Review Of Literature

2. REVIEW OF LITERATURE

One of the key characteristics of digital images with a discrete representation is its pliability to manipulation. Therefore, even back in 1989, the sesquicentennial year for photograph when digital images was gaining popularity, 10% of all color photographs published in United States were actually digitally altered and retouched, according to the Wall Street Journal (Amsberry, 1989). The main function of image forensics is to assess the authenticity and the origin of images. Back in 1993, the idea of trustworthy camera has been proposed as a way to make the trustworthiness of digital images accountable (Friedman, 1993).

Verifying the integrity of digital images and detecting the traces of tampering without using any protecting pre-extracted or pre-embedded information have become an important and hot research field of image processing. The field even though is in its infantry stage has attracted several researchers and the works related to the present topic are presented in section.

2.1. NEAR-DUPLICATED IMAGE REGIONS

Detection of near-duplicated image regions may signify copy−move (copy-paste) forgery. In this type of forgery, a part of the image is copied and pasted into another part of the same image typically with the intention to hide an object or a region. Bayram *et al.* (2009) proposed a clone detector based on Fourier Mellin transform of the image's blocks. The Fourier Mellin transform is invariant with respect to scale and rotation. This allows a better behavior of the method when dealing with slightly resized and rotated cloned regions.

Dybala *et al.* (2007) proposed a cloning detection method based on a filtering operation and nearest neighbor search. Fridrich *et al.* (2003) proposed a method detecting copy−move forgery using discrete cosine transform of overlapping blocks and their lexicographical representation. Huang *et al.* (2008) used the SIFT algorithm to detect the cloned regions in

the image. SIFT features are stable with respect to changes in illumination, rotation and scaling. Langille and Gong (2003) proposed a method searching for blocks with similar intensity patterns based on a kd-tree.

Li *et al.* (2007) proposed a duplicated region detection method based on wavelet transform and singular value decomposition. Li *et al.* (2008) using a copy−move detector for JPEG images based on blocking artifacts. Lin *et al.* (2003) proposed a method using radix sort. Luo *et al.* (2006) proposed a copy−move forgery detection method based on seven intensity-based characteristics features. Mahdian and Saic (2007) proposed a method for detecting near-duplicated regions based on moment invariants, principal component analysis and kd-tree. Myna *et al.* (2007) proposed a method using the idea of log-polar coordinates and wavelet transforms. Popescu and Farid (2008) proposed a method based on representing image blocks using principal components analysis. Bravo-Solorio and Nandi (2009) proposed a near-duplication detection method based on log-polar coordinates. The method is invariant with respect to reflection, rotation or scaling. Zhang *et al.* (2006) proposed a copy−move detection method based on wavelet transform and phase correlation.

Although these methods are capable of detecting near duplicates parts of the image, their computational time is very high and typically they produce a high number of false positives. Furthermore, a human interpretation of the results is necessary.

2.2. INTERPOLATION AND GEOMETRIC TRANSFORMATIONS

When two or more images are spliced together to create high quality and consistent image forgeries, geometric transformations are almost always needed. These transformations, typically, are based on the re-sampling of a portion of an image onto a new sampling lattice. This requires an interpolation step, which typically brings into the signal statistical changes.

Detecting these specific statistical changes may signify tampering. Fillion and Sharma (2010) analyzed the detection of content adaptive resizing. They proposed a technique capable of detecting the presence of seamcarving.

Gallagher (2005) proposed a method for detecting digitally zoomed images. The method is based on the periodicity in the second derivative signal of interpolated images. Kirchner (2008) proposed a re-sampling detection method based on linear filtering and cumulative periodograms.

Kirchner and Gloe (2009) analyzed resampling detection in re-compressed JPEG images. Liu and Sung (2009) proposed a method for detecting resized JPEG images. Their work is based on neighboring joint density features of the DCT coefficients and classification relying on support vector machines. The paper shows blocking artifacts can help to increase the resampling detection performance in JPEG compressed images.

Mahdian and Saic (2009a) proposed a method for detecting the traces of interpolation based on a derivative operator and radon transformation. The same authors, in another work Mahdian and Saic (2009b) analyzed the usefulness of cyclostationarity theory in image forensics and proposed a local cyclostationarity detector to find the traces of scaling and rotation. Methods dealing with detection of interpolation have weak results when dealing with JPEG images.

Therefore, Nataraj *et al.* (2009) proposed a method for detecting JPEG resized images. The method is based on addition of a suitable amount of Gaussian noise to the image so that the periodicity due to JPEG compression is suppressed while that due to the resizing is retained. In order to detect the traces of resampling,

Popescu and Farid (2005) analyzed the imperceptible specific correlations brought into the resampled signal by the interpolation step and proposed a resampling detector based on an expectation/maximization

algorithm. Prasad and Ramakrishnan (2006) analyzed several spatial and frequency domain techniques to detect the traces of resampling. Their most promising method is based on zero-crossings of the second difference signal.

Anindya Sarkar *et al.* (2009) proposed a machine learning based framework for detection of seam carving. The framework is based on the Markov features, consisting of 2-D difference histograms in the block-based DCT domain. Wei *et al.* (2008) proposed a method for estimation of the rescaling factor. The method is based on periodic artifacts brought into the signal by the interpolation process.

The mentioned methods mostly are efficient when the image being analyzed is in a non-compressed format. Artifacts of JPEG compression typically conceal the traces of interpolation.

2.3. IMAGE SPLICING

When dealing with the photomontage detection problem, one of the fundamental tasks is the detection of image splicing. Image splicing assumes cut and paste of image regions from one image onto another image. To detect image splicing, Dong *et al.* (2008) proposed a support vector machine based method. Their features are gained by analyzing the discontinuity of image pixel correlation and coherency caused by splicing. Farid (1999) proposed how to detect un-natural higher-order correlations introduced into the signal by the tampering process. The method is based on bispectral analysis. Hsu and Chang (2006, 2007) proposed a method based on camera response function estimated from geometry invariants. Gopi (2007) and Gopi *et al.* (2006) proposed how to detect forgeries using an artificial neural network, independent component analysis and auto-regressive coefficients.

Jing and Hongbin (2006) proposed a method for detecting image splicing based on a Sobel edge detector, a derivative operation and a Hough transform. Johnson and Farid (2007a) proposed how to detect compositing of

two or more people into a single image based on estimating the camera's principal point from the image of a person's eyes. Lin *et al.* (2005) proposed a method based on computing the inverse camera response functions by analyzing the edges in different patches of the image and verifying their consistency.

Ng and Chang (2004) proposed and studied an image-splicing model based on the idea of bipolar signal perturbation. The same authors (Ng *et al.*, 2004) proposed a method for detecting the abrupt splicing discontinuity using bicoherence features. Ng and Tsui (2009) and Ng (2009) proposed an edge-profile-based method for extracting CRF signature from a single image.

Chen *et al.* (2007) and Shi *et al.* (2007) analyzed image splicing. The proposed methods are based on Hilbert–Huang transform, statistics of 2-D phase congruency and a natural image model to classify spliced images from authentic images. Wang *et al.* (2009) proposed an image splicing detection method based on gray level co-occurrence matrix (GLCM) of thresholded edge image of image chroma.

Zhang *et al.* (2008) proposed a splicing detection scheme based on moment features extracted from the discrete cosine transform and image quality features. Many of the mentioned methods work well when the image being analyzed is compressed by a high quality factor. Otherwise, the compression artifacts make the localization of the forgery very difficult.

2.4. COMPUTER GRAPHICS AND PAINTINGS

In today's digital age, high quality computer graphics look so photorealistic that it is difficult to visually differentiate them from real images. Since this technique can also be used to create convincing image forgeries, there is a need to have sophisticated methods distinguishing between computer graphics and photographic images. To detect tampering in computer graphics images, Sankar *et al.* (2009) used the wavelet

decomposition coefficients of natural images. In particular, the fractional lower order moments in the image wavelet domain are extracted and evaluated with the support vector machines.

Dehnie *et al.* (2006) presented an approach by focusing on the imaging sensor's pattern noise. Ng *et al.* (2005) proposed a geometry-based image model motivated by the physical image generation process for classifying photographic images and photorealistic computer graphics. The same authors also deployed an online system for distinguishing photographic and computer graphic images (Ng *et al.*, 2007).

Dirik *et al.* (2007) investigated the problem of identifying photo-realistic computer generated and real images by introducing features to detect the presence of color filter array demosaicking and chromatic aberration. Khanna *et al.* (2008) proposed a method based on residual pattern noise. Leykin and Citzu (2003) offer to use edge properties features for effectively differentiate paintings from photographs. Leykin *et al.* (2005) found that photographs differ from paintings in their color, edge, and texture properties. Based on these features, they trained and tested a classifier for distinguishing paintings from photographs.

Sankar *et al.* (2009) proposed a framework for differentiating between computer graphics and real images based on an aggregate of other existing features and a feature selection procedure. Shi *et al.* (2007) proposed a method using features formed by using statistical moments of characteristic function of wavelet subbands and their prediction errors.

Sutthiwan *et al.* (2009) used statistical moments of 1-D and 2-D characteristic functions to derive image features that can distinguish between computer graphics and photographic images. Wu *et al.* (2008) proposed a method based on zero connectivity and fuzzy membership to detect forged regions in inpainted images.

Methods pointed out in this section work well for noncompressed images or JPEG images with a high quality factor. Otherwise, they typically fail.

2.5. JPEG AND COMPRESSION PROPERTIES

In order to alter an image, typically the image must be loaded onto photo-editing software and after the changes are done, the digital image is re-saved. Sophisticated methods capable of finding the image's compression history can be helpful in forgery detection. Battiato and Messina (2009) experimentally analyzed some of weakness and strength points of the current solutions based on DCT and JPEG properties. Chen and Hsu (2009) proposed a quantization noise model to characterize single and doubly compressed images.

Fan and Queiroz (2003) proposed a method determining whether an image has been previously JPEG compressed. If so, compression parameters are estimated. Specifically, a method for the maximum likelihood estimation of JPEG quantization steps was developed. Farid (2009) proposed a method for detecting composites created by JPEG images of different qualities. The method detects whether a part of an image was initially compressed at a lower quality than the rest of the image.

Feng and Doerr (2010) detect double JPEG images by using periodic artifacts of re-quantization and discontinuities in the signal histogram. Lukas and Fridrich (2003) presented a method for estimation of primary quantization matrix from a double compressed JPEG image. The paper presents three different approaches from which the Neural Network classifier based one is the most effective. Pevny and Fridrich (2008) proposed a method based on support vector machine classifiers with feature vectors formed by histograms of low-frequency DCT coefficients.

Fu *et al.* (2007) proposed a statistical model based on Benford's law for the probability distributions of the first digits of the block-DCT and quantized JPEG coefficients. Based on the assumption that block operation create disparities across block boundaries, Li (2009) proposed a method for analyzing the properties of image's blocks. Lin *et al.* (2007) proposed a method allowing estimating which kind of source encoder has been applied on the input image.

Luo *et al.* (2007) proposed a method for detecting recompressed image blocks based on JPEG blocking artifact characteristics. The same authors also proposed a detection method (Luo, 2008) for identifying the blocking artifacts. The method is based on cross-differential filter and maximum-likelihood estimation.

Mahdian and Saic (2009) proposed a method for detection double compressed JPEG images based on histograms properties of DCT coefficients and support vector machines. Neelamani *et al.* (2006) proposed a method to estimate the JPEG compression history. . Popescu (2005a) proposed a double JPEG Compression technique by examining the histograms of the DCT coefficients. Qu *et al.* (2008) formulated the shifted double JPEG compression as a noisy convolutive mixing model to identify whether a given JPEG image has been compressed twice with inconsistent block segmentation. Sorell (2008) has explored the conditions under which primary quantization coefficients can be identified. Steven Tjoa *et al.* (2007) proposed a method for determining which transform was used during compression. The method is based on analyzing the histograms of coefficient subbands.

Tjoa *et al.* (2007a) proposed a block size estimation scheme making on the nature of prior image compression or processing. Ye *et al.* (2007) proposed a forgery detection method checking image quality inconsistencies based on blocking artifacts caused by JPEG compression. Zhang *et al.* (2008) proposed a method for detecting JPEG 2000. The method is based on the

statistical difference in the sub-band discrete wavelet transform coefficient histograms between single and double JPEG 2000 compression.

A typical advantage of the methods in this group is their good response for detecting re-saved images. The problem is that often images only are rotated, resized, enhanced (e.g. contrast), re-saved, etc. So, only the knowledge that image has been re-saved often is not enough.

2.6. COLOR FILTER ARRAY AND INTER PIXEL CORRELATION

Many digital cameras are equipped with a single charge-coupled device (CCD) or complementary metal oxide semiconductor (CMOS) sensor. The color images are typically obtained in conjunction with a color filter array. At each pixel location only a single color sample is captured. Missing colors are computed by an interpolating process, called Color Filter Array (CFA) Interpolation. The tampering process can destroy the specific correlations brought into images pixels by CFA interpolation.

Cao and Kot (2009) proposed a demosaicing regularity detection method based on partial second-order derivative correlation models which detect both the intrachannel and the cross-channel demosaicing correlation. Dirik and Memon (2009) proposed two features analyzing traces of CFA. The paper shows the successful application of features for tamper detection and for distinguishing between computer graphics and real images.

Fan *et al.* (2009) proposed a neural network based method for analyzing the traces of CFA. Huang and Long (2008) proposed a decision mechanism using BP neural networks and a majority-voting scheme for demosaicking correlation recognition and digital photo authentication. The method also distinguishes the digital camera photographs from computer graphics.

Poilpre *et al.* (2008) described a method for detecting the traces of Bayer CFA interpolation. The method searches for CFA related peaks in the

Fourier domain. Popescu and Farid (2005) described the specific correlations brought by the CFA interpolation into the image and proposed a method capable of their automatic detection. The method is based on an expectation/maximization (EM) algorithm and uses a linear model.

Swaminathan *et al.* (2007, 2009) technique to find the camera's color array pattern and the color interpolation methods. The estimated interpolation coefficients allow to determine the brand and model of the camera from which an image was captured. One of the most important drawbacks of methods pointed out in this section is their weak results for stronger JPEG compression. Otherwise, they are able to localize the doctored parts of the image with a good precision.

2.7. LIGHTING

Different photographs are taken under different lighting conditions. Thus, when two or more images are spliced together to create an image forgery, it is often difficult to match the lighting conditions from the individual photographs. Therefore detecting lighting inconsistencies can propose another proper way to find traces of tampering.

Under certain simplifying assumptions, arbitrary lighting environments can be modeled with a nine-dimensional model based on a linear combination of spherical harmonics. Johnson and Farid (2007b) have shown how to approximate a lower-order five-dimensional version of this model and how to estimate the model's parameters from a single image. Another work from same authors focuses on image forgeries created by splicing photographs of different people (Johnson and Farid, 2007c). Authors suggest how the direction to a light source can be estimated from specular highlights that appear on the eye.

Farid and Bravo (2010) described several computational methods for detecting inconsistencies in shadows and reflections. Gholap and Bora (2008)

proposed a method to find the forgery in digital images by estimation of the illuminant color. Zhang *et al.* (2009) described how image composites can be detected by enforcing the geometric and photometric constraints from shadows. In particular, they explored shadow relations that are modeled by the planar homology and the color characteristics of the shadows measured by the shadow matte.

A very important advantage of this group is that it is not easy to conceal the traces of inconsistencies in lighting conditions. The disadvantage of the group is the necessary human interpretation of the results.

2.8. LOCAL NOISE

Additive noise is a commonly used tool to conceal the traces of tampering and is the main cause of failure of many active or passive forgery detection methods. Often by creating digital image forgeries, noise becomes inconsistent. Therefore, the detection of various noise levels in an image may signify tampering.

Gou *et al.* (2007) proposed a method based on three sets of statistical noise features. Their features are based on an image denoising algorithm, wavelet analysis and a neighborhood prediction. Mahdian and Saic (2008; 2009c) proposed a method for detecting local image noise inconsistencies based on estimating local noise variance using wavelet transform and a segmentation step.

Popescu (2009a) proposed a method based on measuring the local noise variance using the second and fourth moments. Typically, these methods work well when the level of noise is noticeably different in various parts of the image. Their common problem is their high rate of false positives.

2.9. 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. By the tampering process, the aberration can become inconsistent across the image. This can be used as another way to detect image forgeries.

Dirik *et al.* (2007) proposed a simple method for detecting the presence of chromatic aberration. The method is based on an upscaling operation and mutual information. Gloe *et al.* (2010) analyzed the lateral chromatic aberration and proposed a low-cost estimator of this aberration. Furthermore, test results based on an image database are provided. Johnson and Farid (2006a) proposed a model describing 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.

Methods dealing with chromatic aberration work well for non-compressed non-uniform parts of the image. For the uniform regions of the image or typical JPEG images weak results can be expected.

2.10. IMAGE PROCESSING OPERATIONS

When altering an image, very often a combination of basic image processing operations is applied to the images. Detecting traces of these operations can be very helpful in identifying forgeries. Avcibas *et al.* (2004) proposed a method that discriminates between original and processed images. Here, the work is based on training a classifier with image quality features called generalized moments.

Battiato *et al.* (2005, 2009) used several sets of features for detecting various common image processing operations by constructing classifiers using features based on binary similarity measures, image quality metrics,

higher-order wavelet statistics and a feature selection approach. Farid (2006) proposed three techniques for detecting traces of image processing operations in scientific images. Specifically, image segmentation techniques are employed to detect image deletion, healing, and duplication.

Kirchner and Fridrich (2010) analyzed the detection of median filtering in digital images. Lukas (2000) analyzed usefulness of basic filtering techniques for detection of tampering. Some tampering operations can be approximated as a combination of linear and non-linear components. Swaminathan *et al.* (2006; 2008) modeled the linear part of the tampering process as a filter, and obtained its coefficients using blind deconvolution. These estimated coefficients are then used to identify possible manipulations.

Probably the most common problem of the methods in this section is their weak results for stronger JPEG compression.

2.11. BLUR AND SHARPENING

Often forgeries are created by combination of two or more source images. So, finding in an image various regions with different blur characteristics (blur inconsistencies) can be helpful in detection image forgeries. Furthermore, blur operation is one of the commonly used methods to conceal the traces of tampering.

Cao *et al.* (2010) proposed a local blur estimator for measuring the blurriness of pixels along image's edges. The same authors (Cao *et al.*, 2009) previously proposed a method for detecting sharpened images. The method is based on histogram gradient aberration and ringing artifacts metric. Hsiao and Pei (2005) proposed a tampering detection method based on blur estimation (using images DCT coefficients). Li and Zheng (2008) proposed a method based on the local entropy of the gradient.

Qu *et al.* (2009) proposed an image splicing detector based on the sharp splicing boundaries. Stamm and Liu (2008) proposed a method

detecting global contrast enhancement operations. The method uses artifacts introduced into an image's histogram during the enhancement operations. Sutcu *et al.* (2007) proposed a forgery detection method based on regularity properties of wavelet coefficients used for estimating sharpness and blurriness of edges. Wang *et al.* (2008) proposed an image forgery detection system based on the consistency of defocus blur. The method uses local blur estimation at edge pixels.

Zheng and Liu (2008) proposed a method for detecting traces of artificial blur. Their work is based on a wavelet homonorphic filtering and a mathematical morphology procedure. Zhang and Zhang (2007) a forgery detection method based on analyzing the presence of traces of feather operation used to create a smooth transition between the forged region and its surroundings. Zhou *et al.* (2007) proposed a method for detection of blurred edges. The method is based on edge preserving smoothing filtering and mathematical morphology.

Unfortunately, most of the methods pointed out in this section need a human interpretation of the output.

2.12. PROJECTIVE GEOMETRY

When two or more images are spliced together it can often be difficult to keep the appearance of the image's correct perspective. Thus, applying the principles from projective geometry to problems in image forgery detection can be also a proper way to detect traces of tampering.

Johnson and Farid (2007b) proposed three techniques for estimating the transformation of a plane imaged under perspective projection. Using this transformation, a planar surface can be rectified to be front to parallel, providing a useful forensic tool. Zhang *et al.* (2009) described a technique for detecting image composites by analyzing two-view geometrical constrains.

A very important advantage of this approach is that it is hard to conceal the traces of inconsistencies in projective geometry. Difficulties for automation create one of the main drawbacks of this approach.

2.13. SEMANTIC CONTENT OF IMAGE

Analyzing the semantic content of the image can have a crucial role in image forgery detection. Lee *et al.* (2006) suggest to find perceptually meaningful regions using an image segmentation technique and by using a common-sense reasoning techniques to find ambiguities and anomalies within an image.

A disadvantage of this approach is the need of human interpretation of the results.

2.14. ACQUISITION DEVICE ANALYSIS AND IDENTIFICATION

It is important to note that there also are other groups of forensic methods effective in forgery detection. For example, methods analyzing the image acquisition device have been shown to be very helpful. These methods mostly are based on sensor noise and can be used for a variety of image forensic tasks like

- Forgery localization (Fridrich *et al.*, 2009, Chen *et al.*, 2008)
- Demosaicking artifacts (Bayram *et al.,* 2008; Kharrazi *et al.*, 2004, Bayram *et al.,* 2005, Swaminathan *et al.*, 2009) used the traces of demosaicing to analyze the camera)
- Sensor dust characteristics (e.g., Dirik *et al.* (2008) used location and shape of dust specks in front of the imaging sensor and their persistence make dust spots a useful fingerprint for digital single lens reflex cameras)
- JPEG properties (Kee and Farid (2010) proposed to use the quantization tables to distinguish between original and modified photos)

Furthermore, there also are methods dealing with identification of source cell-phones (for instance, Celiktutan *et al.*, 2008, used binary similarity measures, image quality measures and higher order wavelet statistics to achieve this goal). Typically, a common drawback of these methods is that when the origin (acquisition device) of the image being analyzed is unknown, they cannot be applied. If the acquisition device is known, mostly they need have available a set of other images from the same particular device or at least from the same device model. But, inspite of these disadvantages, the major positive point that attracts many researchers is the fact that no extra embedding information is needed and the original image is not required for comparison of tamper detection.

2.15. ONE-CLASS CLASSIFICATION METHODS

For one-class classification several models have been proposed. Most often the methods focus on outlier detection. Conceptually the most simple solution for outlier detection is to generate outlier data around the target set. Then an ordinary classifier is trained to distinguish between the target data and outliers (Roberts *et al.*, 1994). Koch *et al.* (1995) used ART-2A and a Multi-Layered Perceptron for the detection of (partially obscured) objects in an automatic target recognition system. Unfortunately this last method requires the availability of a set of near-target objects (possibly artificial) belonging to the outlier class. The methods scale very poorly in high dimensional problems, especially when the near-target data has to be created and is not readily available.

In classification or regression problems a more advanced Bayesian approach can be used for detecting outliers (Bishop, 1995; MacKay, 1992; Roberts and Penny, 1996). Instead of using the most probable weights for a classifier (or regressor) to compute the output, the output is weighted by the probability that the weights for the classifier or regressor is correct, given the data. The classifier outputs are automatically moderated for objects remote

from the training domain. These methods are not optimized for outlier detection and they require a classification (or regression) task to be solved. They are therefore very suitable to detect the outlier objects in a classification or regression task. When just a set of objects is available, these methods cannot be used. They also tend to be computationally expensive.

Another possible approach is to use a density method which directly estimates the density of the target objects (Barnett and Lewis, 1978). By assuming a uniform outlier distribution and by the application of Bayes rule the description of the target class is obtained. This directly influences the choice where the probability should be thresholded to obtain a target and an outlier region. For instance, in Bishop (1994) and Tarassenko *et al.* (1995) the density is estimated by a Parzen density estimator. Parra *et al.* (1996) used one Gaussian model.

Ritter and Gallegos (1997) not only estimated the target density but also the outlier density. Unfortunately, this procedure requires a complete density estimate in the complete feature space. Especially in high dimensional feature spaces this requires huge amounts of data. Furthermore, it assumes that the training data is a typical sample from the true data distribution. In most cases the user has to generate or measure training data and he (she) might not know beforehand what the true distribution might be. He can only hope to cover the 'normal state' area in the feature space. This makes the application of the density methods problematic. On the other hand, when a large sample of typical data is available, the density method is expected to work well.

In some cases, prior knowledge might be available and the generating process for the objects can be modeled. When it is possible to encode an object x in the model and to reconstruct the measurements from this encoded object, the reconstruction error can be used to measure the fit of the object to the model. It is assumed that the smaller the reconstruction error, the better

the object fits to the model and the more likely that it is not an outlier. These methods will therefore be called the reconstruction methods. Due to the introduction of prior knowledge, it is expected that these methods will work well, and that they will suffer less from poor generalization and low sample size. On the other hand, when the model does not fit the data well, a large bias might be introduced which completely destroys all good characteristics.

Finally, boundary methods have been developed which only focus on the boundary of the data. They try to avoid the estimation of the complete density of the data (which might be impossible from small sample sizes) and therefore also work with an uncharacteristic training data set. These pure one-class classification methods are relatively new and are completely focused on the problem of one-class classification (the other methods were mainly used in conventional classification problems).

The main advantage of these methods is that they avoid the estimation of the complete probability density. This not only gives an advantage when just a limited sample is available, it is even possible to learn from data when the exact target density distribution is unknown. A user might be able to sample the feature space just very sparsely, without knowledge as to, what more typical examples are and what the exceptions are. For the boundary methods, it is sufficient that the user can indicate just the boundary of the target class by using examples. The user does not have to model or sample the complete distribution.

An attempt to train just the boundaries of a data set is made in (Moya and Hush, 1996; Moya *et al.*, 1993). Here neural networks are trained with extra constraints to give closed boundaries. Unfortunately, this method inherits the weak points in neural network training, i.e. the choice of the size of the network, weight initialization and the stopping criterion.

Although in principle the boundary methods are more efficient than the density estimation, it is not directly clear how one should define a boundary

around a target set Xtr, how to define the resemblance of an object x to a target set Xtr and where to put the threshold. In most cases a distance d(x) to the target set is defined which is a function of (Euclidean) distances between objects, between the test object and the target objects, and between the target objects themselves, where it is used for finding outliers in large databases). This requires well defined distances in the feature space and thus well-scaled features.

This section the various methods proposed for tamper detection. Each group has its own advantages and disadvantages. Almost all of the method requires an additional security secret code like watermark or the original image to be available during tamper detection, which is not always possible to acquire. From the survey, it is understood that studies that focus on the characteristics of the acquisition device (camera in the present study) is sparse. This research, focus on this area and use camera features to detect image forgery or tampering. The method used is presented in the next chapter, Methodology.