

A Study to Ascertain and Differentiate between Genuine and Transplanted Documents/Signatures

Manisha Mann*, Shukla SK and Shruti Gupta

Amity Institute of Forensic Sciences, Amity University, Noida, UP-201303, India

Abstract

Digital forgery is very common these days because the use of digital document is increasing day by day and readily available software's for manipulating the documents. It is very necessary to authenticate the integrity of digital document, whether it's genuine or not. Various secure detection systems are made which involves use of algorithm, but this paper deals with examining simple features present in image/document which can be used to check the authenticity of an electronic document, whether they are genuine or forged. These simple features are like Color variation, Font size difference, Pixelate resolution, Range of magnification-low, medium and high range of magnification and extracting hash value-MD4, MD5, SHA-1, RIPEMD-128 and RIPEMD-160. Results are highly reliable.

Keywords: Digital documents; Digital forgery; Image hashing; Pixels

Introduction

Now a days use of electronic documents have been increased tremendously which results in increase in digital forgery. A forgery is an unlawful act of forging a document or item for the reasons of misrepresentation or trickery. Digital forgery is very common now, because digital images are not difficult to manipulate and alter due to easily available image processing and editing software's. These days, it is feasible to include or exclude any important characteristics from a picture without leaving any conspicuous hints of altering. Digital forensics is the current topic which has received attention recently. Digital images play an important role in depicting and transferring the data easily, therefore new techniques for detection of forgery in digital images have been investigated. There is a little difference between image forgery and digital image forgery, as the digital image forgery deals with the digital image as compared to photographs which are used in image forgery. There are many different computer graphic editing software are available like Adobe Photoshop, GIMP (GNU Image Manipulation Program), and Corel Paint Shop, etc.

Areas like legal, criminal, journalism, medical requires the digital document to be authentic. So there is high demand for a dependable, safe and secure detection system, which is capable to determine or check whether the digital image/document is real or altered. As due to presence of easily available image editing software's alteration could be done to the digital document and some modification may be impossible to be seen by human eye, these modification results in some core statistics changes in the digital document which can be detected.

Forgery techniques in digital images are classified into three main groups

- Copy Paste Forgery (Image splicing).
- Image retouching.
- Copy Move Forgery (Image cloning).

Copy paste forgery also known as image splicing: In this type an altered duplicate copy or a document is prepared with the help of an original image along with some additional images, for instance including particular area of the additional image to the original one, just to hide or manipulate the image.

Image retouching: In this type the forger manipulates the image in a way so that the modification in the content of the image becomes unnoticeable.

Copy move forgery also known as Image cloning: In this type of forgery a distinct part of an image is copied and moved to another part of the same image.

Usually there are two types of detection techniques or approaches are used. Following are the two techniques

Active method includes features like watermarking which helps in detecting digital tampering like name, signature, date, etc.

Passive method in this method digital image forgeries are detected without taking any use of the features or information of the original image [1].

Image hashing/Image fingerprinting

Image hashing or image fingerprinting is a procedure of providing value that is specific to particular image by examining its content. This fingerprint is basically a string which is assimilated with other fingerprint for possible matches [2]. There are two types of image hashing techniques:

Perceptual hashing: Perceptual hashing a technique which produces fingerprint of different multimedia like audio, video or image file with the help of algorithm. Perceptual hash functions (PHF) are most commonly used in area of digital forensics and protection against copyright infringement. Through PHF one can assimilate and map source the data with the help of correlation between the hashes. For

***Corresponding author:** Manisha Mann, Amity Institute of Forensic Sciences, Amity University, Noida, UP-201303, India, Tel: +919910471429; E-mail: Manisha.mann001@gmail.com

Received May 21, 2015; **Accepted** July 21 2015; **Published** July 26, 2015

Citation: Mann M, Shukla SK, Gupta S (2015) A Study to Ascertain and Differentiate between Genuine and Transplanted Documents/Signatures. J Forensic Res 6: 293. doi: [10.4172/21577145.1000293](https://doi.org/10.4172/21577145.1000293)

Copyright: © 2015 Mann M, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

instance, Wikipedia contains a database of hashes of online articles and books for which copyright is hold by the authors, when any Wikipedia users upload any document containing data of online books or article registered in that database will have same hash value and will match with each other and this can be used in flagging or pointing out plagiarism [3].

Cryptographic hashing: Cryptography is a technique in which a message authentication code (MAC) is produced with the help of hash function. Even a slight change in input message results in distinct hash value. Following are mostly used hash algorithm:

- MD (Message Digest Algorithm) - MD4, MD5.
- SHA (Secure Hash Algorithm) – SHA-0, SHA-1.
- RIPEMD (RACE Integrity Primitive Evaluation) – RIPEMD-160 [4].

Pixels: Pixel is derived from a word “*picture element*”. In a computer picture, a pixel is the simplest unit of programmable color. The size of the pixel relies on the resolution of the display screen. If the display is at its highest resolution, the physical size of pixel will be equal to dot pitch of display. If the resolution is less than the highest, then size of pixel will be larger than dot pitch [5].

Early studies on this topic were Weihai li; et al. developed a method for detection of copy paste forgery in manipulated JPEG pictures and also locates the position or the area of manipulation. This method works by extracting DCT block artefact grid and determining mismatch of grid [6]. Thirumagal, et al. proposed a forensic technique for detection of contrast enhancement (globally or locally applied) and by identifying the peculiarities of intrinsic fingerprint the histogram equalization in a picture can be detected [7]. Najah Muhammad, et al. proposed an effective non-intrusive technique for detection of copy move forgery. In this technique the image is segmented and the similarity is detected with the help of Dyadic Wavelet Transform (Dy WT) [8]. R. Venkatesan, et al. proposed an image indexing technique which is known as image hash function. Randomized signal processing is used by algorithm for a non-reversible compression of an image which results into arbitrary binary strings [9]. Kelsy Ramirez-Gutierrez, et al. proposed two algorithm to detect authenticity of an image, even if the image is affected by distortion like filtering, compression and other malevolent modification like geometric distortion. The algorithm can also detect tampering and also the localized tampered areas [10].

This paper deals with examining simple features present in image/document which can be used to check the authenticity of an electronic document, whether they are genuine or forged. These simple features are like Color variation, Font size difference, Pixelate resolution, Range of magnification: low, medium and high range of magnification and extracting hash value: MD4, MD5, SHA-1, RIPEMD-128 and RIPEMD-160.

Material and Methods

Sample size

25 samples of documents were created which contained signatures, dates, names and addresses, which have been transplanted from the originals on those documents. Doctor's prescription, list of student selected in any institution, stamp papers, certificates and appointment letters are the type of samples.

Sample collection

The 25 samples were collected from Google Images.

Procedure adopted for analysis

As original disputed documents cannot be gathered due to their authorization and confidentiality, which should be maintained by government forensic laboratories with due reason, such documents were prepared manually for the research with the help of software's. These samples were then analyzed in soft copy format. Then the signatures were cropped from the originals, copied and pasted on the documents to be forged with the help of MS Paint. These forged documents were then examined in soft copy format with the help of Picasa. Following are the features on the basis of the samples were examined:

- Colour variation.
- Font size difference.
- Pixelate resolution.
- Range of magnification: low, medium and high range of magnification.
- Hash value: MD4, MD5, SHA-1, RIPEMD-128 and RIPEMD-160

Software's and application used for analysis

- Microsoft Paint or MS Paint is a simple graphics program that has been included with all versions of Microsoft Windows. This program can be in colour mode or two-colour, black-and-white, but there is no grayscale mode.
- Picasa 3.9 is an image organizer and image viewer for organizing and editing digital photos plus an integrated photo-sharing website, originally created by a company named Lifescape in 2002 and owned by Google since 2004.
- Fileformat.info is a website which reveals the all types of hash value of any type of file like audio, video or any document. Basically it's an online hash value calculator.

Samples (observed features)

Since the number of samples created and observed for this study is very large, it would not be convenient to attach all the samples and pictures of the features observed in each sample. Therefore, only a few samples are being attached with the zoomed in images of their parts showing some difference from the whole of the document, indicating different sources of origin (Figures 1-4).

Result and Discussion

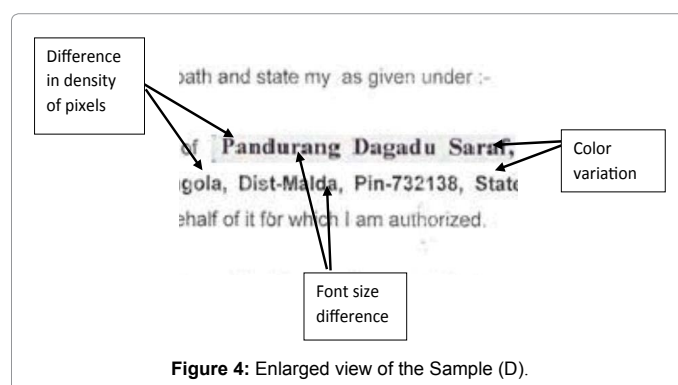
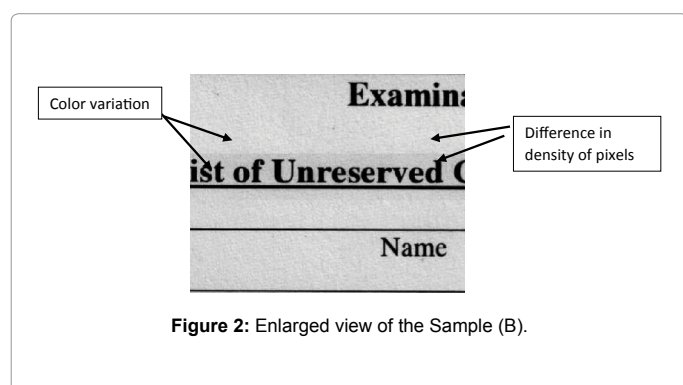
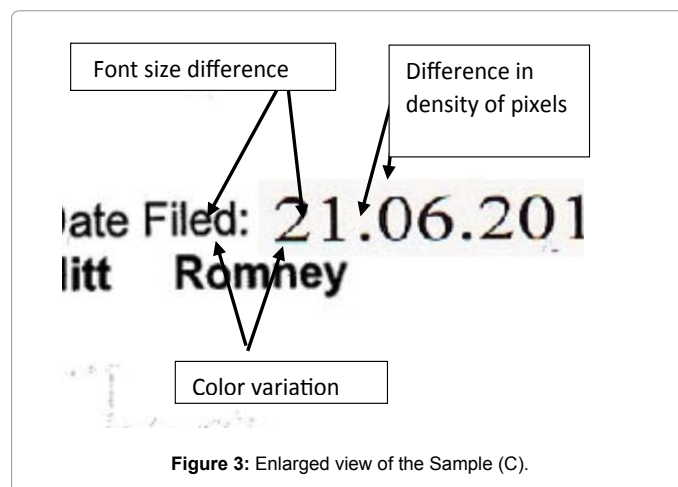
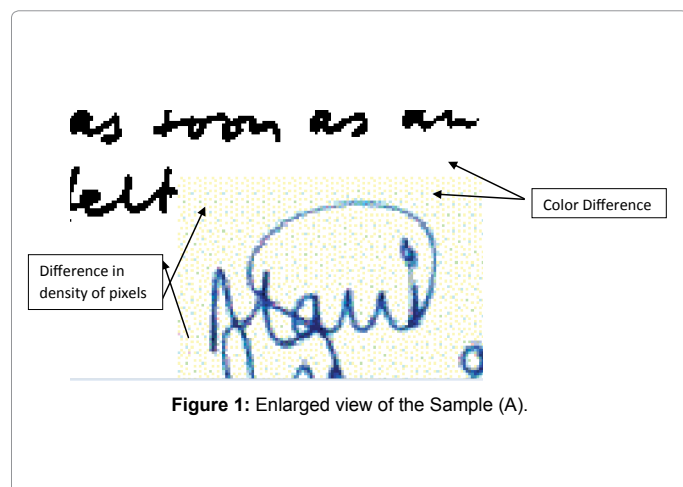
The following 25 samples were examined on basis of different features in Picasa Software:


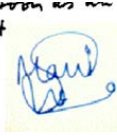
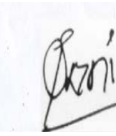
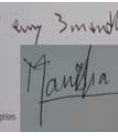


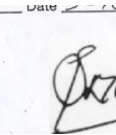
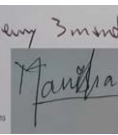
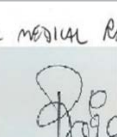
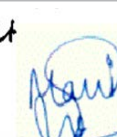

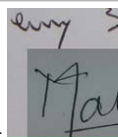
- Range of magnification: low, medium and high range of magnification.
- Colour variation.
- Font size difference.
- Pixelate resolution.


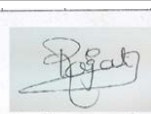
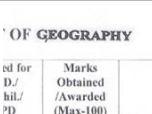
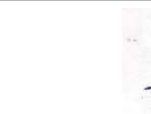
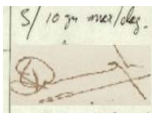
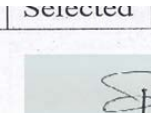
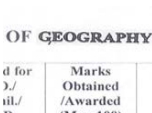

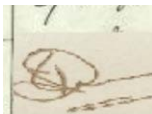
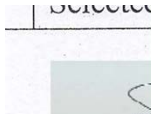
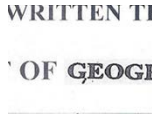

Above mentioned features are given in Table 1.


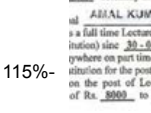
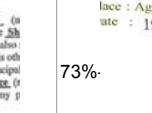


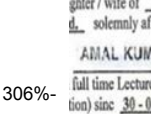

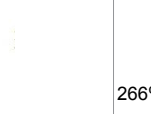




Range of magnification

All the samples were first analyzed by magnifying them. The range



Features/Samples	Sample 1	Sample 2	Sample 3	Sample 4
Color Difference	Present	Present	Present	Present
Font Difference	-	-	-	-
Pixelate Difference	Present	Present	Present	Present
Low-range magnification (%)	105% 	152% 	231% 	26% 
Mid-range magnification (%)	160% 	266% 	352% 	40% 
High-range magnification (%)	369% 	465% 	535% 	106% 
Features/Samples	Sample 5	Sample 6	Sample 7	Sample 8
Color Difference	Present	Present	Present	Present
Font Difference	-	-	Present	-

Pixelate Difference	Present	Present	Present	Present
Low-range magnification (%)	132%- 	45%- 	34%- 	119%- 
Mid-range magnification (%)	175%- 	68%- 	59%- 	208%- 
High-range magnification (%)	465%- 	181%- 	157%- 	419%- 

Features/Samples	Sample 21	Sample 22	Sample 23	Sample 24
Color Difference	Present	Present	Present	Present
Font Difference	Present	Present	Present	Present
Pixelate Difference	Present	Present	Present	Present
Low-range magnification (%)	67%- 	115%- 	73%- 	175%- 
Mid-range magnification (%)	201%- 	306%- 	100%- 	266%- 
High-range magnification (%)	814%- 	1076%- 	535%- 	814%- 
Features/Samples	Sample 25			
Color Difference	Present			

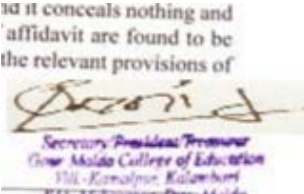
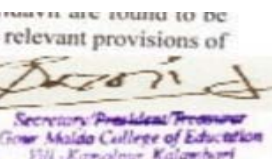
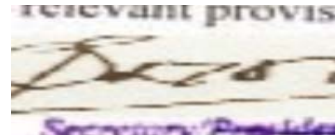
Font Difference	-
Pixelate Difference	Present
Low-range magnification (%)	309%- 
Mid-range magnification (%)	541%- 
High-range magnification (%)	1252%- 

Table 1: Shows the observation made on examination of samples with transplanted signatures on basis of Range of magnification, Color variation, Font size difference and Pixelate resolution.

Sample	Original Sample		Simulated Sample	
1	MD4	34e60c92ce5cc5b80ffc86dff636cc9	MD4	195522f215886931bcc957710df7b9da
	MD5	547626b3ad63062faa47fc09b3a4bdf8	MD5	cc9399c13d221a29634609bb79a30068
	RipeMD128	2fc0a2e99761f9b01823de9d078daf86	RipeMD128	caa0ab9adeb53aec7530b2f5907729d0
	RipeMD160	a443a857fd7388a58ca874bfc15e070f96de83f8	RipeMD160	90b548d6de1dd0f5853e9415bebc44c13706785
	SHA-1	c22c1e1ca93797651c4c12bf56c51d3b1774d409	SHA-1	8027cc878585510c8d01341bc3efa4c71e259125
2	MD4	c09f10c73a2e970c0d67dc8a5fca3f24	MD4	89f7abce148018daa6486be49077658e
	MD5	627f26cb5e2b19fb7ac42c4f82c7344a	MD5	ffbad43edfd1fb34e52afa33cbd35de9
	RipeMD128	f1f804fdb6145b980ade9470126f249f	RipeMD128	ce2259b7a0bf06db701adc4f280f8798
	RipeMD160	190ebc7e27430377c71c6b45659bd3f4d46cb79	RipeMD160	4d43ed392826af8d2dab585dc43156c3858b38ca
	SHA-1	59313a8a1e2a68b9d977b47958e22f77d7297750	SHA-1	5200700a04963b99756762a1d3328cb8e10bbeed
3	MD4	8585091a427cd5f9f081543460e827e1	MD4	8cd64294745f69f6a2d495e52a844c57
	MD5	11484a3e1860019303556a081835b194	MD5	ed9523be8127b46ab567f52fb0c69fd9
	RipeMD128	029e1c3da86efd9ea16cff4e4364e2c3	RipeMD128	625eb0d620764900b533bba4c1d83cbd
	RipeMD160	e2518f62b450f4118b015f38fa3e3ad9d6f5ae63	RipeMD160	c5c7fa61f56dc67a4effba113c4df113685db545
	SHA-1	eaaf2ca6a0b54be2d073910e7dd7a30af7699d2b	SHA-1	03381802d2537ef7a29fd9afe2b4f49b39271746
4	MD4	e485f01784b65e14210f7a19554ec68a	MD4	8cd64294745f69f6a2d495e52a844c57
	MD5	8ed4c472c40c71ef797a9ec695209d5fb	MD5	ed9523be8127b46ab567f52fb0c69fd9
	RipeMD128	0ef26f21f9588cf54678c63536c13cfc	RipeMD128	625eb0d620764900b533bba4c1d83cbd
	RipeMD160	e0740defb804dbcfb3c21d5ca743f18f5a533817	RipeMD160	c5c7fa61f56dc67a4effba113c4df113685db545
	SHA-1	4311a59f99f976147b162f669adbed2c45765346	SHA-1	03381802d2537ef7a29fd9afe2b4f49b39271746

5	MD4	041c280894f1091486b0cfd2d2be414b	MD4	0f0da986daf5865624009206b10c1a34
	MD5	4808eb48de6d2b662a2e375f391f57b	MD5	04cd4fc2232ba8a759f6d9e8c7c0d3d8
	RipeMD128	b7249ecf67f84d3c1ede3e910169dd40	RipeMD128	6818554234382e32d62fafb14c778a0e
	RipeMD160	729b5d6b47d92d6cc6d307dea3a76d95815b861a	RipeMD160	fc57668923491d7ce97f30bece618ef9705920ee
	SHA-1	d1a6c16dcb1667e5d2b2bcb619df989f1d1bcadc	SHA-1	0367a2ec5e28408b77a6574f84eb9b8a09a7ad49
6	MD4	9288683235d520c02507cb5655e2228c	MD4	dffd3733cc02b63ff1be9e30ca6967b8
	MD5	d6ee7ad6027cfc7702f73fe660fcb87	MD5	92a3fda4be8161cc080c0d25e9f02385
	RipeMD128	2a6ffe4b2120a90ce3fa4eb53a6c5e24	RipeMD128	60499ea7222a4a34577b6cdd8690082c
	RipeMD160	aff2963e22689179e0c43d1917ca9508c9472249	RipeMD160	c43c0375e08b8261d4bfd0b14528a88974771574
	SHA-1	745b01a751e03c38839df2f1b7148a19dc85f6b4	SHA-1	88a05fddb16fb0e3c94349e4c43fc41d86e8d477
7	MD4	f89a32a2680ed68bd8a3ec6a2d2f6fe5	MD4	0a4cff65dda898c8b51c53510f8d7cd9
	MD5	2aee4249a0b8c3aeb11c32d9e273e56b	MD5	613bed51df2c2116c5936a92517087b1
	RipeMD128	33a29d570a66710ac3ece05fdae2f0ee	RipeMD128	014a1ff8044956e1ef7e6793db953e43
	RipeMD160	54ff2e94aee2bb490222cb6a0ff95d23fa81f40c	RipeMD160	771630a9b19a11afd492b44ce16800e0c0700531
	SHA-1	2957a2ad9758fb45441f2f0dd2c032bdc715c5c8	SHA-1	4aa833ca95d0a626b6ac64e98d12e9de4911968c
8	MD4	3cf16e1061bbeac1e03e3a88ed903356	MD4	39b7ff75415e890377c6f8a87fa532a0
	MD5	146ab572a19d38a5da9aa64817848545	MD5	1c4d242baf9cb7adabef4dfd449bab38
	RipeMD128	6ea515264b83c420cce4af0170ce8e85	RipeMD128	6a42a2a81bfce86a2503a3c66c4254ae
	RipeMD160	988b74357a2da53c85bb0b4101e3ee3dd291c22e	RipeMD160	f56f7d2b4c900d9c218a3553915c6d8588dea231
	SHA-1	7068093c2882207c6aca6e24f7b78deac23f97c	SHA-1	03fe735eed05ecf895a8e9b60d2b2f8b6def24
9	MD4	c5674db9ce02104f8efcfe9d217fe710	MD4	fc72f82425725abe7b5188c8214b66d5
	MD5	45f48e5a5d032da442365d6abd41ca40	MD5	aff3fcffc4338db57dd3fa7d4c2c6914
	RipeMD128	5c1a46be4319ecf35322d2ec31e01d0	RipeMD128	d6a904119b1dbc17ff7acc3f3edd8397
	RipeMD160	64c6df70b0159a61d7aa11daf478f933e8c40dc9	RipeMD160	2b2b26fa59a1bca0fd94bae1134fa69cdb481710
	SHA-1	e8fbdebbe85d028cd32266840aeb770cc58cb54	SHA-1	4cc9c72d46449365b49a4d1d0a7d234bde6073d3
10	MD4	467c5fa689a5d3ae39783979661a2c85	MD4	77937ee57e3ad8f161db9ae6e99ed9e5
	MD5	d32295aa1408cff73d6187dbce68b1d1	MD5	34a490991d24fb8b1d3cd476ccc4eae9
	RipeMD128	21a46949f3f17d9b324c3efa60a5c9d0	RipeMD128	4d6ba7204830a717cb87a261c2710763
	RipeMD160	9719019953db085cb42eb72bd9e0ab60a33368fc	RipeMD160	760cd8955df58e8460833819a0d73a93c1291fdd
	SHA-1	215abf28bc44ccecbfd515c8bca55ed0de59c1a5	SHA-1	e69a8a6acfd9c98e8c6dc6646beb35354779f939d

Table 2: Shows the observation made on examination of samples with transplanted signatures on basis of Hash values.

at which transplantation was first observed was considered as lowest range of magnification for the specific sample. Simultaneously the range of magnification was increased and it was observed that on increase of magnification the transplantation was more prominently seen. All the samples were analyzed between 26 to 1252 ranges of magnification.

Color variation and font size difference

All the samples were analyzed to check color variation in two aspects: color variation in paper background and color variation in ink. Difference in the font size of the samples was examined and observation was noted down in the Table 1. The samples which were transplanted had noticeable difference in their background/ink color as well as font size.

Pixelate resolution

All the samples were analyzed to detect difference in the pixelate resolution. It was found that the transplanted area had difference in their pixelate resolution as compared to the document on which they were transplanted.

All the samples (including the original document and simulated document) were analyzed using this application. Five type of hash values were calculated: MD4, MD5, RIPEMD-128, RIPEMD-160 and SHA-1. The change in the hash values of original and the simulated document were observed and the values were noted down in the Table 2.

Conclusion

The present research reveals some simple methods which can help document experts as well as common man to establish whether the document is authentic or not. The results are extremely beneficial and reliable as after an appropriate magnification, *color variation* can be seen between transplanted area and the original document. Along with it on increasing the range of magnification the differences in *font size* and *pixelate resolution* are very much evident. The *range of magnification* was found to be between 26-1252%, which clearly determines the difference between the transplanted area and original document. The five types of *hash values* were also calculated to determine the difference between simulated document and the original one, the hash values were: MD4, MD5, RIPEMD-128, RIPEMD-160 and SHA-1. By adopting the material and methods used in this research, experts can be benefitted in dealing with softcopy transplanted forgery cases up to a huge extent.

References

- Hajeel S, Sulong G (2013) State of the art of copy-move forgery detection techniques: a review. International journal of computer science 10: 174-183.
- Rosebrock A (2014) Fingerprinting images for near-duplicate detection.
- Zauner C (2010) Implementation and benchmarking of perceptual image hash functions.
- Wen C, Yang K (2006) Image authentication for digital image evidence. Forensic science journal 1-11.
- <http://whatis.techtarget.com/definition/pixel>.
- Li W, Yuan Y, Yu N (2008) Detecting copy-paste forgery of jpeg image via

-
- block artifact grid extraction. International workshop on local and non-local approximation in image processing.
7. Thirumaga S, Allwin D (2012) Image manipulation detection using intrinsic statistical fingerprints. International journal of advanced research in computer science and software engineering 2: 207-212.
 8. Muhammad N, Hussain M, Muhammad G, Bebis G (2011) Copy-move forgery detection using dyadic wavelet transforms. Computer graphics, imaging and visualization (cgiv), 2011 eighth international conference 103-108.
 9. Koon RV, Jakubowski M, Moulin P (2000) robust image hashing. Image processing. Proceedings. international conference on, 3: 664-666.
 10. Ramirez-gutierrez K, Nakano-miyatake M, Meana H (2013) Image authentication using perceptual hashing. Academic journals 8: 447-455.