MAJOR PROJECT REPORT

ON

FACE DETECTION UNDER OCCLUSION

Submitted for the Partial Fulfillment of the Degree

MASTER OF TECHNOLOGY

IN

SIGNAL PROCESSING AND DIGITAL DESIGN

BY

AJIT EMMANUEL
2K15/SPD/04

Under the Guidance of

Sh. RAJESH ROHILLA

DEPARTMENT O F ELECTRONICS AND COMMUNICA TION ENGINEERING
DELHI TECHN OLOGICAL UNIVERSITY, DELHI-110042 (SESSION 2015-2017)

CERTIFICATE

This is to certify that the thesis entitled “Face Detection Under Occlusion” being submitted by Ajit Emmanuel, 2K15/SPD/04 for partial fulfillment of the degree “Master of Technology” in “Signal Processing and Digital Design” from Delhi Technological University, is based on work carried out by Ajit Emmanuel under my guidance and supervision. The matter contained in this thesis has not been submitted elsewhere for award of any other degree to the best of my knowledge and belief.

Sh. Rajesh Rohilla
Associate Professor
Deptt.of E&C Engg.
Delhi Technological University
Delhi-110042

ACKNOWLEDGEMENT

I would like to express my heartily gratitude and thanks to my project guide Prof. Rajesh Rohilla, Associate Professor in Department of Electronics and Communication Engineering, Delhi Technological University, for continuous inspiration, encouragement and guidance in every stage of preparation of this thesis work.

I am also extremely thankful to Dr. S. Indu, Head of the Department of Electronics and Communication Engineering, Delhi Technological University, for the support provided by her during the entire duration of degree course and especially in this thesis.

I would like to thank my batch mates for their support throughout the entire duration of the degree.

Ajit Emmanuel
2K15/SPD/04
Department of Electronics and
Communication Engineering
Delhi Technological University

ABSTRACT

In this thesis we are concentrating on the surveillance application which can be used to identify no of faces in a surveillance area. One of the main area is finding the attendance in the class and for marking the attendance. In this paper where the faces are occluded under varied occlusions can also be detected using the algorithm used in attaining the intended result.

This research represents a robust approach in obtaining the in- tended result. Here for this research I use the Speed up Robust Feature (SURF) algorithm to obtain the intended result. For the specific surveillance application we should have the database of the individuals in a screen. Then the images of persons stored in the database is evaluated against the current surveillance image for personal identification. A thorough search is done by the proposed SURF algorithm to obtain the output.

Features of both the images in the database and the surveillance image is extracted using the said algorithm and the features are matched to identify the matched features, from this matched features depending on the threshold value the output can be obtained. The result shows higher accuracy compared to other methods such as Voila Jones, SIFT, etc.

Contents

List of Figures

CHAPTER 1

Introduction

1.1 Biometrics

Biometrics is mainly related to human characteristics. Biometrics is used in computer authentication in order to identify a person for accesses of control of any surveillance or any place where a large group of people is working where authenticity of the person is important.

Biometric identifiers are very much distinctive where their characteristics are used to label and describe persons in particular. Biometric identifiers are of two types 1.Behavioral characteristics 2.Physiological characteristics. Behavioral characteristics are described often to the particular pattern of behavior or attitude of an individual but are not restricted to the voice or gait or any rhythm. Many of the researchers have termed behavioumetrics to represent the latter class of biometrics. Physiological characteristics mainly points to the body shape. Examples of which mainly points to face recognition, palm print, DNA, palm veins, iris recognition, retina and are not limited to fingerprint.

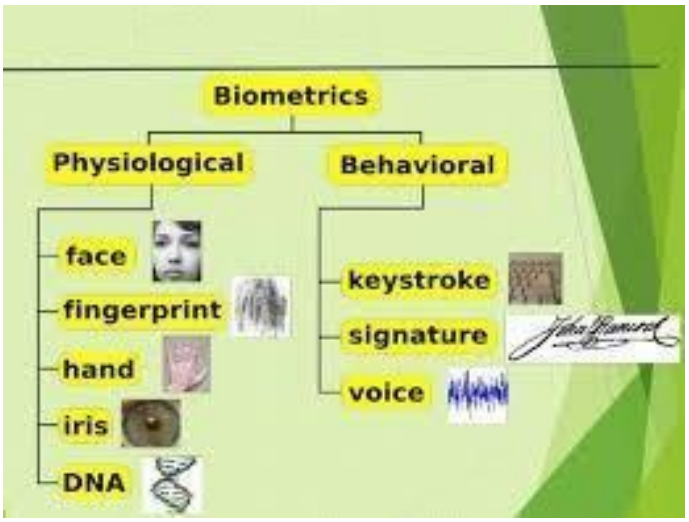


Figure 1.1: Types of Biometric System

Figure 1.2: Types of Biometric System

Most of the many traditional ways of access of control contains identifications which are based on token systems such as passport, driver's license and knowledge based systems like password or any personal number used for identification. These biometric identifiers are very unique to certain particular persons that are very much reliable to verifying the identity of the particular person than token based systems. However the use of biometrics raises the concerns of privacy and the use of information.

1.1.1 Biometric Functionality

The different aspects of human behavior chemistry and physiology is used in this biometric authentication system. The use and selectivity of a rather more particular biometric for a specific application involves taking weights of many several factors. There are said to be seven such factors

Universality–It refers to the fact that each and every particular person using a particular system that should possess a particular trait.

Uniqueness–It refers that the trait of a particular person should be used and be sufficiently dissimilar for particular individuals in a small admissible population so that they can be eminent between people.

Permanence–It identifies with the way in which a specific characteristic can differ after some time. Particularly a characteristic with great execution will be sensibly invariant with time regarding the particular coordinating algorithm.

Measurability–It identify with procurement or the estimation of trait. Specifically the information obtained ought to be in the passable allow with ensuing preparing and extraction of the pertinent sets of feature.

Performance–It identifies with the speed, strength, speed and exactness of the innovation utilized.

Acceptability–It identifies with how people in the important populace that acknowledges the technology utilized to such an ex- tent that they are ready to the biometric quality caught and surveyed.

Circumvention–It relates without hardly lifting a trait with which an attribute may be utilizing an artifact or substitute.

Proper biometric is very application dependent. Certain biometrics will be superior to anything others in view of the levels of com- fort and security. No specific single biometric will meet the prerequisites of each conceivable applications.

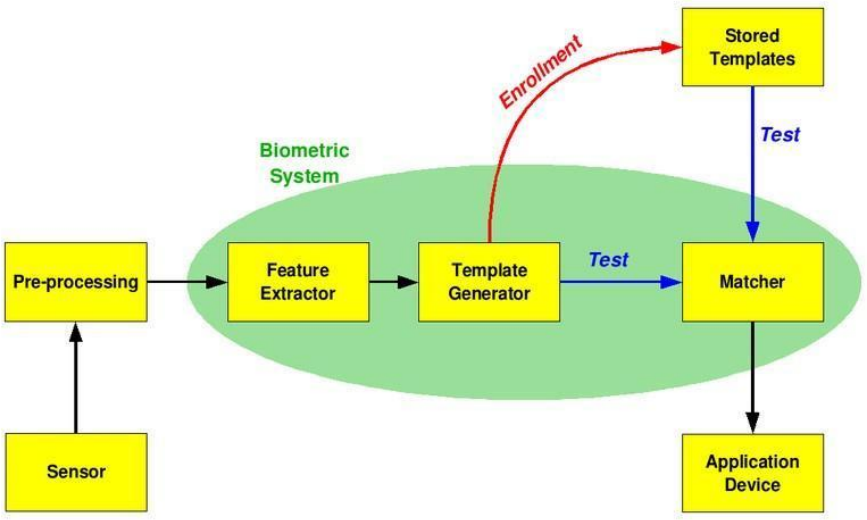


Figure 1.3: Block Diagram of Biometric System

The block diagram represents essential modes in biometric frame- work. Initially in the confirmation procedure the method of the framework plays out a balanced of a caught biometric with a layout put away in database to check the individual they claim to be. Three stages are required in the check of a specific individual. The initial step implies reference models for all the specific clients can be produced and put away in the database display. In the second step the examples are prepared to be coordinated with the reference model to make the real and faker and the edge is ascertained. Test- ing phase is the third step. The process involves the use of personal identification number or a particular username in which a particular template is used for comparison. Verification should be taken care in such a way such to prevent multiple use of same identity.

Secondly the recognizable proof mode in which the frameworks play out an examination with one to numerous against a biomet- ric database to build up a character of an obscure individual. The framework will just prevail with regards to recognizing the specific individual in contrast with specific biometric test to the layout which is put away in the database which falls inside the limit of the dataset. ID mode is utilized for giving constructive or contrary acknowledgment of the individual where the set up framework in which the specific individual denies to be. The last capacity can be diminished through biometrics since strategies utilized PINs and passwords are inadequate.

The first time a particular individual is using a biometric system is considered as enrollment. During the time of enrollment any information about biometric system or of an individual is stored in database. In the later uses the information in the biometric is detected and compared with the stored information during the time of enrollment. It is crucial in storage and retrieve of systems themselves for the robustness of the biometric system.

The first block is the sensor is the interface part between the system and the outside world. It needs to collect the necessary data. In most times to get the data of the image in the acquisition system but in can change according to the image characteristics. Second block is used for performing all the necessary preprocessing so that it can remove artifacts

to boost the input data for example removal of noise for the process of normalization. In third block features necessary for data are extracted which is the most important step in such a way that the correct features are extracted in the way it is optimized. The vector of a particular image is used to make a template. The template is used in the relevant characteristics of the source. The elements which are not used in the algorithm for comparison are discarded to decrease and protect the identity of the person.

1.1.2 Multimodal Biometric System

It is mainly used to overcome unimodal biometric system limitations. For example in the iris recognition system which consists of aging iris. And where scanning system such as finger scanning may be worn out or prints may be cut. The unimodal biometric system is limited by integrity of the identity of the particular person. It is very rare that where several unimodal systems have to suffer from limitations of identity. This is case where multimodal biometric systems is used to obtain data sets from the same individual or information obtained from different biometrics.

The unimodal biometric system can used stacked sequentially in such a way that a multimodal biometric system can be formed. The multimodal biometric system can be fused in other methods such as parallel hierarchical and serial modes of integration. The fusion of such biometric information occurs in the different stages in the recognition system. In different levels of feature fusion the data itself or are extracted from a recognition system. The matching of different level fusion finally results of multiple classifiers which are combines by techniques pertaining to different modes.

The last process includes distinctive levels of combination in the last consequences of different classifiers which are consolidated by systems in a dominant part voting. Combination of various levels of combination is accepted to be more as a result than alternate levels of combination in which the arrangement of highlight contains data which is rich about the biometric information in the coordinating set for choosing the yield in the classifier. Along these lines highlight level combination can give better outcomes.

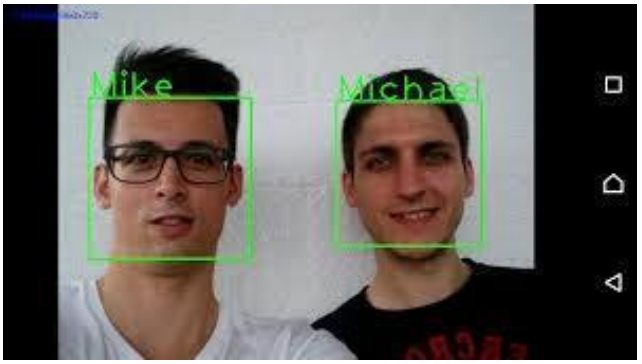


Figure 1.4: Face Recognition

1.2 Face Recognition

The structure of face is a complex multidimensional structure and it needs high computational methods for recognition. One of our primary focus of attention in our personal and social life lies in identifying the role and identity of individual. We have the ability to recognize the number of faces recognized and learned through our span of lives and identify the faces at a glance even after years. The variations in face is due to distractions like beards, hair style change, glasses and mainly aging.

The recognition of face is an important and integral part of bio- metric face recognition. In these basic biometric traits of human the traits are matched to the existing data and result depend on the identification on the particular person who is traced. The features of the face is extracted and is used in the implementation through different algorithms which are efficient and the modifications are done to improve the existing algorithm models.

Frameworks and PCs that perceive face can be connected with an assortment of numerous handy applications which can incorporate distinguishing proof of criminal, security frameworks and personality check. The identification and acknowledgment of face is utilized as a part of many places, for example, in facilitating pictures in various sites and in long range interpersonal communication locales. The innovations in software engineering is utilized to accomplished in confront identification and recognition.

The extracted features from the face is processed and is compared with the database where the similarity with the different face are processed with similarity in the database. From the database if a particular face is recognized it shows that the system shows the similar existing face in the database else the data is unknown. In the surveillance system in most common systems if an unknown dace repeatedly comes in a system where it is unknown when the first time an unknown face comes it stores the database for that particular face so that the next time the face is caught in the surveillance system it is stored so that it can recognized when it comes in system for further reasons. It is mainly used in identification of criminals. In particular the face recognition techniques are divided into two different groups based on appearance and hostile features for either whole or specific regions in the appearance for identification which are mainly based on geometrical features and the relationships between them.

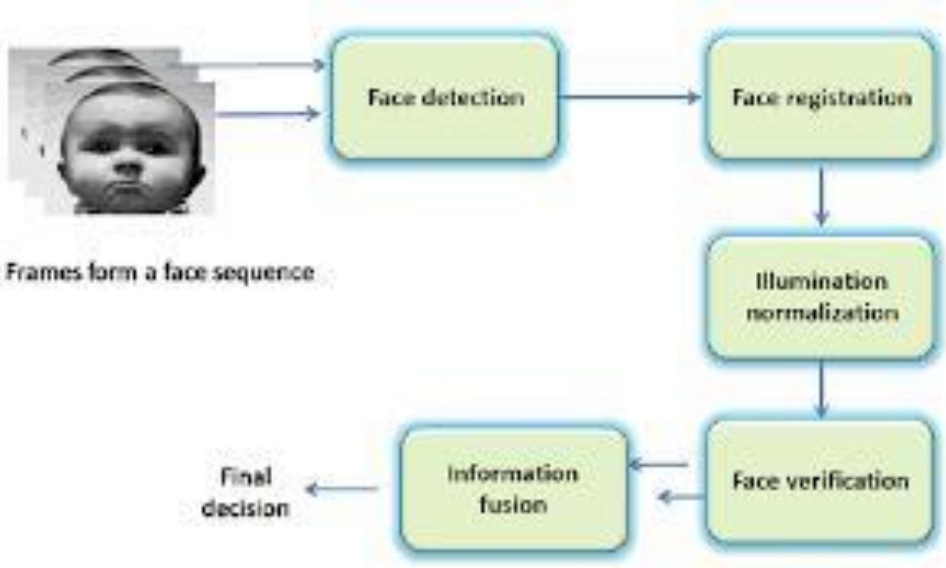


Figure 1.5: Face Recognition System

1.2.1 Face Recognition Systems

The applications which are mainly related to computer vision are for commercial and for law enforcement applications. Face match- ing, verification of the user and the control of access of any user or crowd surveillance or in enhanced computer interaction will be-

come easily possible in an effective face recognition system which is implemented. The research in computer vision starts from the early 1960s and other is only now that the results are acceptable such that the errors are minimized. However in this outside world face recognition is still an area of research which is very active and the works are still improving exponentially and still not a complete algorithm is not developed to detect a face with full efficiency and zero error.

The automated face recognition systems inadequacy is apparent and compared for our own ability of face recognition system. We hope in a computer generalized world a face recognition is extremely complex task to get instantaneous output to detect a face and the recognition ability is more robust now than computers can be to. By the introduction of computer vision the challenging situations in the current scenario such as angles or viewpoints and lightening conditions and scaling differences, the different background do not effect on the ability of computer vision to recognize particular individuals.

1.3 Challenges of Face Recognition

The research for decades have been going in face recognition but its still solved completely and the main problem related to it are image acquisition and the conditions in which images are obtained. The system which currently in place cannot handle all the challenges in the current scenario or the current world and none of the algorithms is not completely error free and the performance of the cur- rent system which recognize face is still very suspect to the sensitive and the variability in the facial areas. In particular the face recognition is set to represent the sets of challenges in scientific reasons. For a particular instance the no of more popular recognition methods particularly assume that there are more than one samples for every particular person available from the feature extraction during the training phase. In the real world different face recognition applications which include ID cards and law enhancement so this assumes that hold because of the single sample in a per person recorded in these systems.

Extricating the discriminant and strong components which in- corporate the characteristics of individual conservative and in developing the edge between various

people which is vital and basic and the troublesome issue in the face acknowledgment field. Besides in this we can coordinate or surpass the normal execution of a human face in this present reality in confront acknowledgment targets. Our primary goals is to make an ideal model on the face in a modest number, furthermore on how we can accomplish the vigor level shown for the human face acknowledgment, thirdly on making the impacts of maturing in the appearances and these destinations can be in the always expanding interest of the innovation.

Figure 1.6: Illumination Problem

1.3.1 Image Quality

Image quality-The fundamental and essential necessity of any face acknowledgment framework is the suspects, nature of face pictures and a decent quality picture is the one which is under quality conditions. For deciding the picture includes the nature of picture is essential and without the accumulation calculations of facial elements the strong approach will be lost. Therefore even the best algorithm will be of no use if the quality of image declines.

1.3.2 Illumination Problem

There occurs a situation where some faces appears differently due to the lighting change. We know that the change in lighting can change the appearance of the object drastically.

What we have to overcome is the lighting which is irregular. This is where the image processing tool is used to provide a method to avoid lighting invari- ance and methods used to avoid this is these methods Normalizing, Histogram equalization, Order-statistic filtering.

1.3.3 Approaches on Illumination Problem

The illumination problem can be approached by these methods.

Heuristic approach

It is an approach which depends on the symmetry of appearances on human. This calculation gives almost consummate precision for acknowledgment in frontal face picture under the distinctive lighting conditions.

Statistical approach

In this statistical approach each of the image in which features of image is represented in vector form in the dimensional space. So therefore our goal is to choose and post the statistical tool which is right used for the extraction and analysis for the underlying manifold.

Light modeling approach

Some of the recognition methods are to model a template which is lighting in order to stack the illumination algorithms are invariant.

Model based approach

The recent model which is approached is to try to build a 3D model. This gives us the idea to make intrinsic shape and texture which is fully independent from the parameters of light variations. The 3D model is used in building a model which is right fit for the input images. The directions of light and the shadows can be estimated automatically.

Multi spectral imaging approach

Using multi spectral images that can capture an image data at different specific wavelengths. We can separate the wavelengths using filters or any other particular wavelets. The separation of spectral information about illumination from other spectral information can be done by MSI.

1.3.4 Pose Variation

The training data usually used by face recognition systems are the front view images of particular images of individuals. front view of face images contains information of the particular face than other poses or angles of images. When the system happens to recognize a rotated face which is a problem which is tackled by frontal view training data. the user needs to have multiple views of the particular individual in a face d.

Figure 1.7: Pose Variation

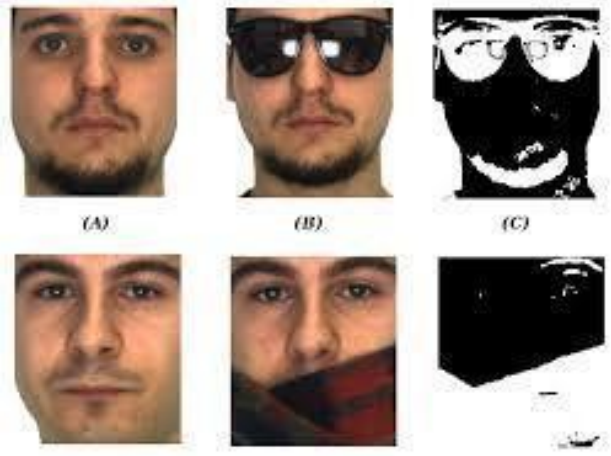


Figure 1.8: Different Occlusions in Face

1.4 Occlusion

Occlusions are mainly common in outside world and in one of the major obstacle in face recognition. When many images which are rich in texture can be easily detected under hard occlusions with many different local features such as SURF. Many manmade objects have very different non uniform regions. These objects which are texture less can be characterized by the structure of contour which are ambiguous and are without occlusions. The instant detection of the objects which are texture less objects have ambiguity and require recognition under different viewpoint and contains severe occlusions. While many researches have addressed each part of the recognition component separate.

CHAPTER 2

LITERATURE SURVEY

2.1 Introduction

The main objective and the task of extracting the differences be- tween the two images of a particular scene or an object is the part of many applications in computer vision applications. Reconstruction of image, calibration of camera, reconstruction of a 3D object and the recognition of object and a few like that. There are main steps in this. Firstly finding the interest points that are said to the distinctive locations in an object or an image. These interest points can be junctions, blobs, corners etc. the distinctive property of interest point detector is repeatability that is whether it can reliably find the same points of interest in different viewing conditions. Secondly the neighborhood of the interest points which is represented by the feature vector. This has to be distinctive and should have the property of robustness which is immune to noise, errors and the geometric differences. Thirdly which is matching which is done by matching of distinctive vectors to different images. The matching is usually done by Euclidean distance method. Preferably lower dimensions are considered for easiness.

Our main goal should be to develop the descriptor and detector which is very much in the state of the art and which is easy and fast to make computing calculations easier while the performance of the algorithm is not sacrificed. For succeeding our main aim should be to balance between the above said requirements which is reducing the dimensions of the descriptor and reducing the complexity and keeping it sufficiently distinctive. There are a wide variety descriptors and detectors are considered in this literature. While we are constructing the detector and descriptor.

When we are working with local features the main and initial issue that needs to be settled is the level of invariance required. This depends clearly on the geometric and the image metric deformations required which is expected and are determined to the changes in the conditions of viewing. One of the main focus is the rotation and invariance of the

descriptors and the detectors. One more thing we have to consider is the complexity of the feature and the robustness in the deformations that occur. Anisotropic scaling and perspective effects are said to be second order effects. They can be covered by the descriptor robustness.

Its claimed that the complexity additional which is full of affine features which should be invariant often have a negative impact in the working and its not bad unless a really large changes in the viewpoints are made. In some rare cases the invariance in rotation can be discarded which results in scale invariant version of any descriptor which we call upright SURF. In quite an applications the camera angle changes vey less one such application is the navigation of robot or a visual guide where the camera rotates only along the vertical axis. One such benefit of avoiding the absurd use of rotation invariance it can affect the speed and also the power which is discriminative. Photometric deformations are concerned in a simple linear model with particular scale factor and the offset required. Here in these descriptors we don't use color.

2.2 Point Detectors

The locator most likely utilized as a part of prior days is the Harris corner indicator which is proposed in 1988 which is presumably in light of the Eigen vales and the minute framework. Harris corner are fundamentally not scale invariant. The researcher Lindeberg presented the idea of programmed scale determination which permits to recognize the intrigue focuses in the picture with which each of their own qualities scale. He at that point does a try different things with the determinant of the hessian framework subterranean insect the laplacian to identify structures like blobs. Mikolajczyk and Schmid have discovered a technique which is strong and scale invariant which distinguish highlights with the repetability which is high and they have made Harris Laplace and Hessian Laplace. In this they utilized a measure of harris and the determinant of the Hessian framework to choose the area and the scale is chosen by Laplacian. Concentrating on speed he approximated the Laplacian of Gaussian by a Difference of Gaussians filter.

There are other scale invariant feature point detectors are pro- posed. The examples of the interest point detectors is proposed by Kadir and Brady in which where he maximizes the entropy within a particular region. From this we can infer that Hessian-based identifiers are probably going to be more steady and repeatable than their Harris-based partners. Furthermore, utilizing the determinant of the Hessian matrix instead of its follow appears to be exceptionally beneficial and in limited structures. And furthermore it approximations like the DoG can convey speed to minimal effort as far as lost exactness.

2.3 Feature Descriptors

A large variety of feature descriptors are proposed like derivatives of Gaussian, invariance moments, complex structures, stable filters and descriptors representing the distribution in the small scale features and the interest points. This is best explained in the fact that they can extract a substantial amount in the information about the spatial intensity patterns at the time being robust to the deformations and the localization errors.

The descriptor SIFT can compute a histogram of local oriented gradients around the interest points and it can store bins in a 128 dimensional vector. Eight orientations bins is used in each of the 4x4 location. Some refinements in this basic scheme is proposed in which a PCA is applied on the gradient image. A 36 dimensional descriptor is yielded by PCA SIFT which is fast for matching but which is proved to be less distinctive more than SIFT in a study and the matching is reduced by slower feature computation.

The descriptor which is SIFT still seems very appealing descriptor in practical purposes and they are used nowadays widely. Its relatively and distinctively fast which is very useful for online applications. We have also said SIFT on FPGA and we could see the very much improved speed by the order of magnitude. The high dimensionality in the descriptor is one of the drawback in the SIFT during the matching step. For particular online applications in a PC in each of the three steps which is basically detection, description, matching and should be faster.

The detector which we utilized as a part of this paper which is SURF depends on the Hessian framework however which utilizes an exceptionally essential estimation like a DoG in the Laplacian based locator. It primarily depends on how essential pictures are utilized to diminish the ideal opportunity for calculation so we call it a quick Hessian locator. The descriptor depicts a conveyance of Haar Wavelet reactions inside the intrigue point neighborhood. The speed is abused by fundamental pictures and just 64 measurements are utilized by diminishing the element of time calculation and the coordinating and are expanding the vigor is all the while expanded. Here we likewise display an ordering step which depends on the indication of Laplacian which nor just builds the coordinating velocity additionally in the descriptor strength.

Inorder to make the method more self proportionate we can dis- cuss the integral image concept. This allows fast implementation in the box made of convolution filters. The entry of the image can be represented as

$$I_{\Sigma}(\mathbf{x}) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j)$$

where I(x) at a location x=(x,y) represents the sum of all pixels in the input image I of a rectangular region which is formed by the point x and the origin.With I calculated now it only takes four additions to calculate the sum of the intensities over any upright, rectangular area, independent of its size.

2.4 Fast Hessian Detector

We based in detector in the Hessian matrix because of the performance which is good in accuracy and the time of computation time. However rather than using a measure for selecting the particular location and the scaling so we rely on the hessian determinant of both. Given a point x=(x,y) on the image in an image I, the Hessian matrix H(x, σ) in x at

$$\mathcal{H}(\mathbf{x}, \sigma) = \begin{bmatrix} L_{xx}(\mathbf{x}, \sigma) & L_{xy}(\mathbf{x}, \sigma) \\ L_{xy}(\mathbf{x}, \sigma) & L_{yy}(\mathbf{x}, \sigma) \end{bmatrix}$$

scale is defined as follows

where Lxx(x, σ) is the convolution of the Gaussian second order derivative, with the image I in point x, and similarly for Lxy(x, σ) and Lyy(x, σ).

2.5 SURF Descriptor

We know that SIFT gives a good performance which can now be compared to other descriptors and is remarkable. In this method we mix a crudely localized information and the gradient which is distributed which is related to features which seems to yield very good results while clouding the effects in the localization errors in terms of space and scale. By using comparative strengths and in the orientations of gradients which reduces the effect on the images metric changes. In this proposed SURF descriptor which is based on the similar properties and the complexity cut down even further.

The first step is supposed to be fixing and reproducing orientation which is based on the information from a circular region around some particular interest points. Then the next process is to construct a region of square which is aligned to the orientation selected and extracting the SURF descriptor point from the image. These two steps now are explained sequentially. Further we also explain an upright SURF which is invariant to rotation and more faster which is mainly suited for horizontal based applications.

2.5.1 Orientation Assignment

Inorder for the process to be invariant to rotation we can identify a reproducible sort of orientation for the interest points. For this we use Haar Wavelet responses both the directions of an image. From this in a neighborhood which is circular of a particular radius around the particular interest point with the corresponding scale at which an interest point is detected was detected is selected. And the next step is the sampling step which is scale dependent and is chosen to be a particular s. in keeping with the rest the wavelet responses are computed. And accordingly the high scales of the wavelets are big too. Therefore for this purpose we have to use integral images for the fast filtering. There are 6 operatoins needed to compute the response in direction of any scale.

Once these wavelet responses are to be calculated and weighted with a particular Gaussian value centered at the particular point so these responses are represented as vectors in this space with different horizontal responses and are represented as vectors in the particular space with the horizontal responses with strength along the horizontal and vertical response strength along the different ordinate. The orientation which is dominant and estimated by calculating the sum of all the particular responses within an orientation within covering an angle of 60.

The vertical and horizontal responses with where the window are added. The two added responses that can yield a new vector. And the orientation in which the vector lends to the interest point which is the longest vector. The sliding window which depends on the size in which a parameter which has been chosen experimentally. Small sizes fire on the single dominating different wavelets responses which yields maximum in vector length that are not taken. Both will result in an unstable orientation in the region of interest. In USURF it skip this step.

2.5.2 Descriptor Components

For the descriptor that is extracted. The step consists of constructing a region of a square centered around a particular interest point and it is oriented along the orientation

selected in the earlier para. For the particular version this is not a transformation which is necessary. The size of the window 20 times size of the dimension.

The region is divided into small regions of 4x4 small regions. This can keep the important information which is important. For each small region we can really compute the sample features in the 5x5 spaced points. For a reasons of simplicity we call the Haar wavelet response in the both directions. Horizontal and vertical here is de- fined in terms of the selected point orientation interest. To increase in the robustness to deformations geometry and the localization errors the responses are weighted at centered at the interest point.

CHAPTER 3

PROBLEM FORMULATION

Identifying the particular individual face in a particular environment relates to a difficult task. The most difficult task in the current scenario is that to detect face under occlusion. Occlusions can be any object or any natural or artificial object which causes disturbance in the actual particular object we have to detect. Face detection under occlusions can be a difficult and major task.

Occlusions can be any particular object they are wearing such as glasses, scarfs, shall, etc. It can be due to any artificial object obstructing the field of a face. It can be due to the side pose or the illumination problem mentioned earlier and the methods also have been discussed to overcome the problems. Occlusions may not be objects itself, it can be an another individual covering a face which is to be recognized particularly.

The main areas where the system is used nowadays is in the applications of recognizing an in a crowd where many faces are there, basically it refers to a surveillance system which is nowadays used in offices and areas where security is a main concern. It is mainly used in the military applications for identifying enemies.

The growth in the field is biometry is fast advancing where the technology is replaced in a quick manner when an advanced algorithm is developed. Here we have focus on a particular problem such that to identify the no of people in a particular room such that the people are under occlusions discussed earlier. And a solution is found to fight this particular problem. The objective and the method proposed is explained down further.

3.1 Proposed Work

In this proposed work we are implementing an attendance system which is basically to identify the no of persons in a crowded image which can also be occluded. Here basically the occlusion means one face of the student covering the other face of the student or any

other object in the class or any particular room. Different pose of an individual can also be considered an occlusion. So we can basically say we are implementing an attendance system or to identify or to find the no of persons which we know ie the faces stored in the database, in a particular surveillance area.

In this system of approach we have used one of the advanced methods in identification. The method which we used is the SPEED UP ROBUST FEATURES (SURF) method which gives us enough accuracy to obtain the intended output. The SURF method here we use is a more advanced method than SIFT but using SURF we can obtain the output faster with more efficiency.

3.2 Objective

Our main objective is to overcome the challenges we have mentioned earlier. To overcome this hazardous task we have used a suitable algorithm which fits perfectly for the proposed system we mentioned earlier. Our main objective is to identify the faces accurately and to find the number of faces in the area. One of the main factors which affects so determined purpose is the human occlusions ie when the faces of the persons are overlapped but by decreasing the threshold value that we use to obtain the output more faces are obtained in the output.

Despite the success of earlier methods still there are areas of problems which need to be addressed drastically. Here in this method I have tried to solve a particular problem of occlusion which is dis- cussed earlier. Among those prominent issues discussed the pose variation and illumination problem have been addressed drastically and it was one of the main objective, ie to improve the system so that the output is error free and have a higher efficiency compared to other methods.

In the second stage I describe the process in which the features of the search image and query image are extracted and are used for matching the features and are used to identify the face under occlusion which is shown in Figure 4.2

CHAPTER 4

METHODOLOGY

In the proposed system we mentioned in the abstract there are two parts which is discussed below. The first which is basically used to extract the features. The second which is used in the matching of the face.

In the first stage which is basically a feature extraction stage. Here we use an algorithm which is basically SURF to extract the features. Different steps involved in the extraction of the SURF features. The steps are shown in the block which follows and is shown in the block diagram in Figure 4.1

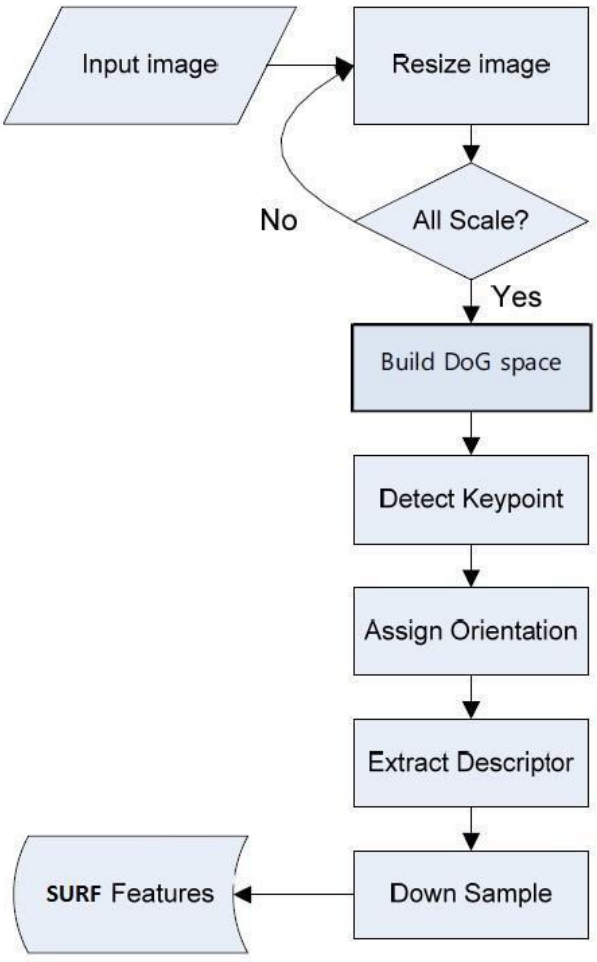


Figure 4.1: Block Diagram for feature extraction

4.1 Stage 1

Input Image

The input image can be any image from which the SURF features can be extracted. These image can be both the query image or the search image. Where the query image is the particular individual which is stored in the database and the search image is the image of the scene where the individual is identified.

Resize

Resize is one of the main steps which is important. Its is used in the process such that the dimension is made equal to all the query image.

Scaling

For changing the visual appearance of the particular image scaling is used so that to alter the quantity of stored information in the representation of the scene or like preprocessor in the image processing which is operated in the particular scale.

Build Gaussian

Building the Gaussian space for the extraction of the keypoints is oone of the primary step. To obtain the Hessian matrix the formation of Gausssian is very important. The hessian matrix is formed from the Laplace of Gaussian.

Detect Keypoint

SURF uses a square shaped filters for approximation of Gaussian smoothing. Filtering the image in a square is much more faster if the integral image can be used. The determinant of the Hessian is used for selecting the scale so that keypoints can be extracted.

Assign Orientation

For the image to be invariant to rotation what it does is it identifies an area of interesting points which is reproducible and we calculate the Haar wavelets responses in the horizontal and vertical directions. Once these wavelet responses are calculated and

weighted with a Gaussian which is centered at the interest points and these responses are represented in terms of vectors in the space.

Extract Descriptor

In the extraction of the descriptor the first step consists of a square region centered around the interest point and is oriented along the orientation selected. From these descriptors are extracted.

Down Sample

Down sampling or scaling which are both in this case synonymous with the down sampling of descriptors such that the dimensionality is reduced considerably without loosing much information about the particular object.

SURF Features

Once the descriptors are down sampled the features remaining constitutes the SURF features which is basically used in our algorithm for the purpose of face recognition.

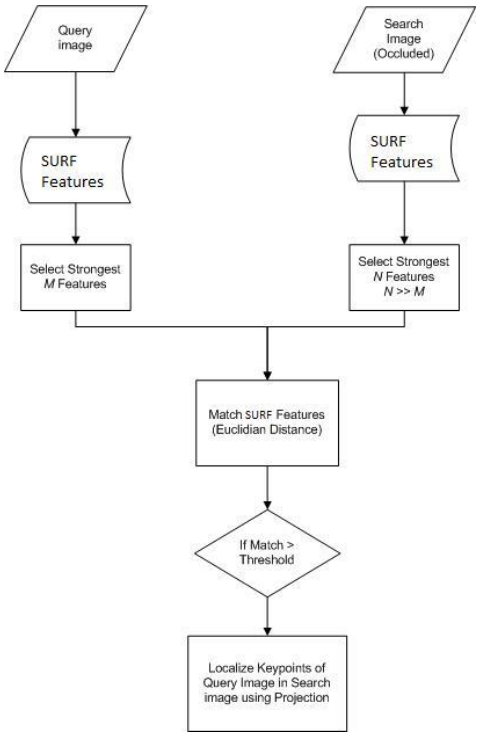


Figure 4.2: Block Diagram for face matching

4.2 Stage 2

Query Image

It is the image which is stored in database. It is the faces of different individuals that is stored in the database for identifying different persons in the scene. Query image can be any image which can be stored in the database for the purpose of identification. sup- pose in an office where the attendance is marked based on the face recognition system, the faces of all the employees are stored.

Search Image

It is the image in which where a particular individual is to be identified from a particular scene. The image referred may be occluded or be interfered by different hazardous objects. From this particular search image we are identifying the individual we need to identify.

SURF Features

Using the SURF feature algorithm we discussed in the stage.1 we extract the features of both the query image and search image. We use these features of both search image and query image for identification process which is to to identify descriptors for matching problem which in turn identifying the individual.

Strongest Features of Query Image

From the features we extracted using the SURF algorithm we select the strongest features which can be any particular number which depends on the problem and the applications we are using. In this problem we are left with so we choose strongest 100 points. Here we choose 100 strongest points because only a face is in the database so only less number of features are prominent so we choose less number compared to search image.

Strongest Features of Search Image

From the features we extracted using the SURF algorithm we select the strongest features which can be any particular number which depends on the problem and the applications we are using. In this problem we are left with so we choose strongest 300 points. Here we have to choose the strongest 300 features because the scene from which we are identifying an individual from a large scene which have more strong features compared to the query image.

Match SURF Features

From the extracted features of both the search image and query image where we have selected the strongest features from both the images. From the strongest features selected from both the search and query image we match the matching points or the SURF features extracted. The matched features will be less compared to the features obtained from the strongest features of query image. so these matched features is used to identify the particular person from the search image.

Threshold

Threshold value generally pertains to a particular value where we can start interpreting depending on a particular number. Here we use threshold value which is a count of the features(SURF) extracted and the strongest selected. The threshold value is number which is very much less than the strongest features extracted from the query image. And if the matching is greater than the threshold value we localize the keypoints.

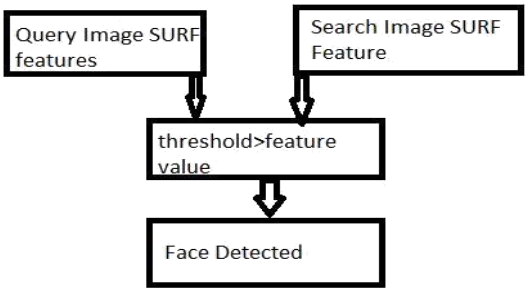


Figure 4.3: Block Diagram for detecting no of faces

Localizing Keypoints

From the features obtained after the matching of the features we localize the keypoints such that it is projected from the search image. These projected points are used for segmenting the face in the search image for identifying the face.

Stage 3

In this stage where keypoints after matched are thresholded to the feature values between the query image features and the search im- age feature. The number of faces detected depends on the threshold value given to the identified feature value.

CHAPTER 5

ANALYSIS AND RESULTS

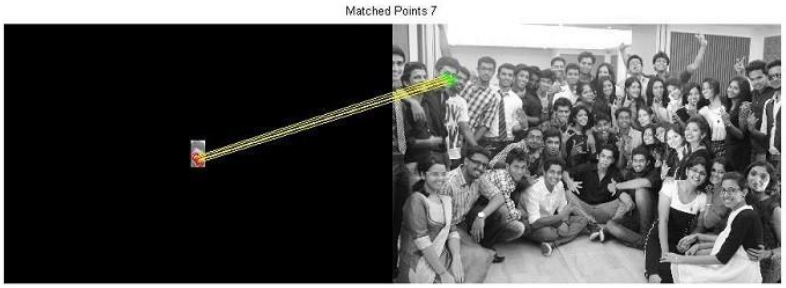
5.1 Scenario 1

Faces Stored in the Database

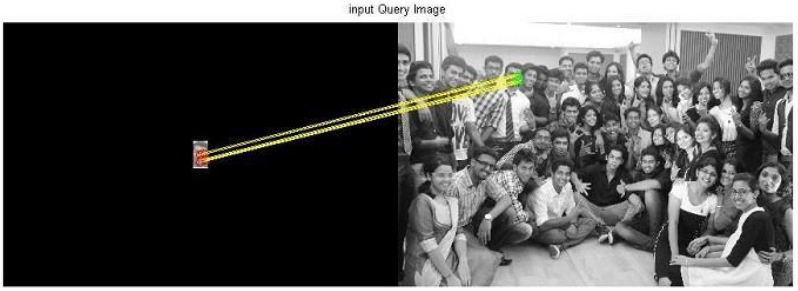
Search Image

Strongest features of Search Image

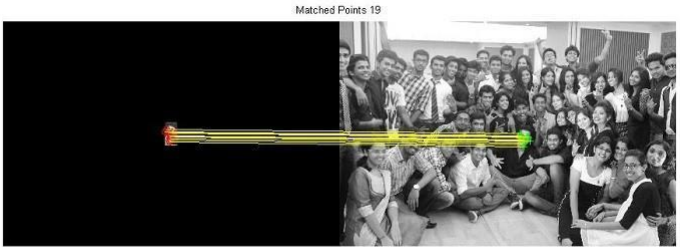
Feature matching with face 1



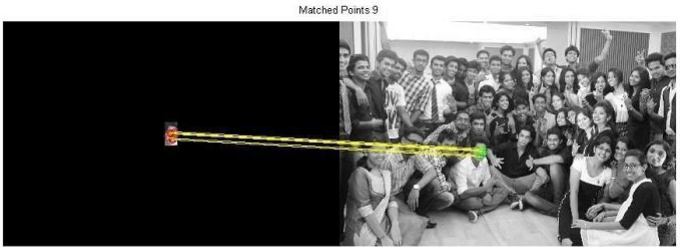
Feature matching with face 2



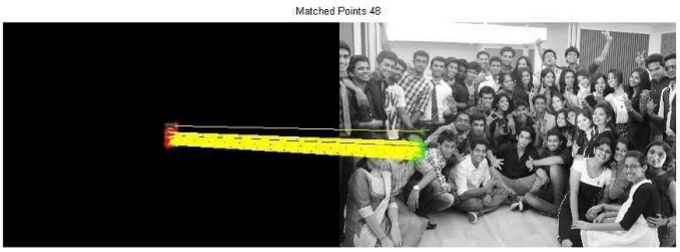
Feature matching with face 3



Feature matching with face 4



Feature matching with face 5



Threshold Value = 7

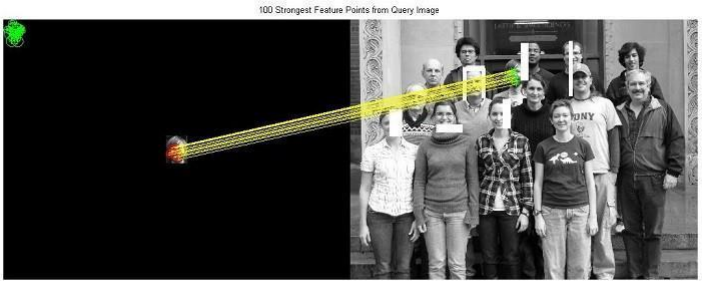
Number of faces matched = 5

5.2 Scenario 2

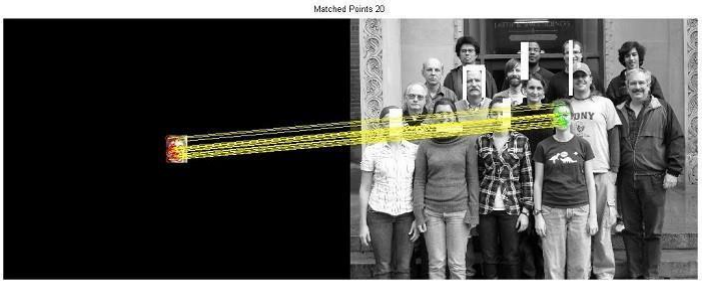
Faces Stored in the Database

Occluded Search Image

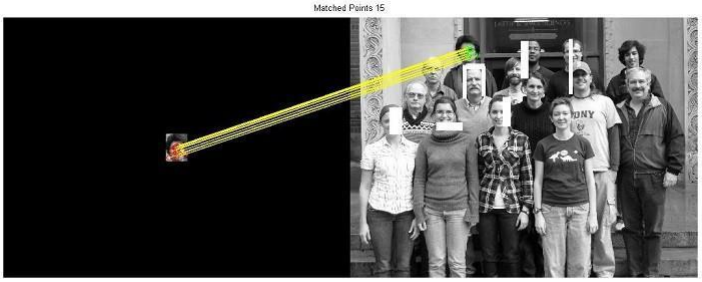
Feature matching with face 1



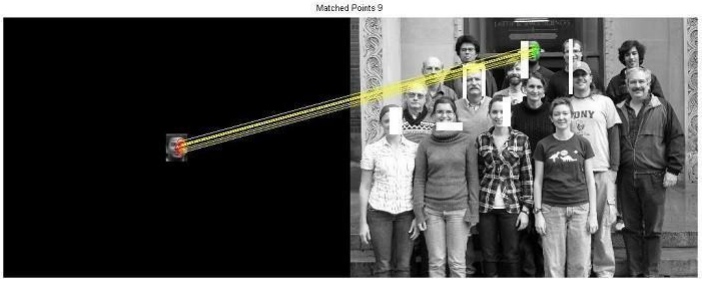
Feature matching with face 2



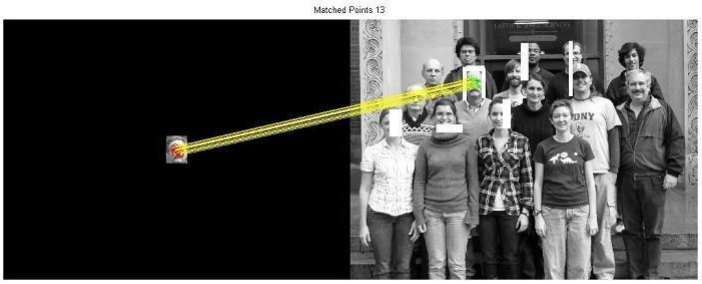
Feature matching with face 3



Feature matching with face 4



Feature matching with face 5



Threshold Value = 20

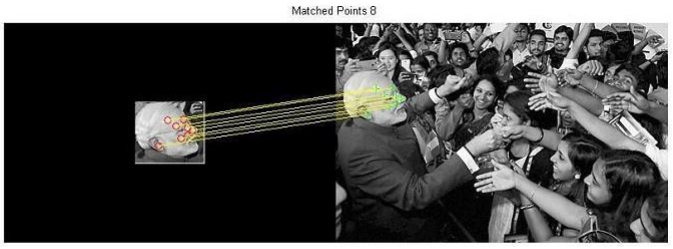
Number of faces matched = 3

5.3 Scenario 3

Face Stored in the Database

Search Image

Feature matching with face

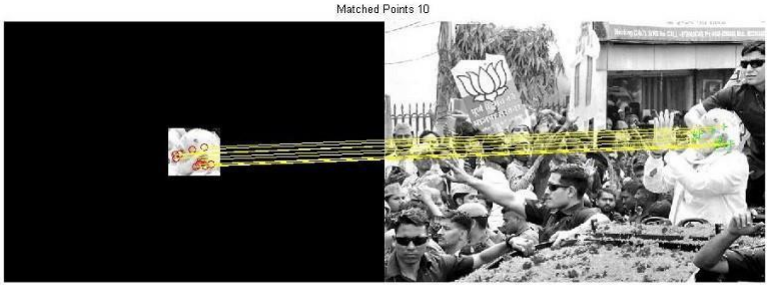


5.4 Scenario 4

Face Stored in the Database

Search Image

Feature matching with face

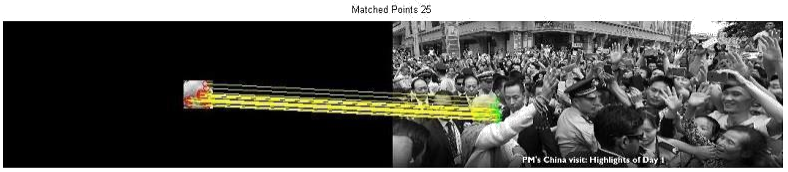


5.5 Scenario 5

Face Stored in the Database

Search Image

Feature matching with face


Matched Points 25
PM's China visit: Highlights of Day 1

CHAPTER 6

CONCLUSION

This thesis presents a robust approach in detecting face under occlusion. Which further can be used to identify faces in surveillance area and to take the count of individuals in the surveillance area. The method described in this thesis incorporates the SURF algorithm for identifying faces. Our main area of focus was the surveillance under occluded and challenging scenarios. Here we basically used the SURF algorithm to extract features of the corresponding images(ie the query image and the search image) and the features are respectively matched so that a particular individual can be recognized in the surveillance area.

The proposed method in this thesis gives a good detection rate of faces in the surveillance image under challenging and occluded circumstances. Experimental results indicate it gives a good efficiency than other methods like Voila Jones, SIFT, etc The prescribed methodology gives an efficiency around 98 percentage regardless of the occlusions and constrained scenarios like cluttered scenes pose variations,etc.

The proposed method is fast and it reduces mathematical computation. This novel method of face recognition is a significant contribution in recognizing faces. This method can be extended to real time applications in which indoor and outdoor surveillance is also taken into account

References

[1] Sonu Lamba, Neeta Nain, and Harendra Chahar, A Robust Multi Modal Approach for Face Detection in Crowd, in 2016 12th International Conference on Signal Image Technology and Internet Based Systems, Malviya National Institute of Technol- ogy.

[2] Rui Min and Jean-Luc Dugelay, Inpainting of Sparse Occlusion in Face Recognition EURECOM, Department of Multimedia Communications 2229 Route Crtes, 06560, Sophia Antipolis, France.

[3] Nese Alyuz, Berk Gokberk, and Lale Akarun, Senior Member, IEEE, 3-D Face Recognition Under Occlusion Using Masked Projection,IEEE Transactions on Information Forensics and Se- curity, VOL. 8, NO. 5, may 2013.

[4] Niall McLaughlin, Ji Ming, and Danny Crookes, Largest Match- ing Areas for Illumination and Occlusion Robust Face Recogni- tion IEEE Transactions on Cybernitics, VOL. 47, NO. 3, march 2017.

[5] Baofeng Zhang; Yingkui Jiao; Zhijun Ma; Yongchen Li; Jun- chao Zhu, An efficient image matching method using Speed Up Robust Features,2014 IEEE International Conference on Mechatronics and Automation.

[6] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, Speeded Up Robust Features, Katholieke Universiteit Leuven.

[7]. Jianbo Shi and C. Tomasi, "Good features to track," Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEE
E Computer Society Conference on, Seattle, WA, 1994, pp.593-600.

[8] B. Babenko, M. H. Yang and S. Belongie, "Visual tracking with online Multiple Instance Learning," Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, Miami, FL, 2009, pp. 983-990.

[9]. Jianbo Shi and C. Tomasi, "Good features to track," Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on, Seattle, WA, 1994, pp.593-600.

[10]T. Agui, Y. Kokubo, H. Nagashashi, and T. Nagao, “Extraction of FaceRecognition from Monochromatic Photographs Using NeuralNetworks,”Proc. Second Int’l Conf. Automation, Robotics, andComputer Vision, vol. 1, pp. 18.8.1-18.8.5, 1992.

[11]N. Ahuja, “A Transform for Multiscale Image Segmentation by Integrated Edge and Region Detection,”IEEE Trans. PatternAnalysis and Machine Intelligence,vol. 18, no. 9, pp. 1211-1235,Sept. 1996.

[12]Y. Amit, D. Geman, and B. Jedynak, “Efficient Focusing andFace Detection,”Face Recognition: From Theory to Applications,H. Wechsler, P.J. Phillips, V. Bruce, F. Fogelman-Soulie, andT.S. Huang, eds., vol. 163, pp. 124-156, 1998.

[13]T.W. Anderson,An Introduction to Multivariate Statistical Analysis.New York: John Wiley, 1984.

[14]M.F. Augusteijn and T.L. Skujca, “Identification of Human Facesthrough Texture-Based Feature Recognition and Neural NetworkTechnology,”
Proc. IEEE Conf. Neural Networks,pp. 392-398, 1993.