jpeg image than the bmp image.

IV. CONCLUSION

The proposed non-overlapping block based model for detection of copy-paste tampering within an image is based on the analysis of statistical parameters and is implemented on a wide variety of digital images having different formats and dimensions. Using this model, three parameters have been analysed using specific threshold values. These threshold values have been determined by performing special different tests for each of the selected parameters. More than 400 images taken from different domains having different formats and dimensions have been tested using MATLAB interface. The proposed model is tested for three different formats and for seven different selected block sizes. Manual tampering has been induced at different locations within the given image thus generating a tampered image.

The results observed for different parameters and for different images having different types, formats and dimensions along with their blocks depict the tampering status, its extent and location as per careful selection of t and l values. Out of three parameters, variance and kurtosis are the ideal parameters showing maximum variation in original and tampered images for different formats and for different selected block sizes. The results obtained can be used as the initial verification of the images for tampering and to enhance the tampering detection process by identifying most likely cases of possible image tampering.

Moreover, it is deduced from the generated results that the use of non-overlapping blocks instead of overlapping blocks also reduces the time complexity and the value of the statistical parameters is dependent on the size of the image blocks. With the increase in size of image block, value of parameter also increases. The final verdict about tampering is given after comparing the block-wise values of the features of original image and tampered image. Moreover, it is deduced from the results that the proposed model works well with low to moderate copy-past tampering.

The limitation of this model is that the image formats gif and tiff are not considered for testing because during normalization the color and image quality of gif image reduces whereas tiff image is having 4 dimensions instead of 2 as in grayscale images or 3 as in colored images which is not supported by the present proposed experimental model.

The proposed model generates reasonable results and is further expanded for generating and analyzing the

impact of transformation over a wider range of images having different types, formats and dimensions.

V. REFERENCES

[1]. Kusam, P. Abrol, and Devanand, "Digital Tampering Detection Techniques: A Review", BVICAM's International Journal of Information Technology, vol. 1, no. 2, pp. 125-132, 2009.

[2]. X. Pan and S. Lyu, "Region duplication detection using image feature matching", IEEE Transactions of Information Forensics and Security, vol. 5, no. 4, pp. 857-867, Dec. 2010.

[3]. B. L. Shivakumar and S. S. Baboo, "Detection of region duplication forgery in digital images using SURF", International Journal of Computer Science Issues, vol. 8, no. 1, pp. 199-205, Jul. 2011.

[4]. Li Kang and X. Cheng, "Copy-move forgery detection in digital image", IEEE 3rd International Congress on Image and Signal Processing, vol.5, Oct. 2010, pp. 2419-2421.

[5]. G. Li, Q. Wu, D. Tu, and S. Sun, "A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD," IEEE International Conference on Multimedia and Expo, Jul. 2007, pp. 1750-1753.

[6]. D. Sharma and P. Abrol, "SVD Based Noise Removal Technique: An Experimental Analysis", International Journal of Advanced Research in Computer Science, vol. 3, no. 5, pp. 214-218, Sept. – Oct. 2012.

[7]. S. Bayram, H. T. Sencar and N. Memon, "An efficient and robust method for detecting copy-move forgery," Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing, Apr. 2009, pp. 1053-1056.

[8]. H. J. Lin, C.W. Wang and Y.T. Kao, "Fast copy–move forgery detection", WSEAS Transactions on Signal Processing, vol. 5, no. 5, pp. 188–197, 2009.

[9]. X. Wang, X. Zhang, Z. Li and S. Wang, "A DWT-DCT based passive forensics method for copy-move attacks", IEEE Third International Conference on Multimedia Information Networking and Security, Nov. 2011, pp. 304-308.

[10]. Y. Huang, W. Lu and D. Long, "Improved DCT-based detection of copy-move forgery in images", Forensic Science International, vol. 206, issues 1-3, pp. 178-184, Elsevier, March 2011.

[11]. Y. Huang, W. Lu, W. Sun and D. Long, "Improved DCT-based detection of copy-move forgery in images", Forensic Science International, vol. 206, pp.178–184, Elsevier, 2011.

[12]. Y. Cao, T. Gao, L. Fan and Q. Yang, "A robust detection algorithm for copy-move forgery in digital images", Forensic Science International, vol. 214, pp.33–43, Elsevier, 2012.

[13]. J. Zhao and J. Guo, "Passive forensics for copy-move image forgery using a method based on DCT and SVD", Forensic Science International, vol. 233, issues 1-3, pp. 158-166, Elsevier, Dec. 2013.

[14]. W. Luo, J. Huang, and G. Qiu, "Robust Detection of Region Duplication Forgery in Digital Images", In Proceedings of the 18th International Conference on Pattern Recognition, vol. 4, Aug. 2006, pp. 746-749.

[15]. A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting duplicated image regions," Department of Computer Science, Dartmouth College, Tech. Rep. 2004-515, 2004.

[16]. J. Zhang, Z. Feng and Y. Su, "A new approach for detecting copy-move forgery in digital images", 11th IEEE International Conference on Communication Systems, Nov. 2008, pp. 362-366.

[17]. A. Sharma and P. Abrol, "Research Issues in Designing Improved Eye Gaze Based HCI Techniques for Augmentative and Alternative Communication", International Journal of Emerging Technologies in Computational and Applied Sciences, vol. 6, no. 2, pp. 149-153, September-November 2013.

[18]. V. Kumar and P. Gupta, "Importance of statistical measures in digital image processing", International Journal of Emerging Technology and Advanced Engineering, vol. 2, Aug. 2012.

[19]. M. Tajrobekar. (2014, March 8). "Where must we use variance and mean of image?," [Online]. Available:http://www.researchgate.net/post/Where_must_we_use_variance_and_mean_of_image.

[20]. Dattatherya, S.V. Chalam and M.K. Singh, "A generalized image authentication based on statistical moments of color histogram", ACEEE International Journal on Recent Trends in Engineering and Technology, vol. 8, no. 1, Jan 2013.
[21]. D.J. Wheeler. (2011, July 29) "Problems with Skewness and Kurtosis, Part One," [Online]. Available:http://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-one.html.
[22]. D.J. Wheeler. (2011, January 08) "Problems with Skewness and Kurtosis, Part Two," [Online]. Available:http://www.qualitydigest.com/inside/quality-insider-article/problems-skewness-and-kurtosis-part-two.html.