

Robust Method for Detection of Copy-Move Forgery in Digital Images

Garima Mishra¹, A.K Singh²

Department of Electronics Engineering, K.N.I.T. Sultanpur (U.P.)^{1,2}

Abstract: There is various forgeries possible on digital images such as image tampering, copy-move forgery and image compositing. Out of these, copy-move forgery is a type of image forgery, in which a part of original digital image is copied and pasted to another part in the same original image to make it, forged one. This paper describes blind image forensics approach for detecting copy-move forgery. In this technique forged image is reduced in dimension using 3-level DWT (Discrete Wavelet Transform) [1]. Then the compressed image is divided into overlapping blocks of fixed size. These blocks are sorted using lexicographic sorting and phase correlation is used as similarity criterion. Now duplicated blocks are identified and copied part is clearly mentioned in the original image. Detected forgery is displayed with the help of duplication map that gives count of pixels forged. This approach drastically reduces the time needed for the detection process and improves the accuracy of detection.

Keywords: Copy-Move forgery; digital signature; digital image forensics; DWT; phase correlation.

I. INTRODUCTION

Manipulation and tampering of digital images without leaving any obvious clues became very easy with the widely available, easy to use and extremely powerful digital image processing tools such as Photoshop, GIMP, coral and Freehand. As a result, there is a rapid increase of the digitally manipulated forgeries in mainstream media and on the Internet. This trend indicates serious vulnerabilities and decreases the credibility of the digital images. Therefore, developing techniques to verify the integrity and the authenticity of the digital images became very important, especially considering the images presented as evidence in a court of law, as news items, as a part of a medical record, or as a financial document. In this sense, image tamper detection is one of the primary goals in image forensics.

There are various kind of image forgeries but copy-move forgery is a type of image forgery in which a portion of the digital image is copied from one place and pasted somewhere in the same image. This process can be done with/without any modifications on the duplicated regions. In this tampered region would exhibit the same characteristics (like atmospheric condition, noise, distortion and various other) as the rest of the image which makes it hard to identify forged region by normal vision.



Fig. 1. Left is manipulated, right is the original image.

Hence, to detect copy-move forgeries, we need techniques that can detect the image regions which occurs more than once in the image. An example for this type of forgery can be seen in Fig. 1, where a group of soldiers are duplicated to cover President George W. Bush. It is impossible to detect forgery by naked eyes so there is need of a robust method of forgery detection. To make the computation quicker, J.Fridrick proposed an effective blocking approach, in which the image blocks are represented by quantized DCT (Discrete Cosine Transform) coefficients, and a lexicographic sort is adopted to detect the Copy-Move blocks [4]. A.C.Popescu developed a similar method by applying the Principal Component Analysis (PCA) to yield a reduced dimension representation [5]. G. Li proposed a method which is called sorted neighborhood method. This method is based on DWT (Discrete Wavelet Transform) and SVD (Singular Value Decomposition) [6]. The DWT and SVD method suffers from the drawback that the computation of SVD takes lot of time and it is computationally complex. S.A Fattah used 2D DWT method for forgery detection [7].

In this paper, a 3-level wavelet based approach is presented where the wavelet transform is used for compression of tampered image and then phase correlation is used as similarity checking criterion for identifying duplicity of overlapping blocks formed.

This paper is divided into four sections. Section 2 describes the proposed algorithm, section 3 shows experimental results and section 4 gives conclusion and future work.

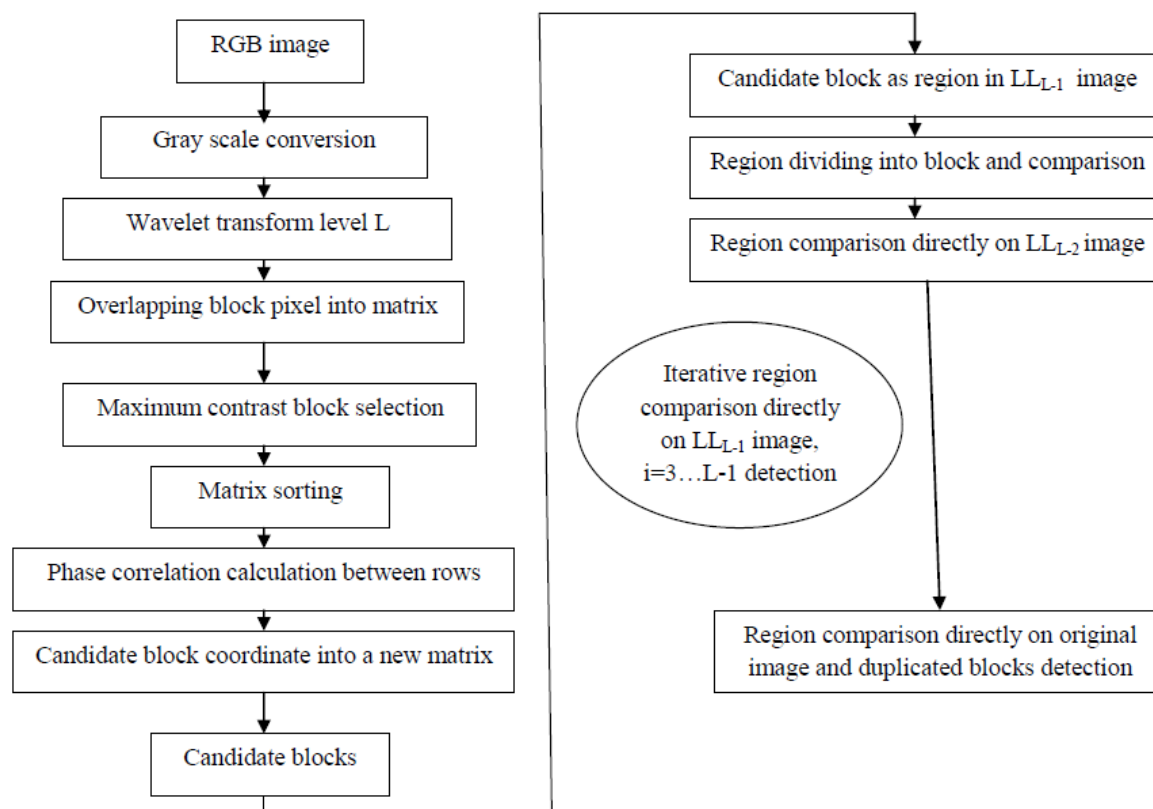


Figure 2. Block diagram of proposed method

II. PROPOSED ALGORITHM

A. Block diagram of proposed method

This phase deals with detection of candidate blocks on the lowest level of wavelet transform compressed image. Fig 2. Shows the block diagram of the proposed method. In the proposed method, candidate blocks are considered as regions in LL_{L-1} level image, divided into blocks and then compared. After this level, they are compared directly until original image is encountered. On the original image, final match is performed.

B. Discrete Wavelet Transform

The transform of a signal is just another form of signal representation. During transform there is no change in the information content present in the signal. The Wavelet Transform provides a time-frequency representation of the signal. DWT is used because it reduces the size of image at each level, e.g., a square image of size $2j \times 2j$ pixels at level L reduces to size $2j/2 \times 2j/2$ pixels at level $L+1$. One advantage of using the DWT with decimation is that the size of the transformed data will be reduced to one half at each level.

However, in this paper, only level one decomposition is used, which results in four decomposed sub images: LL, LH, HL, and HH. The HL band indicates the variation along X-axis while the LH band indicates the Y-axis variation. The power is more compact in the LL band that corresponds to the approximation image and this LL band image is used in the proposed scheme. These sub images can be combined together to restore the original image which was decomposed. An example image along with its wavelet transform applied up to level 3 is shown in Fig. 3. If the number of levels used for decomposition is 'L', then the matching is performed on the LL image at level 'L' denoted by LL_L . Fig.4 shows the image pyramid [10].

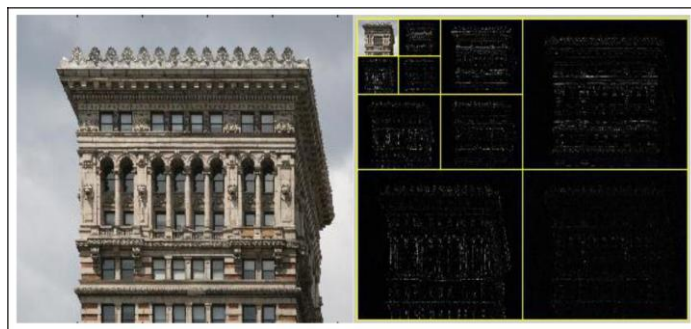


Figure 3. An example image along with its wavelet transform applied up to level 3: the left is an original image, and the right it's Wavelet

At each iteration, the images used for matching of overlapping blocks are $LL_L, LL_{L-1}, \dots, LL_1$. LL_L image is the image at the lowest (coarse) resolution. LL_L image is used for matching of blocks and then these matched blocks are carried to the next higher level. Final match is performed on the original image itself. [13]

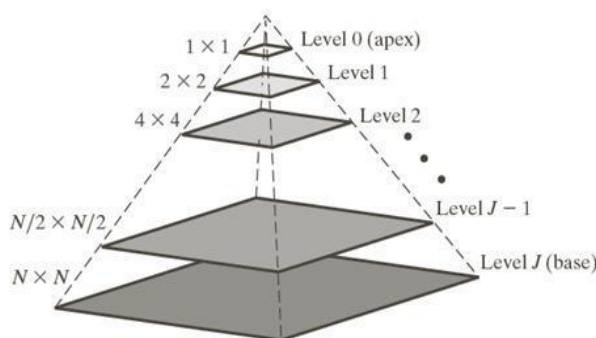


Figure 4. A pyramidical structure of an image

C. Phase Correlation

The correlation between the two images produces an image with top intensities at locations where the two images match the most. Correlation can be achieved by gliding one image over the other and taking the outcome of the products of the corresponding samples at each point. The idea behind correlation is that the resulting image will have the greatest value at the point where the images of hills and bottoms match up. In other words, this match will be complete when the contents of the images match up exactly.

This is a method of image registration. This can be used for template matching [7]. The ratio R between two images 'img1' and 'img2' is calculated as follows:

$$R = \frac{F(img1) \times conj(F(img2))}{\|F(img1) \times conj(F(img2))\|}$$

where 'F' is the fourier transform, and 'conj' is the complex conjugate. The inverse Fourier transform of 'R' is the phase correlation

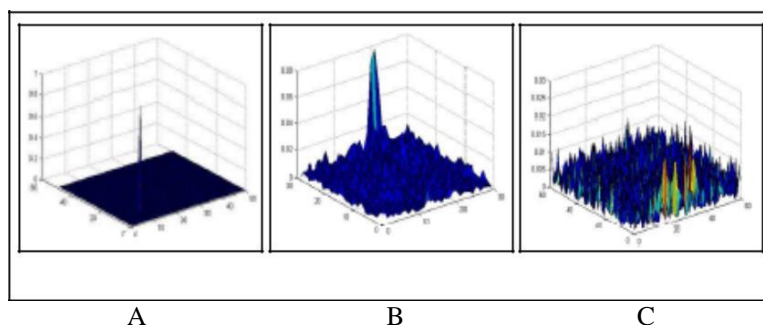


Figure 5 Phase correlation between two images: "A" in ideal case, "B" in real case, and "C" in no matching case.



D. Execution Details

Detection of Reference and Match Blocks

1. Read the image selected by the user as input.
2. If the input image is not a gray scale image then convert it into a gray scale image.
3. Apply wavelet transform up to specified level "L" to the gray image.
4. Extract all overlapping $b \times b$ block in the LL_L image
5. Form a matrix "A" of dimension b^2 columns and $(M-b+1) \times (N-b+1)$ rows by extracting the resulting pixel values by rows into a row of "A".
6. Form another matrix "B" same as "A" with two additional columns for storing top-left co-ordinates.
7. Sort matrix "A" lexicographically. For each row of "A":
 - a. Compute the phase correlation for the block corresponding to the current row with the blocks corresponding to "p" rows below the current row.
 - b. If the computed maximum phase correlation value exceeds a preset threshold value "t", then store the top left coordinates of the corresponding reference block and the matching block from B matrix in a new row of a matrix.
8. End

Algorithm for Comparison of Reference and Match Blocks

1. For each LL_{L-i} level in the image pyramid, where $i = 3, \dots, L-1$.
 - 1.1 For each row of the matrix
 - a. Form a reference region by padding "m" pixels on all the sides of the $b \times b$ reference block.
 - b. Form a matching region by padding "m" pixels on all the sides of the $b \times b$ matching block.
 - c. Get all overlapped block in reference region by extracting the resulting pixel values by rows into a row of matrix "refReg", and store the top-left co-ordinates "indRef".
 - d. Get all overlapped block in matching region by extracting the resulting pixel values by rows into a row of matrix "matchReg", and store the top-left co-ordinates "indMatch".
 - e. Merge two matrices "refReg" and "matchReg" in "A", by labeling each block of them with unique value in order to indicate the reference block and matching one in the merge.
 - f. Sort matrix "A" lexicographically.
 - g. Compare the current row with block corresponding to "p" rows below the current row. If the computed maximum phase correlation value exceeds a preset threshold value "t", then store the top left coordinates of the corresponding reference block and the matching block from the "index" matrix in a new row of a matrix.
 - 1.2 End
 2. End
3. Plot the blocks as copied and pasted regions on the given input image.

III. RESULTS AND PERFORMANCE ANALYSIS

The forgery detection performance of the proposed algorithm has been tested on several images taken from publicly available databases[9],[11]. For the purpose of testing, forged images are prepared by copying certain portion of the image and pasting it to another portion of that image and then saving the forged image in a standard format (JPEG).

At the beginning of the testing phase, it is assumed that only the forged image in a standard format (JPEG) is available. For performance evaluation, most widely used parameters, hit rate, miss rate, and false detection rate (FDR) are used, which can be defined as:

$$\text{HIT RATE} = \frac{\text{Image detected as forged being forged}}{\text{Forged image}} \times 100$$

$$\text{MISS RATE} = \frac{\text{Image detected as not forged being forged}}{\text{Forged image}} \times 100$$

$$\text{FDR} = \frac{\text{Image detected as forged being original}}{\text{Forged image}} \times 100$$

It is expected that the performance will improve with smaller block size. In Table I comparison is shown between two methods. In proposed method block size is 8×8 used and feature dimension is 64.

TABLE I Comparison of avg. forgery detection performance

Methods	Avg. Hit Rate	Avg. Miss Rate	FDR
Method[6]Li	73%	29%	10.9%
Proposed	95%	5%	7%

The forgery detection performance of the proposed algorithm has been tested on several images taken from publicly available databases [9], [11]. We have compressed images using DWT as compression method and used phase correlation as similarity checking criterion [8]. We have used block size based on image size and this block size will be doubled as we move to next higher DWT level. The detected results over tampered image for all DWT levels are shown in Fig. 3.

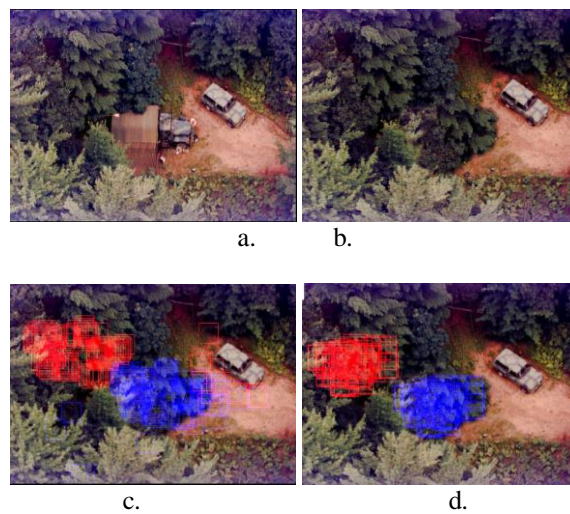


Figure 6. Forgery detection result (1). (a)original image (b) Tampered image (c) detection result on LL_{L-1} level image (d) detection result on LL_{L-2} level image

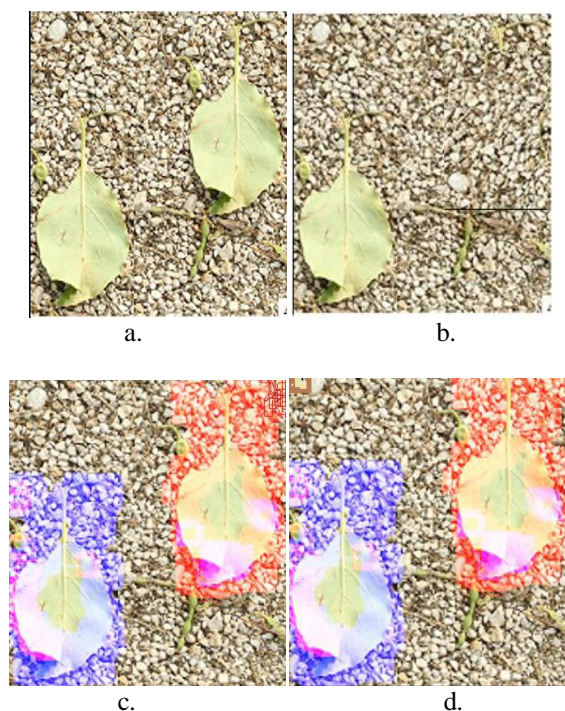


Figure 7. Forgery detection result (2). (a)original image (b) Tampered image (c) detection result on LL_{L-1} level image (d) detection result on LL_{L-2} level image



IV. CONCLUSIONS AND FUTURE WORK

In this paper an algorithm for detecting copy move forgery using 3-level discrete wavelet transforms is proposed. This algorithm has lower computational complexity and reduced time of detection. since exhaustive search for identical blocks is performed only on the image at the lowest resolution.

If image is affected by noise and JPEG quality level change then also this algorithm gives satisfactory performance. Although duplicated regions with rotation through angles and scaled regions cannot be detected. In the future, we would like to search for some mechanism to deal with these problems. In addition to this, the same can be extended to work on videos.

REFERENCES

- [1] Myna.A.N. , M.G.Venkateshmurthy , C.G.Patil "Detection of Region Duplication Forgery In Digital Images Using Wavelets and Log-polar Mapping", in Proc. of International Conference on Computational Intelligence and Multimedia Applications, Volume 3, 13-15 ,pp.371 – 377, July 2-6, 2007.
- [2] H. Farid, "A survey of image forgery detection," IEEE Signal Processing Magazine, vol. 26, no. 2, pp. 16–25, March 2009
- [3] V. Christlein, C. Rieves, J. Jordan, C. Rieves, E. Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," IEEE Trans. Information Forensics and Security, vol. 7, no. 6, pp. 1841-1854, December 2012.
- [4] J. Lukas, J. Fridrich, and M. Goljan, "Detecting digital image forgeries using sensor pattern noise," in proc. Security, Steganography, Watermarking of Multimedia Contents VIII, part of EI SPIE 2006, January, 2006.
- [5] A.C.Popescu and H.Farid, "Exposing digital forgeries by detecting duplicated image regions," Dartmouth College, Hanover, New Hampshire, USA: TR2004-515, 2004.
- [6] G.Li, Q.Wu, D.Tu, and Shaojie Sun, "A sorted neighborhood approach for detecting duplicated regions in image forgeries based on DWT and SVD," IEEE International Conference on Multimedia & Expo, 2007.
- [7] S. A. Fattah, M. M. I. Ullah, M. Ahmed, I. Ahmed and C. Shahnaz, "A scheme for copy-move forgery detection in digital images based on 2D-DWT," IEEE 57th International Midwest Symposium on Circuits and Systems (MWSCAS), College Station, TX, 2014, pp. 801-804.
- [8] Stephane G. Mallat "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation" IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.11, no.7, pp. 674-693, July 1989.
- [9] I. Amerini, L. Ballan, R. Caldelli, A. Bimbo and G. Serra, "A SIFT based forensic method for copy-move attack detection and transformation recovery," IEEE Trans. Information Forensics and Security, vol. 6, no. 3, pp. 1099-1110, September 2011.
- [10] Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins, "Digital Image Processing using MATLAB", Second Edition, Pearson Publications, 2004
- [11] G. Muhammad, M. Hussain, K. Khawaji, and G. Bebis, "Blind copy move image forgery detection using dyadic undecimated wavelet transform," in proc. Int. Conf. Digital Signal Processing (DSP), July 2011.
- [12] Rafael C. Gonzalez, Richard E. Woods, "Digital Image Processing", Second Edition, Pearson Publications, 2002.
- [13] Khan, S. and Kulkarni, A. (2010), Reduced Time Complexity for Detection of Copy-Move Forgery Using Discrete Wavelet Transform, International Journal of Computer Applications IJCA, vol. 6, no. 7, pp. 31–36.