

DETECTING DIGITAL IMAGE FORGERIES USING RE-SAMPLING BY AUTOMATIC REGION OF INTEREST (ROI)

P. Subathra¹, A. Baskar² and D. Senthil Kumar³

^{1,2}Department of Computer Science and Engineering, Amrita School of Engineering, India
E-mail: ¹p_subathra@cb.amrita.edu and ²a_baskar@cb.amrita.edu

³Department of Computer Science and Engineering, Anna University of Technology Tiruchirappalli, India
E-mail: chandsent@yahoo.co.in

Abstract

Nowadays, digital images can be easily altered by using high-performance computers, sophisticated photo-editing, computer graphics software, etc. It will affect the authenticity of images in law, politics, the media, and business. In this paper, we proposed a Resampling technique using automatic selection of Region of Interest (ROI) method for finding the authenticity of digitally altered image. The proposed technique provides better results beneath scaling, rotation, skewing transformations, and any of their arbitrary combinations in image. It surmounts the protracted complexity in manual ROI selection.

Keywords:

Digital Forensics, Interpolation, Resampling Detection, Auto-ROI

1. INTRODUCTION

The preamble of high-resolution digital cameras and modish photo-editing tools become straightforwardly available to most of the people in a reasonable price. This effect in easily creating digital forgeries that is rigid to differentiate from genuine photographs. These forgeries, if used in the mass media or courts of law, can have a significant impact on our society.

Existing digital forgery detection methods are divided into active and passive (blind) approaches. The active approach needs the prior information about the image, technique used is digital watermarking. In contrast to active approaches, passive approaches do not need any prior information about the image. They work in the absence of any digital watermark or signature. Techniques available are block-based, CFA (Color Filter Array) based, Re-Sampling Technique etc.

1.1 WATERMARKING

The idea of digital watermarking is to embed information into an image that can be extracted later to verify authenticity. Watermarking uses specialized camera to generate an image-specific digest and bundle it with the image at the time of recording. The image can be authenticated at a later date by regenerating a digest and checking against the original; a difference indicates that the image was modified since recording. The main drawback of this method which limit this approach is watermarking requires specialized cameras, currently only a few cameras obtainable have this features with pricey. Further, these systems do not allow modifications to an image, including modifications that could improve the image, such as sharpening or enhancing contrast.

1.2 JPEG BLOCK BASED TECHNIQUE

JPEG image format is popularly used in most digital cameras and image processing software. Usually JPEG compression introduces blocking artifacts. Manufacturers of digital cameras and image processing software typically use different JPEG quantization table to balance compression ratio and image quality. Such differences will also cause different blocking artifacts in the images acquired. When creating a digital forgery, the resulted tampered image may inherit different kind of compression artifacts from different sources.

1.2.1 Methodology:

The basis for JPEG compression is the block DCT transform. Because each 8x 8 pixel image block is individually transformed and quantized, artifacts appear at the border of neighboring blocks in the form of horizontal and vertical edges. When an image is manipulated, these blocking artifacts may be disturbed. When an image is cropped and recompressed, a new set of blocking artifacts may be introduced that do not necessarily align with the original boundaries. The tampered area in an image may have come from an original image with a higher or lower JPEG Quality Factor than the other parts of the image. The difference between within and across block boundaries is calculated using the threshold value(t), since artifacts appear at the border of neighboring blocks in the form of horizontal and vertical edges. But the limitation of the JPEG Block technique is it accurately works with the images saved in the JPEG format and not on other image formats such as PNG (Portable Network Graphics) and BMP (Windows Bitmap.) If image tampering occurs in an uncompressed image and then that image is converted to the JPEG image format, the JPEG Block Technique will fail to capture evidence of tampering. Fig.1 shows the result of block based techniques.

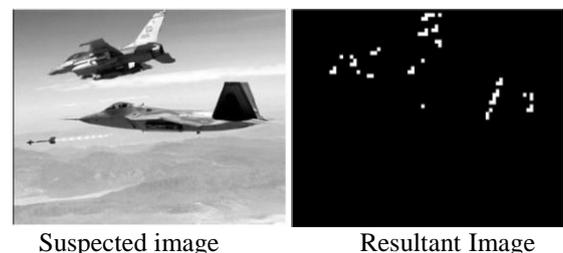


Fig.1(a) Forged Image (b) Output image from block based technique, the white part shows the tampered portion of the input(suspected) image and this result is obtained by setting threshold (t=60)

1.3 DIGITAL IMAGE AUTHENTICATION FROM JPEG HEADER

A camera signature (Header) [11] is extracted from a JPEG image consisting of: information about quantization tables, Huffman coding, thumbnails and EXIF format. This method will not give information which region of image is tampered.

1.4 RESAMPLING TECHNIQUE

In contrast to these approaches, we describe a technique called resampling, generally the digital forgery is created by splicing together of two or more images. Sequentially to create a realistic match, it is often necessary to resize, rotate, or stretch portions of the images. This process requires resampling the original image onto a new sampling lattice using some form of interpolation. Although this resampling is often imperceptible, it introduces specific correlations into the image, which when detected can be used as evidence of digital tampering. The different correlations methods are used in this technique to identify the digital tampering are:

DFT:

The linear and cubic interpolated signals introduce periodicity in variance function of their second-order derivative [3]. This periodicity is simply investigated by computing the discrete Fourier transform (DFT) of an averaged signal obtained from the second derivative of the investigated signal. Furthermore, based on the fact that the mentioned periodicity is directly related to the resampling rate, then we can easily estimate the resampling rate. The major weakness of the method is that it cannot be applied to rotated or skewed images.

SECOND DERIVATIVE:

Another work concerned with the detection of resampling and interpolation this method is similar to above method [2], in which that the second derivative of an interpolated signal produces detectable periodic properties. The periodicity is simply detected in the frequency domain by analyzing a binary signal obtained by zero crossings of the second derivative of the interpolated signal. The disadvantage of this method is it fails to detect in rotated image.

EM ALGORITHM:

In this method the imperceptible specific correlations is brought into the resampled signal by the interpolation step. Their method is based on the fact that in a resampled signal, it is possible to find a set of periodic samples that are correlated in the same way as their neighbors. The core of the method is an expectation/maximization (EM) algorithm [1]. The main output of the method is a probability map containing periodic patterns if the investigated signal has been resampled. The major disadvantage of this method is to select Region of Interest (ROI) manually, and it is time consuming process.

In this work we propose resampling technique using automatic selection of Region of Interest (ROI). This will improve the time and performance comparatively the existing algorithm.

The rest of this paper organized as follows. In section 2 resampling algorithm and our auto ROI were discussed. In section 3 our experimental results were discussed.

2. RE-SAMPLING ALGORITHM USING AUTOMATIC ROI

Re-Sampling can be of three types Up-sample, Down-sample and Interpolate. Bilinear, bicubic and nearest neighborhood are the most widely used interpolation filters, so a signal is re-sampled by a known amount, it is possible to find a set of periodic samples that are correlated to their neighbors nor the specific form of the correlations are known.

Our system consists of three modules namely segmentations, its segments object in the given suspected image followed by automatic ROI selection, it's used to select ROI in the segmented image automatically finally detecting re-sampling module gets input from automatic ROI selection and detects the forgery part. This has been illustrating in Fig.2.

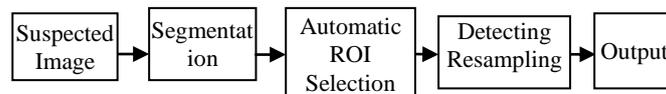


Fig.2. Flow Diagram

2.1 SEGMENTATION

In most of the image tampering essentially objects are spliced from one image and copied or pasted on the other image, later in order to make a realistic match only these portion goes for the geometrical transformation i.e. scaled, rotated. This resampling will introduces specific correlations into the image. With this assumption, we introduce segmentation. The Investigated or suspected image inputs the segmentation module. There are various classical segmentation algorithms like sobel, K-means, Watershed, Region growing is avails and it gives better results. In this work we applied these classical approach to segment the objects followed by dilation is done to dilate the image and fill the interior gaps finally erosion is done to smoothen the objects.

2.2 AUTO REGION OF INTEREST SELECTION

Segmented objects inputs to the auto ROI selection, automatic ROI is done by following steps:

Step 1: The centroid is estimated for the given segmented object.

Step 2: A 100x100 block is selected by keep this centroid as a center.

The above steps is extends to other objects and individual ROI is estimated for the given suspected image.

2.3 DETECTING RE-SAMPLING

In order to determine if a signal has been re-sampled, we employ the expectation-maximization algorithm (EM) to simultaneously estimate a set of periodic samples that are correlated to their neighbors, and the specific form of these correlations. We begin by assuming that each sample belongs to one of two models. The first model, M1, corresponds to those samples $y(i)$ that are correlated to their neighbors, the second model, M2, corresponds to those samples that are not correlated to their neighbors.

The EM algorithm is a two-step iterative algorithm: (1) in the E-step the probability that each sample belongs to each model is

estimated; and (2) in the M-step the specific form of the correlations between samples is estimated. More specifically, in the E-step, the probability of each sample $y(i)$ belonging to model M1 can be obtained by,

$$R(i) = |y(i) - \sum_{K=-N}^N \alpha_n(k)y(i+k)|.$$

Note that the E-step requires an estimate of alpha, which on the first iteration is chosen randomly. In the M-step, a new estimate of alpha is computed using weighted least-squares then normal distribution of each sample is calculated by,

$$P(i) = \frac{1}{\sigma_n \sqrt{2\pi}} e^{-R^2(i)/2\sigma_n^2}.$$

The estimation of the weighting coefficients represents the amount of contribution from each pixel in the interpolation kernel.

In the M-step, a new estimate of alpha is computed using weighted least-squares, that is, minimizing the quadratic error function and solving for alpha yielding,

$$\bar{\sigma}_{n+1} = (Y^T W Y)^{-1} Y^T W \bar{y}$$

where, the matrix Y is Input image and W is a diagonal weighting matrix along the diagonal.

The E-step and the M-step are iteratively executed until a stable estimate of alpha is achieved. The final alpha has the property that it maximizes the likelihood of the observed samples and then probability map P is obtained from the above step. Then p is high-pass filtered to remove undesired low frequency noise. The probability map P is Fourier transformed and only the re-sampled images yield periodic maps in the magnitude of the Fourier transform.

The periodic pattern introduced by re-sampling depends on the re-sampling rate. As a result, it is possible to find traces of re-sampling.

3. ALGORITHM

Input: Suspected Image
Output: Tampered portion
The suspected image $I(x, y)$ is given as the input to segmentation module.

Step 1: Segmentation

- The edges of the image $I(x, y)$ are detected by using sobel operator.
- Then Dilation is performed to dilate the image and fill the interior gaps of the objects.
- The segmented image $S(x, y)$ is obtained by smoothening the objects using erosion.

Step 2: Automatic ROI Selection

- Next take an object in segmented image $S(x, y)$ and estimate the centroid of the object.
- Now keep this centroid as center. A 100x100 blocks is selected so that a Region of Interest (ROI) is obtained.

Step 3: Detecting Re-Sampling

- For the obtained Region of Interest (ROI), the probability map $P(x, y)$ is estimated using EM algorithm.

b. Then high-pass filtered is applied to remove undesired low frequency noise
c. Now the Fourier transform of $P(x, y)$ is taken.
Repeat the step 2 and step 3 until all the object in $S(x, y)$ is Fourier transformed.

4. RESULT

The experiments were carried out in a system having Intel® CoreTM2 Duo processor T6600 with speed 2.2 GHz and 4GB RAM. The implementation was done by using MATLAB 7.6. To searching for periodic traces of interpolation, we absorb the pattern of the output in waveform format. If any portion of the image is tampered, then the pattern of the waveform tends to be different compared to other regions. The proposed method is also compared with JPEG Block technique. The following are different dataset taken and applied for our algorithm.



Fig.3(a) and (b). Tampered image from the net (c). Tampered image using Photoshop (d). Non-tampered image

Automatic Re-Sampling Technique

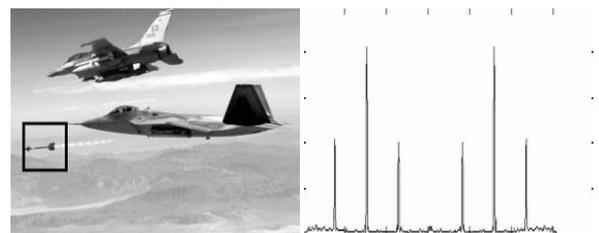


Fig.4(a). Non-Tampered Region

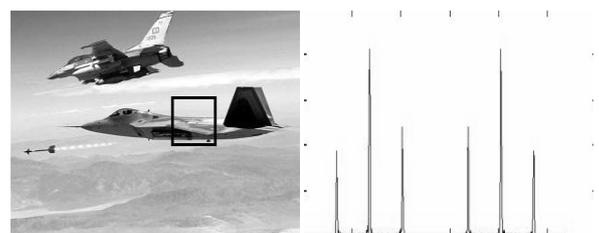


Fig.4(b). Non-Tampered Region

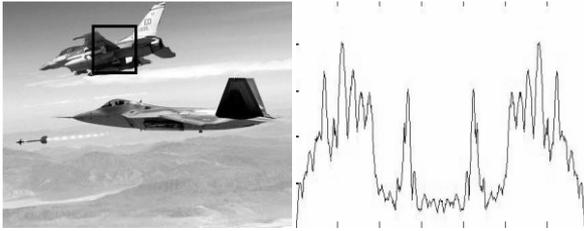


Fig.4(c). Tempered Region

In the above Figures a black box showed which are Automatic ROI Selection and their corresponding output. Fig.4(a) and Fig.4(b) shows the similar output which means that the investigate region is not been tampered and Fig.4(c) shows that the output pattern differs from the other investigate regions, so that region is the tempered portion.

JPEG Block Technique

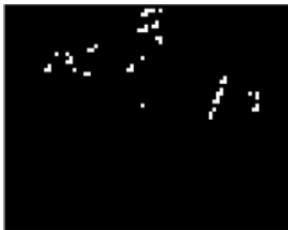


Fig.4(d). Output image from JPEG Block Technique algorithm, the white part shows the tampered portion of the input (suspected) image and this result is obtained by setting threshold ($t=60$)

Automatic Re-Sampling Technique

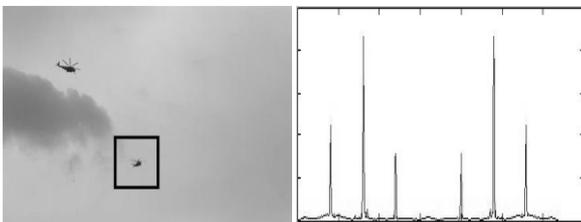


Fig.5(a). Non Tempered Region

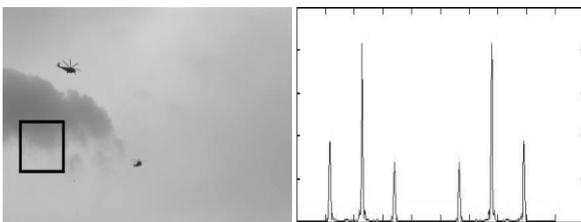


Fig.5(b). Non Tempered Region

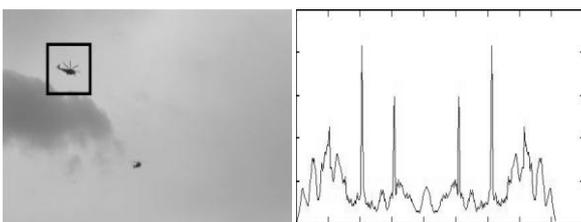


Fig.5(c). Tempered Region

Fig.5(a) and Fig.5(b) shows the similar output means that the investigate region not been tampered and in Fig.5(c) the output pattern differs from the other investigate regions so that region is the tempered portion

JPEG Block Technique



Fig.5(d). Output image from our JPEG Block Technique, the white part shows the tampered portion of the input (suspected) image and this result is obtained by setting threshold ($t=70$)

Automatic Re-Sampling Technique

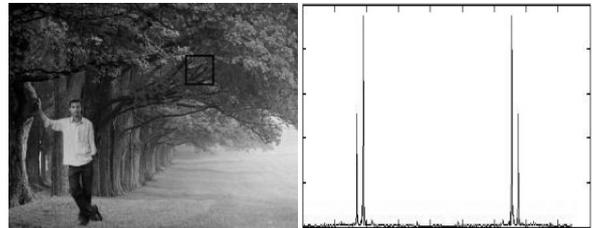


Fig.6(a). Non Tempered Region

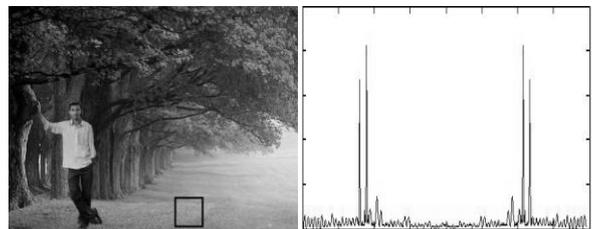


Fig.6(b). Non Tempered Region

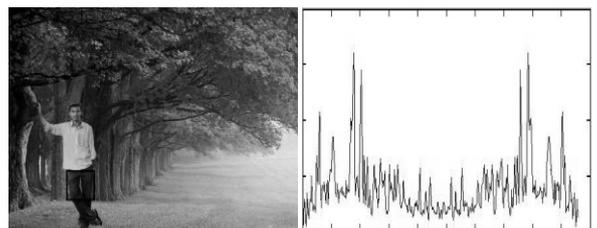


Fig.6(c). Tempered Region

Fig.6(a) and Fig.6(b) shows the similar output means that the investigate region not been tampered and in Fig.6(c) the output pattern differs from the other investigate regions so that region is the tempered portion

JPEG Block Technique



Fig.6(d). Output image from our JPEG Block Technique, the white part shows the tampered portion of the input (suspected) image and this result is obtained by setting threshold ($t=70$)

Automatic Re-Sampling Technique

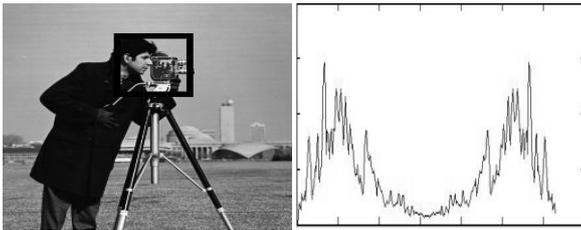


Fig.7(a). Non Tempered Region

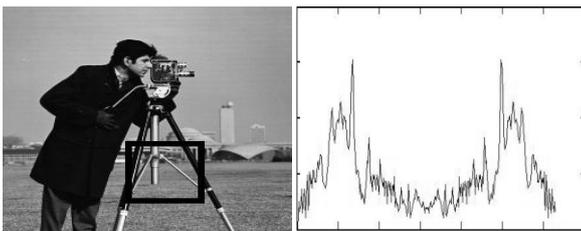


Fig.7(b). Non Tempered Region

In these case regions from the non-tampered image have been investigated (in order to check our algorithm works correctly) Fig.7(a) and Fig.7(b) shows the similar output means that the investigate region not been tampered

The proposed method is applied to various cases of tampered images like rotating, up-sampling, and down-sampling, resized & rotated a portion of images and the results are examined. The system gives good results and the same are shown in the graph Fig.8.

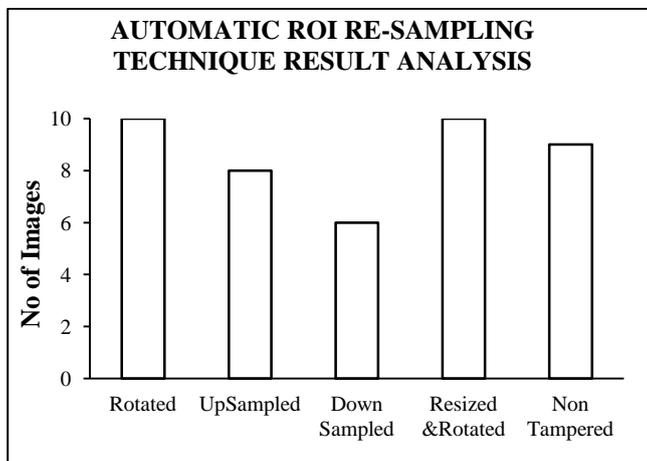


Fig.8. Result chart

5. CONCLUSION

In this paper we presented a Resampling technique using automatic selection of Region of Interest (ROI) method. The proposed method is based on a statistical approach. This technique is easily detecting the traces of tampering region of scaling, rotation, skewing transformations, and any of their arbitrary combinations in image. This method is fast, blind, and efficient. It works for a wide variety of resampling factors or rotation angles.

REFERENCES

- [1] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of re-sampling", *IEEE Transaction on Signal Processing*, Vol. 53, No. 2, pp. 758-767, 2005.
- [2] S. Prasad and K. R. Ramakrishnan, "On resampling detection and its application to image tampering", *Proceedings of IEEE International Conference on Multimedia Exposition*, pp. 1325-1328, 2006.
- [3] C. Gallagher, "Detection of linear and cubic interpolation in JPEG compressed images", *Proceedings of Canadian Conference on Computer Robot Vision*, pp. 65-72, 2005.
- [4] J. Fridrich, "Methods for tamper detection in digital images", *Proceedings of ACM Workshop on Multimedia and Security*, pp. 19-23, 1999.
- [5] B. Mahdian and S. Saic, "Detection of copy-move forgery using a method based on blur moment invariants", *Journal of Forensic Science International*, Vol. 171, No. 23, pp. 180-189, 2007.
- [6] M. K. Johnson and H. Farid, "Exposing digital forgeries by detecting inconsistencies in lighting", *Proceedings of ACM Workshop on Multimedia and Security*, pp. 1-10, 2005.
- [7] C. Rey and J. L. Dugelay, "A survey of watermarking algorithms for image authentication", *EURASIP Journal on Applied Signal Processing*, Vol. 2002, No. 1, pp. 613-621, 2002.
- [8] G. K. Rohde, C. A. Berenstein, and D. M. Healy, "Measuring image similarity in the presence of noise", *Proceedings of SPIE Medical Imaging*, Vol. 5747, pp. 132-143, 2005.
- [9] C. Rey and J. L. Dugelay, "A survey of watermarking algorithms for image authentication," *EURASIP Journal on Applied Signal Processing*, Vol. 2002, No. 6, pp. 613-621, 2002.
- [10] E. H. W. Meijering, W. J. Niessen, and M. A. Viergever, "Piecewise polynomial kernels for image interpolation: A generalization of cubic convolution", *Proceedings of International Conference on Image Processing*, Vol. 3, pp. 647-651, 1999.
- [11] Kee E. Johnson, M. Farid and H. E. Kee, "Digital Image Authentication from JPEG Headers", *IEEE Transactions on Information Forensics and Security*, Vol. 6, No. 3, pp. 1066-1075, 2011.