CHAPTER 2

LITERATURE REVIEW

2.1 Preamble

Images and videos have become the major information carriers in the present era. Doctoring images is becoming more frequent as a way to influence people and alter their feelings and opinion in response to various events. Hence, image authentication has become a necessity. In recent years, digital forensics has emerged as a discipline to authenticate digital images or detect forgery in them and passive forgery detection is an active topic of research. Image forensics is a burgeoning research field as there is a never ending competition between image forgery creators and image forgery detectors [8]. A review of proposals made by the forgery detection researchers is made the subsequent sections.

2.2 Categories of Forgery Detection Techniques

Several techniques have been attempted to detect different types of forgery in digital images with varying success. They have been classified by different researchers based on different characteristics such as type of forgery, type of detection strategy, and levels of processing [5, 8, 9, 13, 14, 15, and 16]. The classification based on detection strategy seems to be reasonable. There are five categories of detection techniques [9]:

A. **Pixel-based Techniques** that detect statistical anomalies introduced at the pixel level and analyze pixel-level correlations arising from tampering.

B. **Format-based Techniques** that leverage the statistical correlations introduced by a specific lossy compression scheme.

C. **Camera-based Techniques** that exploit artifacts introduced by the camera lens, sensor, or on-chip post-processing.

D. **Physically-based Techniques** that explicitly model and detect anomalies in the three-dimensional interaction between physical objects, light, and the camera.

E. **Geometry-based techniques** that make measurements of objects in the world and their positions relative to the camera.

In the subsequent paragraphs salient features of the contributions made by researchers to this subject are recorded.

2.3 Pixel-based Techniques

Pixel is the building block of an image. When an image is tampered, anomalies are introduced into the pixel level correlations which can be analyzed and detected by statistical techniques. Copy-move forgery introduces a correlation between the original image segment and the pasted one which can be used as a basis for a successful detection.

Jessica Fridrich, David Soukal, and Jan Lukas [10] investigate the problem of detecting the Copy-Move forgery. They have reviewed the possibility of detecting forgery by exhaustive search. The image and its circularly shifted version are overlaid looking for closely matching image segments. Let us assume that x_{ij} is the pixel value of a grayscale image of size M×N at the position i, j. In the exhaustive search, the following differences are examined:

$$| x_{ij} – x_{i+k \bmod(M) j+l \bmod(N)} |, k = 0, 1, \ldots, M–1, l = 0, 1, \ldots, N–1 \text{ for all i and j}$$

Computational complexity of this method is more and can be reduced by 4 by inspecting only those shifts [k,l] with $1\le k \le M/2$, $1\le l \le N/2$.

For each shift [k,l], the differences $\Delta x_{ij} = | x_{ij} – x_{i+k \bmod(M)\ j+l \bmod(N)} |$, are calculated and thresholded with a small threshold t. Thresholding is difficult, because in natural images, a large amount of pixel pairs will produce differences below the threshold t.

Although simple and effective, this method is computationally expensive. The comparison and image processing require the order of MN operations for one shift. Thus, the total computational requirements are proportional to $(MN)^2$. Computational requirements for an image that is twice as big are 16 times larger. This makes the exhaustive search a viable option only for small images.

Autocorrelation has been tried based on the idea that the original and copied segments will introduce peaks in the autocorrelation for the shifts that correspond to the copied-moved segments. However, because natural images contain most of their power in low-frequencies, if the autocorrelation r is computed directly for the image itself, r would have very large peaks at the image corners and their neighbourhoods. Hence autocorrelation was computed not from the image directly but its high-pass filtered version. Best performance was obtained using the 3x3 Marr filter.

The authors have worked on two Block Matching methods – Exact match and Robust Match - which are significantly better and faster. In Exact match, a block of pixels of size bxb 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 block, the pixel values from the block are extracted by columns into a row of a two-dimensional array A with b^2 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 bxb blocks. To identify the identical rows, the rows of the matrix A are lexicographically ordered. The matching rows are searched by looking for two consecutive identical rows.

The matching blocks will be disconnected if retouching has been done on the pasted segment to cover the traces of the forgery. Also, if the forged image has been saved as JPEG, majority of identical blocks would disappear as the match becomes only approximate and not exact.

In the Robust match the robust representation of the blocks that consists of quantized DCT coefficients are ordered and matched. The quantization steps are calculated from a user-specified parameter Q. The Q-factor determines the quantization steps for DCT transform coefficients. As 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 some false matches.

The image is scanned from the upper left corner to the lower right corner while sliding a bxb block. For each block, the DCT transform is calculated and the DCT coefficients are quantized and stored as one row in a matrix A. The matrix will have (m–b+1)(n–b+1) rows and b^2 columns for the exact match. The rows of A are lexicographically sorted.

As the quantized values of DCT coefficients are compared, the algorithm might find too many matching blocks and many of them may not be really forged.

Alin C. Popescu and Hany Farid [17] employ Principal Component Analysis (PCA) to small fixed size image blocks to yield a reduced dimension representation. This representation is robust to minor variations in the image due to additive noise or lossy compression. Duplicated regions are then detected by lexicographically sorting all of the image blocks. Detection is possible even in the presence of significant amounts of corrupting noise.

In practice, neither the re-sampling amount nor the specific forms of the correlations are typically known. Alin C. Popescu and Hany Farid [18] employ the Expectation/ Maximization algorithm (EM), to simultaneously estimate a set of periodic samples that are correlated to their neighbours, and the specific form of these correlations. The algorithm assumes that each sample belongs to one of two models. The first model, M1, corresponds to those samples that are correlated to their neighbours, and the second model, M2, corresponds to those samples that are not. 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. Assuming a linear interpolation model, the expectation step reduces to a Bayesian estimator, and the maximization step reduces to weighted least-squares estimation. The estimated probability is then used to determine if a portion of the image has been resampled. But the method is applicable only to uncompressed TIFF images, and JPEG and GIF images with minimal compression.

A. C. Gallagher [19], 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 investigated by computing the Discrete Fourier Transform (DFT) of an averaged signal obtained from the second derivative of the investigated signal.

Weiqi Luo, Jiwu Huang and Guoping Qiu [20] compute seven characteristic features for each overlapping block - c_1, c_2, c_3 as the average of red, green, and blue components respectively. In the Y channel (Y=0.299R+0.587G+0.114B), a block is divided into 2 equal parts in 4 directions and compute c_4, c_5, c_6 and c_7 according to the formula c_i=sum(part(1))/sum(part(1)+part(2)) where i=4,5,6,7. For each block B_i, a block characteristics vector V(i) =(c_1(i),c_2(i),c_3(i),c_4(i),c_5(i),c_6(i),c_7(i)) is computed and saved in an array A. Then the array A is lexicographically sorted and similar block pairs are searched using some thresholds. This technique has lower computational complexity and is claimed to be more robust against stronger attacks and various types of after-copying manipulations, such as lossy compressing, noise contamination, blurring and a combination of these operations.

Babak Mahdian and Stanislav Saic [21] use blur moment invariants representation of the overlapping blocks. Blur moment invariants are suitable to represent image regions due to the fact that they are not affected by the blur degradation present in the region. Another advantage is that they are computed by a summation over the whole image, so they are not significantly affected by additive zero-mean noise. For determining similar blocks instead of lexicographic sorting, k-d tree representation is used. Like other existing methods, this method also has problem with uniform areas in images. Since identical or similar areas in the image are being searched, the method will logically label not duplicated parts also as duplicated in uniform areas, such as the sky. Thus, a human interpretation of the output of any duplication image regions detection method is obviously necessary. Another disadvantage of this method is its computational time.

To create a convincing composite, it is often necessary to resize or rotate portions of an image. For example, when creating a composite of two people, one person may have to be resized to match the relative heights. This process requires resampling the original image onto a new sampling lattice, introducing specific periodic correlations between neighbouring pixels. Because these correlations are unlikely to occur naturally, their presence can be used to detect this specific manipulation.

B. Mahdian and S. Saic [22] 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. They also propose 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.

Matthias Kirchner [23] 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 detector based on cumulative periodograms.

S. Prasad and K. R. Ramakrishnan [24] 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.

When splicing is performed carefully, the border between the spliced regions can be visually imperceptible. In [25] and [26], the authors show that splicing disrupts higher-order Fourier statistics, which can subsequently be used to detect splicing.

Photographs contain specific statistical properties. Some researchers have exploited the statistical regularities in natural images to detect various types of image manipulation [27, 28, 29]

M. K. Bashar et. al. [30] have proposed a wavelet based feature representation scheme for detecting duplicated regions in images. Multi-resolution wavelet decomposition is applied to small fixed-sized image blocks first. Normalized wavelet coefficients are then stacked in an order from lower to higher frequencies. This kind of representation appears robust to block matching. Duplicated regions are then detected by lexicographically sorting all of the image blocks and applying threshold to the desired frequency of the offsets of the block coordinates. A semi-automatic technique that detects accurate number of duplicated regions is also proposed. Impressive results have been obtained compared to linear PCA based representation. The wavelet-based method has achieved higher precision with the comparable recall rates compared to PCA based method. However, the feature strength in the noisy and compressed domains has not been explored. To deal with the noisy environment the well-known algorithms for wavelet-based de-noising may be used. It may be necessary to integrate various characteristics of multiple wavelets or other information for more robust feature representation. Unlike other existing algorithms this algorithm works even when the doctored image is truncated.

Hwei-Jen Lin, Chun-Wei Wang and Yang-Ta Kao [31] propose a method for detecting Copy-Move forgery. In this, the given image is divided into overlapping blocks of equal size. Feature for each block is then extracted and represented as a vector. The vectors are then sorted using the radix sort instead of lexicographic sorting to improve the computational complexity at the expense of a slight reduction in the robustness. The difference (shift vector) of the positions of every pair of adjacent feature vectors in the sorting list is computed. The accumulated number of each of the shift vectors is evaluated. All the feature vectors corresponding to the shift vectors with large accumulated numbers are detected whose corresponding blocks form a tentative result. Median filtering and connected component analysis are performed on the tentative result to obtain the final result. The algorithm can detect duplicated regions rotated through some fixed angles only but not with arbitrary rotations.

Sevinc Bayram, Husrev Taha Sencar, and Nasir Memon [32] propose an approach for detecting copy-move forgery which is more robust to lossy compression, scaling and rotation type of manipulations compared to the previously proposed schemes which use DCT coefficients and Eigen values as features. They have used Fourier-Mellin Transform (FMT) to extract the features and counting bloom filters as an alternative to lexicographic sorting to improve the computational complexity. The method is very accurate, even when the copied region was undergone severe image manipulations. Use of counting bloom filters offered a considerable improvement in time efficiency at the expense of a slight reduction in the robustness.

Sergio Bravo-Solorio and Asoke K. Nandi [33] propose a forensic method to detect duplicated regions, even when the copied portion have experienced reflection, rotation or scaling. To achieve this, overlapping blocks of pixels are re-sampled into log-polar coordinates, and then summed along the angle axis, to obtain a one-dimensional descriptor invariant to reflection and rotation. Moreover, scaling in rectangular coordinates results in a simple translation of the descriptor. This approach allows an efficient search of similar blocks, by means of the correlation coefficient of its Fourier magnitudes.

Saiqa Khan and Arun Kulkarni [34] describe a method to detect Copy-Move forgery in digital images. First DWT (Discrete Wavelet Transform) is applied to the input image to get a reduced dimensional representation. The reduced image is divided into overlapping blocks which are then sorted and duplicated blocks are identified using Phase Correlation as similarity criterion. Since detection is carried out on lowest level image representation the time for detection is reduced and accuracy of detection is increased. But duplicated regions with rotation through angles and scaled regions cannot be detected.

Seung-Jin Ryu, Min-Jeong Lee, and Heung-Kyu Lee [35] propose a detection method for Copy-Move forgery that localizes duplicated regions using Zernike Moments. The magnitude of Zernike moments is invariant against rotation. The method is resilient to distortions such as additive white Gaussian noise, JPEG compression, and blurring. To evaluate the performance Precision, Recall, and F1-measure which are often-used measures in information retrieval techniques are used. However, detection errors have occurred due to the quantization and interpolation error. The method is reported to be weak against scaling or the other tampering based on affine transform. Also appropriate data structure are required to represent nearest neighbours

Vincent Christlein, Christian Riess and Elli Angelopoulou [36] present a rotation-invariant selection method for detecting Copy-Move forgery, which they call Same Affine Transformation Selection (SATS). It provides the benefits of the shift vectors at an only slightly increased computational cost. Also, the method explicitly recovers the parameters of the affine transformation applied to the copied region. SATS outperforms shift vectors when the copied region is rotated, independent of the size of the image.

Weihai Li and Nenghai Yu [37] have proposed an algorithm based on the Fourier-Mellin Transform with features extracted along radius direction. Also, to reduce computational cost, a link processing is introduced instead of hash value counting in the counting bloom filters. Moreover, a vector erosion filter is designed to cluster distance vectors, which is usually achieved through vector counters in existed Copy-Move detection algorithms. The algorithm can detect duplicated regions rotated through any angle whereas earlier algorithms [32] can handle only slight rotation.

Ghulam Muhammad, Muhammad Hussain, Khalid Khawaji, and George Bebis [38] have developed a blind Copy-Move image forgery detection method using Dyadic Wavelet Transform (DyWT) which is shift invariant and therefore better than discrete wavelet transform (DWT). The key idea is that the similarity between the copied and moved blocks from the LL1 sub-band should be high, while the one from the HH1 sub-band should be low. Pairs of blocks are sorted based on high similarity using the LL1 sub-band and high dissimilarity using the HH1 sub-band. Using thresholding, matched pairs are obtained from the sorted list as copied and moved blocks.

B.L. Shivakumar and S.Santhosh Baboo [39] present a technique to detect Copy-Move forgery based on Speeded up Robust Features (SURF) and KD-Tree. SURF algorithm is used to extract features and KD-tree is used for used for searching the nearest neighbours. The proposed method can detect forgery with minimum false match for images with high resolution. However, a few small copied regions have not been successfully detected.

Mengyu Qiao, Andrew H. Sung, Qingzhong Liu, Bernardete M. Ribeiro [40] have used multi-resolution and multi-orientation curvelet transform for the detection of Copy-Move forgery. Curvelet transform is applied to the segmented overlapping blocks of the input image and statistics of sub-bands are extracted which are then sorted. By similarity comparison, duplicated blocks are identified. Forgeries have been detected in JPEG compressed images also despite rotation and scale manipulations.

S. Murali, Basavaraj S. Anami, and Govindraj B. Chittapur [41] have proposed a method which considers an image in the YUV colour space and detects its edges. Then combining the horizontal and vertical edges, the feature map is generated using which forgery regions can be detected.

Anil Dada Warbhe and R. V. Dharaskar [42] present a method based on Independent Component Analysis (ICA) to detect the Copy-Move forgery in digital images. ICA algorithms are good for separating an instantaneous mixture of the non-Gaussian sources. But, the method needs the forged image as well as the original image.

Pravin Kakar and N. Sudha [43] have presented a transform-invariant technique. These are obtained by using the features from the MPEG-7 image signature tools which have been developed for robust content-based image retrieval, in order to detect Copy-Move forgeries. The MPEG-7 image signature tools have been modified to deal with copied regions in a single image. The feature matching process used utilizes the inherent constraints in matched feature pairs to improve the detection of cloned regions. The descriptors used are invariant to a lot of common image processing operations (scaling, rotation, flipping, noise addition, JPEG compression and blurring.

2.4 Format-based Techniques

The most popularly used image format in most digital cameras and image processing software is JPEG. This lossy compression scheme employs a quantization table that controls the amount of compression achieved. Different cameras typically employ different tables. The quantization scheme of an image can be searched in a database of known cameras for confirming or denying the source of an image. Similarly, comparison to a database of photo-editing software can be used in a forensic setting to determine if an image was edited after its original recording [44] and [45].

For performing forgery, an image is opened by a photo-editing software program and saved after tampering it. If the original image is in JPEG format, the forged image will have compression done on it twice. In case of splicing, the pasted portion from another image will likely exhibit traces of only a single compression while the rest of the image will exhibit signs of double compression [46]. Double compression introduces specific artifacts not present in singly compressed images which be used as evidence of some manipulation. Periodicity of the artifacts introduced is exploited to detect double JPEG compression [47]. Junfeng He, Zhouchen Lin, Lifeng Wang, and Xiaoou Tang extended this approach to detect localized traces of double compression [48].

The basis for JPEG compression is the block DCT transform. Because each 8x8 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.

Weiqi Luo, Zhenhua Qu, Jiwu Huang, Guoping Qiu [49] have found that 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. Pixel value differences are computed from 4-pixel neighborhoods that are spatially offset from each other by a fixed amount, where one neighbourhood lies entirely within a JPEG block and the other borders or overlaps a JPEG block. A histogram of these differences is computed from all 8x8 non-overlapping image blocks. An 8x8 Blocking Artifact Matrix (BAM) is computed as the average difference between these histograms. For a compressed image, this matrix has a specific pattern which is disrupted when the image is cropped and recompressed. Supervised pattern classification is employed to discriminate between authentic and inauthentic BAMs.

Shuiming Ye, Qibin Sun and Ee-Chien Chang [50] describe how to detect more localized manipulations from inconsistencies in blocking artifacts. From a region of the image which is presumed to be authentic, the level of quantization is first estimated for each of 64 DCT frequencies. Inconsistencies between the DCT coefficients D and the estimated amount of quantization Q are computed. Variations of this value across the image are used to detect manipulated regions.

Weihai Li, Yuan Yuan and Nenghai Yu [51] have proposed a novel blind approach to detect copy-paste trail in doctored JPEG images. The approach works well even when a JPEG image is truncated or multi-compressed, by extracting the Discrete Cosine Transform (DCT) and Block Artifact Grid. The DCT grid and Block Artifact Grid (BAG) match together in un-doctored images. This approach can detect copy-paste forgery effectively whether the copied area came from the same image or not, if only the copied image is JPEG compressed. It works even when the doctored image is truncated. The authors opine that BAG marking algorithm may be improved to achieve clearer grid map and more efficiency. Also the alignment of BAG can be made automatic.

A. Garg, A. Hailu, and R. Sridharan [52] exploit the fingerprints of JPEG compression which are obtained by estimating the quantization matrix and use them to detect different forgeries such as copy-paste, cropping, rotation, and brightness changes. This approach assumes that only one type of forgery is applied per image. Testing has been done only grayscale images and the methods here are easily extendible to colour images. Algorithms can detect forgeries on a single JPEG image and also when the forged image is recompressed. Performance of this algorithm improves with increasing quality factor unlike block-artifact based detection schemes.

Tiziano Bianchi and Alessandro Piva [53] have proposed an image forensic algorithm under the hypothesis that a tampered image presents either aligned or nonaligned double JPEG compression. Based on a statistical model characterizing the artifacts that appear in double JPEG compression, the algorithm automatically computes a likelihood forgery map. The algorithm does not require manual selection of a suspect region to test the presence or the absence of double compression artifacts. The proposed Bayesian approach can be easily extended to work with traces left by other kinds of processing.

2.5 Camera-based Techniques

Several techniques have been developed that model artifacts introduced by various stages of the imaging process. Camera artifacts can be modelled and estimated and inconsistencies among them can be used as proof of forgery.

Optical systems deviate from ideal models in that they fail to perfectly focus light of all wavelengths. Lateral chromatic aberration manifests itself as a spatial shift in the locations where light of different wavelengths reaches the sensor. Micah K. Johnson and Hany Farid [54] describe how to estimate lateral chromatic aberration in order to detect manipulation.

Most digital cameras employ a single sensor and capture colour images using a Colour Filter Array (CFA). Missing colour samples are obtained by interpolation and a three channel colour image is formed. The estimation of the missing colour samples is referred to as CFA interpolation or demosaicking. CFA interpolation introduces specific statistical correlations between a subset of pixels in each colour channel. Since the colour filters in a CFA are typically arranged in a periodic pattern, these correlations are periodic. It is unlikely that the original recorded pixels will exhibit the same periodic correlations. As such these correlations can be used as a type of digital signature.

Alin C. Popescu and Hany Farid, [55] utilize the correlations introduced by the Colour Filter Array (CFA) in digital cameras, which are likely to be destroyed when an image is tampered. They quantify the specific correlations introduced by CFA interpolation. Lack of these correlations in any portion of an image can be detected.

They have demonstrated this approach in revealing traces of digital tampering in both lossless and lossy compressed colour images interpolated with several different CFA algorithms. They describe how to determine both the form of the correlations and which pixels are and are not CFA-interpolated using the Expectation/ Maximization (EM) algorithm. It is a two-step iterative algorithm: In the Expectation step the probability of each pixel being correlated with its neighbours is estimated and in the Maximization step the specific form of the correlations among pixels is estimated. By modelling the CFA correlations with a simple linear model, the expectation step reduces to a Bayesian estimator and the maximization step reduces to weighted least squares estimation. In an authentic image, it is expected that a periodic pattern of pixels will be highly correlated with their neighbours; deviations from this pattern is a proof of forgery.

Most digital camera sensors are very nearly linear and hence a linear relationship between the amount of light measured by each sensor element and the corresponding pixel value is expected. But in most cameras, a point wise nonlinearity termed response function is applied to enhance the final image. Z. Lin, R. Wang, X. Tang, and H-V Shum [56] describe the estimation of response function. Differences in the response function across the image are then used to detect tampering. Y.-F. Hsu and S.-F. Chang, [57] also have presented a similar approach.

After being captured from the sensor, the digital image undergoes a sequence of processing, such as quantization, white balancing, demosaicking, colour correction, gamma correction, filtering and, usually, JPEG compression. This sequence of operations introduces a distinct signature into the image.

Ashwin Swaminathan, Min Wu and K. J. Ray Liu [58] propose a method based on the observation that many tampering operations can be approximated as a combination of linear and non-linear components. They model the linear part of the tampering process as a filter, and obtain its coefficients using blind deconvolution. These estimated coefficients are then used to identify possible manipulations. They estimate the model parameters such as CFA and colour interpolation coefficients by employing component forensic methodologies. These parameters are used to estimate the camera output and to find the coefficients of the tampering block.

Hongmei Gou, Ashwin Swaminathan and Min Wu, [59] model this processing with a generic additive noise model and use statistics from the estimated noise for image forensics.

Ashwin Swaminathan, Min Wu and K. J. Ray Liu [60], model camera processing with a series of in-camera processing operations and a second filtering. The parameters of this camera processing are then used to determine if an image has undergone any form of subsequent processing.

Jan Lukas, Jessica Fridrich, and Miroslav Goljan [61] have assumed that either the camera that took the image or another image taken by that camera is available. The method is based on detecting the presence of the camera pattern noise, which is a unique stochastic characteristic of imaging sensors, in individual regions in the image. The forged region is determined as the one that lacks the pattern noise. The presence of the noise is established using correlation as in detection of spread spectrum watermarks. The method may not provide sufficiently conclusive statistical evidence for regions with naturally low presence of pattern noise (mainly saturated areas or very dark areas). Also forgeries in images which have undergone geometrical processing, such as cropping or resizing may not be successfully detected.

Jessica Fridrich, M. Chen, and M. Goljan [62] model camera processing with an additive and multiplicative noise model. The parameters of the noise model are estimated from the original camera or a series of images from that camera. Correlations between the estimated camera noise and the extracted image noise are then used to authenticate an image.

A non-forged image will have uniformly distributed noise in it. During forgery random noise is added locally resulting in inconsistency of noise distribution. Babak Mahdian, and Stanislav Saic [63] have proposed a method that divides an image under investigation into various partitions with homogenous noise levels. Assuming that the additive noise is white Gaussian they use segmentation method for detecting changes in noise level. When two or more images from different sources are spliced together, the forged image may then contain several regions with various noise levels. The authors have considered gray-level images. The method can be adopted for RGB images also like, applying to each channel separately.

Xunyu Pan, Xing Zhang and Siwei Lyu [64] have attempted a method to expose image forgeries by detecting the noise variance differences between original and tampered parts of an image. The image is first segmented into non-overlapping image blocks. Noise variance at each local image block is computed using an effective noise estimation method. Using k-means algorithm all image blocks are grouped into two clusters. The cluster with fewer blocks is treated as tampered region, assuming that the tampered region is usually smaller than their authentic counterparts. But image regions with complex textures or edges may also have different noise levels, and can result in false detections. Then a refined detection with smaller blocks is done to improve detection accuracy and to reduce false positives.

2.6 Physically-based Techniques

When splicing is done, it is often difficult to exactly match the lighting effects under which different images were originally photographed. If different properties of the lighting environment under which a photograph was taken can be estimated, differences in these properties can be used as proof of forgery.

The amount of light striking a surface is proportional to the surface normal and the direction to the light. P. Nillius and J. O. Eklundh [65] have shown that the direction to the light source can be estimated with knowledge of three-dimensional (3-D) surface normal. Because 3-D surface normals usually cannot be determined from a single image, Micah K. Johnson and Hany Farid [66] consider only the two-dimensional (2-D) surface normal and estimate two of the three components of the light source direction. Although an ambiguity in the estimated light direction remains, the two components of light direction are still useful. They have shown a technique to estimate the direction of a point light source from only a single image using the tools from the field of computer vision and thus detect any forgery.

The same authors [67] describe how to estimate the 3-D direction to a light source from the light's reflection in the human eye. Normally a simplified lighting model consisting of a single dominant light source is assumed. But in reality there can be more sources of lights, creating a complex effect. The same authors [68] describe a method to estimate a representation of such complex lighting environments. They demonstrate that under some assumptions, arbitrarily complex lighting environments can be approximated with a low-dimensional model and estimate the model's parameters from a single image. Inconsistencies in the lighting model are then used as evidence of tampering. The lighting environment coefficients have been estimated from only the green channel. For some objects, other colour channels could provide more reliable estimates.

Yingda Lv, Xuanjing Shen and Haipeng Chen [69] have attempted to detect forgery in an image by the inconsistency in the light source direction of different areas in the image. Search means and the Hestenes–Powell multiplier method are used to calculate the light source direction for local and infinite light source images, respectively. The method has been shown to perform better than the method by Micah K. Johnson and Hany Farid [66].

The luminance of an image is the measurement of the perceived brightness levels S. Murali, Basavaraj S. Anami, and Govindraj B. Chittapur [70] have recognized the fact that in the copy-paste and splicing forgery the luminance of the tampered image and the image block from a second image that is inserted are not the same. Applying techniques such as Luminance level, Hue-Saturation-Value (HSV) and Alternative Filtering Masks, they have been successful in detecting forgeries.

2.7 Geometric-based Techniques

In authentic images, the principal point - the projection of the camera center on to the image plane - is near the center of the image. When an object is modified in the image, the principal point is altered. Differences in the estimated principal point across the image can therefore be used as evidence of tampering. Micah K. Johnson and Hany Farid [71] have described how to estimate a camera's principal point from the image of a pair of eyes or other planar geometric shapes. They have shown how translation in the image plane is equivalent to a shift of the principal point. Inconsistencies in the principal point across an image can then be used as proof of forgery.

Several tools from projective geometry that allow for the rectification of planar surfaces and, under certain conditions, the ability to make real-world measurements from a planar surface have been reviewed by Micah K. Johnson and Hany Farid [72]. They have reviewed three techniques for estimating the transformation, H, of a plane imaged under perspective projection. With this transformation, a planar surface can be rectified to be front parallel. Each technique requires only a single image, and each exploits different geometric principles. In the first and most direct approach, the shape of a known polygon on the world plane is used to directly estimate H. In the second approach, vanishing lines and known angles or length ratios are used to estimate a projective and affine matrix, whose product yields H. In the third approach, a pair of coplanar circles is used to estimate the projective and affine matrix. In each approach, the estimation of H requires some user assistance in selecting lines, polygons, circles, etc. The authors have shown that the removal of planar distortion can be helpful in certain forensic settings. In addition, under certain conditions real-world metric measurements can be made on the planar surface, providing another useful forensic tool.

2.8 Video Forgery Detection Techniques

Recently, scientific research is focusing on detection of forgery in digital video also. All the possible modifications that can be applied to still images are applicable to the individual frames of a video sequence. These modifications are replicated in the temporal dimension increasing the number of degrees of freedom in the alterations of a video signal. Hence, video forensics proves to be extremely hard than still image forensics as the recovering of the signal processing history could be much more complex. In addition to this, video data is mostly always available in compressed formats and strong compression ratios may destroy the footprints of forgery [73].

Despite the significant literature available on digital image forensics, video forensics still presents many unexplored research issues, because of the peculiarities of video signals with respect to images and the wider range of possible alterations that can be applied on this type of digital content [74].

Motion Picture Expert Group (MPEG) standards are normally used for compression to reduce both spatial redundancy within individual video frames and temporal redundancy across video frames.

Michihiro Kobayashi and Takahiro Okabe [6] have developed an approach to detect suspicious regions in a video of a static scene on the basis of the noise characteristics. An image signal contains irradiance-dependent noise the variance of which is described by a Noise Level Function (NLF) as a function of irradiance. Forged pixels in the regions clipped from another video camera can be differentiated by using maximum aposteriori estimation for the noise model when the NLFs of the regions are inconsistent with the rest of the video. Parameters of NLFs and the authenticity of each pixel are estimated by using the expectation maximization (EM) algorithm.

D. Labartino et. al. [75] have proposed a method to determine which parts of a frame in a MPEG-2 video have been altered. The method works by first locating frames that have been intra-coded twice, and then applying a double quantization analysis to them, based on a proposed model specific for MPEG-2.

Yongjian Hu et. al. [76] have presented an algorithm to detect duplicated frames based on video sub-sequence fingerprints. The fingerprints employed are extracted from the DCT coefficients of the Temporally Informative Representative Images (TIRIs) of the sub-sequences.

Chih-Chung Hsu et. al. [77] propose an approach for locating forged regions in video using correlation of noise residue. Block-level correlation values of noise residual are extracted as a feature for classification. They model the distribution of correlation of temporal noise residue in a forged video as a Gaussian Mixture Model (GMM) and propose a two-step scheme to estimate the model parameters. A Bayesian classifier is used to find the optimal threshold value based on the estimated parameters.

2.9 Summary

Several approaches to detect different types of forgery in digital images have been surveyed to understand the evolution of this research area and the state-of-the-art.

There is an urgent and growing need for digital image tampering detection techniques. Copy-Move and Copy-Paste forgeries are the most frequently chosen forgery techniques. A large number of Pixel-based techniques have been attempted. Format-based and Camera-based are the next. A few attempts have been made in the Physically-based and Geometry-based category. Some methods for detecting forgery in digital video are also available in literature.

Most of the methods aim at detecting inconsistencies in an image to detect forgery. Many algorithms utilize strengths from different transforms to reduce the number comparisons and make the algorithm robust. However, no method has been able to achieve 100% robustness against all types of forgery or different post-processing operations. Also selection of block size is a critical factor. If it is too small, false positives appear or if too large, some forged areas go undetected. In addition to this, the threshold parameter selection is manual and has to be wisely set to avoid false positives. There is abundant opportunity for developing more robust, accurate and efficient algorithms.