Analysis Of Different Techniques Of Image Forgery Detection

G.R.Talmale
G.H.Raisoni College OF Engg.
Nagpur, Maharashtra,India

R.W.Jasutkar
G.H.Raisoni College OF Engg.
Nagpur, Maharashtra,India

ABSTRACT

Multimedia Forensics has become important in the last few years. There are two main interests, namely source identification and forgery detection. Source identification focuses on identifying the source digital devices (cameras, mobile phones, camcorders, etc) using the media produced by them, while forgery detection attempts to discover evidence of tampering by assessing the authenticity of the digital media (audio clips, video clips, images, etc) . Digital images have seen increased use in applications where their authenticity is of prime importance Digital images can be forged easily with today's widely available image processing software In this paper we describe a passive approach to detect digital forgeries by techniques of image forgery.

Keywords

Multimedia Forensic,Image Forgery,CopymoveBlind

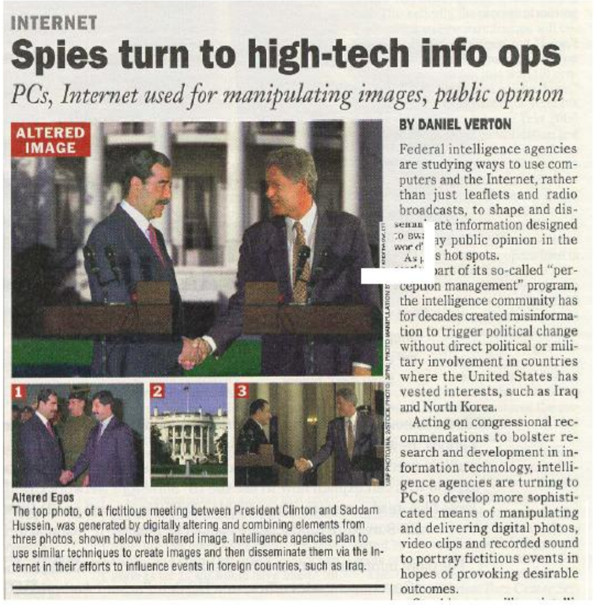
1. INTRODUCTION

Given the fast and widespread penetration of multimedia into all areas of life, the need for mechanisms to ensure reliability of multimedia information has become important. Today, digital media is relied upon as the primary way to present news, sports, entertainment, and information regularly that captures current events as they occur. They are introduced as evidence in court proceedings and commonly used in processing, analysis, and archiving of financial and medical documents. The long-term viability of these benefits requires the ability to provide certain guarantees about the origin, veracity, and nature of the digital media. For instance, the ability to establish a link between a camera and the digital image is invaluable in deciding the authenticity and admissibility of a digital image as legal evidence. Similarly, doctoring images is becoming more frequent as a way to influence people and alter their attitudes in response to various events [1], [2]. Hence, for conventional and online media outlets, the capability to detect doctored images before they are published is important to maintain credibility. Recent research efforts in the field of media forensics have begun to address these issues [3]–[5].In this paper discuss different techniques of image forgery detection.

2. THE NEED FOR DETECTION OF DIGITAL FORGERIES

The availability of powerful digital image processing programs, such as PhotoShop, makes it relatively easy to creat digital forgeries for one or multiple images.

An example of a digital forgery is shown in Figure 1. As the newspaper cutout shows, three different photographs were used in creating the composite image: Image of the White House, Bill Clinton, and Saddam Hussein. The White House was rescaled and blurred to create an illusion of an out-of-focus background. Then, Bill Clinton and Saddam were cut off from two different images and pasted on the White House image. Care was taken to bring in the speaker stands with microphones while preserving the correct shadows and lighting. Figure 1 is, in fact, an example of a very realisticlooking forgery. Another example of digital forgeries was given in the plenary talk by Dr. Tomaso A. Poggio at Electronic Imaging 2003 in Santa Clara. In his talk, Dr. Poggio showed how engineers can learn the lip movements of any person from a short video clip and then digitally manipulate the lips to arbitrarily alter the spoken content. In a nice example, a video segment showing a TV anchor announcing evening news was altered to make the anchor appear singing a popular song instead, while preserving the match between the sound and lip movement. The fact that one can use sophisticated tools to digitally manipulate images and video to create non-existing situations threatens to diminish the credibility and value of video tapes and images presented as evidence in court independently of the fact whether the video is in a digital or analog form. To tamper an analogue video, one can easily digitize the analog video stream, upload it into a computer, perform the forgeries, and then save the result in the NTSC format on an ordinary videotape. As one can expect, the situation will only get worse as the tools needed to perform the forgeries will move from research labs to commercial software. Despite the fact that the need for detection of digital forgeries has been recognized by the research community, very few publications are currently available. Digital watermarks have been proposed as a means for fragile authentication, content authentication, detection of tampering, localization of changes, and recovery of original content [1]. While digital watermarks can provide useful information about the image integrity and its processing history, the watermark must be present in the image before the tampering occurs. This limits their application to controlled environments that include military systems or surveillance cameras. Unless all digitalacquisition devices are equipped with a watermarking chip, it will be unlikely that a forgery-inthe- wild will be detectable using a watermark. It might be possible, but very difficult, to use unintentional camera "fingerprints" related to sensor noise, its color gamut, and/or its dynamic range to discover tampered areas in images. Another possibility for blind forgery detection is to classify textures that occur in natural images using statistical measures and find discrepancies in those statistics between different portions of the image ([2], [3]). At this point, however, it appears that such approaches will produce a large number of missed detections as well as false positives.



INTERNET

Spies turn to high-tech info ops

PCs, Internet used for manipulating images, public opinion

BY DANIEL VERTON

Federal intelligence agencies are studying ways to use computers and the Internet, rather than just leaflets and radio broadcasts, to shape and disseminate information designed to sway public opinion in the world's hot spots.

As part of its so-called "perception management" program, the intelligence community has for decades created misinformation to trigger political change without direct political or military involvement in countries where the United States has vested interests, such as Iraq and North Korea.

Acting on congressional recommendations to bolster research and development in information technology, intelligence agencies are turning to PCs to develop more sophisticated means of manipulating and delivering digital photos, video clips and recorded sound to portray fictitious events in hopes of provoking desirable outcomes.

ALTERED IMAGE

Altered Egos

The top photo, of a fictitious meeting between President Clinton and Saddam Hussein, was generated by digitally altering and combining elements from three photos, shown below the altered image. Intelligence agencies plan to use similar techniques to create images and then disseminate them via the Internet in their efforts to influence events in foreign countries, such as Iraq.

Figure 1 Example Of Digital Forgery

In the next section, we introduce one common type of digital forgeries – the copy-move forgery – and show a few examples. Possible approaches to designing a detector are discussed we describe the detection method based on approximate block matching. This

approach proved to be by far the most reliable and efficient. The method is tested in the last.

3. COPY MOVE FOREGERY

copied areas will likely blend with the background and the human eye cannot easily discern any suspicious artifacts. Because the copied parts come from the same image, its noise component, Because of the extraordinary difficulty of the problem and its largely unexplored character, the authors believe that the research should start with categorizing forgeries by their mechanism, starting with the simple ones, and analyzing each forgery type separately. In doing so, one will build a diverse Forensic Tool Set (FTS). Even though each tool considered separately may not be reliable enough to provide sufficient evidence for a digital forgery, when the complete set of tools is used, a human expert can fuse the collective evidence and hopefully provide a decisive answer. In this paper, the first step towards building the FTS is taken by identifying one very common class of forgeries, the Copy-Move forgery, and developing efficient algorithms for its detection. In a Copy-Move forgery, a part of the image itself is copied and pasted into another part of the same image. This is usually performed with the intention to make an object "disappear" from the image by covering it with a segment copied from another part of the image. Textured areas, such as grass, foliage, gravel, or fabric with irregular patterns, are ideal for this purpose because the color palette, dynamic range, and most other important properties will be compatible with the rest of the image and thus will not be detectable using methods that look for incompatibilities in statistical measures in different parts of the image. To make the forgery even harder to detect, one can use the feathered crop or the retouch tool to further mask any traces of the copied-and-moved segments.Examples of the Copy-Move forgery are given in Figures 2–4. Figure 2 is an obvious forgery that was created solely for testing purposes. In Figure 3, you can see a less obvious forgery in which a truck covered with a position truck was covered with a portion of the foliage left of the truck (compare the forged image with its original). It is still not too difficult to identify the forged area visually because the original and copied parts of the foliage bear a suspicious similarity. Figure 4 shows another Copy-Move forgery that is much harder to identify visually. This image has been sent to the authors by a third party who did not disclose the nature or extent of the forgery. We used this image as a real-life test for evaluating our detection tools. A visual inspection of the image did not reveal the presence of anything suspicious.

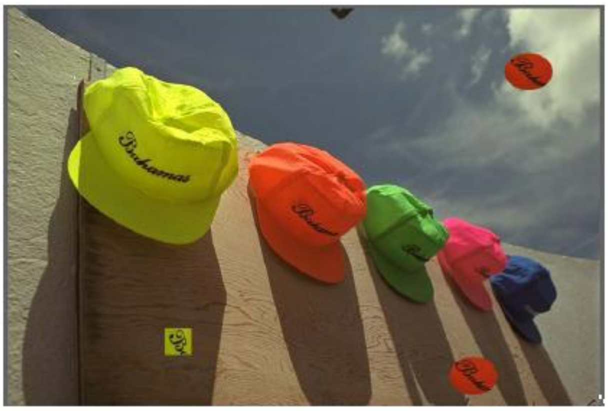


Figure 2 Test image "Hats"

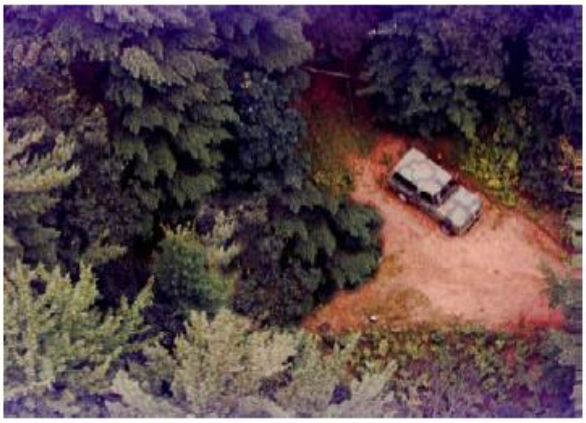
Any Copy-Move forgery introduces a correlation between the original image segment and the pasted one. This correlation can be used as a basis for a successful detection of this type of forgery. Because the forgery will likely be saved in the lossy JPEG format and because of apossibluse Thus we can formulate the following requirements.

- The detection algorithm must allow for an approximate match of small image segments
- It must work in a reasonable time while introducing few false positives (i.e., detecting incorrect matching areas).
- Another natural assumption that should be accepted is that the forged segment will likely be a connected component rather than a collection of very small patches or individual pixel.

3.1 Detection Of Copy Move Forgery By Block Matching

3.1.1 Exact Match:

- The first algorithm described in this section is for identifying those segments in the image that match exactly. Even though the applicability of this tool is limited, it may still be useful for forensic analysis. It also forms the basis of the robust match detailed in the next section. In the beginning, the user specifies the minimal size of the segment that should be considered for match. Let us suppose that this segment is a square with $B\times B$ pixels. The square is slid by one pixel along the image from the upper left corner right and down to the lower right corner. For each position of the $B\times B$ block, the pixel values from the block are extracted by columns into a row of a two-dimensional array A with $B2$ columns and $(M–B+1)(N–B+1)$ rows. Each row corresponds to one position of the sliding block. Two identical rows in the matrix *A* correspond to two identical $B\times B$ blocks.



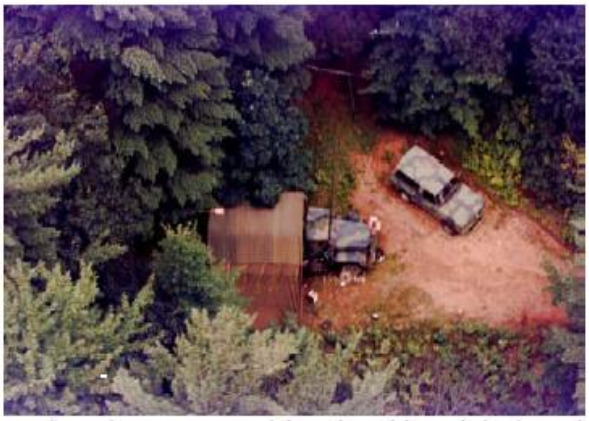


Figure 3 Forged Test Images jeep above and its original version below

Figure 4 Test image Golf with unknown origin

To identify the identical rows, the rows of the matrix *A* are lexicographically ordered (as $B\times B$ integer tuples). This can be done in *MN*log2(*MN*) steps. The matching rows are easily searched by going through all *MN* rows of the ordered matrix *A* and looking for two consecutive rows that are identical.

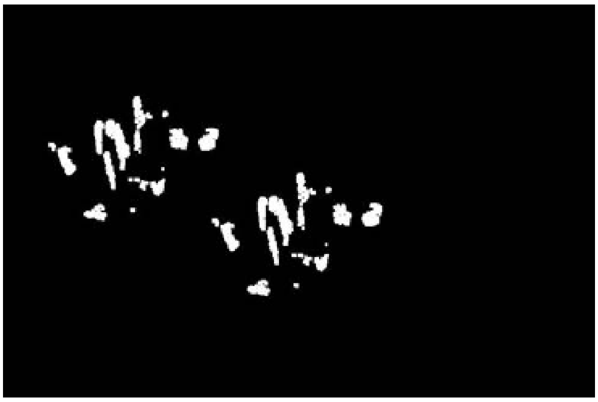


Figure 5 Result Of Block Match Copy Dtetection Algorithm

The matching blocks found in the BMP image of Jeep (Figure 3) for *B*=8 are shown in Figure 5 The blocks form an irregular pattern that closely matches the copied-and-moved foliage. The fact that the blocks from several disconnected pieces instead of one connected segments indicates that the person who did the forgery has probably used a retouch tool on the pasted segment to cover the traces of the forgery. Note that if the forged image had been saved as JPEG, vast majority of identical blocks would have disappeared because the match would become only approximate and not exact (compare the detection results with the robust match in Figure 8). This also why the exact match analysis of images from Figures 2 and 4 did not show any exactly matching blocks. In the next section, the algorithm for the robust match is given and its performance evaluated.

3.1.2 Robust Match

The idea for the robust match detection is similar to the exact match except we do not order and match the pixel representation of the blocks but their robust representation that consists of quantized DCT coefficients. The quantization steps are calculated from a user-specified parameter *Q*. This parameter is equivalent to the quality factor in JPEG compression, i.e., the Qfactor determines the quantization steps for DCT transform coefficients. Because higher values of the Q-factor lead to finer quantization, the blocks must match more closely in order to be identified as similar. Lower values of the Q-factor produce more matching blocks, possibly somefalse matches. The detection begins in the same way as in the exact match case. The image is scanned from the upper left corner to the lower right corner while sliding a $B\times B$ block. For each block, the DCT transform is calculated, the DCT coefficients are quantized and stored as one row in the matrix *A*. The matrix will have (*M*– *B*+1)(*N*–*B*+1) rows and $B\times B$ columns as for the exact match case. The rows of *A* are lexicographically sorted as before. The remainder of the procedure, however, is different. Because quantized values of DCT coefficients for each block are now being compared instead of the pixel representation, the algorithm might find too many matching blocks (false matches). Thus, the algorithm also looks at the mutual positions of each matching block pair and outputs a specific block pair only if there are many other matching pairs in the same mutual position (they have the same shift vector). Towards this goal, if two consecutive rows of the sorted matrix *A* are found, the algorithm stores the positions of the matching blocks in a separate list (for example, the coordinates of the upper left pixel of a block can be taken as its position) and increments a shift-vector counter *C*. Formally, let ($i1$, $i2$) and ($j1$, $j2$) be the positions of the two matching blocks. The shift vector s between the two matching blocks is calculated as $s = (s_1, s_2) = (i_1 – j_1, i_2 – j2)$. Because the shift vectors $–s$ and s correspond to the same shift, the shift vectors s are normalized, if necessary, by multiplying by –1 so that $s1 \geq 0$. For each matching pair of blocks, we increment the normalized shift vector counter *C* by one:$C(s_1, s_2) = C(s_1 , s_2) + 1$ .shift vectors $s_{(1)}$,

$s(2)$, …, $s(K)$, whose occurrence exceeds a user-specified threshold T: $C(s(r)) > T$ for all $r = 1, …, K$. For all normalized shift vectors The shift vectors are calculated and the counter *C* incremented for each pair of consecutive matching rows in the sorted matrix *A*. The shift vector *C* is initialized to zero before the algorithm starts. At the end of the matching process, the counter *C* indicates the frequencies with which different normalized shift vectors occur. Then the algorithm finds all normalized, the matching blocks that contributed to that specific shift vector

are colored with the same color and thus identified as segments that might have been copied and moved. The value of the threshold *T* is related to the size of the smallest segment that can be identified by the algorithm. Larger values may cause the algorithm to miss some not-so-closely matching blocks, while too small a value of *T* may introduce too many false matches. We repeat that the Qfactor controls the sensitivity of the algorithm to the degree of matching between blocks, while the block size *B* and threshold *T* control the minimal size of the segment that can be detected

4. BLIND METHODS FOR DETECTING IMAGE FORGERY

The blinds methods are regarded as a new direction and in contrast to active methods, they work in absence of any protecting techniques and without using any prior information about the image or the camera that took the image. To detect the traces of tampering, blind methods use the image function and the fact that forgeries can bring into the image specific detectable changes (e.g., statistical changes). In recent years various methods for detecting image fakery appeared. In this paper we focus on blind methods using the detection of traces of

- near–duplicated image regions,
- interpolation and resampling,
- inconsistencies in chromatic aberration,
- noise inconsistencies,
- double JPEG compression,
- inconsistencies in color filter array (CFA) interpolated images,
- inconsistencies in lighting.

4.1 Detection of Near–Duplicated Image Regions:

In a common type of digital image forgery, called copy– move forgery, a part of the image is copied and pasted into the another part of the same image, typically with the intention to hide an object or a region (for an example see Figure 2). The copy–move forgery brings into the image several near–duplicated image regions. So, detection of such regions may signify tampering. It is important to note that duplicated regions mostly are not identical. This is caused by lossy compression algorithms, such as JPEG, or by possible additional use of retouch tools. Existing near–duplicated regions detection methods mostly haveseveral steps in common: tiling the image with overlappingblocks,

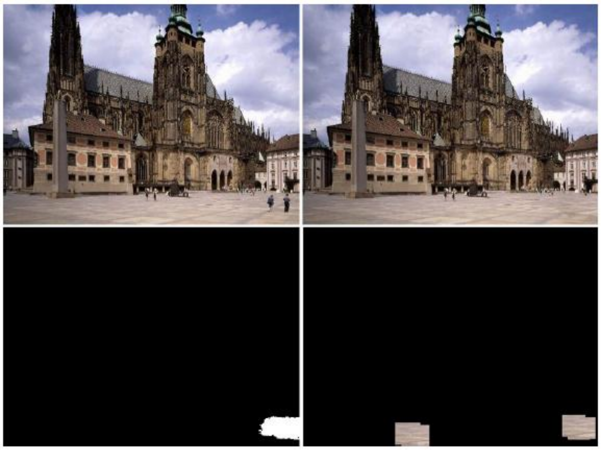


Figure 6. Shown are: original image (top left), an example of a copy–move forgery (top right), the difference between the original image and its fake version (bottom left), and the duplicated regions map created by application of the near–duplicated image regions detection method to the top right image.

Feature representation and matching of these blocks. The first copy–move detection method has been proposed by Fridrich et al. [4]. The detection of duplicated regions is based on matching the quantizied lexicographically sorted discrete cosine transform (DCT) coefficients of overlapping image blocks. The lexicographically sorting of DCT coefficients is carried out mainly to reduce the computational complexity of the matching step. The second

method has been proposed by Popescu and Farid and is similar to [4]. This method differs from [4] mainly in the representation of overlapping image blocks. Here, the principal component transform (PCT) has been employed in place of DCT. The next copy–move detection method has been proposed by B. Mahdian and S. Saic . In this work, overlapping blocks are represented by 24 blur moment invariants up to the seventh order. This allows successful detection of copy–move forgery, even when blur degradation, additional noise, or arbitrary contrast changes are present in the duplicated regions. The blocks matching phase is carried out using a kd–tree representation.

4.2 Detection of Traces of Resampling and Interpolation

When two or more images are spliced together (for an example see Figure 3), to create high quality and consistent image forgeries, almost always geometric transformations such as scaling, rotation or skewing are needed. Geometric transformations typically require a resampling and interpolation step. Therefore, by having sophisticated resampling/interpolation detectors, altered images containing resampled portions can be identified and their successful usage significantly reduced. Existing detectors use the fact that the interpolation process brings into the signal specific detectable statistical changes. In [10], A. C. Popescu and H. Farid have analyzed the imperceptible specific correlations brought into the resampled signal by the interpolation step. Their interpolation detection 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. The main output of the method is a probability map containing periodic patterns if the investigated signal has been resampled. In , B. Mahdian and S. Saic have analyzed specific periodic properties present in the covariance structure of interpolated signals and their derivatives. Furthermore, an application of Taylor series to the interpolated signals showing hidden periodic patterns of interpolation is introduced. The paper also proposes a method capable of easily detecting traces of scaling, rotation, skewing transformations and any of their arbitrary combinations. The method works locally and is based on a derivative operator and radon transformation. In [7], Matthias Kirchner gives an analytical description about how the resampling process influences the appearance of periodic artifacts in interpolated signals. Furthermore, this paper introduces a simplified resampling detectoe based on periodograms. In , A. C. Gallagher in an effort to detect interpolation in digitally zoomed images has found that linear and cubic interpolated signals introduce periodicity in variance function of their second order derivative. This periodicity is simply investigated by computing the DFT of an averaged signal obtained from the second derivative of the investigated signal. Another work concerned with the detection of resampling and interpolation has been proposed by S. Prasad and K. R. Ramakrishnan . Similar to, the authors have noticed 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.

4.3 Detection of Inconsistencies in Chromatic Aberration

Optical imaging systems are not ideal and often bring different types of aberrations into the captured images. Chromatic aberration is caused by the failure of the optical system to perfectly focus light of all wavelengths. This type of aberration can be divided into longitudinal and lateral. Lateral aberration happens by a spatial shift in the locations where light of different wavelengths reach the sensor. This causes various forms of color imperfections in the image. As shown in [6], when an image is altered, the lateral chromatic aberration can become inconsistent across the image. This may signify tampering. It is possible to model the lateral aberration as an expansion/contraction of the color channels with respect to one another. In [6], M. K. Johnson and H. Farid approximate this using a low-parameter model. The model describes the relative positions at which light of varying wavelength strikes the sensor. The model parameters are estimated using an automatic technique based on maximizing the mutual information between color channels

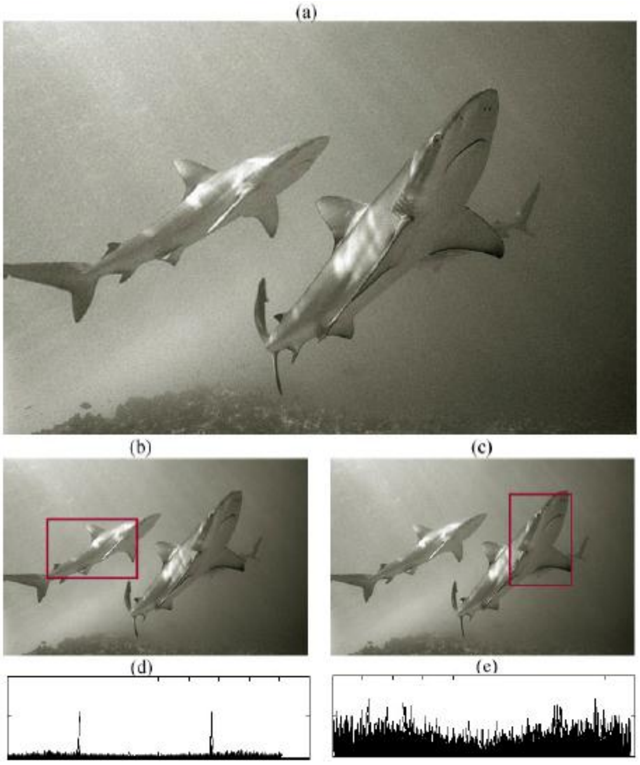


Figure 7 Shown are: an image containing a resampled region (a). In this image, the shark on the left side has been resized by factor 1.4 using the bicubic interpolation. Output of the resampling detector described in [5] is shown in (d). Peaks clearly signify the presence of interpolation. The method has been applied to the denoted region shown in (b). The output of [12] applied a non–resampled region is shown in (c). The testes region is shownin (c).

4.4 Detection of Image Noise Inconsistencies

A commonly used tool to conceal traces of tampering is addition of locally random noise to the altered image regions. Generally, the noise degradation is the main cause of failure of many active and passive image forgery detection methods. Typically, the amount of noise is uniform across the entire authentic images. Adding locally random noise may cause inconsistencies in the images noise (for an example see Figure 4). Therefore, the detection of various noise levels in an image may signify tampering. A. C. Popescu and H. Farid have proposed in [7] a noise inconsistencies detection method based on estimat ing the noise variance of overlapping blocks by which they tile the entire investigated image. The method uses the second and fourth moments of the analyzed block to estimate the noise variance. The proposed method assumes white Gaussian noise and a non-Gaussian uncorrupted image. Another method capable of detecting image noise inconsistencies is proposed in [4] by B. Mahdian and S.Saick this method

Saic. The method is based on tiling the high pass diagonal

wavelet coefficients of the investigated image at the highest resolution with non–overlapping blocks. The noise variance in each block is estimated using a widely used medianbased method. Once the noise variance of each block is estimated, it is used as a homogeneity condition to segment the investigated image into several homogenous subregions.

5. CONCLUSIONS

Our focus in this paper has been addressed to digital image forensics. Digital image forensics is a new and rapidly growing research field. We have introduced various existing copy move forgery and blind methods for image tamper detection. Probably the main drawback of existing methods is highly limited usability and reliability. But it should be noted that the area is growing rapidly and results obtained promise a significant improvement in forgery detection in the neverending competition between image forgery creators and image forgery detectors.

6. ACKNOWLEDGMENT

The authors would like to thank to hod and guideof Computer Science and Engineering department of Raisoni College Of Engineering India for suggesting this line of research.

7. REFERENCES

[1] J. Fridrich, "Methods for "Methods for Tamper Detection in Digital Images", Proc. ACM Workshop on Multimedia and Security, Orlando, FL, October 30−31, 1999, pp. 19−23.

[2] S. Saic, J. Flusser, B. Zitová, and J. Lukáš, "Methods for Detection of Additional Manipulations with Digital Images", Research Report, Project RN19992001003 "Detection of Deliberate Changes in Digital Images", ÚTIA AV ČR, Prague, December 1999 (partially in Czech).

[3] J. Lukáš, "Digital Image Authentication", Workshop of Czech Technical University 2001,

[4] M. Arnold, M. Schmucker, and S. D. Wolthusen. Techniques and Applications of Digital Watermarking and Content Protection. Artech House, Inc., Norwood, MA, USA, 2003.

[5] H. Farid. Creating and detecting doctored and virtual images: Implications to the child pornography prevention act. Department of Computer Science, Dartmouth College, TR2004-518:13, 2004.

[6] J. Fridrich and J. Lukas. Estimation of primary quantization matrix in double compressed jpeg images. In Proceedings of DFRWS, volume 2, Cleveland, OH, USA, August 2003.

[7] J. Fridrich, D. Soukal, and J. Lukas. Detection of copy– move forgery in digital images. In Proceedings of Digital Forensic Research Workshop, pages 55–61, Cleveland, OH, USA, August 2003. IEEE Computer Society Prague, Czech Republic, February 2001.