

SQUINT EYE DETECTION USING IMAGE PROCESSING

Submitted in partial fulfillment of the requirements for the award of
Bachelor of Engineering Degree in
Electronics and Communication Engineering

by

M.BANUMATHI(37130046)

B.BANUPRIYA(37130047)



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION
ENGINEERING**

**FACULTY OF ELECTRICAL
AND ELECTRONICS ENGINEERING**

SATHYABAMA UNIVERSITY
JEPPIAAR NAGAR, RAJIV GANDHI SALAI,
CHENNAI – 600119. TAMILNADU.

APRIL 2021



SATHYABAMA UNIVERSITY
(Established under Section 3 of UGC Act, 1956)
Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai- 600 119
www.sathyabamauniversity.ac.in



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION
ENGINEERING**

BONAFIDE CERTIFICATE

This is to certify that this Project Report is the bonafide work of **.M.BANUMATHI(37130046)** and **B.BANUPRIYA(37130047)** who carried out the project entitled “**SQUINT EYE DETECTION USING IMAGE PROCESSING**” under our supervision from November 2020 to April 2021.

Dr, T.VINO.M.E.,(Ph.D)
Internal Guide

Dr.T.RAVI .M.E.,(Ph.D).
Head of the Department

Submitted for Viva voce Examination held on _____

Name :

Signature:

INTERNAL EXAMINER

EXTERNAL EXAMINER

DECLARATION FORMAT

We, **M.BANUMATHI (37130046)** and **B.BANUPRIYA (37130047)** hereby declare that the Project Report entitled “**SQUINT EYE DETECTION USING IMAGE PROCESSING**” done by us under the guidance of **Dr.,T.VINO. M.E., (Ph.D).**, at Sathyabama Institue of Science and Technology is submitted in partial fulfillment of the requirements for the award of Bachelor of Engineering degree in **ELECTRONICS AND COMMUNICATION..**

1 .M.BANUMATHI (37130046)

2. B.BANUPRIYA (37130047)

DATE:

PLACE:

SIGNATURE OF THE CANDIDATES

ACKNOWLEDGEMENT

We are pleased to acknowledge our sincere thanks to Board of Management of SATHYABAMA for their kind encouragement in doing this project and for completing it successfully. I am grateful to them.

We convey our thanks to **Dr. N. M. NANDITHA, M.E., Ph.D.**, Dean, School of Electrical and Electronics and **Dr. T. RAVI, M.E., Ph.D.**, Head of the Department, Dept. of Electronics and Communication Engineering for providing me necessary support and details at the right time during the progressive reviews.

We would like to express our sincere and deep sense of gratitude to our Project Guide **Dr. T. VINO, M.E., Ph.D.**, for her valuable guidance, suggestions and constant encouragement paved way for the successful completion of my project work.

We wish to express our thanks to all Teaching and Non-teaching staff members of the Department of Electronics and Communication Engineering who were helpful in many ways for the completion of the project.

ABSTRACT

This project goal is to detect the squint, which means the improper alignment of eyes. Eye is unable to intend the object because one of eye looks at the object but the other somewhere else like in or out, or turn up or turn down. So, the brain is unable to fuse the image from the both eyes in order to produce binocular vision or 3D. As a result, reduction in concentration power and loss of vision occurs. Over years, many methods are developed for the detection of squint eye but they are cost effective. In this work hough transform is used to extract the features from the face to determine squint eye. Two types of phases are used for extraction. phase 1: The detection of face, iris, nose edge points. phase 2: The distance between the nose edge point to the two eye centers are calculated for various positions of iris. Based on the difference between these distance one can determine the squint.

TABLE OF CONTENTS

CHAPTER No.	TITLE	PAGE No
	ABSTRACT	v
	LIST OF ABBREVIATIONS	viii
	LIST OF FIGURES	ix
	LIST OF TABLES	x
1	INTRODUCTION	1
	1.1 TYPES OF SQUINT	3
	1.1.1 ESOTROPIA	3
	1.1.2 EXOTROPIA	3
	1.1.3 HYPOTROPIA	3
	1.1.4 HYPERTROPIA	4
	1.2 CAUSES OF SQUINT	4
	1.3 LITERATURE REVIEW	5
	1.4 PROBLEM DEFINITION	10
	1.5 OBJECTIVE	11
2.	TECHNIQUES TO DETECT SQUINT	11
	2.1 DETECTION BY NONLINEAR ENHANCEMENT AND SEGMENTATION	11
	2.2 AUTOMATIC DETECTION BY MODELING	12
	2.3 USING PROJECTION FUNCTION	12
	2.4 DIGITAL IMAGE PROCESSING	12
	2.5 DETECTION USING HOUGH TRANSFORM	13
3.	PROPOSED SYSTEM	14
	3.1 BLOCK DIAGRAM	
	3.2 READING AN IMAGE	14
	3.3 FACE DETECTION	15
	3.3.1 Feature type and evaluation	15
	3.3.2 Haar features	16
	3.3.3 Cascading classification	16
	3.3.4 properties	17

CHAPTER No	TITLE	PAGE No
	3.3.5 scaling and rotation	18
	3.3.6. stage property	19
	3.3.7 cascade training	20
	3.3.8 limitation	22
3.4	IRIS DETECTION	23
	3.4.1 HOUGH TRANSFORM	24
	3.4.2 CIRCLE HOUGH TRANSFORM	25
	3.4.2.1 Feature extraction with known radius	26
	3.4.2.2 Feature extraction for multiple circles with known radius	27
	3.4.2.3 Accumulator matrix and voting	27
	3.4.2.4 Extracting circle parameter with unknown radius	28
	3.4.2.5 Limitations	29
	3.4.2.6 Applications	29
3.5	NOSE POINT DETECTION	30
	3.5.1 Viola-Jones nose detection algorithm	31
3.6	CALCULATION OF DISTANCE BETWEEN NOSE POINTS AND IRIS CENTERS	31
	3.6.1 Euclidean distance	31
	3.6.2 One dimension	33
	3.6.3 Two dimensions	33
	3.6.4 Three dimensions	34
	3.6.5 N dimensions	34
	3.6.6 Squared Euclidean distance	34
4.	RESULT & DISCUSSION	35-38
5.	CONCLUSION	
	5.1 SUMMARY	45
	5.2 CONCLUSION	46
	REFERENCES	47

LIST OF SYMBOLS AND ABBREVIATIONS

HT	Hough Transform
NLES	Non linear Enhancement and Segmentation
IR	InfraRed
IPF	Integral Projection Function
VPF	Variance Projection Function
HPF	Hybrid Projection Function
GPF	Generalised Projection Function
MM	Milli Meter
LC	Left Eye Center
RC	Right Eye Center
JPEG	Joint Photographic ExpectGroup
MATLAB	Matrix Laboratory
PNG	Portable Network Graphics
TIFF	Tagged Image File Format
BMP	Bitmap
GIF	Graphic Interchange Format
AA	Accumulator Array
CHT	Circle Hough Transform
CC	Cascading Classifier
ROI	Region of Interest
PD	Pupillary Distance
DWD	Dark white Dark
ED	Euclidean Distance
RBM	Restricted Boltzmann Machine
HOG	Histogram of Oriented Gradients
SVM	Support Vector Machine
LFW	Labeled faces in the wild
HALT	Histogram Based Segmentation Technique

LIST OF FIGURES

FIGURE No.	TITLE	PAGE No.
1.1	Esotropia /Exotropia	3
1.2	Hyoyropia/Hexotropia	4
3.1	Haar feature selection	19
3.2	working of cascade classifier	20
3.3	The Hough transforms maps a point into a line	24
3.4	The Hough transform: a line and its parameters in the polar coordinate system	26
3.5	Feature Extraction with known Radius	27
3.6	Feature Extraction for multiple circle with known Radius R	28
4.1	Input image	35
4.2	Face Detection	36
4.3	Iris Detection	37
4.4	Final output	38
4.5	Verified Image	39-41

LIST OF TABLES

TABLE NO	TITLE	PAGE No
1	Distance calculation	

CHAPTER 1

1. INTRODUCTION

The medical term of squint eye termed to be 'abnormal condition of an eye' which is otherwise known as Heterotropia. Misalignment of eye states the eye's unfocused direction; where an eye focuses the object but the other somewhere else. When the brain is unable to detect the image to be recognized from the eyes, they are noted as two different images which cannot combine with the image obtained and thus they result in interrupting the 3D or binocular vision. The squint may be permanent or occur causational, common symptoms used for identification are the words in a book look blurred, straight lines are degraded and a dark or empty area appeared in the center of the vision.

Squint is nothing but the eyes are not pointing in the same direction, one eye may look upward, downward, inward or outward. The medical name for squint is crossed eye or lazy eye. Squint refers to the misalignment of an eyes, squint also called as the Strabismus. For proper vision the eyes should aligned with respect to each other. The one eye focuses on the object which intend to look at ,but the other eye not focusing on the object which intend to look at ,the both the eye are sending different signals to the brain but the brain is not able to fuses these two different signals and produces the blurred image . This leads to the reduction in the total field of three dimensional vision or binocular vision and alleviation in the concentration.

Early treatment of strabismus in infancy may reduce the chance of developing amblyopic and depth perception problems. When strabismus is congenital or develops in infancy, it can cause amblyopic, in which the brain ignores input from the deviated eye. Even with therapy for amblyopic, stereo blindness may occur. The appearance of strabismus may also be a cosmetic problem. One study reported 85% of adult strabismus patients " reported that they had problems with work, school, and sports because of their strabismus".

Squint is the term is used when the eyes are not pointing in the same direction or eyes are not aligned with respect to each other they towards different directions. In most cases one eye looking inwards, outwards, upwards or downwards .The signal from the both the eye are unable to combine by the brain this leads to the loss of concentration. The methods used to detect the squint eye are Nonlinear Enhancement and Segmentation Automation Detection by Modeling Using Projection function and Hough Transform.

The squint eye detection is determined in two phases: In first phase the detection of face, iris center, nose detection is carried out. In second phase the distance from right iris center to nose and from left eye iris center to nose is calculated. We can determine whether squint is present or not by comparing these two distances.

Squint is the term used when the eyes are not pointing in the same direction or eyes are not aligned with respect to each other by different directions. In most cases, eye looks at different directions i.e. inwards, outwards, upwards or downwards. The signals from both the eyes are unable to combine by the brain this leads to the loss of concentration.

1.1 TYPES OF SQUINT

Eyes should be in normal horizontal position. Iris diameter shouldn't be very small in respect to the size of image. The two circles of irises must have similar radiuses. For a normal human subject, the irises are not different having diameter. The distance between the two centers of the circles are divided by the average radius of one circle had to be bounded between some values. In a normal human, the inter eye distance is of about 63 mm with an iris diameter of about 12 mm, this means that the value of the fraction has to be around 5. We could also use the distance between the centers of the ellipses for similar considerations.

1.1.1 *Esotropia*

An eye that turns towards the nose bridge is called an esotropia.

1.1.2 *Exotropia*

An eye that turns towards the ear is called an exotropia.



Fig1.1. Esotropia /Extropia

1.1.3 *Hypertropia*

An eye that turns upwards is called a hypertropia.

1.1.4 Hypotropia

An eye that turns downwards is called a hypotropia.

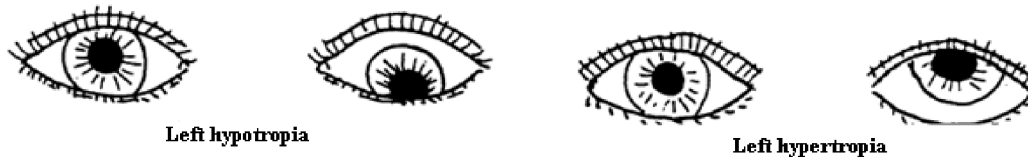


Fig 1.2.Hypotropia/ Hexotropia

1.2 CAUSES OF SQUINT

Squint affected people are suffered by the damage of the muscles controlling the eye. The damaged nerve controls the muscles. The poor development eye muscles are controlled by the center of the brain. Due to the poor development of vision in eye can stop the brain being able to keep the eyes together. This occurs in adults who have had a squint from their childhood. So, if we detect the squint early, it is possible to recover. This project provides how to detect the squint detection based on the Hough transformation. The detection of squint is done in two phase. The detection of face from the original image and finds the left eye center and right eye center and nose edge point in phase 1 and the calculation of the distance between the nose edge point to the left eye center and the distance between the nose edge point to the right eye center and the angles between the left eye center,

Right eye center and nose edge point is determined. Images are collected from squint patients at different positions of iris of an eye, the distance ratios and angles are calculated for those images. We can determine the whether the squint is present or not by comparing the ratio of distance of two images and angles of two images.

First determine the ratio between two distances LC1 and LC2 and then find the ratio between the LC2 and LC. If both the ratios are equal, we can say that it is not squint and if it is not, we can say it is a squint.

Let the centers of the left eye and right eye be LC and RC and the center point of nose be NC, let us consider the three images of a normal person. The features are LC1, RC1 and NC1, from image21 and the LC2, RC2 and NC2, from image2 and LC, RC and NC from image are extracted. First determine the ratio between two distances LC1 and LC2 and then find the ratio between the LC2 and LC. If both the ratios are equal, we can say that it is not squint and if it is not, we can say it is a squint.

1.3 Literature Review

The gaze is a product of two contributing factors eye locations and the head pose. The estimation of these two factors is achieved using expensive, bulky, or limiting hardware. The *eye location* algorithms are found in commercially available eye trackers share the problem of sensitivity to head pose variations and require the user to be either equipped with a head-mounted device or to use a high-resolution camera combined with a chinrest to limit the allowed head movement.

The appearance-based methods that make use of standard low-resolution cameras are considered to be less invasive and more desirable in a large range of applications. The reported results evidence the appearance of eye based center localization for *improved* accuracy of tracking. The integrity of skin-tone edge-based detector tracks the hidden-Markov-model describes a coarse-to-fine pose estimation method by combining facial asymmetry and 3-D model. The method provides free tracking by maintaining a dynamic set of key frames with views of the head under various poses and registers.

Several head pose or eye location methods show the underlying assumption to estimate the eye location or head pose. The IR illumination is used to detect the pupils and derive the head pose by building a feature space. Although their compound tracks the property promote they separate against the methods. The practical limitations need to improve the accuracy of less attractive in comparison with monocular low-resolution implementations. The head tracker initializes the use of localizing and the orientation of the eyes, where the normalized eye obtained from the head tracker. The employed mechanism evaluates the tracking quality.

The field of view while displacement of the vectors between the pose-normalized eye locations and their resting positions are used to adjust the gaze estimation.

Sequential combination of proposed unified framework that provides a deep integration of the used head poses tracker and eye location estimation methods. The working range of the eye locator is extended for the shortcomings of the reported eye locators due to extreme head poses are compensated. The head tracker provides better pose accuracy and recovers the correct pose when the head tracker is lost. The eye location and head pose are in a multimodal visual gaze estimation system by the head pose.

Building automatic face recognition system has been a hot topic of computer vision and pattern recognition for decades. In general, an automatic face recognition task is accomplished in two steps, face detection and face recognition. The former determines any face image present and return the image location and extent of each face.

The geometric feature-based matching algorithm such as eyes is detected before any other processing can take place. In the second class, the algorithms based on template matching, faces are correctly aligned before recognition, which is usually performed based on the detection of eyes.

In fact, the position of the eyes and the interocular distance are relatively constant for most people, detecting the eyes serves as an important role in face normalization, and further localization of facial landmark. Therefore, the detection of the eyes is a vital component in automatic face recognition systems. According to Huang and Wechsler, there are two major approaches for eye detection. The first approach is the holistic approach, attempts to locate the eyes using global representations. The second approach is the abstractive approach, extracts and measures discrete local features, and then employs standard pattern recognition techniques to locate the eyes using these features. The representatives are deformable template-based algorithms presented and extended in both holistic and abstractive approaches after obtaining the eye windows. The image regions include eyes, image projection functions which can be used to locate eye landmarks that are then used to guide the accurate detection of the eye position and shape. Among the image projection functions, the integral projection function (IPF) is the most popular one. The variance projection function (VPF) is usually more sensitive to the variation in the image than IPF, in some cases it cannot reject the variation of the image where IPF works. Both IPF and VPF are viewed as special cases of GPF. Another special case of GPF is the hybrid projection function (HPF), which inherits the robustness of IPF and the sensitiveness of VPF, is empirically developed. Experiments on three face databases show the special case of GPF are effective in eye detection. HPF is better than both IPF and VPF, while VPF is better than IPF. Moreover, it is found that IPF is more effective on the occidental face databases than on the oriental face databases, and VPF is more effective on the oriental face databases than on the occidental face database. An analysis of the detections reveals that such a race effect may owe to the shadow caused by the noses and eyeholes the different race of the people.

The conceptual and intellectual challenges of such a problem, causes the faces to non-rigid and have a high degree of variability in size, shape, color and texture. Numerous techniques have been developed and designed typically to deal with single images. Detection of human appearance in image tracks the humans in videos.

From both neurological and computational point of view symmetry plays an important role in facial part dimensions as it has been demonstrated that an exceptional dimension reduces the facial symmetry. The present method usually returns the image location of a rectangular bounding box containing a face in focusing to isolate the human body in a relatively clean background from input images or videos containing people roughly localized and applying a list of human-based measuring units on the

human bodies. This proposed algorithm goes beyond the limits of all existing technologies as it obtains the unique functional features by enabling the proposed model to work with different skin color tone, applying it to low-quality images, detecting faces with eye glasses, determining the position of facial parts and detect several faces on one image is typically designed to deal with single images.

An algorithm insight to construct a framework for robust and rapidly visualize the detection. Toward the end of the construction a frontal face detection system which achieves detection and false positive rates which are equivalent. This face detection system is most clearly distinguished from previous approaches in its ability to detect faces. In other face detection systems, auxiliary information, such as image differences in video sequences, or pixel color in color images, have been used to achieve high frame rates. Our system achieves high frame rates working only with the information present in a single grey scale image. These alternative sources of information can also be integrated with our system to achieve even higher frame rates. There are three main contributions of our face detection framework.

The notion behind the focus of the approach is often possible to rapidly determine where in an image a face might occur. More complex processing is reserved only for these promising regions. The key measure of such an approach is the “false negative” rate of the attention process. It must be the case for all the face instances for the selected attention filter.

A face detection operator can be learned with filter of over 50% of the image will be preserved 99% of the faces. This filter is exceedingly efficient; it can be evaluated in 20 simple operations per location/scale. Those sub-windows which are not rejected by the initial classifier are processed by a sequence of classifiers, each slightly more complex than the last. If any classifier rejects the sub-window, no further processing is performed. The complete face detection cascade has 38 classifiers, which total over 80,000 operations. Nevertheless the cascade structure results in extremely rapid average detection times. On a difficult dataset, containing 507 faces and 75 million sub-windows, faces are detected using an average of 270 microprocessor instructions per sub-window. In comparison, this system is about 15 times faster than an implementation of the detection system constructed by Rowley. An extremely fast face detector will have broad practical applications. These include user interfaces, image databases, and teleconferencing. This increase in speed will enable real-time face detection applications on systems where they were previously infeasible. In applications where rapid frame-rates are not necessary, our system will allow for significant additional post processing and analysis. In addition our system can be implemented on a wide range of small low power devices, including hand-helds and embedded processors. In our lab we have implemented this face detector on a low power 200 mips *Strong Arm* processor which lacks floating point hardware and have achieved detection at two frames per second.

The Viola–Jones object detection framework is the first object detection framework to provide competitive object detection rates in real-time proposed in 2001 by Paul Viola and Michael Jones. It was motivated primarily by the problem of face detection, although it can be trained to detect a variety of object classes. This algorithm is implemented in Open CV as CV Haar Detect Objects (). Viola Jones detector become famous due to its open source implementation in the OpenCV library. In order to find an object of an unknown size is usually adopted to work this field that possesses a high efficiency and accuracy to locate the face region in an image.

Toward this end we have constructed a frontal face detection system which achieves detection and false positive rates which are equivalent to the best published results. This face detection system is most clearly distinguished from previous approaches in its ability to detect faces extremely rapidly. Operating on pixel images, faces are detected at frames per second on a conventional. In other face detection systems, auxiliary information, such as image differences in video sequences, or pixel color in color images, have been used to achieve high frame rates. Our system achieves high frame rates working only with the information present in a single grey scale image. These alternative sources of information can also be integrated with our system to achieve even higher frame rates. There are three main contributions of our face detection framework. We will introduce each of these ideas briefly below and then describe them in detail in subsequent sections.

1.4 PROBLEM DEFINITION

Squint eye is meant to be the abnormal condition of the eye which means the eyes are not aligned with respect to each other or crossed eye, also called as Strabismus. Strabismus is classified as various types according to their eye by which one can determine the problem and that can be resolved by means of the Image processor. In the oculomotor behavior, cognitive visual function, and vision deficiencies have been studied from a qualitative eye-movement analysis approach. As the dynamics of eye movement have been subjectively and qualitatively analyzed, it has been hard to compare related data to a particular diagnosis. As a result, impairments cannot be easily identified. Eye movement studies are subjective and manual way to the variable and unreliable. Standardized but manual processes are not time or cost efficient. The method that has been categorized as the gold standard is the scleral search coil technique. SSC provides accurate but expensive and invasive method to study eye movements.

1.5 OBJECTIVE

To determine the squint eye by using the image processor is the objective of our project. PYTHON is used as the resolving key source of the image to the affects squint eye. An

Eye tracker measures high sampling rates and the position of the eyes with high accuracy. This enables the calculation of head movements and extends calculation of the true eye movements within the orbit of the eye. The dynamics of eye movement are studied using data.

CHAPTER 2

2. TECHNIQUES TO DETECT THE SQUINT

There are different methods available to detect the squint eye. In this project we have taken the Hough transform to detect the squint eye. The some of the following methods described below.

2.1. Detection by Nonlinear Enhancement and Segmentation

This algorithm is purely based on histogram-techniques for the problem of squint eye evaluation. The detection of anomalies in human eye's retina is a biomedical problem, appropriate for image processing and automated segmentation, whose solution is intended to help the doctors in their decision making process. Use of the proposed detector may reduce false negatives and give reliable detection accuracy in both position and mass size. Histogram- based enhancement technique (MLE), uses histogram equalization as its core operator and a histogram- based segmentation technique (HALT) to segment areas that differ slightly from their background regions. This method is able to detect actual drusen in all cases. Even in hard-to-diagnose cases, many small and vague drusen.

2.2. Automatic detection by Modeling

This is the simplest yet efficient method for squint eye detection. Modeling the eye by means of its light refracting layers is calculated. We modeled the human eye as a circle circumscribed in an ellipse, where circle represents the iris of human eye and the ellipse represents the eye lashes. This is not the only model for an eye; there are infinite models possible and they can be applied as per requirement.

2.3. Using Projection Function

Due to its indifference against disturbances of the contour of an object Mean and variance projection functions are utilized in each eye pair window to validate the presence of the eye for locating the corner and center point of the eye., So in this technique possible eye areas are localized using a simple thresholding in colour space followed by a connected component analysis to quantify spatially connected regions and further reduce the search space to determine the contending eye pair windows.

2.4. Digital image processing

Detection of squint eye may be achieve by measuring the landmark points (eye corners, iris center points), from which the approximate eyelid contours are estimated and thus the exact determination of the squint angle is achieved and by processing and extracting the information from the digital image captured by the camera amount of squint may be calculated. At a first step, the position of the iris is defined as a region of interest (ROI).

2.5. Detection using Hough Transform

The Hough-Transformation (HT) is also suited for the detection of the squint eye. The HT is based on the idea of transforming all contour points belonging to the structure into one point of transformation space (accumulator array (AA)). This point is identified as the global maximum of the AA. It becomes more prominent the greater the number of contour points belonging to the transformed object is. The HT can be calculated for any curve described by parameters, which turn into the axes of the AA.

Centre, corner and radius of the circular iris in space can thus be read as coordinates of the global maximum of the AA which in this case is three-dimensional. The binary edge image is produced by calculating the modulus of the gradient of the original image.

CHAPTER 3

3. PROPOSED SYSTEM

The proposed method has been carried out in two phases: In first step the face detection, pupil's center and nose point is determined; in second phase the distance between the pupil's center to the nose point are calculated.

3.1. BLOCK DIAGRAM

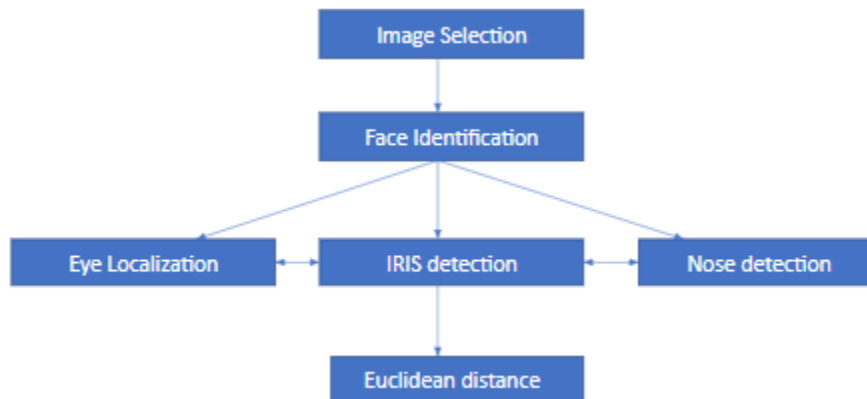


Fig:BLOCK DIAGRAM

3.2 READING AN IMAGE

PYTHON's IPT has a built-in function to open and read the contents of image files in most popular formats (e.g., TIFF, JPEG, BMP, GIF, and PNG).

The `imread` function allows you to read image files of almost any type, in virtually any format, located anywhere. This saves you from a (potentially large) number of problems associated with file headers, memory allocation, file format conventions, and so on, and allows to focus on what you want to do to the image once it has been

Read and stored into a variable in the PYTHON workspace.

The syntax to read an image is given by

```
img=cv2.imread("test.jpg")
```

3.3 FACE DETECTION

Dlib is a modern C++ toolkit containing machine learning algorithms and tools for creating complex software in C++ to solve real world problems. It is used in both industry and academia in a wide range of domains including robotics, embedded devices, mobile phones, and large high performance computing environments. Dlib's open source licensing allows you to use it in any application, free of charge.

Major Features

Image Processing

Routines for reading and writing common image formats.

Automatic color space conversion between various pixel types

Common image operations such as edge finding and morphological operations

Tools for detecting objects in images including frontal face detection and object pose estimation.

High quality face recognition.

Facial Landmarks



- Jaw Points = 0–16
- Right Brow Points = 17–21
- Left Brow Points = 22–26
- Nose Points = 27–35
- Right Eye Points = 36–41
- Left Eye Points = 42–47
- Mouth Points = 48–60
- Lips Points = 61–67

Many features are there in Dlib C++ Library like Documentation, High quality portable code, Machine learning algorithm, numerical algorithm, graphical model interface algorithm, image processing, threading, networking, testing, etc.

Example of facial landmarks,



Face feature is considered to be a key requirement in many applications such as Biometrics, Facial recognition systems, video surveillance Human computer interface etc. Therefore reliable face detection is required for the success of these applications. The task of human facial feature extraction is not easy. Human face varies from person to person. The race, gender, age and other physical characteristics of an individual have to be considered thereby creating a challenge in computer vision.

Facial feature detection aims to detect and extract specific features such as eyes, mouth and nose. The goal of the face detection in this project was only to provide a focus-of-attention mechanism for the eye-detection algorithm used later. There are many face extraction methods reported by researchers in the field of face recognition. To combine image information and knowledge of the face is the main trend. Four different methods were evaluated for determining the face region. The first was a neural network based face detection package described by Jones and Rehg. This method did not provide acceptable results on infant faces, and the time and resources to train our own neural network were not available.

The next two methods, mixture model and histogram-based color detection, both described in Rowley et al., were also found to be inadequate due to unavailability of the dataset large enough to build a quality model.

In our project in order to detect the face region The Viola–Jones object detection framework method was used. This method was proposed in 2001 by Paul Viola and Michael Jones. This detector uses a Haar-like cascade classifier.

3.3.1.Feature types and evaluation

The characteristic feature of Viola–Jones algorithm which makes a good detection algorithm is stated below. Robust is the feature having very high detection rate & very low false-positive rate always.

Real time for practical applications at least 2 frames per second must be processed. Face detection is the only distinguishing goal to the faces from non-faces is evaluated. The four stages of algorithm are in detail.

1. Haar Feature Selection
2. Creating an Integral Image
3. Adaboost Training
4. Cascading Classifiers

The features sought by the detection framework universally involve the sums of image pixels within rectangular areas. As such, they bear some resemblance to Haar basis functions, which have been used previously in the realm of image-based object detection. However, since the features used by Viola and Jones all rely on more than one rectangular area, they are generally more complex. The figure on the right illustrates the four different types of features used in the framework. The value of any given feature is the sum of the pixels within clear rectangles subtracted from the sum of the pixels within shaded rectangles. Rectangular features of this sort are primitive when compared to alternatives such as steerable filters. Although they are sensitive to vertical and horizontal features, their feedback is considerably coarser.

3.3.2. Haar Features

All human faces share some similar properties. These regularities may be matched using Haar Features. Few properties common to human faces are the eye region which is darker than the upper-cheeks. The nose bridge region is brighter than the eyes. Composition of properties forming matchable facial features which are localized and eyes size, mouth, bridge of nose. The value of the oriented gradients pixel intends the four features matched by this algorithm are then sought in the image of a face. The rectangle features poses value;

Rectangle feature= Σ (pixels in black area) - Σ (pixels in white area)

Three types of rectangle features are two, three, and four-rectangles, and Viola & Jones used two-rectangle features. The difference in brightness between the white & black rectangles over a specific area of each feature is related to a special location in the sub-window..

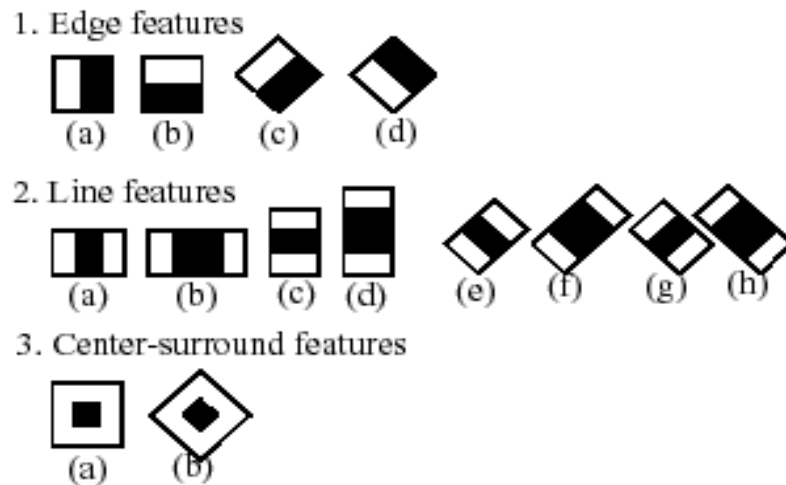


Fig 3.1 Haar feature selection

.An image representation called the integral image evaluates rectangular features in constant time, which gives them a considerable speed advantage over more sophisticated alternative features. Because each feature's rectangular area is always adjacent to at least one other rectangle, it follows that any two-rectangle feature can be computed in six array references, any three-rectangle feature in eight, and any four-rectangle feature in nine. The integral image at location (x, y) , is the sum of the pixels above and to the left of (x, y) , inclusive.

3.3.3. Cascading Classifier (CC)

Cascading is a particular case of ensemble learning based on the concatenation of several classifiers, using all information collected from the output from a given classifier as additional information for the next classifier in the cascade. Unlike voting or stacking ensembles, which are multiexpert systems, cascading is a multistage one.

The first cascading classifier is the face detector of Viola and Jones. The requirement was that the classifier be fast in order to be implemented on low CPU systems, such as cameras and phones.

3.3.4. Properties

The properties of Cascading Classifier (CC) are mentioned as follows. They are Scaling and rotations, Stage properties and Cascade training are explained in detail.

3.3.5. Scaling and Rotations

If the first stage is very fast, we have eliminated 60% of rectangles not containing a face very quickly.

The training procedure for one stage is therefore to have many weak learners, train them as a group, but be mindful of having only a few active weak learners so the computation time remains low.

The first detector of Viola & Jones had 38 stages, with 1 feature in the first stage, then 10, 25, 25, 50 in the next five stages, for a total of 6000 features. The first stages remove unwanted rectangles rapidly to avoid paying the computational costs of the next stages, so that computational time is spent analyzing deeply the part of the image that have a high probability of containing the object.

3.3.6. Stage Properties

To have good overall performance, the following criteria must meet each stage validate all faces and can produce many false positives. If stage 1 were to mark as 'does not contain a face' 20% of rectangles containing a face, then the total performance of the chain cannot be higher than 80% true positive, whatever the next stages are, since 20% of faces have already been rejected.

This suggests that a good stage must have 100% true positive and for example 40% false positive, that is accept all rectangles containing faces and erroneously mark many rectangles as potentially containing a face, to be eliminated by later stages. It should be noted that for a first stage, 100% true positive and 40% false positive still gives a lot of false negative, if only 1 in a 1000 rectangles in an image contain a face, there will still be 400 to 1 false possible faces after the first stage.

If the first stage is very fast, we have eliminated 60% of rectangles not containing a face very quickly. The first detector of Viola & Jones had 38 stages, with 1 feature in the first stage, then 10, 25, 25, 50 in the next five stages, for a total of 6000 features. The first stages remove unwanted rectangles rapidly to avoid paying the computational costs of the next stages, so that computational time is spent analyzing deeply the part of the image that have a high probability of containing the object.

3.3.7. Cascade Training

Cascades are usually done through cost-aware ADABOOST. The sensitivity threshold (0.8 in our example) can be adjusted so that there is close to 100% true positives and some false positives. The procedure can then be started again for stage 2, until the desired accuracy/computation time is desired. After the initial algorithm, it was understood that training the cascade as a whole can be optimized, to achieve a desired true detection rate with minimal complexity. Examples of such algorithms are RCBoost, ECBoost or RCECBoost.

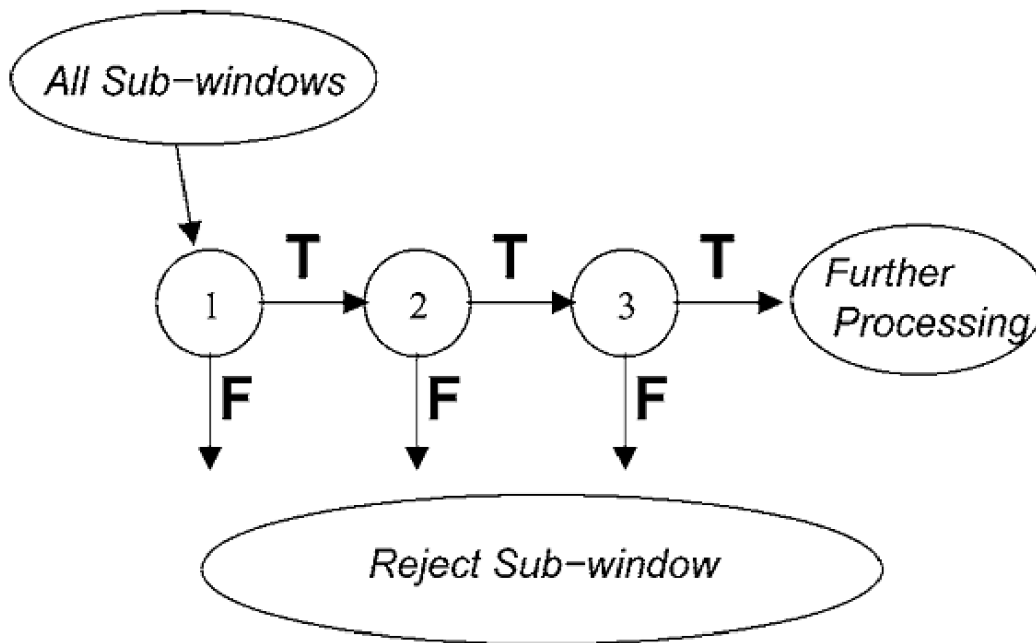


Fig 3.2. Working of Cascade Classifier

Cascade classifiers are available in OpenCV, with already trained cascades for frontal faces. Training a cascade on different object is possible. We have taken a haar-like cascade classifier. After evaluating the performance of all four face classifiers included with OpenCV, the “frontalface_alt” classifier was found to be the best. Detection is done in a two-stage process. The detector is run once, and then it is run again with lower accuracy if no faces are detected. Because the input photos are studio pictures, with the face as the dominant feature, a large rectangle size was used in the detection.

This greatly increases the speed of the operation without affecting accuracy. The face can be safely assumed to be the main feature so false positives could be easily culled by discarding matches smaller than 25% of the image. This number could potentially be made larger, but it was decided to err on the side of caution.

3.3.8. Limitation

One unusual characteristic exhibited by this classifier is that it generally fails to detect the face in pictures that had been cropped fairly closely around the face. Because most studio pictures were not expected to be cut this close, this behavior was deemed acceptable.

3.4. IRIS DETECTION

Eye Detection is one of the very essential steps in many applications, such as face expression analysis, gaze estimation, face detection or recognition, criminal investigation and human interactions. Eye position detection assists eye contour detection. Eyes are the salient and relatively stable features on the face when compared with other facial features. The position of other facial features can be detected with the help of eye position.

There are two purposes of eye detection. One is to detect the existence of eyes, and another is to accurately locate eye positions. Under most situations, the eye position is measured with the pupil center.

To detect eye pupil the algorithm isolate the region of interest (ROI) in area of interest to detect eye by processing this region as shown in "figure" by dividing this region into two parts; left half contain right eye and a right half contain the left eye .the study apply Viola - Jones right eye detector on right eye region to detect the right eye, and apply Viola - Jones left eye detector on left eye region to detect the left eye .Our last detection step is determining the location of eye pupil and determine the distance between them . The study use the Iris recognition system which consists of an automatic segmentation system that is based on the Hough transform to be able to localize the circular iris pupil detection is based on the fact that pupil will have the highest black value relative to everything else in the iris region .

3.4.1. Hough Transform

The technique used here to detect the circles is Hough Transform. Hough Transform is a globally accepted method for detecting edges, it is feature extraction technique which stands out in image processing. Hough Transform was basically used for the purpose of identification of lines which was later improved upon in identifying positions of arbitrary shapes which include circles and ellipses.

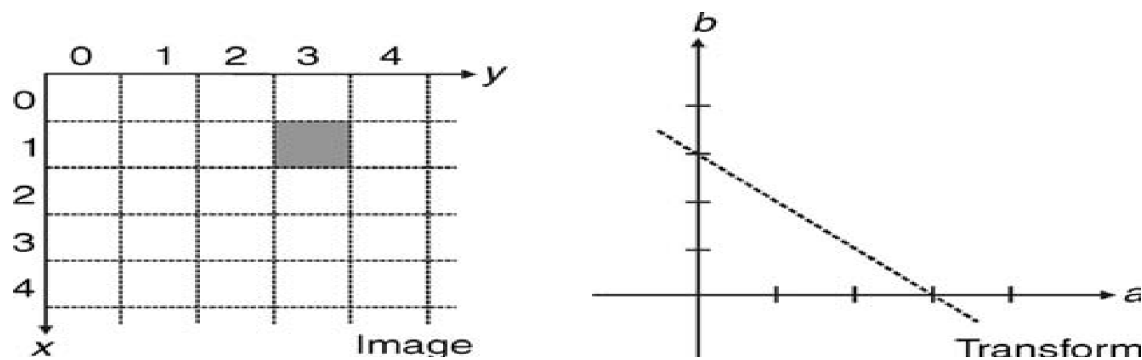


Fig 3.3. The Hough transforms maps a point into a line

Let (x, y) be the coordinates of a point in a binary image. The Hough transform stores in an *accumulator array* all pairs (a, b) that satisfy the equation $y = ax + b$. The (a, b) array is called *transform array*.

Let the centers of the left eye and right eye be LC and RC and the center point of nose be NC, let us consider the three images of a normal person. The features are LC1, RC1 and NC1, from image21 and the LC2, RC2 and NC2, from image2 and LC, RC and NC from image are extracted.

The meaning of two or more lines intersecting in the transform domain is that the points to which they correspond are aligned in the image. The points with the greatest number of intersections in the transform domain correspond to the longest lines in the image.

Describing lines using the equation $y = ax + b$ (where a represents the gradient) poses a problem, since vertical lines have infinite gradient. This limitation can be circumvented by using the *normal representation* of a line, which consists of two parameters: ρ (the perpendicular distance from the line to the origin) and θ (the angle between the line's perpendicular and the horizontal axis). In this new representation is shown in figure vertical lines will have $\theta = 0$. It is common to allow ρ to have negative values, therefore restricting θ to the range $-90^\circ < \theta \leq 90^\circ$.

The relationship between ρ , θ , and the original coordinates (x, y) is

$$\rho = x \cos \theta + y \sin \theta$$

Under the new set of coordinates, the Hough transform can be implemented as follows:

1. Create a 2D array corresponding to a discrete set of values for ρ and θ . Each element in this array is often referred to as an *accumulator cell*.
2. For each pixel (x, y) in the image and for each chosen value of θ , compute $x \cos \theta + y \sin \theta$ and write the result in the corresponding position— (ρ, θ) —in the accumulator array.
3. The highest values in the (ρ, θ) array will correspond to the most relevant lines in the image.

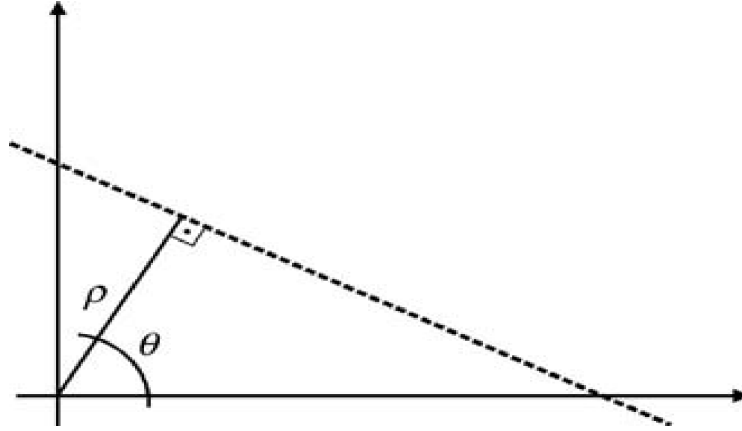


Fig 3.4. The Hough transform: a line and its parameters in the polar coordinate system

3.4.2. Circle Hough Transform (CHT)

The CHT is a feature extraction technique for detecting circles. It is a specialization of Hough Transform. The purpose of the technique is to find circles in imperfect image inputs. The circle candidates are produced by “voting” in the Hough parameter space and then select the local maxima in a so-called accumulator matrix.

Theory:

In 2D space, the circle can be represented as

$$(X-a)^2 + (Y-b)^2 = R^2 \quad \dots (1)$$

Where,

(a, b) represents the center of the circle,

R represents the radius of the circle

If (X, Y) is fixed, the parameters of the circle are extracted according to the equation (1). The three dimensional representation is given by (a, b, r) . And all the parameters that satisfying (x, y) would lie on the surface of an inverted right-angled cone whose apex is at $(x, y, 0)$.

In the 3D space, the circle parameters can be identified by the intersection of many conic surfaces that are defined by points on the 2D circle. This process can be divided into two stages. The first stage is fixing radius then find the optimal center of circles in a 2D parameter space. The second stage is to find the optimal radius in a one dimensional parameter space.

The CHT is used to obtain the circles features from the image of known radius and unknown radius.

3.4.2.1 Feature Extraction with known Radius

If the radius is fixed, then the parameter space would be reduced to 2D (the position of the circle center). For each point (x, y) on the original circle, it can define a circle centered at (x, y) with radius R according to (1). The intersection point of all such circles in the parameter space would be corresponding to the center point of the original circle.

Consider 4 points on a circle in the original image (left). The circle Hough transforms is shown in the right. Note that the radius is assumed to be known. For each (x, y) of the four points (white points) in the original image, it can define a circle in the Hough parameter space centered at (x, y) with radius r . An accumulator matrix is used for tracking the intersection point. In the parameter space, the voting number of points through which the circle passing would be increased by one.

Then the local maxima point (the red point in the center in the right figure) can be found. The position (a, b) of the maxima would be the center of the original circle.

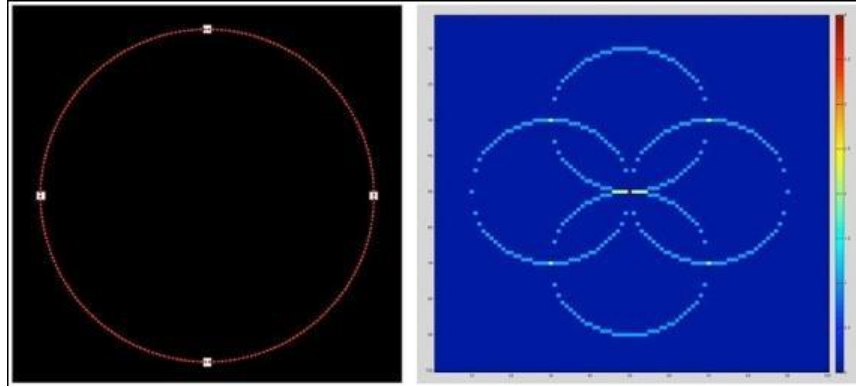


Fig 3.5. Feature Extraction with known Radius

3.4.2.2. Feature extraction for multiple circles with known Radius R

Multiple circles with same radius can be found with the same technique.

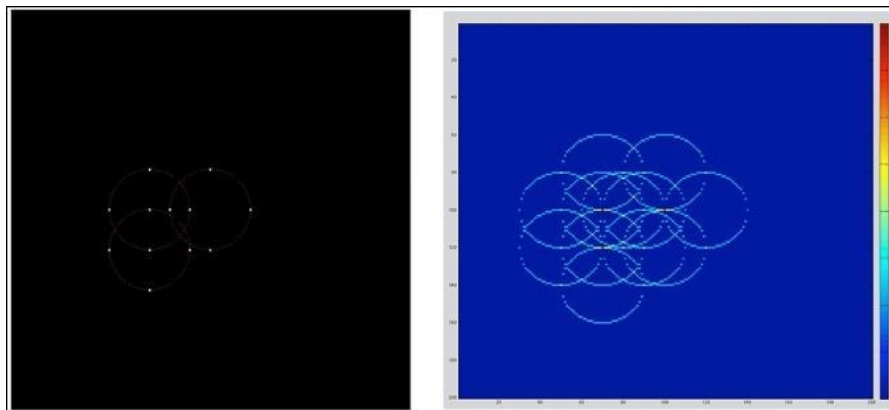


Fig 3.6. Feature Extraction for multiple circle with known Radius R

3.4.2.3 .Accumulator matrix and voting

In practice, an accumulator matrix is introduced to find the intersection point in the parameter space. First, we need to divide the parameter space into “buckets” using a grid and produce an accumulator matrix according to the grid. The element in the accumulator matrix denotes the number of “circles” in the parameter space that passing through the corresponding grid cell in the parameter space.

The number is also called “voting number”. Initially, every element in the matrix is zeros. Then for each “edge” point in the original space, we can formulate a circle in the parameter space and increase the voting number of the grid cell which the circle passing through. This process is called “voting”.

After voting, we can find local maxima in the accumulator matrix. The positions of the local maxima are corresponding to the circle centers in the original space.

3.4.2.4. Extracting circle parameter with unknown radius

Since the parameter space is 3D, the accumulator matrix would be 3D, too. We can iterate through possible radiuses and for each radius, we use the previous technique. Finally, find the local maxima in the 3D accumulator matrix.

3.4.2.5. Limitations

Since the parameter space of CHT is three dimensional, it may require lots of storage and computation. Choosing bigger grid size can ameliorate this problem. However, choose an appropriate grid size is difficult. Since too coarse a grid can lead to large values of the vote being obtained falsely because many quite different structures correspond to a single bucket. Too fine a grid can lead to structures not being found because votes resulting from tokens that are not exactly aligned end up in different buckets, and no bucket has a large vote.

Also, CHT is not very robust to noise. The boundary of pupil is not quite circle and a small error in detecting this boundary will lead to lose some information surrounding the pupil.

Another problem that makes most of the iris segmentation method fail to detect the pupil boundary is the head rotation or the eye rotation.

3.4.2.6. Applications

People counting - Since the head would be similar to a circle in an image. CHT can be used for detecting heads in a picture, so as to count the number of persons in the image.

3.5 NOSE POINT DETECTION

3.5.1. Viola-Jones nose Detection Algorithm

Nose have three different local characteristics used as follows (i) Similarity of both sides: The left and right sides of nose are similar in a front-view face as shown in Figure, this property of similarity can be measured using Euclidean distance between both sides.(ii) Dark-White-Dark (DWD) property: the lower part of nose region is characterized by two dark nostrils and a light sub region due to the reflection of light on the nose as shown in Figure .



The two nostrils regions are less than the average of middle lighter sub region containing nose tip. (iii) The variation in lower/upper parts property:

When the face is rotated some degrees for these two properties are despaired and the only clear property is the variation between lower part and upper part as shown in "Figure 2c". This variation can be measured by the variance in each part. Based on this analysis, a search is done for a certain region among the ten highest regions detected due to reflection of light at nose tip.

Nose detection is highly important for face part detection in range images and is consider as simpler task than detecting a face because a nose is not as complex and variable as a face. In this paper, the proposed nose detection algorithm has the following three characteristics. First, it is based on The cascade object detector which uses the Viola-Jones algorithm to detect people's noses This model is composed of weak classifiers,

based on a decision stump, which uses Haar features to encode nose details. Second, it can deal with both frontal and non-frontal poses. Finally, it is quite fast and accurate because the algorithm searching in a small region of the image (b) area of interest to detect the nose not the whole image.

EUCLIDEAN DISTANCE

Euclidean distance is the ordinary distance between the two points in an Euclidean space. Euclidean distance also called as Euclidean metric. The Euclidean distance between the two points P and Q is the length of the line connecting segment. In Cartesian Coordinates, if $p = (p_1, p_2, \dots, p_n)$ and $q = (q_1, q_2, \dots, q_n)$ are two point in Euclidean n space then the distance (d) from p to q, or from q to p is given by the Pythagorean theorem.

$$\begin{aligned} d(\mathbf{p}, \mathbf{q}) &= d(\mathbf{q}, \mathbf{p}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + \dots + (q_n - p_n)^2} \\ &= \sqrt{\sum_{i=1}^n (q_i - p_i)^2}. \end{aligned}$$

Euclidean norm or Euclidean length, or magnitude of a vector measures the length of the vector:

$$\|\mathbf{p}\| = \sqrt{p_1^2 + p_2^2 + \cdots + p_n^2} = \sqrt{\mathbf{p} \cdot \mathbf{p}},$$

Where the last equation involves the dot product.

A vector can be described as a directed line segment from the origin of the Euclidean space (vector tail), to a point in that space (vector tip). If we consider that its length is actually the distance from its tail to its tip, it becomes clear that the Euclidean norm of a vector is just a special case of Euclidean distance: the Euclidean distance between its tail and its tip.

The distance between points p and q may have a direction (e.g. from p to q), so it may be represented by another vector, given by

$$\mathbf{q} - \mathbf{p} = (q_1 - p_1, q_2 - p_2, \cdots, q_n - p_n)$$

In a three-dimensional space ($n=3$), this is an arrow from p to q , which can be also regarded as the position of q relative to p . It may be also called a displacement vector if p and q represent two positions of the same point at two successive instants of time.

The Euclidean distance between p and q is just the Euclidean length of this distance (or displacement) vector:

$$\|\mathbf{q} - \mathbf{p}\| = \sqrt{(\mathbf{q} - \mathbf{p}) \cdot (\mathbf{q} - \mathbf{p})}.$$

$$\|\mathbf{q} - \mathbf{p}\| = \sqrt{\|\mathbf{p}\|^2 + \|\mathbf{q}\|^2 - 2\mathbf{p} \cdot \mathbf{q}}.$$

3.6.2. One dimension

In one dimension, the distance between two points on the real line is the absolute value of their numerical difference. Thus if x and y is two points on the real line, then the distance between them is given by:

$$\sqrt{(x - y)^2} = |x - y|.$$

In one dimension, there is a single homogeneous, translation-invariant metric, up to a scale factor of length, which is the Euclidean distance. In higher dimensions there are other possible norms.

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

3.6.3. Two dimensions

In the if Euclidean plane $\mathbf{p} = (p_1, p_2)$ and $\mathbf{q} = (q_1, q_2)$ then the distance is given by

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

$$d(\mathbf{p}, \mathbf{q}) = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}.$$

This is equivalent to the Pythagorean Theorem

Alternatively, it follows from (2) that if the polar coordinates. If the point $\mathbf{p}$ are (r_1, θ_1) and those of $\mathbf{q}$ are (r_2, θ_2) , then the distance between the points is

$$\sqrt{r_1^2 + r_2^2 - 2r_1r_2 \cos(\theta_1 - \theta_2)}.$$

3.6.4. Three dimensions

In three-dimensional Euclidean space, the distance is

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + (p_3 - q_3)^2}.$$

3.6.5. n dimensions

In general, for an n -dimensional space, the distance is

$$d(p, q) = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_i - q_i)^2 + \cdots + (p_n - q_n)^2}.$$

3.6.6. Squared Euclidean distance

The standard Euclidean distance can be squared in order to place progressively greater weight on objects that are farther apart. In this case, the equation becomes

$$d^2(p, q) = (p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_i - q_i)^2 + \cdots + (p_n - q_n)^2.$$

Squared Euclidean Distance is not a metric as it does not satisfy the triangle inequality. However, it is frequently used in optimization problems in which distances only have to be compared.

CHAPTER 4

4.RESULTS AND DISCUSSION

Running our Picture Perfect program on multiple testing images and analyzing the results allowed us to see where we could make improvements to enhance the accuracy of our classifier algorithm

Step 1

First the image is read by the function of imread function. After the read operation, the image should be converted into gray scale image for further processing.



Fig 4.1 Input image

Step 2

The face features are extracted from the image with the help of Viola-Jones algorithm for face detection.



Fig 4.2Face Detection

Step 3

Then the iris detection and the nose detection is achieved by the Viola-Jones algorithm, because this algorithm is suitable for the detection of the objects like circles, ellipses, etc, in an image. The pupil center and the nose edge point are determined.

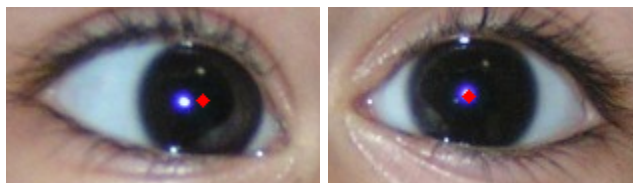


Fig 4.3 Iris Detection

Step 4

Finally, the distance between the left eye center to the nose point and right eye center to the nose are calculated.



Fig 4.4 Final output

We have proposed a face location and accurate eye detection, which is based on multiple evidences, including facial components structure, texture similarity, component feature measurement and contour matching. It can be faster if the color information is used to segment the image first upon the captured face tracking can be implemented. The accuracy of the detection is defined by the difference between the one detected automatically and the one fixed manually. Noticeable improvements have been noted, especially in the rotation and scale normalizations. Another use of this algorithm is to evaluate the detected face is suitable normalization for a successful recognition. To a specific hardware configuration, if the detected eyeballs are judged too small, too close or too far from each other, we expect that the normalization will be very unstable. Here the probability of confusion with another face is reduced because of the increased rejection rate. A detailed comparison of the approach states different lighting conditions and complex backgrounds with three algorithms on same images. Experimental results show that the algorithm is Face part division; based on physical approximation of location of eyes, nose and mouth on face. The proposed algorithm clearly reduces false negative rates and detects the face poses from the images at different imaging resolution and different lighting conditions.

The experimental part in the study was implemented using Matlab environment, presenting a set of experiments on a difficult face detection dataset which has been Widely studied. This research work was initiated as a part of research project for Human Actions Detection System.

Verified Images

Input Images



output images







Table of distance calculation

Image	Left eye to nose	Right eye to nose	SQUINT
1	145	147	YES
2	149	141	NO
3	53	46	NO
4	74	75	YES
5	210	184	YES
6	193	172	YES
7	107	138	YES

The results show RBM and HOG feature extraction methods using SVM as classifier. We first show the eye localization errors in Table 1 and the rotation angle estimation errors in Table 2. The average localization errors and rotation estimation errors were computed on the natural data without doing any additional artificial rotation. Instead we computed the average errors from all the images we selected for the datasets. Table 1 shows the results for localizing the eyes. The two feature extraction methods perform similarly. The average localization errors are very small. This also makes the angle estimation errors in Table 1 very small, although the rotation errors are quite sensitive to small errors in eye localization Table 2. When we examine Table 2, the average rotation errors are quite small. Meanwhile, although a correlation can be seen between Table 1 and Table 2, lower position errors do not directly imply lower rotation errors.

This observation suggests that calculation of rotation angles is sensitive to stability of position information. In this way, we can say that the RBM feature extraction method gives more stable position information than the HOG method. The results on the LFW dataset are quite promising when compared to previous results. We only found one paper describing localization errors on LFW, in average eye localization errors on LFW are 0.081 for the left and 0.084 for the right eye. In this study, we obtained lower error rates as can be seen in Table 1. As for a general comparison with other works, the survey paper presents a lot of eye detection results obtained with many other possible methods. Our method outperforms some of these methods, although the results of the best methods presented in are better than the results obtained with our method. To compare to those results, we want to mention that our best method obtained 95% (96.9%) correctly detected eyes on Feret with a eye localization threshold of 0.1 (0.25), and 85.4% (99%) on LFW with a threshold of 0.1 (0.25). We also show the plot of the average angle estimation errors in Figure 6. To construct the plot in Figure 6, we first rotated every single face image in one of the experimental datasets (IMM) to 0_ degrees using the manually annotated coordinates of the eye centers.

Then, we rotated every image from -30° to 30° in steps of 2° and for each angle we computed the average rotation estimation error. The error rates of the method are lowest between -20° and 20° which corresponds to the range of angles encountered in the training set for the eye detector. Besides, a similar observation already seen in Table 1 and Table 2 about the performance of the two feature extraction methods can also be noticed here. To conclude from all of these observations, the RBM seems to better handle angle estimations than HOG.

CHAPTER 5

5.CONCLUSION

5.1SUMMARY

Detection of the iris becomes indifferent to disturbance created by the iris. Through the automatic analysis of images, strabometry becomes suitable for screening. The high sensitivity of the system leads to an early recognition of symptoms, and thus to the desired treatment of the illness at an early stage of its development.

Since IPF, VPF, and HPF are all special cases of GPF, it is easy to get the comparison of their detection rates on the three experimental face databases. IPF, VPF, and HPF are all effective in eye detection because their detection rates on the experimental face databases are all beyond 90%. The performance of VPF is significantly better than that of IPF on all three databases, while the performance of HPF is significantly better than that of IPF on all three databases and is significantly better than that of VPF. This is not strange because in fact HPF is obtained through experiment determining the optimal parameters of GPF. Analysis of the detections shows that it may owe to the facial shadow. Some detection results are in such a race effect has not been exposed this may cause the fact that the faces with little variety of illumination, pose, and face size so that the eyes are relatively easy to be detected, which is supported by the fact that the detection rates of all the projection function are better than 96.5%. IPF prefers occidental eyes to oriental eyes while VPF prefers oriental eyes to occidental eyes, the performance of VPF is still better than that of IPF on either occidental or oriental eyes, because VPF is more powerful than IPF.

5.2 CONCLUSION

Face detection is an important step to obtain good results with a face recognition system. In this paper, we have presented a Viola-Jones algorithm. In this method, the first is the introduction of a new image representation called the “Integral Image” which allows the features used by our detector to be computed very quickly.

The second is a simple and efficient classifier which is built using the AdaBoost learning algorithm to select a small number of critical visual features from a very large set of potential features. The third contribution is a method for combining classifiers in a “cascade” which allows background regions of the image to be quickly discarded while spending more computation on promising face-like regions. Then the location of eyes and center’s of eyes are detected by the circle Hough Transform. Second is the detection of nose point. This location information is subsequently used to align faces by using a simple geometrical formula. For the eye detector, we also compared results of two feature extraction techniques in eye localization and rotation angle estimation. The results on three different datasets show that the RBM feature extraction technique is better at handling rotation angle estimation than HOG. This is also supported by the angle estimation error plot created by using artificially created angles. We finally examined the effect of rotational alignment in a face recognition experiment in which we compare the use of rotationally aligned and nonaligned faces in a simple face recognition system. The results show that the RBM feature extraction method gives the best angle estimation performance and this in-turn results in better performance in a face recognition system. In future work we will primarily focus on optimizing the face recognition algorithm, which will make use of the rotation alignment method presented in this paper.

REFERENCES

1. Eye Feature Detection towards Automatic Strabismus Screening
2. International Journal of Advanced research in computer engineering & technology volume 1, issue, may 2012
3. International Journal of Engineering and Advanced Technology (IJEAT)
4. Face and Eye tracking algorithm based on Digital Image Processing
5. S. Barattelli, L. Sichelschmidt and G. Rickheit, Eye movements as an input in human computer interaction: exploiting natural behavior, *Annual Conference of IEEE Industrial Electronic Society*, Vol. 4, pp. 2000-2005, 31 Aug – 4 Sept. 1998.
6. T. N. Cornsweet, H. D. Crane, Accurate two-dimensional eye tracker using first and fourth Purkinje images, *Journal of the Optical Society of America*, Vol. 63, No. 8, pp.921-928, Aug. 1973.
7. J. G. Daugman, High confidence visual recognition of persons by a test of statistical independence, *IEEE Transactions on PAMI*, Vol. 15, No. 11, pp.1148-1161. November 1993.
8. A. Fitzgibbon, M. Pilu and R. B. Fisher, Direct least square fitting ellipse, *IEEE Transactions on PAMI*, Vol. 21, No.5, pp. 476-480, 1999.
9. D. Forsyth, J. L. Mundy, A. Zisserman, C. Coelho, A. Heller and C. Rothwell, Invariant descriptors for 3-D object recognition and pose, *IEEE Transactions on PAMI*, Vol. 13, No. 10, pp. 971-991, 1991.
10. A. Gee and R. Cipolla, Determining the gaze of faces in images, *Image and Vision Computing*, Vol. 12, No. 10, pp. 639-647, December 1994.
11. T. E. Hutchinson, K. P. White, Jr., W. N. Martin, K. C. Reichert and L. A. Frey, Human-computer interaction using eye-gaze input, *IEEE Transactions on System, Man and Cybernetics*, Vol. 19, No. 6, pp.1527-1533, 1989.
12. A. E. Kaufman, A. Bandopadhyay and B. D. Shaviv, An eye tracking computer user interface, *Proceedings of Research Frontier in virtual reality workshop*, IEEE Computer Society Press, October 1993, pp. 78-84.
13. K.-N. Kim and R. S. Ramakrishna, Vision-based eyegaze tracking for human

- computer interface, *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, 12-15, Oct. 1999, Vol. 2, pp.324-329.
14. Y. Matsumoto and A. Zelinsky, An algorithm for realtime stereo vision implementation of head pose and gaze direction measurement, *Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 499-504.
 15. G. A. Myers, K. R. Sherman and L. Stark, Eye Monitor, *IEEE Computer Magazine*, Vol. 3, pp. 14-21, 1991.
 16. R. Newman, Y. Matsumoto, S. Rougeaux and A. Zelinsky, Real-time stereo tracking for head pose and gaze estimation, *Proceedings of Fourth International Conference on Automatic Face and Gesture Recognition*, pp. 122-128, 2000.
 17. N. Otsu, A threshold selection method from gray-level histogram, *IEEE Transactions on Systems, man and Cybernetics*, Vol. SMC-9, pp.62-66, Jan. 1979.
 18. T. Rikert, M. Jones, Gaze estimation using morphable models, *Proceedings of International Conference on Automatic Face and Gesture Recognition*, 1998, pp.436-441.
 19. R. Safaee-Rad, I. Tchoukanov, K. C. Smith, B. Benhabib, Three-dimensional location estimation of circular features for machine vision, *IEEE Transactions on Robotics and Automation*, Vol. 8, No. 5, pp. 624-639, 1992.
 20. H. S. Sawhney, J. Oliensis and A. R. Hanson, Description and reconstruction from image trajectories of rotational motion, *Proceedings of IEEE International Conference on Computer Vision*, 1990, pp. 494-498.
 21. R. Stiefelhagen and J. Yang, Gaze tracking for multimodal human-computer interaction, *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP-97*, 1997, Vol. 4, pp. 2617 -2620.
 22. K. Talmi and J. Liu, Eye and gaze tracking for visually controlled interactive stereoscopic displays, *Signal Processing: Image communication* Vol.14, pp. 799-810, 1999.

23. Jian-Gang Wang, *Head-pose and eye-gaze determination for human-machine interaction*, PhD thesis, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, 2001.
24. X. Xie, R. Sudhakar, and H. Azhang, on improving eye feature extraction using deformable templates, *Pattern Recognition* Vol. 17, pp.791-799, 1994.
25. L. R. Young and D. Sheena, Survey of eye movement recording methods, *Behavior research methods and instrumentation*, Vol. 7, No. 5, pp.397-429, 1975.