DIGITAL SPLICING DETECTION

USING LOCAL INVARIANT

FEATURES

A Dissertation submitted in partial fulfillment of the requirement for the

Award of degree of

MASTER OF TECHNOLOGY IN

INFORMATION SYSTEM

Submitted By :

MALLIKA PANT

(2K13/ISY/12)

Under the guidance of

Ms Ritu Agarwal

Assistant Professor

Department of Information Technology

Delhi Technological University

Bawana Road, Delhi-110042

2013-2015

CERTIFICATE

This is to certify that Ms. Mallika Pant (2K13/ISY/12) has carried out the major project titled "Digital Splicing Detection using local invariant features" as a partial requirement for the award of Master of Technology degree in Information Systems by Delhi Technological University.

The major project is a bonafide piece of work carried out and completed under my supervision and guidance during the academic session 2013-2015. The matter contained in this report has not been submitted elsewhere for the award of any other degree.

(Project Guide)

Ms. Ritu Agarwal

Assistant Professor

Department of Computer Science and Engineering

Delhi Technological University

Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I express my gratitude to my major project guide Ms. Ritu Agarwal, Assistant Professor, lT Dept., Delhi Technological University, for the valuable support and guidance she provided in making this major project. It is my pleasure to record my sincere thanks to my respected guide for her constructive criticism and insight without which, the project would not have shaped as it has.

I humbly extend my words of gratitude to Dr. O. P. Verma Sir, Head of CSE Dept. and other faculty members of department for providing their valuable help and time whenever it was required.

Mallika Pant
Roll No. 2K13/ISY/12
M.Tech (Information System)
E-mail: mallikapant09@gmail.com

ABSTRACT

Digital images have become the most important source of information. Due to presence of various image editing tools, images can be easily changed and altered. Therefore, the authentication of digital images has become an important issue. Forgery is performed by copying one part of an image somewhere else in same image. Copied part can be rotated, scaled or cropped while duplicating. So it is necessary to distinguish between authentic and forged images. These techniques are divided into two varieties- one being active i.e. intrusive. It means one needs to embed something in image example watermark, if the image is modified then the embedded data is also modified. Another one is passive i.e non-intrusive. It is a signature based technique. The work presents and compares feature selection algorithms for the detection of image forgery in tampered images. Various features are extracted from normal and spliced using spatial gray level dependence method and many more. Support vector machine has been used for classification. A very difficult problem in classification techniques is the choice of features to distinguish between classes. The feature optimization problem is addressed using a genetic algorithm (GA) as a search method. Classical sequential methods and floating search algorithm are compared against the genetic approach in terms of the best recognition rate achieved and the optimal number of features.

Table of Content

List of Figures

List of Tables

1 Chapter 1- INTRODUCTION

1.1 INTRODUCTION

In today's era, the use of digitalization and computerization can be seen everywhere because technology have reached heights of development. Information is mostly conveyed through digital images and videos as the growth of digital devices and technology have incremented. Various fields like weather forecasting, forensic investigation, journalism and many more have raises the need of digital images.

Digital images doesn't hold the unique stature due to wide available range of image-editing software and advancement in processing techniques, As a result, image forgeries have increased at an alarming rate. Images are manipulated in diverse areas as in journalism, forensic investigation, medical imaging etc which have earn lots of attention by researchers. These have resulted into development of techniques to find images that are forged. Creation of many kinds of image forgery has become easy with the help of techniques used in image processing. Image Splicing is performed by combining and cutting two or more images and is common amongst all. Human visual system is hardly able to perceive spliced images. Therefore, it calls for development of more new and efficient methods to detect digital image splicing.

Image authentication is a vital and important issue of multimedia security which has attracted lots of attention. Among the available technologies for forgery detection, the division can be done into two broad categories, one being active and other passive. The former needs a prior knowledge to detect the authenticity of an image like digital watermarking where we feed some information prior to detection, whereas the latter performs the same work without any information. It utilizes the distinct properties of original images or changed traces to perform authentication.

1.2 DIGITAL IMAGE FORGERY

Image forgery is defined as changing, deleting, or adding some crucial and important features to an image from another image without leaving any obvious trace. There exist various methods used for tampering an image. Based on the methods used to create forged images, digital image forgery is categorized into three main types: Copy-Move forgery, Image splicing, and Image Re-sampling.

1.2.1 Copy-Move Forgery

This type of forgery involves the cutting a region of an image and then pasting somewhere else in the same image to create duplicate portions of a image or to hide some crucial information. Unlike image splicing, source as well as the target image is same in this forgery. In Fig1.1. The missile part is copied and then pasted creating a different look than the original image . Different post- processing techniques like re-sampling, blurring are applied to make forgery disappear

Figure 1.1 Copy Move Forgery

1.2.2 Image Splicing

Image splicing uses cut-and-paste techniques to produce a new fake image from two or more original images. The borders linking the spliced portions can visually be imperceptible when splicing is performed cautiously.

Figure1.2 Image Splicing

1.2.3 Image Resampling

To create a high quality forged image, some selected regions have to undergo geometric transformations like rotation, scaling, stretching, skewing, flipping etc. It is basic technique to make other forgery operations more effective and imperceptible. Resizing is the most commonly used operation applied in any forgery. For example creating a composite image of two people not from same images, resizing is necessary to maintain their relative heights.

Figure 1.3 Image Resampling

1.3 DIGITAL IMAGE FORGERY DETECTION METHODS

Image forensics approaches falls under two categories.

Active approach [1] [2]uses a digital signature or a watermark which is embedded inside the original image to prove or reject the authenticity of the image. But this approach holds a strong limitation that the watermark which is embedded must be performed either by an authorized person processing the image or by the acquisition device.

Passive approach or blind approach [3] [4] uses the facts that there exists statistical changes in images or some marks are left by the camera during the process of creation to detect the forgery attack. Unlike active methods, blind approach doesn't require any information regarding authentic image. Blind forensics detection can be classified into six categories [5] i.e. pixel-based, camera based, format-based, geometric-based, physics-based and source camera identification-based.

1.3.1 Pixel-Based Techniques

Statistical changes introduced at the pixel level during the forgery process can also be used for detection. Pixel based techniques uses these anomalies. At pixel-level, various correlations that occur due to specific type of forgery are analyzed and examined in this technique either in spatial domain or in some transformed domain.

A forensic method may use intrinsic fingerprint of each encoder belonging to. Image coding detector can detect exact image encoder among sub band encoders, DPCM encoders and transform based encoder by the use of intrinsic fingerprint. A group of forensic methods came into existence which was capable to detect local as well as global contrast enhancement. It may also identify where to use histogram equalization and when to not. Key point based method like that of SIFT features are executed very efficiently. They are sensitive to regions of low contrast and image content which is repeated. Therefore, methods of block based criteria came into picture and they clearly brought improvement in performance. Techniques to encode words that are visual and features of indexing give a efficient solution to problems of detecting duplicate images. Histogram of oriented gradients which is a block based technique uses quantization

which is non uniform to create a feature for a single image block for detecting forgery purposes. This method executes by using method of encoding the distribution of the features belonging to image to work with patterns that are highly textured.

1.3.2 Camera-Based Techniques

During different stages of image processing, numerous artifacts are produced and need to be exploited in order to detect tampering traces. Examination of these camera artifacts may include color filter array, camera response, and chromatic aberration and sensor noise imperfections. Various digital cameras make use of a single sensor with CFA (color filter array) .They then interpolate the color samples those are missed in order to get a color image of three channels. The interpolation produces correlations which are specific and which get destroyed when forgery is performed with any image. Interpolation of CFA is done to understand these correlations and quantify them in any region of image. The processing of single chip camera is related to demos icing regularity. It follows a correlation model which is based on partial second order to detect intra-color as well as demos icing correlations in cross channel. Also a reverse technique of classification is employed for classification of demos iced samples into smaller ones which reveals the original grouping accurately. For this, an EMRC algorithm is executed to resolve the ambiguous demosaiced axes. There also exists unified approach to detect source camera from its images and identifying any alterations using photo-response non uniformity noise (PRNU). It is stochastic fingerprint which is unique to imaging sensors. An estimator to estimate maximum likelihood is used to obtain PRNU. Both the above mentioned tasks of digital forensics are achieved by detecting the existence of sensor PRNU under regions of investigation. The process of detection of manipulations in images works by analyzing a patch of image to decide whether it belongs to a category of authentic or spliced ones.. Fusion Boost was incorporated by learning the weight of each classifier in order to construct a strong ensemble detector. The forgery detection based on PRNU works by using the Bayesian Framework to recast the problem, and then accounting for spatial dependencies of decision variables when they are modeled as Markov random field.

1.3.3   Format-Based Techniques

Various correlations that occur during any compression methodology is examined to detect forgery. A method to examine low-quality JPEG images and then to detect the region of image which is compressed at a lower quality level than the existing JPEG quality of image. This type of region is detected when the image is resaved at a multitude of qualities of JPEG and then detects the local minima between image and its JPEG compressed version. 1-D feature was introduced to detect whether a given bitmap image is JPEG compressed previously. The blind approach for detection which is based on estimation of quantization is executed in three steps: pre-screening, selection of candidate region, and then identification of tampered region. Prescreening test is done to detect if the image is JPEG compressed or not. Then , in order to reduce the impact of forged regions on quantization table, regions which act as candidate are selected for estimation of quantization table.

1.3.4   Geometric-Based Techniques

Here principles of projective geometry are applied to develop robust algorithm for detection. Abnormality in relative positions of an object or a person used to detect objects moved or translated in case of any forgery.

1.3.5   Source Camera Identification-Based Techniques

The specifications of source camera are noticed to identify any kind of tampering. These specifications include color filter array interpolations, sensor noise and lens aberrations

1.3.6   Physics-Based Techniques

While merging different photographs, matching lighting conditions is difficult. Light variations are used as evidence to detect tampering. Inconsistency is detected in order to expose fakeness and used as an evidence of forgery. It performs approximation of lightning environment in low dimensional model and then estimates the parameters.

1.4   Related Background

Image splicing detection is used to detect if a given image is a composite one created by combining or separating different portions of two or more images. In previous years, some blind techniques for detection of image tampering came into existence. In series of paper [6] [7], higher order moment spectra, bi-coherence are treated as features in order to detect images that are spliced. When tested, bi-coherence was found sensitive to quadratic phase coupling which is caused by quadratic phase coupling. Detection accuracy of 72% was found when these features are performed on proper dataset of images. Hilbert Huang Transform [8] was used to examine disturbance in linearity caused by image splicing. Also a statistical model using wavelet decomposition for calculating moments of characteristic function was proposed. Detection accuracy was 80.15% was found when these two features are used collectively. Hsu and Chang [9] used camera response function and geometry invariants to detect splicing in semi- automatic manner. In order to convert this work into fully automatic manner, image segmentation is incorporated [10].

The measure of blurriness [11] and local sharpness [12] was taken as advantage to check if image under detection is blurred/smoothed or not. A scheme to detect tampered regions in JPEG images was proposed in [13]. In [14] a scheme used for steganalysis was used to check image splicing. The scheme used was Run- Length based scheme which was in [15] by He et al. Statistical moments obtained from wavelet characteristic function and 2D phase congruency [16] was used as features to differentiate spliced images from original ones. On detection of these features on Columbia Image Splicing Evaluation Dataset [17] using SVM classifier a accuracy of 82.32% was attained. Shi et al. [18]checked two different types of statistical features proposing a natural image model on Columbia Image Splicing Dataset attaining detection accuracy of 91.87%. The two features rely on 1-D and 2-D moments of characteristic function. Another different variety of features based on Markov random process (Transition Probability Matrix) in DCT domain [19] also performed better in terms of detection accuracy.

2 Chapter 2: LOCAL FEATURES

2.1 LOCAL BINARY PATTERNS

Local Binary Patterns is a simple and efficient texture feature. It works by labeling the pixels of an image by applying threshold in the neighborhood of each pixel and then the result is represented as a binary number. It has become popular due to its numerous advantages. It has a discriminative power and is robust to monotonic gray scale changes which enable it to be as a operator in real world applications [6]. It is used to analyze images as it is computationally simple. It is also considered as a approach to unite various diverse structural and statistical models used in texture analysis traditionally

The LBP feature vector is generated by following steps

- The window to be examined is divided into cells (e.g. 16x16 pixels for each cell).
- Each pixel is compared in a cell to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). The pixels are followed along a circle, i.e. clockwise or counter-clockwise.
- Where the neighbor's value is smaller than the center pixel's value, then value considered is "1". Otherwise, "0". This gives an 8-digit binary number.
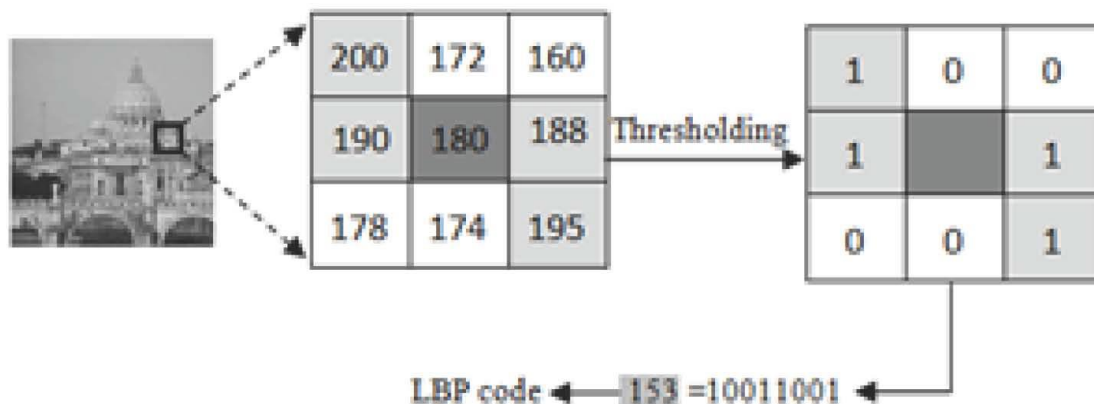
Figure 2.1 Local Binary Pattern

- The histogram is computed, over the cell, of the frequency of each "number" that occurs.
- The histogram is normalized.

- Finally, the histograms of all cells. This represents the feature vector.

This feature vector is processed to classify images using the Support vector machine or other machine-learning algorithm.

2.2   SCALE INVARIANT FEATURE TRANSFORM

SIFT features as the name suggests, these are the features invariant to scaling and rotation of images. These are the local features widely used in gesture recognition, video tracking, and object recognition. It give reliable results while matching images for forgery [7] as it is also robust against affine transformation and illumination variance. The algorithm works in stages defined below:

2.2.1   Scale Space Extrema detection

This stage starts with the convolution process of image with the Gaussian filters at multiple scales. Interest points are the maxima or minima of the difference of Gaussian that occur at different scales.

L is the convolution of the original image I with the Gaussian blurred image G at scale $k\sigma$

$$L(x, y, k\sigma) = G(x, y, k\sigma) * I(x, y) \qquad \text{2.1}$$

and the difference of Gaussian image is given by

$$D(x, y, \sigma) = L(x, y, k_i\sigma) - L(x, y, k_j\sigma) \qquad \text{2.2}$$

Hence the DOG image is the difference of the Gaussian blurred image at different scales
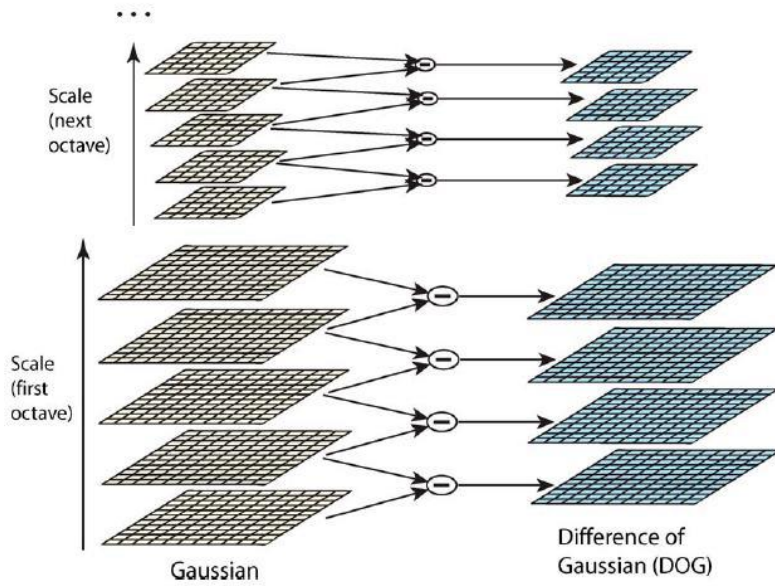
Figure 2.2 The blurred images at different scales and the remaining DoG images.

2.2.2 Key point Localization

At the same scale, each pixel in the DoG image is compared to its 8 neighbors and at neighboring scales to its 9 corresponding neighbors. In case the pixel is either local minima or maxima, it is considered as a key point. Interpolation of neighboring data is applied to determine the position of each candidate key point more accurately.
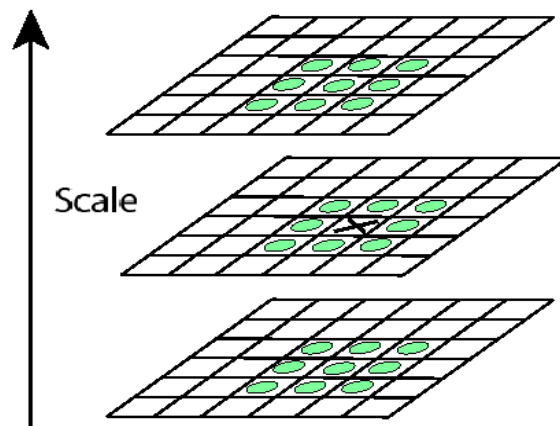
Figure 2.3 Local Extrema Detection

2.2.3   Orientation Assignment

Based on local image gradient directions, each key point is assigned orientation in order to achieve robustness against rotation of image. For a Gaussian smoothed image gradient magnitude m and orientation θ is calculated by following equations.

$$m(x,y) = \sqrt{\left(L(x+1,y) - L(x-1,y)\right)^2 + \left(L(x,y+1) - L(x,y-1)\right)^2} \qquad 2.3$$

$$(x,y) = atan2\left(L(x,y+1) - L(x,y-1), L(x+1,y) - L(x-1,y)\right) \qquad 2.4$$

2.2.4   Key Point Descriptor

Neighborhoods of 4*4 pixels with 8 bins each are chosen to create a set of orientation histograms. The computation of histogram are done from magnitude and orientation values of 16 x 16 region samples around the key point so that every histogram have samples from a 4 x 4 sub region of the original neighborhood region. Gaussian function with scale equal to 1.5 times of the width of the descriptor window is used to weigh the magnitude. The descriptor is then a vector of all the histograms values.
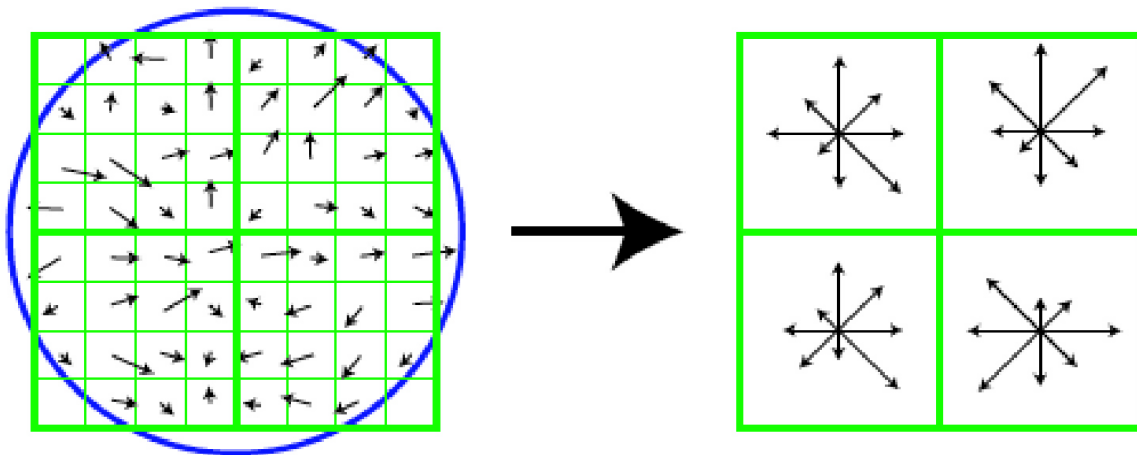
Figure 2.4 SIFT Feature Description

11

2.3   HISTOGRAM OF ORIENTED GRADIENTS

These descriptors rely on the fact that the occurrences of gradient orientation play a vital role in the process of localization of parts in image [8]. Unlike SIFT, it performs on dense grid structure of cells spaced uniformly. It performs local contrast normalization on blocks that are overlapped. It is widely used in object detection. This feature is calculated according to following steps:

2.3.1   Gradient computation

Firstly, the gradient values are computed. To compute these values 1-D centered, point discrete derivative mask is applied in horizontal and vertical directions. Following filter kernels are used to filter the intensity values:

$$[-1,0,1]   \text{And}   [-1,0,1]^T$$

2.3.2   Orientation binning

Next step is to create cell histogram where cells can be radial or rectangular in shape. Each pixel within the cell gives a weighted vote based on the values of gradients for a histogram channel. The histogram channel is orientation based and spread evenly over 0 to 360 degrees or 0 to 180 degrees which depends on whether the gradient is signed or unsigned.
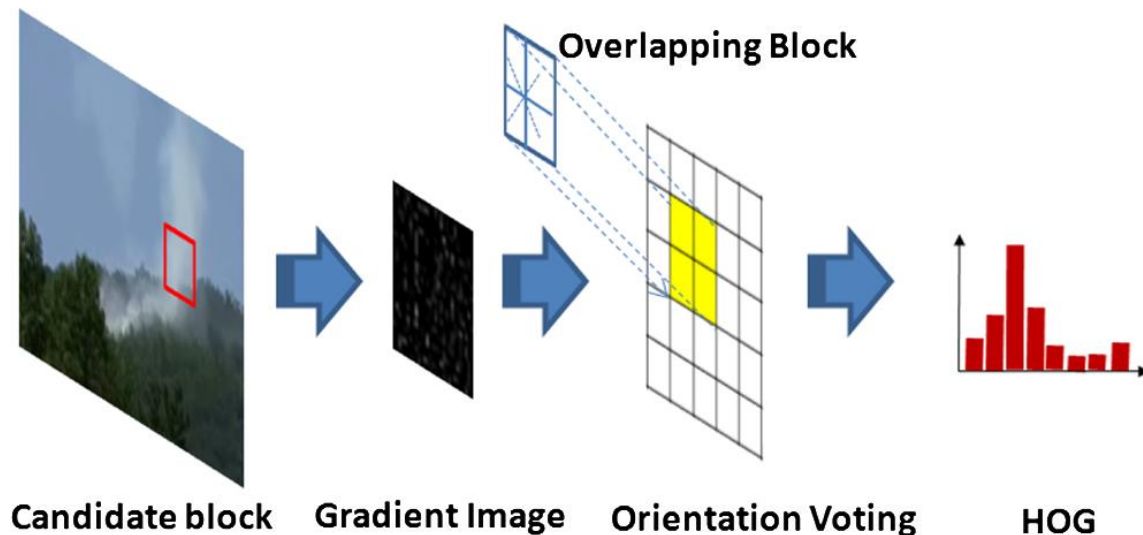
Figure 2.5 Histogram of oriented gradients

2.3.3 Descriptor blocks

In order to detect changes in contrast and illumination, the gradient strengths should be normalized locally. It requires grouping the cells into larger connected blocks. The descriptor is a vector formed by concatenating the normalized cell histogram components from all blocks. There exists two main blocks: rectangular R-HOG blocks and circular C-HOG blocks. R-HOG blocks are generally square grids, consisting of three parameters: the number of pixels per cell, the number of channels per cell histogram and the number of cells per block. Circular HOG blocks are represented with four parameters: the number of radial and angular bins, the expansion factor for the radius of additional radial bins and the radius of the center bin.

2.3.4 Block normalization

Normalization of blocks was performed by following methods. Then the normalization factor can be one of the following:

L2-norm: $f = \dfrac{v}{\sqrt{\|v\|_2^2 + e^2}}$ 2.5

L2-hys: Renormalizing is done after L2-norm is clipped

L1-norm: $f = \dfrac{v}{(\|v\|_1 + e)}$ 2.6

L1-sqrt: $f = \sqrt{\dfrac{v}{(\|v\|_1 + e)}}$ 2.7

Where v is the non-normalized vector containing all histograms in a given block, $\|v\|_k$ be its k-norm for k=1,2 and e be some small constant .

2.4   SPEEDED UP ROBUST FEATURES

SURF which stands for Speeded up Robust Features is a local descriptor used for applications of object recognition. It is more invariant to transformations and faster than SIFT. It uses the multi-resolution pyramid technique to extract features in images which is transformed into coordinates [7]. This algorithm has similar principles and steps as SIFT but details in each step are different. The algorithm has two main steps: interest point detection, local neighborhood.

2.4.1   Interest Point Detection

Unlike SIFT algorithm which uses cascaded filters to detect scale-invariant characteristic points,. SURF algorithm uses square-shaped filters as an approximation of Gaussian smoothing. If the integral image is used, image filtering with a square is much faster .Therefore it is defined as

$$S(x, y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i, j) \qquad\qquad 2.8$$

To find interest points, SURF uses a blob detector based on the Hessian matrix .

The measure of the local change around the points is the determinant of the Hessian and then points are chosen where this determinant is maximal. SURF also uses the same determinant for scale selection. The Hessian matrix H (p, σ) at point p =(x, y) in an image I, and scale σ, is defined as follows:

$$H(p, \sigma) = \begin{pmatrix} L_{xx}(p, \sigma) & L_{xy}(p, \sigma) \\ L_{xy}(p, \sigma) & L_{yy}(p, \sigma) \end{pmatrix} \qquad\qquad 2.9$$

where values in the matrix are the second-order derivatives of the grayscale image.

The interest points are calculated in different scales, as search may require comparison at different scales. Scale space is considered as an image pyramid. Gaussian filter is used for smoothing the images which are then subsample to form next pyramid of higher level. Several floors are calculated as:

$$\sigma\ approx = Current\ filter\ size * \left(\frac{(Base\ filter\ scale)}{Base\ filter\ size} \right) \qquad\qquad 2.10$$

The scale space consists of a number of octaves. Scale spaces are created when box filters of various sizes are applied. Analysis of the scale space is performed by increasing the filter size

rather than decreasing the size. Scale space interpolation is especially important issue here, as the scale difference is relatively large between the first layers of every octave.

2.4.2 Local neighborhood descriptor

A orientation is fixed which relies on information around the interest point. A square region is constructed which is aligned to selected orientation and then SURF features are extracted from it. To achieve robustness against rotation, orientation of interest points are found. To estimate orientation, Haar wavelet responses in x and y directions within a circular neighborhood are weighed by a Gaussian function and then plotted in two dimensional space where abscissa is horizontal response and ordinate is vertical response. The dominant orientation is found when all the responses within a sliding window are summed up.

3 Chapter 3: GENETIC ALGORITHM

3.1 INTRODUCTION

Genetic algorithm is a technique of artificial intelligence that copies the process of natural evolution to generate solutions to optimization problems. Genetic algorithm generate these solutions using techniques like crossover and mutation [23] .These algorithms are widely used in computational science, bioinformatics, economics, physics and many more.

In this algorithm, initially a candidate solution to the problem is taken which evolve towards better solutions. These solutions are taken in any encoding format but usually binary form of 1s and 0s is preferred. This encoding is in chromosomal manner. The process starts from a initial population of individuals which are randomly generated and in each iteration the population is termed as a generation. Fitness function value is evaluated for each generation. The solution to the optimization problem largely depends on the fitness function selection. Based on fitness value individuals are selected for crossover and mutation in order to create a new generation. The same process goes in an iterative manner   until a required fitness value is achieved or maximum number of iterations has been performed.

3.2 METHODOLOGY

The entire procedure of genetic algorithm is:

3.2.1 Initialization

  The initial population in genetic algorithm is generated randomly from within the overall range of solutions which are possible. This solution range is termed as search space. The size of the population depends on type of the problem.   Once the population is initialized or an offspring population is created, the fitness values of the candidate solutions are evaluated.

3.2.2 Selection

During each iteration of generation process, some proportion of population is selected to create new generation. Selection procedure totally follows the survival of the fittest mechanism. It provides more copies of those solutions who acquire higher fitness values. The main idea of selection is to choose better solutions than worse ones. For this, various selection procedures

including roulette-wheel selection, stochastic universal selection, ranking selection and tournament selection

A generic selection procedure may be implemented as follows:

1. Fitness value is calculated for each individual and these values are then normalized. Normalization is performed to make the sum of all fitness values equals 1.
2. The population is arranged in descending order of fitness values.
3. Cumulative fitness values are computed.
4. A random number N between 0 and 1 is chosen.
5. The individual selected is the first one whose cumulative value is greater than N.

- **Roulette wheel selection:**

The chance of selection of a chromosome depends on its fitness value. Steps involved are:

Step 1: Use the fitness function to calculate value (fv) of each chromosome.

Step 2: Sum up (Sf) all the fitness values of all chromosome in population.

$$Sf = \sum_{i=1}^{n} fv_i \qquad\qquad 3.1$$

Step3: Average fitness value (Af)

$$Af = \frac{Sf}{n} \qquad\qquad 3.2$$

Step4: Expected fitness value (Ef) for each chromosome

$$Ef_i = \frac{fv_i}{Af} \qquad\qquad 3.3$$

Step5: Sum of expected fitness of each chromosome

$$SumEf = \sum_{i=1}^{n} Ef_i \qquad\qquad 3.4$$

Step6: Choose a random number (R) in range [0, SumEf].

Step7: Select the chromosome whose (SumEf $>$ = R).

Step 8: Repeat from step 6 for n times, where n represents population size.

- **Elitism Selection:**

Here the chromosomes are arranged in decreasing order as per the fitness values. Then the selection is applies with each pair of chromosomes in an arranged manner. This is done to apply genetic algorithm between both same types of chromosomes, i.e. between strong-strong or between weak- weak chromosomes. It wouldn't apply between weak and strong chromosomes.

- **Rank Selection:**

Rank of a chromosome is decided according to their fitness function values. New fitness values are computed using different fitness function. Roulette wheel method is used to select the chromosome. The steps are:

Step 1: According to fitness values, arrange the chromosomes in non-increasing order.

Step 2: According to the chromosome set arrangement, assign a rank value to it.

Step 3: Calculate the new fitness value for each chromosome using

$$F = max - (\text{max} - \text{min}) * \frac{\text{rank}-1}{\text{Npop}-1} \qquad\qquad 3.5$$

Where 1<max<=2 & min = 2-max.

- **Stochastic Universal Sampling :**

Unlike roulette wheel selection method, here the roulette wheel is spinned just once instead of n times after n points in the wheel are determined. Finally choose n chromosomes that lies in front of the points determined.

- **Binary Tournament Selection:**

Repeat n times the following step:

  A. Randomly choose two chromosomes from the population.

  B. Select the chromosome with the highest fitness value.

3.2.3 Recombination.

This process deal to creating new and better solutions from two or more existing solution. In other words creating new offspring from the parents. In genetic algorithms, crossover is a genetic operator used to create new generation. Different types of crossover are:

- **One point crossover:**

A point on both parents' organism strings is chosen and data beyond that point is swapped between the two parent organisms in either organism string. The points is known as single crossover point
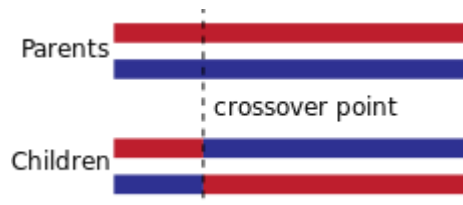
Figure 3.1 One Point Crossover

- **Two point crossover:**

The process is same as single crossover; the only change here is two crossover points are chosen.
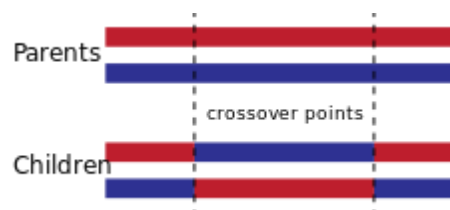
Figure 3.2 Two Point Crossover

- **Cut and Splice:**

Here parents have different choice on points of crossover.
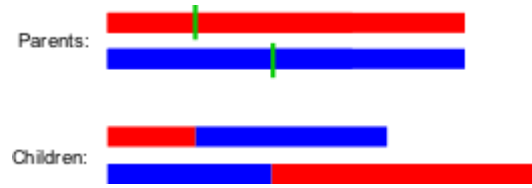
Figure 3.3 Cut and Splice

- **Uniform crossover and Half Uniform Crossover:**

In uniform crossover, individual bits from parent strings are compared. Then swapping of bits is done with a fixed probability. In half uniform crossover, swapping is done between exactly half of the non matching bits.

3.2.4   Mutation.

Similar to biological mutation where gene values are altered in chromosome, mutation is used to maintain diversity from generation to generation. It is a genetic operator to provide new solutions which are better. Mutation occurs during evolution based on mutation probability. This probability should be set low. On setting it too high, a primitive search which is random is performed.

Different mutation types are:

- **Bit string mutation**

The mutation is performed by flipping the bits at random positions in strings.

Example:

1 0 1 0 0 1 0

↓

1 0 1 0 1 1 0

- **Flip Bit**

Here a chromosome is chosen and the bits are inverted.

- **Boundary**

This operator randomly replaces genome with lower or upper bound.

- **Non Uniform**

The probability that amount of mutation will reach zero with the next generation is incremented by using non-uniform mutation operator.

- **Uniform**

Here value of chosen genome is replaced by a uniform random value which is selected by the lower and upper bounds specified by the user.

3.2.5 Termination

This algorithm terminates in case a solution is obtained or maximum number of generations are produced.

4 Chapter 4: CLASSIFIER

Classification is the method of identifying the group to which a new observation belongs given a set of groups or classes. This process is based on training data set which contains observation whose class label is known. In terms of machine learning, classification belongs to supervised learning methods. Any algorithm that does classification is a classifier. It refers to the mathematical function, where input data is mapped into a category. Linear classifier includes perception, logistic regression naive bayes classifier and many more. Classification algorithms are widely used in speech recognition, handwriting recognition, pattern recognition etc.

4.1 Support Vector Machine

SVM has become popular for it's widely used in the fields of pattern recognition and regression in previous years. It belongs to supervised learning methods which run by analyzing data and recognizing patterns. Given a set of training examples with known class labels, SVM develop a model which assigns new examples a class label based on patterns that are recognized. It is a non probabilistic linear classifier. It not only performs linear classification but also non linear one. SVM constructs one or a set of hyper-planes in high dimensional space. Basically, a good classification is done by the plane that covers maximum margin as it reduces the generalization error.
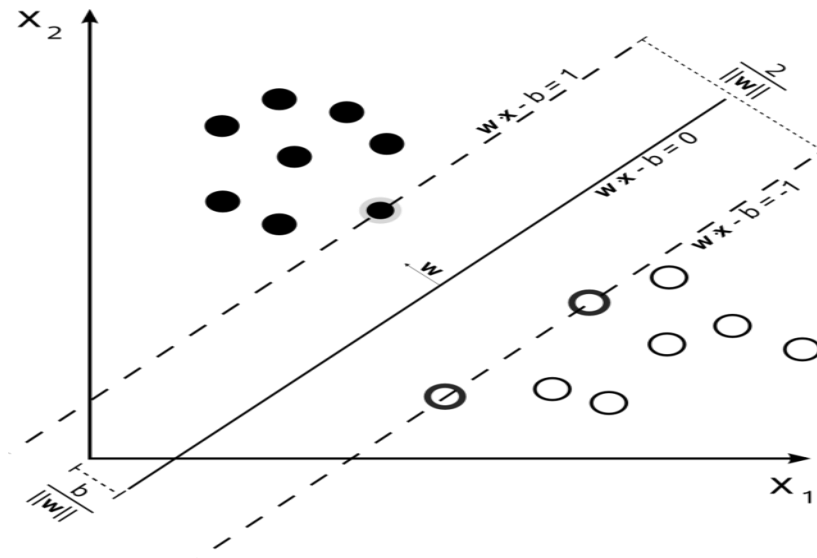
Figure 4.1 Maximal Margin Hyper-plane and margins for an SVM.

22

A set of n points comprising training data D where y is either 1 or -1 which indicates the class to point x belongs. Any hyper-plane can be represented as the set of points x that satisfy equation:

$$w.x - b = 0 \tag{3.6}$$

. Denotes the dot product and w is the normal vector to hyper plane.

Two hyper planes that separate data are selected if the training data is linearly separable. Then the distance between them called margin is maximized. Hyper planes are represented as:

$$w.x - b = 1 \tag{3.7}$$

And

$$w.x - b = -1 \tag{3.8}$$

It can be rewritten as:

$$y_i(w.x_i - b) \geq 1, \quad \text{For all } 1 \leq i \leq n$$

The maximum margin classifier, which is the standard SVM, is found by maximizing the margin and is equivalent to the following problem

Minimize

$$min_{(w,b)} \quad \frac{1}{2}\|w\|^2 \tag{3.9}$$

Subject to

$$y_i(w.x_i - b) \geq 1 \quad \text{For all } 1 \leq i \leq n \tag{4.0}$$

4.2   Twin Support Vector Machine (TWSVM)

It is a binary SVM classifier that solves two related SVM type problems and obtains two non-parallel planes. Both problems are smaller than in a standard SVM .The TWSVM classifier [24] is obtained by solving pair of quadratic programming problems as below

TWSVM1:

$$min_{w^{(1)},b^{(1)},q} \quad \frac{1}{2}\left(Aw^{(1)} + e_1 b^{(1)}\right)^T\left(Aw^{(1)} + e_1 b^{(1)}\right) + c_1 e_2^T q \tag{4.1}$$

$$subject\ to \quad -\left(Bw^{(1)} + e_2 b^1\right) + q \geq e_2, q > 0 \tag{4.2}$$

TWSVM2:

$$min_{w^{(2)},b^{(2)},q} \quad \frac{1}{2}\left(Bw^{(2)} + e_2 b^{(2)}\right)^T\left(Bw^{(2)} + e_2 b^{(2)}\right) + c_2 e_1^T q \tag{4.3}$$

$$subject\ to \quad \left(Aw^{(2)} + e_1 b^2\right) + q \geq e_1, q > 0 \tag{4.4}$$

For each class, it finds two hyper planes and then classifies point. The first term in the objective function of (14) or (15) is the sum of squared distances from the hyper plane to points of one class. Therefore, minimizing it tends to keep the hyper plane close to points of one class. The second term of the objective function minimizes the sum of error variables. In short, TWSVMs consists of a pair of quadratic programming problems. In each QPP, the objective function corresponds to a particular class and the constraints are determined by patterns of the other class. TWSVM classify approximately four times faster than the usual SVM.

5 Chapter 5: Proposed Approach

We proposed an efficient methodology to detect image splicing. The following flowchart illustrates the architecture of our approach.

```
┌──────────────────────────────────────────────────────────────────────┐
│                                                                        │
│                          ┌──────────────────────────────────────────┐ │
│                          │ Extracted              Best feature       │ │
│                          │ features                 subset           │ │
│   ┌──────────────┐  Input│    ┌─────────────────────┐                │ │
│ → │ Feature      │ ──────┼──→ │ Feature   Selection │ ──────────────→│→│
│   │ Extraction   │       │    │ Module              │                │ │
│   │ Module       │       │    │ (Genetic approach)  │                │ │
│   └──────────────┘       │    └─────────────────────┘                │ │
│                          │ Assessment of          Feature            │ │
│                          │ feature subset           subset           │ │
│                          │    ┌─────────────────────┐                │ │
│                          │    │ Evaluation Module   │                │ │
│                          │    │ (SVM classifier)    │                │ │
│                          │    └─────────────────────┘                │ │
│                          │           Training/Testing                │ │
│                          └──────────────────────────────────────────┘ │
│                                                                        │
│                          ┌──────────────────────────────────────────┐ │
│                          │      ┌─────────────────┐                  │ │
│                          │      │   Evaluation    │ ←────────────────│←│
│                          │      └─────────────────┘                  │ │
│                          │            Validation                     │ │
│                          └──────────────────────────────────────────┘ │
└──────────────────────────────────────────────────────────────────────┘
```
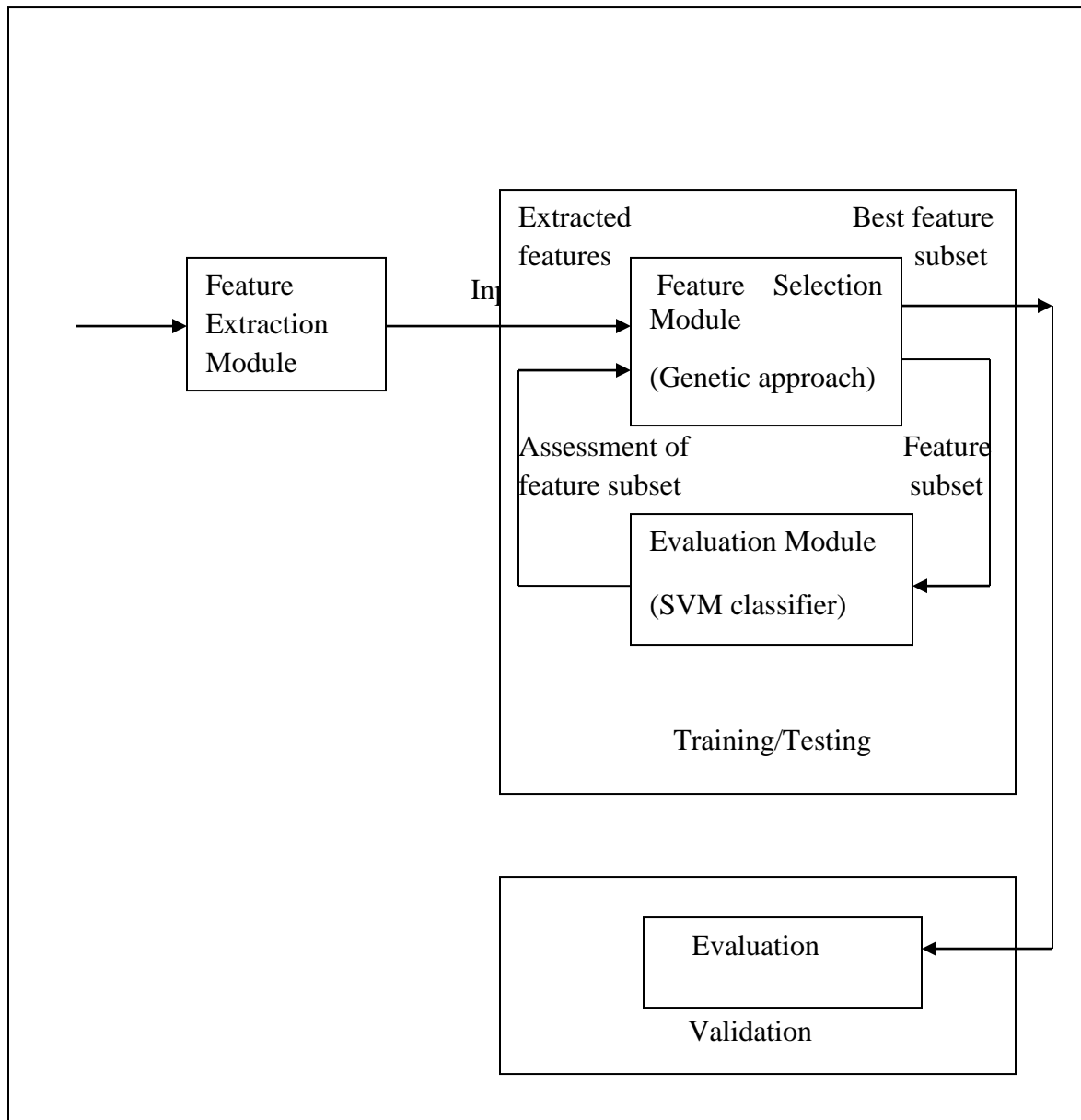
Figure 5.1 Architecture of the detection algorithm

The approach starts by extracting features from given input images. To achieve high recognition rate, a subset of best features need to be selected. Exhaustive search to evaluate best feature subset is not feasible as it increases required computational effort. Genetic approach, being a heuristic search provides a feasible approach for such optimization problem. Therefore, genetic algorithm is used for feature selection. Classical sequential methods are compared with genetic algorithm approach in context with recognition rate and cardinality of feature subset chosen.

5.1 Feature Extraction

A set of ten features are extracted from authentic as well as spliced images. They are:

A. Local Binary Pattern(LBP):

It is an efficient texture feature which works by labeling the pixels of an image and then thresholding each of the neighboring. Finally the result is displayed as a binary number.

B. Entropy:

$$ENT = -\sum_{i=1}^{N}\sum_{j=1}^{N} C(i,j) \log\big(C(i,j)\big) \qquad 5.1$$

Where C denote the co-occurrence matrix and N is the number of levels that are gray in quantized image.

$$C_x(i) = \sum_{j=1}^{N} C(i,j) \qquad 5.2$$

$$C_y(i) = \sum_{i=1}^{N} C(i,j) \qquad 5.3$$

C. Contrast:

$$CON = \sum_{n=0}^{N-1} n^2 \left\{ \sum_{i=1}^{N} \sum_{j=1}^{N} \{C(i,j)\} \right\} \qquad\qquad 5.4$$

D. Histogram of oriented gradients(HOG):

This descriptors use the fact that the occurrences of gradient orientation helps to localize the parts in image. Unlike SIFT, it performs on dense grid structure of cells spaced uniformly. It performs local contrast normalization on blocks that are overlapped.

E. Inverse difference moment:

$$IDM = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{1}{1+(i-j)^2} C(i,j) \qquad\qquad 5.5$$

F. Speeded up robust features:

It uses the multi-resolution pyramid technique to extract features in images which is transformed into coordinates.

G. Inertia:

$$INER = \sum_{i=1}^{N} \sum_{j=1}^{N} (i-j)^2 * C(i,j) \qquad\qquad 5.6$$

H. Cluster shade:

$$CS = \sum_{i=1}^{N} \sum_{j=1}^{N} (i+j-u_x-u_y)^3 * C(i,j) \qquad\qquad 5.7$$

I. Cluster prominence:

$$CP = \sum_{i=1}^{N} \sum_{j=1}^{N} (i+j-u_x-u_y)^4 * C(i,j) \qquad\qquad 5.8$$

27

J. Angular second moment:

$$ASM = \sum_{i=1}^{N} \sum_{j=1}^{N} \{C(i,j)\}^2 \qquad\qquad 5.9$$

5.2   Feature selection:

The main reason to performing feature selection is to use smaller number of features while maintaining the accuracy of classification.

5.2.1   Deterministic search method

Feature selection algorithms are divided into following categories, i.e. Exponential, randomized and sequential algorithms. The algorithms of sequential search method have $O(d^2)$ complexity. They use hill-climbing strategy to select or reject features. Sequential forward selection (SFS) starts by evaluating all feature subsets which contains one feature and select the feature with best performance. This subset is then combined with the feature that gives best performance for larger size subsets. The process repeats until no improvement is done by extending the current subset. Sequential backward selection (SBS) starts with set which contains all features and repeatedly removes a feature whose rejection gives maximal improvement in performance. Both these methods are sub optimal methods that creates a nesting of features in straight forward manner. Due to this previous additions and removal of features cannot be corrected. To overcome this problem, these algorithms are modified and are termed as floating sequential algorithms [11]. Sequential floating forward selection (SFFS) works by applying backward steps after each forward step until the corresponding subsets are better than previous ones.

5.2.2   Genetic algorithm approach

In order to select best set of features, genetic algorithm which is based on global search method is used. Based on Darwin principle, the algorithm states that initial population of individuals have high probability of survival. Here each feature subset is encoded in form of a binary string which is called as chromosome. Each bit is associated with a feature in the binary vector. If the i^t bit of this vector is 1 then it means that i^{th} feature is allowed to participate in classification. In case the bit is a 0, then the feature is not selected. Fitness value is assigned based on a

28

classification performance. Here genetic approach has been used to select the best features by tournament approach. Diagnostic accuracy, sensitivity, and specificity measure are considered for evaluation. Also, the number features that are not selected (zeros in a chromosome) with a coefficient are added.

Fitness function:

$$fitness = Acc + Se + Sp + 0.05 * Number \qquad\qquad 5.10$$

where Acc, Se, and Sp are accuracy, sensitivity and specificity, respectively.

In order to maintain diversity in solution, replacement technique is used and whole population is replaced with new generation. Parents having high selection rate are selected. For simplicity, single point crossover is adopted here. The feature subset among all the generations with the highest classification rate is regarded as the optimum.

5.1 Classification

The approach used here is to improve the usefulness of machine learning techniques that are used for classification. The main criteria are to improve accuracy of classification and reduce the size of feature subsets using genetic algorithm approach. To save time, the population size and number of generations used by genetic algorithm are relatively small. To achieve better results, it is possible to allow more iterations or larger population size. SVM and Twin SVM are used as the classifier.

The entire procedure of feature selection is illustrated below:

```
                    ┌──────────────────┐
                   (   Full   feature   )
                   (     subset          )
                    └──────────────────┘
                              │
                              ▼
                    ┌──────────────────────┐
                    │ Chromosome Decoding   │
                    └──────────────────────┘
                              │
                              ▼
                    ┌──────────────────────┐
                    │ Initial Population    │
                    └──────────────────────┘
                              │
                              ▼
                    ┌──────────────────────┐
                    │ Fitness Evaluation    │◄──────────┐
                    └──────────────────────┘           │
                              │                         │
                              ▼                         │
                          ◇ Stop ◇        Yes           │
                          ◇ Criteria ◇ ──────────┐      │
                              │                   │      │
                              ▼ No                │      │
                    ┌──────────────────────┐      │      │
                    │ Genetic Operators     │      │      │
                    │  ┌─────────────────┐  │      │      │
                    │  │ Selection       │  │      │      │
                    │  └─────────────────┘  │      │      │
                    │  ┌─────────────────┐  │      │      │
                    │  │ Crossover       │  │      │      │
                    │  └─────────────────┘  │      │      │
                    │  ┌─────────────────┐  │      │      │
                    │  │ Mutation        │  │      │      │
                    │  └─────────────────┘  │      │      │
                    └──────────────────────┘      │      │
                              │                    │      │
                              ▼                    │      │
                    ┌──────────────────────┐       │      │
                    │ New Population        │───────┼──────┘
                    └──────────────────────┘       │
                              │                     │
                              ▼                     │
                    ┌──────────────────┐            │
                   (  Selected Features )◄──────────┘
                    └──────────────────┘
```
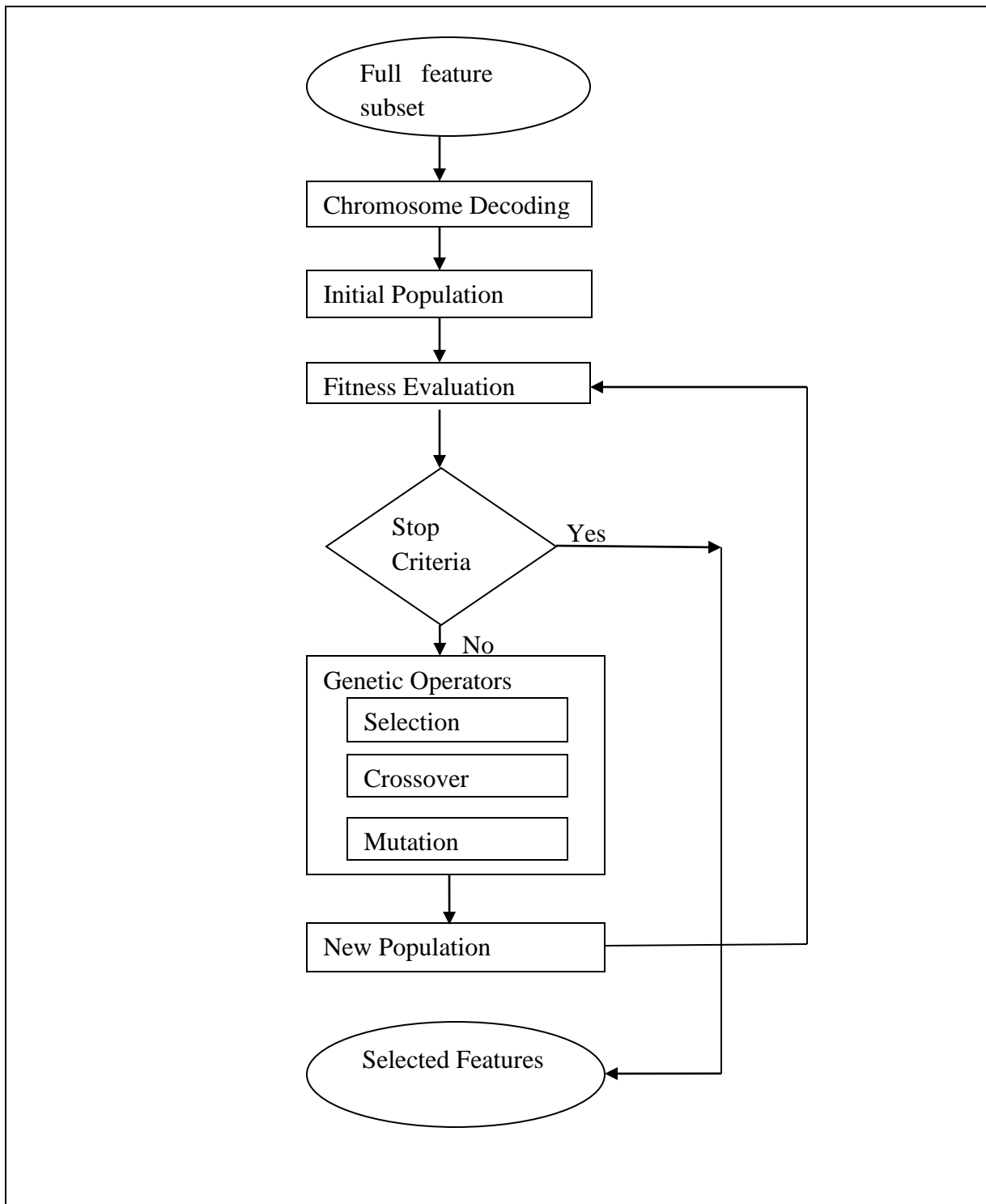
Figure 5.2 Feature selection process.

6 Chapter 6: Results and Analysis

The results for the proposed method is evaluated and determined for its performance and robustness. For evaluation of the results, two datasets are evaluated. One of them is the most commonly available dataset for splicing detection which is Columba Image Splicing Detection. It consists of 933 authentic and 912 spliced images with size of 128 x 128 pixels. Some example images of the dataset can be seen in Fig.6.1, of which the first row are authentic images and the second row are spliced images. Another dataset of 220 images named MICC − F220 is taken. The dataset MICC-F220consists of 220 images: 110 are tampered images and 110are originals. The image resolution varies from 722 x 480 to800 x 600 pixels. The forged images are obtained in the dataset by randomly selecting (both as location and as dimension) an image area (square or rectangular) and copy-pasting it over the image after having applied a number of different attacks such as translation, rotation, scale (symmetric/asymmetric) or a combination of them.

Figure 6.1 Images from Columbia Splicing Dataset

6.1 Performance Measures

Performance is evaluated by the Sensitivity, Specificity and Accuracy measures. Sensitivity measures the proportion of positives which are correctly identified as such and is complementary to the false negative rate. Specificity measures the proportion of negatives which are correctly identified as such and is complementary to the false positive rate.

$$Sensitivity = \frac{TP}{TP+FN}$$ 6.1

$$Specificty = \frac{TN}{TN+FP}$$ 6.2

$$Accuracy = \frac{TN+TP}{TN+TP+FN+FP}$$ 6.3

Where, TP (true positive) is the number of forged key-points that were detected as forged, TN (true negative) is the number of authentic key-points that were detected as authentic, FP(false positive) is the number of authentic key-points detected as forge, and FN(false negative) is the number of forge key-points detected as authentic.

6.2 Results

6.2.1 Data set : MICC F220.

The following table shows the parameters obtained on MICC F220 dataset when the proposed method is applied to it.

Image name	Accuracy	Sensitivity	Specificity	Average
Tampered 1	93.33	100	85.71	93.01
Tampered 2	92.83	100	85.68	92.83
Tampered 3	91.87	100	85.79	92.55
Tampered 4	92.43	100	86.09	92.84

Table 6.1 Performance parameters on MICC F220 using SVM classifier.

When Twinsvm is used as classifier the parameters obtain the values as displayed below:

Image name	Accuracy	Sensitivity	Specificity	Average
Tampered 1	93.33	100	85.71	93.01
Tampered 2	92.83	100	85.68	92.83
Tampered 3	91.87	100	85.79	92.55
Tampered 4	92.43	100	86.09	92.84

Table 6.2 Performance parameter on MICC F220 using Twinsvm as classifier.

The set of ten features got the values as:

Features Set	Accuracy	Sensitivity	Specificity
10 feature set	93.33	100	85.71

Table 6.3 Performance of entire feature set on MICC F220.

Best Chromosomes Selected by GA	Accuracy
HOG,LBP,CON	93.33
HOG,CON,IDM	93.33
LBP.SURF,CON	92.45
LBP, CS,INER	93.33
SURF,ENT,CS	92.84
HOG,LBP,CON,CS	93.02
HOG,LBP,CON,CP	92.65
HOG,SURF,INER,LBP	91.87
IDM,SURF,CS,ENT	93.01
IDM,ENT,LBP,CON	91.67

Table 6.4 Best Chromosomes selected by GA on MICC F220.

6.2.2 Data Set: Columbia Dataset

When the proposed algorithm is applied on Columbia Dataset of image splicing ,the parameters attain these values:

Image name	Accuracy	Sensitivity	Specificity	Average
Tampered 1	88.63	86.95	90.47	88.68
Tampered 2	78.64	84.76	89.76	84.38
Tampered 3	86.57	89.47	88.65	88.36
Tampered 4	88.63	86.95	90.47	88.68
Tampered 5	86.10	85.76	89.56	87.14
Tampered 6	88.65	86.57	90.34	88.52
Tampered 7	86.57	90.47	88.65	88.56
Tampered 8	88.67	87.45	88.34	88.15
Tampered 9	88.67	87.56	88.45	88.22
Tampered 10	85.67	86.76	89.23	87.06
Tampered 11	88.63	86.95	90.47	88.68

Table 6.5 Performance parameters on Columbia Dataset using SVM classifier.

The analysis when done again on the Columbia dataset using the Twinsvm Classifier, attain these values:

Image name	Accuracy	Sensitivity	Specificity	Average
Tampered 1	88.63	86.75	90.47	88.68
Tampered 2	78.64	84.76	89.76	84.38
Tampered 3	86.57	90.47	88.65	88.56
Tampered 4	88.63	86.75	90.45	88.68
Tampered 5	86.10	85.76.	89.56	89.14
Tampered 6	88.65	86.57	90.34	88.52
Tampered 7	86.57	90.47	88.65	88.56
Tampered 8	88.67	87.45	88.34	88.15

Tampered 9	88.67	87.56	88.45	88.22
Tampered 10	85.67	86.76	89.23	87.06
Tampered 11	88.63	86.95	90.47	88.68

Table 6.6 Performance parameter on Columbia using Twinsvm as classifier

The set of ten features when applied on proposed approach:

Features Set	Accuracy	Sensitivity	Specificity
10 feature set	88.63	86.75	90.47

Table 6.7 Performance of entire feature set on Columbia dataset.

The best chromosomes selected by the genetic algorithm:

Best Chromosomes Selected by GA	Accuracy
HOG, INER, CON	88.65
LBP, HOG, CS	88.67
HOG,INER,LBP	89.56
HOG, CS, CON	85.64
SURF,INER,CON	86.78
SURF,INER,CON,ENT	88.97
LBP,IDM,CON,HOG	86.45
LBP,SURF,CON,CS	85.64
HOG,LBP,CON,CS	85.78
ENT,SURF,IDM,CS	86.76

Table 6.8 Best Chromosomes selected by GA on Columbia dataset.

6.3   Comparative Study and Result Analysis

The same set of feature set belonging to MICC F220 when applied to classical feature selection like sequential forward selection (SFS), sequential backward selection (SFS) and Sequential floating forward selection (SFFS) like methods gave the following feature subset with the accuracy as shown below:

Feature Selection Algorithm	Feature Set	Accuracy
SFS	HOG,INER,CS,LBP	93.33
SBS	HOG, INER ,SURF,CS	93.33
SFFS	HOG,INER,LBP,CS	93.33

Table 6.9 Performance of Sequential Feature Selection algorithms on MICC F220.

Similar to the previous one, the feature set from Columbia dataset gave following results:

Feature Selection Algorithm	Feature Set	Accuracy
SFS	HOG,CON,LBP ,INER	88.63
SBS	LBP,IDM,CS,SURF	88.63
SFFS	LBP, CON,CS,HOG	88.63

Table 6.10   Performance of Sequential Feature Selection algorithms on Columbia.

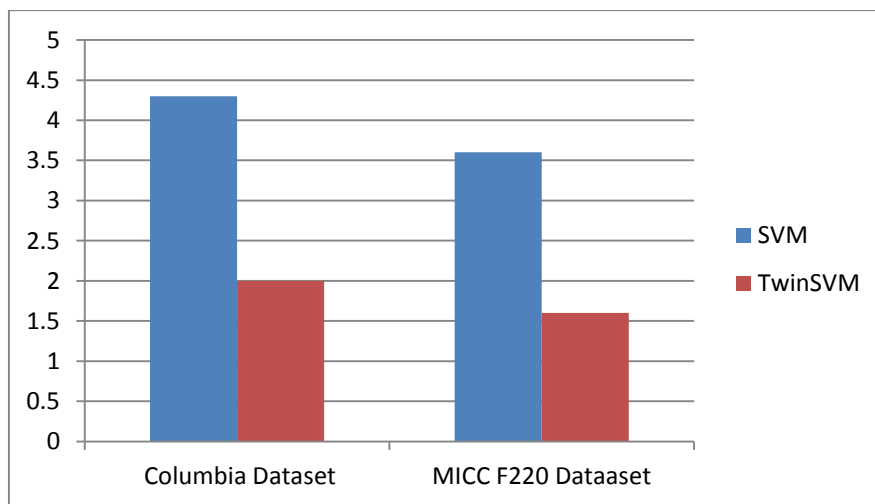The classification time of the algorithm using SVM and TWINSVM as classifier can be compared with the following graph.

Table 6.11 Classification time of SVM and TWINSVM classifier

7 Chapter 7: Conclusion and Future Work

7.1 Conclusion

Image splicing is the common forgery in digital images. A significant amount of research has been done in this field and thus lots of methods exist. However a good approach must have high accuracy and low computational complexity. Also a good approach should be robust to various kinds of manipulations such as rotation, scaling, compression etc.

In this thesis such type of a system is designed with the help of feature selection algorithm to reduce the dimensionality of the feature set and get more optimal results. As we know that different features of the image is used to detect different types of forgery accurately and efficiently. So a methodology is developed to collect the best of all and define a more strong system. To achieve this, set of ten features are selected and from them the best of features are selected for classification to get the final class results and the performance parameters are accuracy, sensitivity and specificity. Better results are being obtained after the proposed methodology as compared to the traditional methodology.

Advancement is done in the above mentioned method and that is feature selection. As we know that complexity of the system increases with the dimensionality increment and even sometimes it leads to over training of the classifier. Among all the features extracted it may be possible that only few are contributing towards the better results and rest are either overtraining the system or just overlapping the results already obtained. In such a case to obtain the best feature subset is very much important task. This task is performed with the help of feature selection method. A fitness function is used for this process of feature selection. The output that is the fitness value is given as a feedback to the classification system. The feature subset corresponding to the highest fitness value is considered for the performance parameter estimation.

7.2 Future Work

There are certain challenges that need to be addressed. The first being the computational complexity. As the dimension of the feature set is large, computational time increases. Though in our approach we tried to reduce the complexity by using twinsvm but it can be enhanced to work more fine.

Another issue to our method is its robustness. Although the method runs on forged images significantly better but improving its performance, detection and accuracy rate is an major challenge faced by any splicing algorithm.

As the technology is growing exponentially and so do the sophistication of the image forensics. It has become growing challenge to detect these forgeries. Hence the field of forgery detection needs exhaustive study and research.

8 References

[1] C Rey and J-L Dugelay, "A survey of watermarking algorithms for image authentication.," *EURASIP Journal on Applied Signal Processing*, pp. 613-621, 2002.

[2] V M Potdar, S Han, and E Chang, "A survey of digital image watermarking," in *3rd IEEE International Conference on Industrial Informatics*, Perth,WesternAustralia, 2005, pp. 709-716.

[3] Granty Regina Elwin J, Aditya T S, and Madhu Shankar S, "Survey on Passive Methods of Image Tampering Detection," in *Proceedings of the International Conference on Communication and Computational Intelligence*, 2010, pp. 431-436.

[4] B Mahdian and S Saic, "A bibliography on blind methods for identifying image forgery," in *Signal Processing:Image Communication*, 2010, pp. 389-399.

[5] M Ali Qureshi and M Deriche, "A Review on Copy Move Image Forgery Detection Techniques," in *IEEE*, 2014.

[6] T T Ng and S F Chang, "Blind Detection of Digital Photomontage using Higher," ColumbiaUniversity, 2004.

[7] T T Ng and S F Chang, "A model for image splicing," in *IEEE International Conference on Image Processing*, Singapore, 2004, pp. 1169–1172.

[8] D Fu, Y Q Shi, and W Su, "Detection of image splicing based on Hilbert–Huang transform and moments of characteristic functions with wavelet decomposition," *Lecture Notes in Computer*, vol. 4283, pp. 177–187, 2006.

[9] Y F Hsu and F S Chang, "Detecting image splicing using geometry invariants and camera characteristics consistency," in *IEEE ICME*, Toronto,Canada, 2006, pp. 549–552.

[10] F Y Hsu and S F Chang, "Image splicing detection using camera response function consistency and automatic segmentation," in *IEEE ICME*, Beijing, China, 2007, pp. 28–31.

[11] Q Zheng, W Sun, and W Lu, "Digital spliced image forensics based on edge blur measurement," in *IEEE International Conference on Information Theory and Information Security*, Beijing,China, 2010, pp. 399–402.

[12] Y Sutcu, B Coskun, H T Sencar, and N Memon, "Tamper detection based on regularity of wavelet transform coefficients," in *IEEE International Conference on Image Processing*, SanAntonio,TX,USA, 2007, pp. 397–400.

[13] Z Lin, J He, X Tang, and C K Tang, "Fast automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," in *PatternRecognition*, 2009, pp. 2492–2501.

[14] J Dong, W Wang, T Tan, and Y Q Shi, "Run-length and edge statistics based approach for image splicing detection," *Lecture Notes in Computer Science*, pp. 76–87, 2009.

[15] Z He, W Sun, W Lu, and H Lu, "Digital image splicing detection based on approximate run length," in *Pattern Recognition Letters*, 2011, pp. 1591–1597.

[16] W Chen, Y Q Sh, and W Su, "Image splicing detection using 2-D phase congruency and statistical moments of characteristic function," in *Imaging Security, Steganography,andWatermarking of Multimedia Contents*, 2007.

[17] T T Ng and S F Chang, "A DataSet of Authentic and Spliced Image Blocks," Columbia University, Technical Report 2004.

[18] Y Q Shi, C Chen, and W Chen, "A natural image model approach to splicing detection," in *MM & Sec*, Dallas,TX,USA, 2007, pp. 51–62.

[19] Zhongwei He, Wei Lu, Wei Sun, and Jiwu Huang, "Digital image splicing detection based on Markov features in DCT and DWT domain," *Pattern Recognition*, pp. 4292–4299, 2012.

[20] Amani A Alahmadil, Muhammad Hussain, Hatim Aboalsamh, Ghulam Muhammad, and George Bebis, "Splicing Image Forgery Detection Based on DCT and Local Binary Pattern," in *IEEE*, 2013, pp. 253-256.

[21] Ramesh Chand Pandey, Sanjay Kumar Singh, K K Shukla, and Rishabh Agrawal, "Fast and Robust Passive Copy-Move Forgery Detection Using SURF and SIFT Image Features," in *IEEE*.

[22] Navneet Dalal and Bill Triggs, "Histograms of Oriented Gradients for Human Detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2005.

[23] Manoj Kumar, Mohammad Husian, Naveen Upreti, and Deepti Gupta, "GENETIC ALGORITHM: REVIEWAND APPLICATION," *International Journal of Information Technology and Knowledge Management*, vol. 2, pp. 451-454, July-December 2010.

[24] Jayadeva , R Khemchandani, and Suresh Chandra, "Twin Support Vector Machines for Pattern Classification," *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE,* vol. 29, pp. 905-910, MAY 2007.

[25] P Pudil, J Novovicova, and J Kittler, "Floating search methods in feature selection," *Pattern Recognition Letters*, pp. 1119-1125, Nov 1994.