Blind Detection of Median Filtering: A Difference domain approach

Joydeep Das
Electronics and Communication Engineering
Indian Institute of Technology Roorkee
Roorkee, Uttarakhand 247667
Email: sayan.uec2014@iitr.ac.in

Dr. Nitin Khanna
Electrical Engineering
Indian Institute of Technology Gandhinagar
Gandhinagar, Gujrat 247667
Email: nitinkhanna@iitgn.ac.in

***Abstract*—In image forensics ,the detection of manipulations which involve content preserving image forgery has received a great deal of interest in the recent past. One such non-linear manipulation of image to hide noises involves median filtering(MF).Anti-forensics method exploits the non-linearity characteristics of median filtered image over other linear algorithms. Henceforth, it becomes necessary to detect median filtering operation in images. This work exploits the difference domain MF image for feature extraction purpose which proves to be better over existing techniques in the detection of MF images. The work also compares the accuracy and fallacy percentage of existing techniques over our proposed method. Hence it is believed that this method will add as a new tool in the detection of MF images.**

***Index Terms*—Median Filtering, Difference domain Images,Error Probability.**

I. Introduction

During the last few decades, with the rapid evolution of software editing tools which manipulate and temper images without any visual indication of forgery, an imperative step has been taken in the field of image anti-forensics to detect such manipulations and provide an authentication for the image. The field of digital forensics is widely studied as such images are widely presented as a proof in judiciary courts. Hence, authenticity about the image is its first measure.

Digital Image forgery can be classified into two major groups: Manipulations which changes contents of the image and tampering which preserves the content of image. Content-changing manipulations copy-move forgery [5] and image splicing[6] which involve modifying, moving ,deleting or adding of new objects in image where as content preserving manipulation focuses on blurring[7], resampling[8], Median filtering[2] and compressing[9] which is done typically to remove visual ambiguity in the image. For the later case ,blind forensics evolved as an optimum solution for anti-forensics research as no prior knowledge of source image or source content is available for authentication. Blind forensics works on the assumption that although effective software editing tools leave no visual artifacts for the tempering detection yet back-end statistics of the manipulated image do get altered to an extent .By performing extensive research to investigate such underlying statistics of image ,blind forensics has helped to curve out the forgery detection to a wide extent.

The following work deals with the blind detection of median filtering operation ,which is used to mitigate noises and smoothing of image. The image manipulating software mostly uses non-linear operator such as median-filtering as compared to other linear operations as in sampling .It is because non-linear operation changes the in-layer statistical property of image which becomes difficult to counter assure about the operation. As a result median filtering detection becomes a prime anti-forensic tool for the knowledge of pre-processing history of image.The present work provides a novel method for detection of median filtered images by revealing the characteristic feature of adjacent pixels in original vs. Median filtered image. From the difference domain properties of an image a set of 48 features has been extracted and experiment has been carried out on the UCID database [10] of 1338 images.

The rest of the paper can be summarized as follows: Section II deals with the statistical property of image before and after median filtering, followed by a literature survey of the existing work on median filtering detection.The proposed feature extraction method for distinguishing median filtered image against original and other filtering operations viz. Gaussian and averaging is described in section IV.Section V and Section VI presents the evaluation of results and comparison against other existing features which shows better accuracy for the proposed method. Finally, Section VII draws the conclusion of the existing work.

II. Characteristics of Median-Filtered Image

In this portion an analytic produce is followed to find out the characteristic feature of median filtered images to distinguish it from other non-filtered image. Analytics for median-filtering in 1D signal is shown below ,which is extended for the 2D images.

Let x be the original signal,x={x(n)}, n=0, 1, 2, ...,where the integer x(n)∈[0, 255] .Suppose the median-filtered signal is y ,which can be written as,

$$y(n) = \underset{i\in[n-r,n+r]}{median}\{x(i)\}. \tag{1}$$

Here,the filter window size is, w = 2r +1, r = 1, 2, 3,... Correspondingly, we have

$$y(n+1) = \underset{i \in [n-r+1, n+r+1]}{median} \{x(i)\} \tag{2}$$

Thus the median value of two neighboring pixels y(n) and y(n+1) are dependent on 2r common elements ,which are C(r)=$\{x(i) | i \in \{n-r+1, n+r\}\}$, brings the neighboring pixels to a common or largely similar value.

Extending the same in 2D signal(image),we present figure(1) which shows a 3X3 median filtered image,where numbers around *R*eference pixel indicates the total number of pixels they share in common during median filtering operation. The colored pixel shows row-column and major-minor diagonals are to be considered for difference-domain approach as they share the highest number of pixels in between thus increasing their probabilities of becoming equal.

1	2	3	2	1
2	4	6	4	2
3	6	Reference Pixel	6	3
2	4	6	4	2
1	2	3	2	1

Fig. 1. A boat.

III. Literature Survey

Because of the highly non-linear and non-trivial nature of median filtering operation ,the relation of the output to input after median filtering shows no common analytical property. That is why research work in this field is only around some of the special characteristics property of MF image.

In [1],Gang Cao et al. mentioned about the probability of zero pixels in first difference image ,considering only the textured portion of that image, as a distinguishing feature for separation of MF against original and other filtering operation. The relevance of textured region was to detect the increased number of zero pixels brought in after median filtered operation. The feature is calculated in vertical and horizontal direction and then fused together to form a 1-D feature vector which depending on varying thresholds distinguishes MF from others. Though the feature serves as a good accuracy measure for un-compressed MF images ,it suffers from huge losses in JPEG compressed image ,as its feature vector is one-dimensional only. In [2],Kirchner and Fridrich proposed the first streaking artifacts of MF images, by taking into consideration the ratio of histogram bins for first difference image. The feature vector h0/h1 ,where h0 denotes the count zero pixels and h1 refers to count one pixels in the 1st difference image. The artifacts revealed that h0/h1 for MF image is significantly large in comparison to original or other filtered images. This artifacts also suffers losses against JPEG compressed image and thus [2] mentions SPAM features as a distinguishing measure against MF versus other manipulations.Other work includes median filtering forensics (MFF) by Yuan et al. ,in [4] which proves similar or somewhat better than SPAM feature detector for JPEG compressed image.

The most recent work of Chen et al. in [3] ,considered the characteristics features for multiple(1st and 2nd) difference domain image to measure Global and Local Characteristics feature(GCF and LCF) ,then combined both the features to extract GLF(GLobal-Local Feature)as a 56-D (44 -Global and 12-Local)feature vector(for B=3) to state its efficacy in both median versus other images which proved to be better than MFF and SPAM features.Our work is an extension of the global-local feature in [3],modified Global and Local feature proved to be better all other existing methods both in the field of compressed and un-compressed image as provided in the section V.

IV. Our Proposed method: Analytical Construction of Feature Sets

For the construction of our 92-dimensional feature vector we divide the feature set into two-halves. First half (44 -D feature vector) takes into account the overall pixel closeness for the entire difference(1st and 2nd order difference) image for textured pixels only ,thus we name it Textured Characteristics while the second feature set takes the expectation of the probability map for pixel values of first and second order difference image.

A. Textured Characteristics

For overcoming the issue of false alarm rates for original images in non-textured region we choose only the textured characteristics of an image. To choose the textured pixels we calculate the variance of its neighbouring pixels as follows ,

$$\sigma(i,j) = \underset{\substack{m \in [i-\lceil d/2 \rceil, i+\lceil d/2 \rceil] \\ n \in [j-\lceil d/2 \rceil, j+\lceil d/2 \rceil]}}{Var} \{I(m,n)\}. \tag{3}$$

The binary textured image is then produced based on the average pixel-wise variance of the overall image ,

$$ImageBinary = 1, if \sigma(i,j) \geq avg\{\sigma(I)\} \tag{4}$$

= 0, otherwise

Where d refers to surrounding block size for variance calculation and *a*vg to the average pixel variance of the overall image.

To detect the Characteristic features in 2D -difference domain image we take into account row, column ,major and minor diagonal difference images as described in Section II as follows,

$$\delta_k^{(p,q)}(n,m) = \delta_{k-1}^{(p,q)}(n,m) - \delta_{k-1}^{(p,q)}(n+p, m+q) \tag{5}$$

where k refers to the k-th difference image with k=0 being the original image and $|p| + |q|$ = i(i=1,2) for row-column and

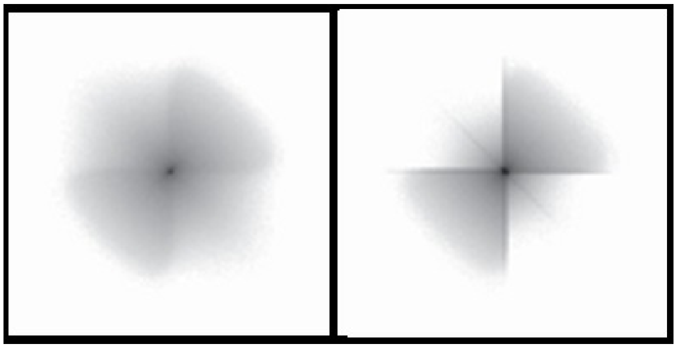

Fig. 2. $P_{1,1}(t_x, t_y)$ for original and MF image

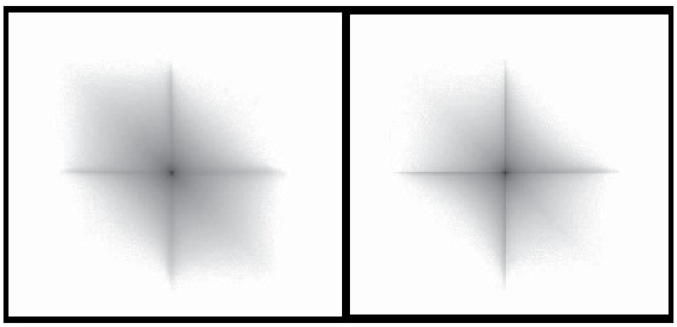

Fig. 3. $P_{1,2}(t_x, t_y)$ for original and MF image

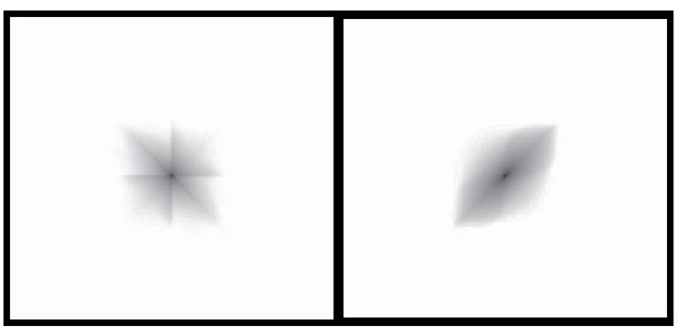

Fig. 4. $P_{2,1}(t_x, t_y)$ for original and MF image

major-minor diagonal to be considered only.We then take the CDF's of the pixel value for difference domain images as a characteristic feature for median filtered images as it is well known that median filtering operation performs smoothing operation thereby increasing the CDF's for smaller pixel values ,t as shown

$$F_k^{(p,q)} = Prob(|\delta_k^{(p,q)}(n,m) \leq t|) \tag{6}$$

We then finally construct our 44-Dimensional *textured feature set (for ImageBinary(i,j)=1 (eqn(4))* by considering $|p| + |q|$ = i(i=1,2) independently with $t \in (0, 10)$ and k=1,2(corresponding to first and second difference image) as shown,

$$P_k^{(i)} = 0.25 \times \sum_{|p|+|q|=i} P_k^{(p,q)} \tag{7}$$

Where $P_k^{(p,q)}$ refers to the 11-dimensional vector for $t \in (0, 10)$.

B. Expectation of the probability map

Our second feature is based on the joint probability of pixels around its locality ,hence we define joint probability of two neighbouring pixels for $\delta_k^{(p,q)}$ as

$$P_{k,l}(t_x, t_y) = Prob(\delta_k^{(p,q)}(n,m) = t_x, \delta_k^{(p,q)}(n, m + l) = t_y) \tag{8}$$

We then plot the joint probability for difference image(for uint8 minimum is -255) to show the plots for median and original images in fig(2-4). The cluster of pixels in first and third quadrant for MF image $P_{1,1}(t_x, t_y)$ is evident from the fact that neighboring pixels in filtered images are highly co-related to each other as explained in section II(B).Thus we then define 16-Dimensional feature sets for such image in *eqn(8)* with $\{k, l\} \in \{1, 2\}$, which takes the expected values of the pixels covering the entire image. The 16-D feature are taken for $P_{1,1}(t_x, t_y)$,$P_{1,2}(t_x, t_y)$ and $P_{2,1}(t_x, t_y)$ making it a 48 dimensional robust feature vector for *probability map expectation.*

V. Experimental Methodology

A. Database Creation

To test the efficacy of the proposed work we conducted our entire experiment on well-known UCID database[10],which comprises of 1338 un-compressed (TIFF)images covering a variety of image aspects like textured ,non-textured, natural and artificial. We convert all the coloured images to gray-scale using rgb2gray and then for creating database with median filtered images of window 3X3,5X5,7X7 ,medfilt2 is used in matlab . Likewise for creating other databases with averaging filter, gaussian filter and bilinear scaling inbuilt matlab functions are used.

B. Training-Testing Pairs and Performance metric

To evaluate the result of our proposed methodology we use LIB-SVM with gaussian and linear kernels. We randomly choose 80 of the entire dataset for training purpose and rest 20 for training in SVM. The SVM output returns the confusion matrix from where we get the percentage of False Positive(FP) and True Negative(TN).We thus define the performance metric as follows, Percentage of error, Pe=average(TN+FP)/2; The average of the Percentage of error for maximum number of iterations is provided in tables.

C. Result

In table 1,We provide the comparison of our Proposed feature with Gang Cao's [1],Streaking Artifacts[2] and GLF[3] for uncompressed TIFF images with measures against Median Filtered (3X3),(5X5),(7X7) as Positive samples (P)with original non-manipulated images as Negative samples (N).Further to distinguish between median filtering (T)with other manipulations like gaussian, averaging and bilinear scaling as Negative Samples(N).The results in table 1 clearly shows the accuracy of our method over first two features [1] and [2].Though GLF features tends to match with our experimental findings.

Table II is an extension of is Just the replica of Table I for post compressed JPEG image where the later means manipulations on TIFF images and then JPEG post compressed with various Quality Factors(QF),the QF being 70 in our case. This table automatically shows the legacy of our method over the GLF feature for median filtered images with original and other filtering operations.

For checking the robustness of our method against the post compressed JPEG images we take all median filtered images viz. (3X3),(5X5),(7X7) as Positives(P) while original and other filtering operations like averaging, Gaussian and scaling as negatives(N).The Percentage error for each method is mentioned in Table III comparing the proposed method with GLF feature.

VI. Tabulation

This section provides below the three tables as mentioned correspondingly in Section V[C].

A. *Table1:*

Shows comparison for Uncompressed TIFF images against three popular feature extractor.

Prob. Error	M3 vs ori	M5 vs ori	M7 vs ori			MED3					MED5		
Features				Avg3	Avg5	gauss0.5	gauss1.5	bsc0.6	Avg3	Avg5	gauss0.5	gauss1.5	bsc0.6
Paper_1	0.56	0.75	1.68	8.58	9.51	2.61	3.36	4.11	8.95	10.08	2.79	2.81	
Paper_2	0.56	0.94	0.74	10.26	11.75	1.31	11.75	4.29	10.26	12.31	1.31	11.38	3.92
GLF	0	0	0	0	0	0	0.3731	0	0	0	0.56	0.37	0.37
Our Local Ft.	0	0.1867	0	0.37	0.37	0.18	0.37	0.18	0.37	0	0.56	0.56	0.37

Fig. 5. Percentage error for uncompressed TIFF images

B. *Table2:*

This table highlights the better error percentage of our method against the remaining three features for JPEG images with *Quality factor(QF):70*.

Features	M3 vs ori	M5 vs ori	M7 vs ori			MED3				MED5	
				Avg3	Avg5	gauss0.5	gauss1.5	Avg3	Avg5	gauss0.5	gauss1.5
GPF	7.09	4.29	4.66	3.79	1.87	8.77	1.67	3.73	1.49	5.6	0.93
LCF	6.34	6.34	4.29	1.49	0.93	6.53	0.93	1.49	0.93	5.97	0.37
GLF	4.86	2.62	1.67	1.7	0.55	6.16	0.56	1.3	0.75	5.97	0.37
Our Local ft.	4.29	4.47	2.98	1.86	0.93	5.24	1.49	1.68	0.56	2.51	0.37
Proposed ft.	3.54	2.34	1.31	1.31	0.37	3.17	0.37	1.51	0.65	3.54	0.18

Fig. 6. Percentage error(Pe) for compressed JPEG images

C. *Table3:*

The following table shows the legacy of proposed method against GLF[3] for the robust feature extractor as mentioned in section V(C).

Feature	Negative	Positive
	99.11	0
GLF	3.11	96.89
	100	0
Our Feature	0.63	99.37

Fig. 7. Percentage error(Pe) for post-compressed JPEG images as a measure of robustness

VII. Conclusion

With the over sophistication of Image-editing tools detection of non-linear median filtering has been a wide interest for researchers and correspondingly many a techniques has been put forwarded over the years for efficient detection of the same. The effort made here correctly classify median filtering versus original and other manipulations and justifies the theory and results obtained by comparing the proposed method with other previously proposed methods.

Acknowledgment

I would like to express my sincere gratitude to Dr. Nitin Khanna for excellent mentorship in making this paper complete. I would further extend my heartiest appreciation to fellow interns and researchers in **Multimedia Forensics and Information Security Lab,IIT Gandhinagar**.The role of entire IIT Gandhinagar and SRIP 2016 sorority, has remained a great inspiration for this paper to bringforth.

References

[1] G. Cao, Y. Zhao, R. Ni, L. Yu, and H. Tian, Forensic detection of median filtering in digital images, in Proc. IEEE Int. Conf. Multimedia Expo, Jul. 2010, pp. 8994.

[2] M. Kirchner and J. Fridrich, On detection of median filtering in digital images, Proc. SPIE, Electronic Imaging, Media Forensics and Security II, vol. 7541, pp. 112, Jan. 2010.

[3] Chenglong Chen, Jiangqun Ni and Jiwu Huang, Blind Detection of Median Filtering in Digital Images: A Difference Domain Based Approach,IEEE transactions on image processing, Vol. 22, No. 12, December 2013.

[4] H. Yuan, Blind forensics of median filtering in digital images, IEEE Trans. Inf. Forensics Security, vol. 6, no. 4, pp. 13351345, Dec. 2011.

[5] S. Bayram, H. T. Sencar and N. Memon, An efficient and robust method for detecting copy-move forgery, in International Conf. on Acoustics, Speech and Signal Processing, Taipei, 2009.

[6] Y.-F. Hsu and S.-F. Chang, Image splicing detection using camera response function consistency and automatic segmentation, in International Conference on Multimedia and Expo,Beijing, 2007.

[7] D. Hsiao and S. Pei, Detecting digital tampering by blur estimation, 1st International Workshop on Systematic Approaches to Digital Forensic Engineering, Washington, 2005.

[8] A. C. Popescu and H. Farid, Exposing digital forgeries by detecting traces of resampling, IEEE Transactions on Signal Processing, vol. 53, no. 2, pp.758-767, 2005.

[9] H. Farid, Exposing digital forgeries from JPEG ghosts, IEEE Transactions on Information Forensics and Security, vol. 4, no. 1, pp.154-160, 2009.

[10] G. Schaefer and M. Stich, UCID - An uncompressed colour image database, in Proc. SPIE, Storage and Retrieval Methods and Applications for Multimedia, San Jose, 2004.