T.R.
USKUDAR UNIVERSITY
INSTITUTE OF SCIENCE

DEPARTMENT OF COMUTER ENGINEERING

MASTER'S DEGREE PROGRAM OF COMPUTER ENGINEERING

MASTER'S DEGREE THESIS

RECOGNISING HUMAN VISION VECTOR IN 3D USING STEREO VISION SYSTEM

AHEBHAMEN SAMUEL OSAHON

Thesis Advisor

Dr. Ihab ELaff

ISTANBUL- 2022

T.R.
USKUDAR UNIVERSITY
INSTITUTE OF SCIENCE

DEPARTMENT OF COMPUTER ENGINEERING

MASTER'S DEGREE PROGRAM OF COMPUTER ENGINEERING

MASTER'S DEGREE THESIS

RECOGNISING HUMAN VISION VECTOR IN 3D USING STEREO VISION SYSTEM

AHEBHAMEN SAMUEL OSAHON

Thesis Advisor

Dr. Ihab ELaff

ISTANBUL- 2022

ABSTRACT

RECOGNISING HUMAN VISION VECTOR IN 3D USING STEREO VISION SYSTEM

Because there is no deterministic technique to discover a face in a given image, human face identification has grown to be a significant area of interest in current research. Furthermore, the algorithms used to recognise faces in photos are quite precise in the types of images they accept as input. Finding faces in the provided, colored class group photo is the challenge. We employ a method that combines established algorithms and heuristics. We can string together several straightforward rejection blocks in order to find faces. The more specifically it can be educated to eliminate non-faces, the deeper the rejection block. Numerous techniques have been tested, including eigenfaces, fisher linear discriminants, maximal rejection, neural networks, template matching, and eigenfaces. Finally, eigenfaces, morphological processes (such as erosion), and skin color segmentation have been combined.

Keywords: Face detection, Face identification, Face recognition.

ABSTRACT

RECOGNISING HUMAN VISION VECTOR IN 3D USING STEREO VISION SYSTEM

Because there is no deterministic technique to discover a face in a given image, human face detection and identification has grown to be a significant area of interest in current research. Furthermore, the algorithms used to recognise faces in photos are quite precise in the types of images they accept as input. Finding faces in the provided, colored class group photo is the challenge. We employ a method that combines established algorithms and heuristics. We can string together several straightforward rejection blocks in order to find faces. The more specifically it can be educated to eliminate non-faces, the deeper the rejection block. Numerous techniques have been tested, including eigenfaces, fisher linear discriminants, maximal rejection, neural networks, template matching, and eigenfaces. Finally, eigenfaces, morphological processes (such as erosion), and skin color segmentation have been combined.

Keywords: Face detection, Face identification, Face recognition.

THANKS TO

First and foremost, I want to thank God Almighty for seeing me through these difficult years. It's been a long and challenging trip, but I'm glad it's come to a rewarding and suitable finale.

I'd want to thank my project supervisor, Dr. Ihab ELAFF, for always being a fantastic mentor and instructor to me. This project would not have been completed without your unwavering support, Dr. Ihab. I will be eternally grateful to you for all of the zoom (online) and face-to-face encounters that shaped the delivery of this thesis. It was a joy and privilege to be closely watched under your supervision, and I pray that God honors you in a variety of ways. In addition, I want to thank the director of the computer engineering department for this amazing chance.

I am grateful to Üsküdar University and all of the faculty members for your help; I could not have achieved without your dedication to providing me with the essential information and resources.

To my parents Mr. Mark Ahebhamen (late) and Mrs. Juliet Ahebhamen, and I am eternally grateful for her love, prayers, sacrifices, compassion, and financial support throughout my life. I am grateful to you, father and mother, and I am confident that with God's aid, I will make you proud. Thank you to all of my siblings Mrs. Cynthia Aghedo, Mr. Moses Ahebhamen and Mr. Shammah Ahebhamen and also to my brother in-law Mr. Vincent Aghedo for your support throughout the years; I am very blessed to have you all in my life.

To my family at large, thank you for your unwavering love and support over the years.

To my wife, Dr. Mrs. Ofure Ahebhamen, it has not been an easy journey, but I am grateful you have stood by me throughout the years. Thank you for standing by me, thank you for all the words of encouragement, and morale support. I pray God grant you all your heart desires and bless you abundantly.

I must appreciate Turkey for opening its heart and inviting me to study in its country.

And to all the friends I've made here in Turkey, you've all been a godsend in some manner, so thank you for sticking with me.

FORM OF DECLARATION

Herewith I declare, that I obtained all the information and documents in this study within the framework of academic rules, presented all visual, auditory, and written information and results in accordance with scientific ethics, did not falsify the data I used, referred to the sources I used in accordance with scientific norms, that my thesis was original except in the cases cited, produced by me and written in accordance with the Thesis Writing Guide of Uskudar University Institute of Health Sciences.

Date: 09/09/2022

Samuel Osahon AHEBHAMEN

Signature

CONTENTS

INDEX OF TABLES

Page

INDEX OF FIGURES

INDEX OF IMAGERY AND ABBREVIATIONS

IST: Information Society Technologies

MV: Machine Vision

ML: Machine Learning

RL: Reinforcement Learning

AI: Artificial Intelligence

PPV: Pars Plana Vitrectomy

2D: Two Dimension

3D: Three Dimension

SLV: Structured-light vision

ASV: Active Stereo Vision

SLS: Structured Light Stereo

BSV: Binocular Stereo Vision

CV: Computer Vision

CMOS: Complementary Metal-Oxide-Semiconductor

SNN: Spiking Neural Network

MVQ: Machine Vision Quality

GUI: Graphical User Interface

NCC: Normalized Cross Correlation

AVI: Automatic Visual Inspection

SVD: Singular Value Decomposition

HPSO: Hyper Particle Scouts Optimization

PSO: Particle Swarm Optimization

DNN: Deep Neural Network

CFR: Conditional Random Field

ANN: Artificial Neural Networks

AFR: Automated Face Recognition

LDA: Linear Discriminant Analysis

HCI: Human Computer Interaction

ISME: Information Science and Management Engineering

IEEE: Institute of Electrical and Electronics Engineering

R-CNN: Region-based Convolutional Neural Network

SSD: Single Shot Detector

SURF: Speeded Up Robust Features

EOG: Electrooculography

LLI: Low-Light Image

RPE: Retinal Pigment Epithelium

LSM: Logistics Sine Map

NCCC: National Computing Colleges Conference

CCTV: Closed-Circuit Television

CCD: Charged Coupled Device

CHAPTER ONE

MACHINE VISION

1.1 INTRODUCING MACHINE VISION.

Early in the 1970s, machine vision technology began to be used in industry, but because of high computer costs, adoption has been slow. The machine vision industry is somewhat fragmented, and tiny technology firms are vying for repeat business in certain industry verticals. Because technology businesses are unwilling to assume all of the associated risks and because end users are accustomed to purchasing tried-and-true technology for their manufacturing facilities, many novel application areas with associated hazards remain unaddressed. The Information Society Technologies (IST) initiative, which is sponsored by the European Commission, has actively supported the transmission, use, and study of machine vision technologies. This marketing has prompted technology providers and end users to collaborate on trial projects and share risks. These solutions have been widely disseminated throughout Europe in parallel, reaching many potential new consumers. (A. J. Soini, "Machine Vision News", *proceedings of Machine Vision Seminar Finnish Society of Automation*, pp. 7-9, 26.-27 5.1998). The technique is called machine vision, and procedures employed to provide applications with automated inspection and analysis based on imaging such as robotics, process control, and automatic inspection guiding, typically within the industrial setting. The term "machine vision" covers a broad range of hardware, software, and technologies items, combined systems, procedures, with knowledge. Machine vision is a subfield of systems engineering, whereas computer vision is a subset of computer science. It makes an effort to use already available technology in creative ways to deal with issues in the real world. The term is usually used for these tasks when industrial automation is involved, but it is also used for related tasks when vehicle guidance is involved. Thanks to machine vision technology, industrial machinery can now "see" what it is doing and make decisions swiftly depending on what it observes. In addition to visual inspection and problem detection, machine vision is most typically utilized for product identification, sorting, and tracking.

Machine vision is one of the essential elements of industrial automation. It has long contributed in improving product quality, speeding up production, and streamlining logistics and manufacturing. The transition to Industry 4.0 is currently being driven by the combination of this tried-and-true technology with artificial intelligence. The manufacturing industry has undergone a cyber-physical shift known as "Industry 4.0."

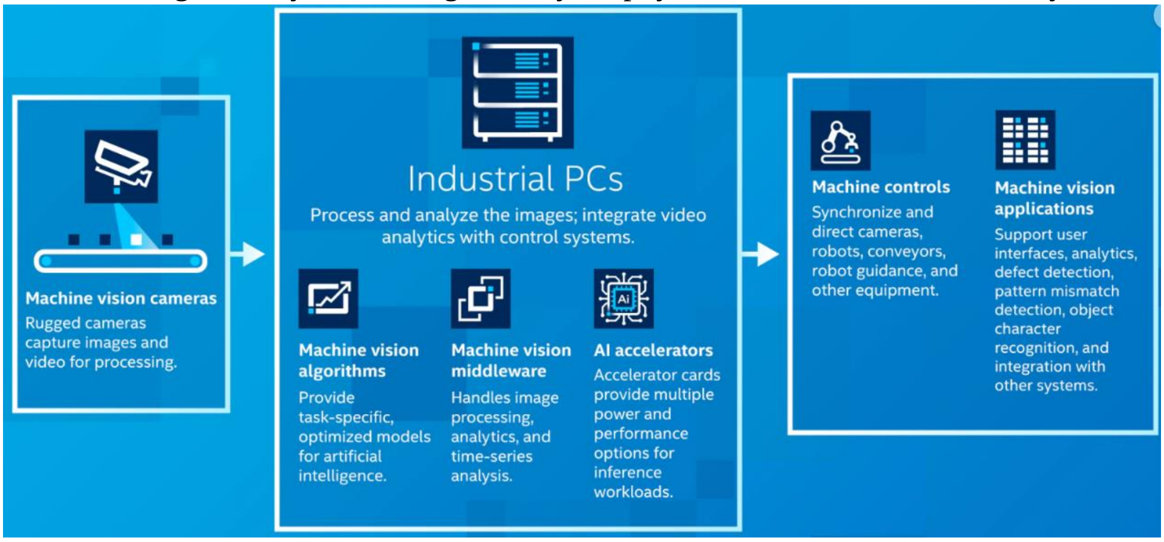


Figure 1.1: Machine Vision System Architecture. (Google search, *http://www.intel.com/content/www/us/en/manufacturing/what-is-machine-vision.html*)

1.1.1 History of Machine Vision

Although the idea of machine vision has been around since the 1930s, some of the most important ideas and discoveries that helped pave the way for the creation of machine vision systems really go back much, much further. The idea of machine vision was first introduced in the 1930s, when New Jersey-based Electronic Sorting Machines offered food sorters based on the use of certain filters and photomultiplier detectors. However, some of the most important discoveries and technologies that influenced the creation of machine vision systems go back much longer. To fully record this, Andy Wilson writes in the Keystone of machine vision systems design from the Vision Systems Design 200th anniversary issue in 2013. One may start by mentioning the invention of early Egyptian optical lens systems dating back to 700 BC, Joseph Marie Jacquard's innovation of punched paper cards in 1801 that allowed a loom to weave complicated patterns mechanically, or Maxwell's 1873 unified theory of electricity and magnetism (Andy Wilson, "Keystone of machine vision systems design", *Vision Systems Design 200th anniversary issue*, 2013). Andy's post delves into the many diverse technologies that have paved the road for today's commercially available machine vision components.

Classic Vision Systems

Before AI and machine learning, machines could "see". In the early 1970s, computers began using certain algorithms to examine images and pinpoint essential components. This traditional machine vision technique can recognize boundaries of an item to place a portion, blobs of linked pixels that suggest a whole, and color variances that indicate defects. Traditional machine vision uses very straightforward processes that don't call for artificial intelligence. Similar to a bar code, text must be clear and concise. Shapes must follow a certain pattern and be predictable. A traditional machine vision system cannot distinguish an apple from an orange, read handwriting, or understand a wrinkled label.

But traditionally, the impact of machine vision has shown significant effect in industry. Machines can detect flaws more quickly and accurately than human sight since they never get tired. Machines are also not constrained by the limitations of human perception. Thermal imaging and X-rays are two techniques that specialized machine vision cameras can utilize to find microscopic defects and metal fatigue.

Inference from Deep Learning and Commercial Machine Vision

Machine vision capabilities are being drastically expanded by A rising array of deep learning AI models, embedded and IoT devices at the network edge, and strong edge computing. The transition to smart factories and Industry 4.0 is being driven by the rapid growth of this capabilities. AI neural network models are used to complement conventional computer vision techniques. When a computer receives an image or a video stream of pictures, machine vision software compares the acquired image data with a neural network model. Thanks to a technique called deep learning inference, computers can now detect tiny disparities like tiny fabric pattern mismatches and tiny circuit board flaws.

For certain purposes, data scientists use tailored neural network models to boost accuracy and speed. A computer analyzes tens of thousands of samples in this process, which is referred to as supervised training, and identifies important patterns, even ones that a person may overlook. There are models for identifying pulled threads in fabric, voids in welds, and dead and off-color pixels in displays. Of course, new and improved models are always being created.

Smart Machine Vision and Autonomous Systems

Machine vision is now being used for much more than just visual inspection and quality controlWith the use of sophisticated machine vision, robots may sense in three dimensions, hold components for one another, and examine one another's work. They can even communicate with their human employees to ensure that they collaborate safely.

Intelligent machines can scan labels and understand signs using natural language processing. Robots with advanced vision systems are able to calculate quantities, identify forms, and pack trucks, shipping containers, and other containers exactly and with the least amount of waste. This shift from robots that can automate straightforward tasks to autonomous machines that can see beyond what the human eye can perceive and reason for themselves to optimize elements over longer times will spark new levels of industrial innovation.

Smart machine vision, which may seem like science fiction, is already in use in factories, warehouses, and shipping centers. It helps human employees by doing the menial tasks so they can focus on the more important ones.

1.1.2 How a machine vision system works.

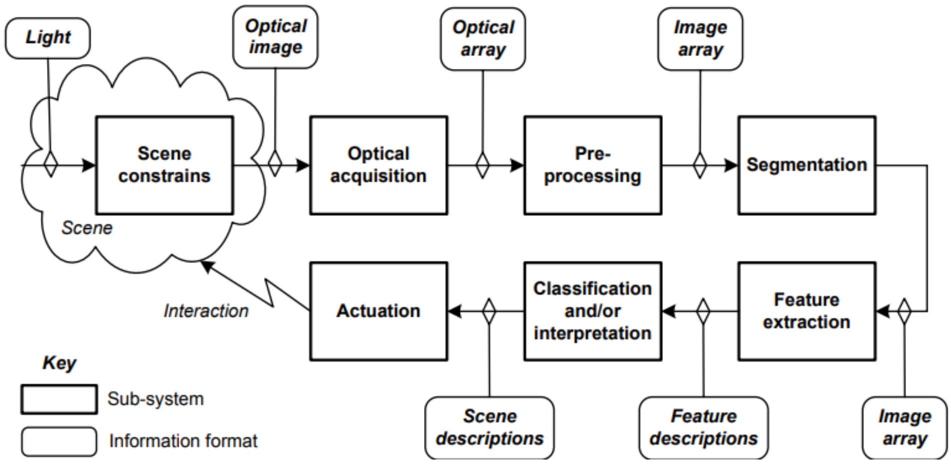


Figure 1.2: Simple block illustration of how a vision system typically operates (*http://www.sciencedirect.com/science/article/abs/pii/S0736584507000233*)

A visual system is capable of gathering images, analyzing them, and recognizing an item or objects among a set of things. As seen in Fig. 1.2, a source of light illuminates the landscape, and image sensors provide an optical representation of the environment. The process of producing an optical picture that can be converted into a digital image using a photo detector is known as image acquisition. Image sensing, image data representation, and digitization are all steps in this process. A digital image's pixel values are modified and prepared during image processing to provide a more acceptable shape for subsequent processes.

1.2 SCOPE OF APPLICATION.

The primary goal of machine vision (MV) is to comprehend the high level content of digital pictures used for a variety of activities, including object detection, object categorization, the finding of geometrical structure, and the exploration of numerous connections between diverse things. As an alternative, MV models the real world utilizing picture and video receptions and visual sensors to identify the surroundings and estimate the attributes.

Research in multimedia retrieval and computer vision is at the cutting edge when it comes to higher dimensional data, such video and 3D. A thorough review and important insights into the current state of the art for higher dimensional features from deep learning as well as conventional methods in this survey is being provided. Present-day methods usually make use of 3D data from sensors or 3D modeling and comprehension of the 3D world (Georgiou T, Liu Y, Chen W, Lew M, "A survey of traditional and deep learning-based feature descriptors for high dimensional data in computer vision". *Int J Multimed Inform Retrieval* 9(3):135–170. 2020).

Understanding the high-level information of pictures that have been received from multiple visual sensors is the initial step in tackling numerous problems in real-world applications, such as the recognition of pedestrians or roads in autonomous driving cars.

Alternatively, specific decision-making problems can be resolved by extracting data from the captured photos about the surrounding environment. As a result, the effectiveness of various judgments based on the resultant comprehension of the environment must be used to measure the effectiveness of picture understanding technology. These decisions are in turn influenced by the system's past action results. As an illustration, a robot's navigational decisions are dependent on its estimated localization, which is the result of a particular MV job called vision-based positioning estimation.

The use of machine learning methods for MV tasks like image registration, image segmentation, image classification/retrieval, three-dimensional image reconstruction, object detection and object tracking, etc. is currently one of the key driving technologies for high-level understanding of the environment from acquired images. These kinds of tasks have been conceptualized as learning issues, and machine learning (ML) is crucial and pervasive for the extraction of patterns or regularities from pictures. As a result, ML is now a crucial part of MV. An emerging ML technology called reinforcement learning (RL) involves learning by interacting with the environment. It also enables accounting for decision outcomes and subsequent actions resulting from the resolution of the related MV tasks (Utton R.S., Barto A.G. "Reinforcement learning: an introduction." *MIT Press, Cambridge*. 2018).

The foundation of smart operations, logistics, and manufacturing is industrial machine vision. Every stage of the production process may benefit from the intelligence, analysis, and efficiency that machine vision cameras, embedded IoT sensors, and industrial PCs can provide.

Machine vision systems are often used to check silicon wafers, microchips, and components including resistors, capacitors, and lead frames in the manufacturing of semiconductors.

In the automobile sector, machine vision systems are used for engine block inspection, painted surface inspection, welding quality control, fast prototyping, and the identification of component issues. The following can be checked when evaluating products and quality control procedures: the presence of components (screws, cables, suspension), regularity of assembly, proper execution and placement of holes and shapes (curves, circular areas, perpendicular surfaces, etc.), correct selection of equipment options for the implementation of the quality of surface markings (manufacturer's numbers and geographical detail), geometrical dimensions (with an accuracy of a single micron), and other factors (location and colour).

Machine vision can be applied and used in the following ways and different fields including:

- Biometrics, is the most reliable automated personal identification technique in demand in a society that is becoming more and more computerized is biometrics. A biometric system is an automated way to confirm or identify a live individual based on a physiological trait, such as a fingerprint or iris pattern, or a behavioral trait, such as handwriting or keystroke patterns (B. Miller, "Biometrics: Vital signs of Identity", *IEEE spectrum*, 1994).
- placement
- extensive industrial production,

- Small-lot production of one-of-a-kind items
- safety measures used in workplaces
- middle examination (e.g. quality control)
- control over mobile robots that can operate on their own, industrial management systems, counting, bar-code scanning, and storage interfaces for digital systems; visual oversight of warehouse inventories
- the monitoring of food products' purity and quality,
- use of bridges
- automation of retail
- railroads are exploited
- agriculture
- assistance for the eyesight of the blind (Artificial Visual Sensing) (For instance, Artificial Eye System, Super Vision System).

1.2.1 Benefits of Machine Vision

1. **Machine vision in manufacturing:** Throughput of your manufacturing line can be increased, labor expenses can be decreased, and personnel can be freed to concentrate on higher-value tasks thanks to machine vision technology used in manufacturing.
   For instance, a car manufacturing company that combined data modeling and machine learning technologies in collaborations with Intel as well as Nebbiolo Technologies into weld inspection and important quality-control processes saw an increase in the number of welds analyzed daily, a decrease in labor costs in their factories, and the ability to transition to greater preventative monitoring, rather than merely responding to problems. Additionally, machine vision offers continuous quality assurance checks on product contents, packaging, and labeling in strictly regulated industries like pharmaceuticals. Machine vision can be used in supply chains to automatically scan and monitor things at every stage of the operation, giving you a precise, real-time inventory account.
2. **Machine vision in operations:** The enhancement of employee health and safety is one important benefit of adopting machine vision in operations. AI-powered computer vision can make sure that employees are keeping their distance from one another and wearing the appropriate safety gear. Robots and other devices having machine vision capabilities can interact with humans and analyze their behaviors to help stop accidents before they happen. They can automatically shut down equipment or alert the operator if a condition is unsafe, lowering risk for your staff and your business.
   By continually analyzing data from cameras, microphones, and other embedded sensors in industrial machinery and equipment, industrial PCs may also utilize AI to identify problems and wear signs prior to failure. This makes it possible to schedule preventive maintenance in advance, avoiding unanticipated downtime and spreading maintenance expenditures out over time.
   To guarantee correct usage and storage in asset management and security, AI (artificial intelligence) can recognize and track things in video feeds. It can

also inform management if assets breach a designated boundary. Security camera systems have the potential to become proactive security partners that can control building access and identify dangerous circumstances.

1.2.2 General Benefits of Machine Vision

The use of machine vision systems has increased throughout time; although appearing to be a straightforward technology, it is actually highly complex. A machine vision system, at its most basic level, is a computer that utilizes one or more video cameras and digital signal processing to perceive its surroundings. Real-world issue solving is the intended outcome, but a vision system is capable of much more.

1. **Eliminates Human Errors**: although extraordinary, the human eye is not perfect. Machine vision can successfully estimate product quantities because of its precision, reproducibility, and speed, even though vision is best for qualitative interpretation. A vision system installed on a production line, for instance, may examine hundreds or even thousands of parts every minute. Machine vision systems can check item features that are too tiny for the human eye to see when combined with high-resolution cameras. Additionally, operator tiredness and differences between individual operators may be eliminated by machine vision. These characteristics can significantly lower, if not completely eliminate, the danger of defective or dismantled items.
2. **Lower Cost:** a vision system can increase production efficiency and reduce the amount of manpower required to run the machinery. It can also lower the scrap rate, which will result in less material waste and, consequently, lower your overhead. Tolerances apply to all dimensions regardless of how pieces are made. Processes and outcomes may be improved with the use of machine vision. Precision manufacturing is especially important for expensive components because a failure might end up costing thousands of dollars. Implementing the checks and balances provided by a vision system is consequently essential.
3. **Reduces Downtime:** A vision system protects against component damage by eliminating physical touch between a test system and produced parts. As wear and tear take their toll on mechanical components, it also cuts down on the time and costs associated with repairs. Machines operate more quickly as they require less care, allowing you to consistently and easily achieve production targets.
4. **Increases Throughput:** This idea builds on the one before it in that less downtime always results in higher throughput. A vision system, however, develops this idea a little farther. Approximately one second quicker than even skilled operators, it can send correction orders. This lessens the need for manual system correction and aids in maintaining production levels.
5. **Enhances Safety:** Machine vision makes the industrial process less dependent on humans, which makes the workplace more secure. Operators of large, powerful machinery are less likely to sustain injuries. Workers' exposure to dangerous components and materials is decreased while also being prevented from contaminating clean rooms.

6. **Identifies Print Defects:** Detecting printing irregularities, such as improper color tones, imperfect prints, or missing characters, may be challenging at best. However, a vision system can easily complete this duty. First, the system receives a master or golden picture. Then, all produced components are compared to that picture. Any departure from the master is signaled for rectification right away.
7. **Takes Accurate Measurements:** This application entails precisely figuring out an object's measurements. The diameter, radius, distance, and depth of an image may all be found and measured by a vision system. This application might be used to measure the liquid fill level in a container or determine the inner diameter of an engine cylinder bore for practical applications. Both 2D (2-Dimension) and 3D (3-Dimension) cameras are capable of gathering this info.
8. **Detects Flaws:** Machine vision can identify product anomalies like surface dents and scratches, as you might have previously assumed. To discriminate between "acceptable" and "unacceptable" defects, it's important to apply detection boundaries properly. Fortunately, machine vision is perfect for these jobs since it relies on examples rather than strict rules.
9. **Locates Objects:** In applications like robotic navigation, locating objects is frequently done using a vision system. To provide an example, the objective may be to find the coordinates or location of a specific object. The object is then picked up or another location-specific action is carried out using this information. Machine vision is versatile and has a variety of uses in industrial operations. The secret is in picking the appropriate technology for your industry. Thankfully, you don't have to make this choice by yourself. We can assist in both streamlining your requirements and easing the move to a vision system.

1.3 MONOVISION.

As defined, Presbyopia, sometimes known as "aging eye," is a condition where it becomes difficult to focus properly on close objects as a result of normal aging of the eyes. According to the American Academy of Ophthalmology, monovision or blended vision can be used as a remedial treatment. Naturally, it gets tougher as you get older to read a newspaper, a menu, or to concentrate on a smartphone screen. Even while readers or magnifying glasses can help you handle such vision issues, having to put them on every time you want to read might be inconvenient. For those who desire to age without reading glasses, monovision is a possible alternative.

Presbyopia can be treated with monovision, a technique in which one eye is adjusted for distance and the other for close work. It provides decreased or perhaps complete independence from the need for spectacle correction for presbyopes. We present a patient whose unilateral pars plana vitrectomy (PPV) adaptation to monovision was not noticed until it was reversed with cataract surgery. (Zia I. Carrim, Nicholas M. Hickley, and Fiona Bishop", Monovision: a refractive consideration in cataract surgery after vitrectomy", *International Ophthalmology* volume 32, pages623–625. 2012).

Figure 1.3: Showing a pair of eyes experiencing imbalance needed to be corrected by Monovision. The imbalance called Presbyopia (*https://drtcarlson.com/eye-health/what-is-monovision/*).

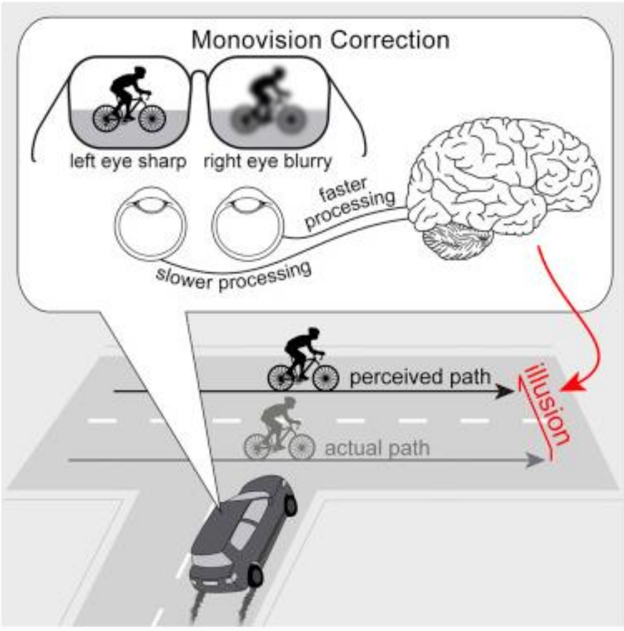


Figure 1.4: Monovision Correction of illusion caused by Presbyopia. (Johannes Burge, Victor Rodriguez-Lopez, and Carlos Dorronsoro. "Monovision and the Misperception of Motion." *Science Direct Current Biology* Volume 29, Issue 15, 5 August 2019, Pages 2586-2592).

1.3.1 HOW MONO VISION WORKS

Refractive surgery, intraocular lenses, contact lenses, and artificial lens implants can all be used to create monovision (like LASIK). It functions under the assumption that many people have a "dominant" eye, or an eye that sees better when the other is closed.

Given that studies have linked handedness with eye dominance, that is, righties tend to have their right eyes as the dominant eyes, and vice versa, with some possible exceptions, determining eye dominance is simple. If you can write with various hands, for instance, it's more difficult to establish eye dominance. Monovision allows a person to view objects up close by correcting the dominant eye's vision for distance while leaving the other eye nearsighted. Still working together, the eyes provide clearer vision at all distances. Since the dominant eye's visual input is more concentrated in the brain, monovision is effective.

In order to improve near vision, a near-powered lens is placed in the non-dominant eye, far vision isn't noticeably affected.

Even while monovision may seem challenging at first, most people quickly get used to it. When both eyes are open, it is typically impossible to distinguish between the eyes that has been configured for close or distance vision.

1.3.2 MONOVISION'S CONTRADICTIONS.

People often need a week or two to become used to monovision. Some people, though, can't fully adjust. Because of this, it is not a general remedy for presbyopia. To avoid making a more permanent decision, such as monovision LASIK surgery, eye experts advise trying out monovision contact lenses first.

An augmentation operation on the eye that was left for close vision can be used to undo the effects of monovision LASIK if the patient finds it difficult to adjust.

Along with issues with adaptation, monovision can impair depth awareness (3D visual perception and distance of objects). For tasks like reading small print or driving at night, you may still need reading glasses on occasion. You may find it challenging to participate in performance sports, drive at night, or conduct other close-up tasks if you have monovision.

The technique might also be preferred differently depending on a variety of criteria, such as professional preferences. Monovision, for example, might be perfect for a retiree of 50 but not as suitable for a driving instructor of the same age.

1.4 STEREO VISION.

Computer stereo vision is the process of extracting 3D information from digital pictures, such as those created by a CCD camera. By analyzing information about a scene from two separate angles and comparing the respective intensities, it is feasible to extract 3D information positions of things in the two panels.

The biological process of stereopsis is comparable to this. (Wikipedia Free Encyclopedia, *http://en.wikipedia.org/wiki/Computer_stereo_vision*). Stereopsis, or binocular vision, is the process through which the brain receives visual stimuli from both eyes at once to create the illusion of depth.

To record two unique viewpoints of a picture in a manner similar to how people perceive while their eyes are closed, traditional stereo vision employs two cameras that are split horizontally from one another. These two photos may be compared to extract the relative depth information and generate a disparity map, which encodes the difference in the horizontal coordinates of the corresponding picture spots. The scene depth at the relevant pixel point is inversely proportional to the values in this disparity map.

The right image from the right camera must be the right picture from the right camera must be displayed to the observer's right eye, and the left image from the left camera a stereoscopic device for a human to compare the two views.

Applications for 3D stereo displays include entertainment, information transmission, and automated systems. In disciplines like robotics, stereo vision is crucial for determining the relative positions of 3D objects close to autonomous systems. Object recognition is one of the additional uses of robotics (Sumi, Yasushi, et al. "3D object recognition in cluttered environments by segment-based stereo vision." *International Journal of Computer Vision* 46.1 (2002): 5-23).

1.4.1 Processes of StereoVision.

A computer vision system needs to go through a number of pre-processing processes.

1. The image must first be corrected for distortion, including tangential and barrel distortion. This guarantees that the picture produced by an ideal pinhole camera matches the image that was observed.
2. Picture rectification, which is required to compare the image pairs, involves projecting the image back into a common plane.
3. A comparison of the two images-based information measure is minimized. This generates a disparity map and provides the most accurate estimate of the location of features in the two photos.
4. It is also possible to project the received disparity map into a three-dimensional input image. The cameras' projective specifications may include used to compute the point cloud in a way that yields measurements at a known scale.

The study of computer vision is a significant area of research right now. It incorporates techniques including picture capture, manipulation, comprehension, and analysis (R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010). The goal of computer vision techniques is to mathematically represent a complicated visual world. One of the goals of computer vision is to reconstruct the qualities of the environment we perceive, such as its lighting, form, and color distributions, based on one or more photographs. The reconstruction of the three-dimensional coordinates of points for depth estimation is a significant research issue, and it is one that stereo vision, a subfield of computer vision, attempts to solve. A stereo camera, or two horizontally positioned cameras, makes up a stereo vision system (i.e., one on the left and the other on the right). The recovery of visual depth information is then performed on the two pictures that were simultaneously taken by these cameras (B. H. Bodkin, *Real-time mobile stereo vision [M.S. thesis]*, University of Tennessee, 2012). The difficulty is in figuring out the most accurate way to approximate the variations between the perspectives depicted in the two photographs in order to map (i.e., plot) the correspondence (i.e., disparity) of the environment. The equivalent pixels that are horizontally moved between the left picture and the right image are represented by a disparity map, intuitively. Every year, new approaches and strategies are created to address this issue, and they all show signs of increasing in precision and efficiency.

1.4.2 VISION HULL

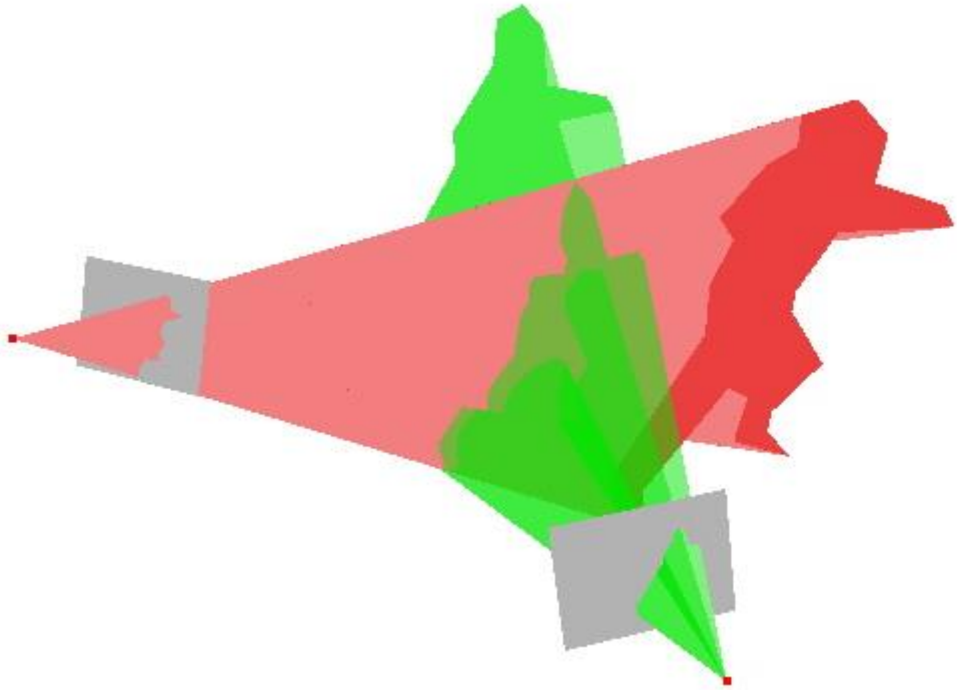


Figure 1.5: Vision Hull (*From Wikipedia, the free encyclopedia***)**

A. Laurentini's shape-from-silhouette 3D reconstruction technique produces a geometric object called a visual hull. This method is predicated on the idea that the foreground and background of an image may be distinguished.

The 2-D outlines of the item serve as the foundation for several algorithms that detect and recreate 3-D objects. In general, adopting a silhouette-based strategy to identify a nonconvex object entails ignoring some surface characteristics that may serve as identification cues. By employing numerous silhouettes of the item, volume intersection algorithms cannot rebuild the same characteristics. The challenge of determining which nonconvex object features are important for comprehending silhouette-based images is dealt with in this study. The geometric idea of a 3-D object's visible hull is presented for this purpose. This is the item that can be swapped for object S without changing any silhouettes and is the closest approximation of object S that can be achieved using the volume intersection technique (A. Laurentini "The visual hull concept for silhouette-based image understanding". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. pp. 150–162, February 1994).

3D reconstruction is the process of recreating the form and appearance of genuine items in three dimensions. There are active and passive ways to complete this task.

1.4.3 Active StereoVision

When an observer is actively controlling the geometric parameters of the sensory equipment, they are said to be active. The activity's goal is to alter the limitations underlying the observed events in order to raise the standard of the perceptual outcomes.

Examples of what we mean by a "active observer" include a monocular observer who travels with a known or unknown velocity and a binocular observer who can rotate his eyes and monitor nearby things (J. Aloimonos,, I. Weiss, and A. Bandyopadhyay,, "Active vision", *Intl J. Computer Vision*, vol. 1, no. 4, pp. 333-356, 1988). Automated robot control presents a tough problem called visual navigation. Automatically detecting and avoiding obstacles is essential for many robot applications, including autonomous mobility and item handling in dangerous areas. The capacity to identify clear space corridors along the robot's track is crucial in this situation and necessitates complex visual processing. The majority of the time, it is feasible to benefit from the cameras' active control (F. Ferrari,, E. Grosso,, M. Magrassi, and G. Sandini,, "A stereo vision systemfor real time obstacle avoidance in unknown environment", *Proc. Intl Workshop Intelligent Robots and Systems*, 1990-July).

The description of a novel control architecture for mobile robots. The robot can function at progressively higher degrees of proficiency thanks to layers of the control system. Asynchronous modules that interact over narrow channels make up layers. Each module is an example of a rather straightforward computational machine. Lower-level layers' functions can be absorbed by higher-level layers by suppressing their outputs. But as higher levels are introduced, lesser levels continue to work. A reliable and adaptable robot control system is the end product. The system has been used to manage a mobile robot that roamed freely across computer rooms and labs (R.A. Brooks,, "A robust layered control system for a mobile robot", *IEEE Trans. Robotics and Automation*, vol. 2, pp. 14-23, Apr. 1986).

In order to make the stereo matching problem simpler, active stereo vision uses a lighting source, such as a light detection and laser. Passive stereo vision is the alternative word.

1. Conventional structured-light vision (SLV)

A laser or modulated light, for example to identify correspondences between projectors and cameras.

2. Conventional active stereo vision (ASV)

Similar to passive, active stereo vision (ASV) often employs a structured light or laser to produce stereo images. but only performs stereo matching for camera-camera correspondences.

3. Structured-light stereo (SLS)

There is a combination technique that applies both projector-camera as well as camera-camera correspondences.

1.4.4 Application of StereoVision

Applications for 3D stereo displays include entertainment, information transmission, and automated systems. In disciplines like robotics, stereo vision is crucial for determining the relative positions of 3D objects close to autonomous systems. Robotics can also be used, in this depth information makes it possible for the system to distinguish between occluding picture components that the robot may not otherwise be able to identify as a separate thing by any other criteria for object recognition, such as one chair in front of another. Binocular Stereo Vision (BSV) system importance has increased recently with the advancement of industrial intelligence in a number of application domains. (R. Anchini, G. Di Leo, C. Liguori and A. Paolillo, "Metrological characterization of a vision-based measurement system for the online inspection of automotive rubber profile", *IEEE Trans. Instrum. Meas.*, vol. 58, no. 1, pp. 4-13, Jan. 2009). Existing BSV systems' location accuracy, however, falls short of meeting all industry standards. This is due to the fact that the majority of modern BSV systems are constructed using engineers' past experiences, which results in poor location accuracy. Because of this, more and more researchers are concentrating on creating the best BSV system possible, whose high accuracy can fulfill the stringent standards (G. Di Leo and A. Paolillo, "Uncertainty evaluation of camera model parameters", *Proc. 2011 IEEE Conf. Instrum. Meas. Technol.*, pp. 1-6, 2011).

It is necessary to study the following two relationships in order to create a BSV system with the highest accuracy possible:

1. How each parameter and the potential inaccuracy are related.
2. The connections between each parameter.

Digital stereo vision can be used in science to extract data from aerial surveys, calculate the derivation of geometry enabling 3D building mapping, and even contour maps. It can also be used to calculate 3D heliographical data, like that from the NASA STEREO experiment, or for photogrammetric satellite mapping.

1.5 LITERATURE REVIEW.

1. Hui Chen and Wen Cui (Hui Chen, Wen Cui 2020) study provides comparative analysis between active structured light and multi-view stereo vision technique for 3D reconstruction of face model surface. The 3D face reconstruction methods of active structured light and multi-view vision - based are contrasted. By illuminating a facial model with coded structured light and obtaining multi-view photos, the comparison is carried out. Depending on the environment, the final point clouds from each sensor are compared in terms of measurements in the scanned items to decide which sensor is most appropriate. This study has little limitation, the high execution time required to compute the models was one thing that might be learned from the trials. More input photographs must be collected in order to produce a more accurate point cloud, however the image numbers are uncertain.
2. Min Young Kim, Shirazi Muhammad Ayaz, Jaechan Park, YoungJun Roh (Min Young Kim, Shirazi Muhammad Ayaz, Jaechan Park and YoungJun Roh 2014), this study provides a knowledge on adaptive 3D sensing system based on variable magnification using stereo vision and structured light. This study discusses the design and construction of a 3D sensing system with configurable magnification. To gather high-resolution data from a far-off object, the suggested system comprises of a stereo camera with an illuminating projector and zoom lenses. The design and execution of the suggested 3D sensor unit with adjustable amplification based on stereo vision and structured light are described in this study. To obtain high resolution data from 3D objects, this system includes two digital cameras with zoom lenses.

3. Melvyn L. Smith, Lyndon N. Smith, Mark F. Hansen (Melvyn L. Smith, Lyndon N. Smith and Mark F. Hansen 2021) study provides the quiet revolution in machine vision - a state-of-the-art survey paper, including historical review and perspectives.
According to the literature, machine vision systems are employed in several sectors for a variety of tasks. The greatest of them is functional testing and certification, which is predicted to expand at a 9.5% annual pace during the following four years (Machine Vision Systems, 2020). Additionally, new industries are emerging, with an increasing need for outside applications especially due to deep learning's increased capabilities. Due to these changing uses, it is predicted that the market for machine vision systems would expand quickly, from $8.6 billion in 2020 to $17.7 billion by 2027 (Machine Vision Systems, 2020). This will increase demand for increasing efficiency and additional features. However, it is also evident from the research that there are still issues, thus it would seem acceptable to try to rank them as potential areas for further study. First, deep learning requires a lot of training data, often labeled data. There may not always access to this and there might not be enough publicly available data to train CNNs. As we've seen, very few researchers really train deep CNNs from start. Instead, a lot of people employ networks that have already been trained on vast amounts of picture data, and then they use the training weights as feature extractors on smaller datasets in the solution domain.
4. Pengcheng Wei, Xiangping Yu, Zhenpeng Di, Xiaojun Dai, Bo Wang, Yushan Zeng (Pengcheng Wei, Xiangping Yu, Zhenpeng Di, Xiaojun Dai, Bo Wang and Yushan Zeng 2022), this study investigates the robot automated navigation model under computer intelligence algorithms and machine vision. Design of Robot automatic navigation under computer intelligent algorithm and machine vision.
In order for robotic systems to best serve all spheres of society, this project. Given the high cost and limited work flexibility of intelligent robots today, this study creatively investigates and enhances the image processing algorithm and control algorithm. The Canny algorithm is coupled with the classic ant colony method to improve it, and a Canny-based ant colony algorithm is presented to identify the trajectory edge. This is done in the navigational line edge detection step. This work does, however, have certain shortcomings. For instance, the navigation line border in the photographs that were gathered is clearly visible when using the image processing approach in this study. Despite the image processing algorithm's excellent efficiency, when the lighting is bad and there are too many distractions, the guidance line may be lost.
5. Tiejun Huang, Yajing Zheng, Zhaofei Yu, Rui Chen, Yuan Li, Ruiqin Xiong, Lei Ma, Junwei Zhao, Siwei Dong, Lin Zhu, Jianing Li, Shanshan Jia, Yihua Fu, Boxin Shi, Si Wu, Yonghong Tian (Tiejun Huang, Yajing Zheng, Zhaofei Yu, Rui Chen, Yuan Li, Ruiqin Xiong, Lei Ma, Junwei Zhao, Siwei Dong, Lin Zhu, Jianing Li, Shanshan Jia, Yihua Fu, Boxin Shi, Si Wu and Yonghong Tian), in this study, this paper highlights the drawbacks and limitations of digital cameras and recommends advanced and better alternative techniques. The significant drawbacks of digital camera is that they are unable to capture the quickly evolving photonic environment due to the picture and video formats they have inherited from film cameras. Here, we introduce vidar, a bit sequence array that records and reconstructs scene radiance at any time by using each bit to indicate if the accumulation of photons has surpassed a threshold. We have created a vidar camera that is 1000 times quicker than traditional cameras by using just consumer-level complementary metal-oxide-semiconductor

(CMOS) sensors and integrated circuitry. By considering vidar as spike trains in biological vision, we have further created a spiking neural network (SNN)-based computer vision (CV) system that combines the speed of the machine with the mechanism of biological vision, delivering high-speed object identification and tracking 1000 times quicker than human vision. By displacing video with vidar, vidar cameras will put camera advancement back on track, having realized the innovation capacity of opto - electronic advanced technologies which has been repressed for generations. By displacing conventional recording devices almost in every fields, specifically for high-speed scenes, vidar cameras will also begin a revolution in the camera industry.

Since optical time and space changes are a natural input for SNNs, the heart of Vidar is a spike stream that captures these changes. In the age of artificial intelligence, Vidar will be crucial as a new generation of the machine vision eye.

6. Safouane El Ghazouli, Alain Vissiere, Loius-Ferdinand Lafon, Mohammed-Lamjed Bouazizi, Hichem Nouira (Safouane El Ghazouli, Alain Vissiere, Loius-Ferdinand Lafon, Mohammed-Lamjed Bouazizi and Hichem Nouira 2022) study provides the improved effect and impact of advanced calibration machine vision in industrailization is being discussed. Industry 4.0 is using camera-based imaging technologies for actual inspection of big mechanical part manufacture. It results in taking preventative measures throughout the process of production and then fabricating mechanical parts accurately and within the required tolerances. Consequently, a prior calibration procedure is required for the usage of camera-based scanners. It entails calculating the intrinsic and extrinsic parameters necessary to connect the 3D world point to its picture plane projection.

The suggested ML-approach is a Z-method optimization technique based on the recognition of ideal positions. It consists of two steps: (1) multidimensional polynomial regression, which is used to determine the association among extrinsic parameters and reprojection errors, and (2) PSO minimization, which entails locating the polynomial's lowest values under both linear and nonlinear restrictions. The effectiveness of the ML-approach was quantified utilizing artificial samples. The measurements obtained by calibrating a virtual camera with a generated pattern were matched to the intrinsic and extrinsic characteristics of a specified pinhole. If compared to the recently published CW- and R-method, the resultant parameters are more accurate when the suggested ML-approach is used.

The ML-approach has been shown to be reliable for camera calibration in settings with regulated illumination.

However, it might take a lot of time, especially when a huge learning base is adopted, which serves as one of its great limitations.

7. Chenglin Yu, Hailong Pei (Chenglin Yu and Hailong Pei 2021), in this study Face recognition framework based on effectiveness computing and adversarial neural network and its implementation in machine vision for social robots, a face recognition system is developed using the neural computing paradigm and the neural network concept. The experimental findings demonstrate the high detection accuracy and speed of processing of the suggested technique.

This article examines how, as a result of recent advancements in computer vision technology, target and object identification accuracy has increased dramatically. In the realm of computer vision, whose applications include mobile payments, secure cities, criminal investigation, and other areas, face recognition is one of the key study areas. Manual feature extraction is required for traditional face recognition techniques. The selected features are time-consuming, labor-intensive, and heavily influenced by subjective factors. The most significant innovation in machine vision right now is deep learning. It can automatically extract more key elements from a face picture than more conventional techniques for recognizing faces. The outcomes of the experiment demonstrate how well this technology can identify the condition of the face. This technique may also distinguish when the face posture changes. The limitation is that, it can only work on scanty and small data set,which hopefully would be improved on so it can take bigger and larger data set to confirm the reliability.

8. Rajiv Soundararajan, Soma Biswas (Rajiv Soundararajan, Soma Biswas 2018), this paper discusses Machine vision quality assessment for robust face detection discusses face detection algorith in respect to machine vision quality (MVQ). Face detection algorithms' evaluation of picture quality using machine vision quality (MVQ) is referred to as MVQ. Contrary to perceptual image quality, which often rely on human observers, MVQ is determined by elements like the face detection algorithm and performance metrics like recall and accuracy. By combining the expected precision and recall of the face identification algorithm with picture attributes based on natural scene statistics, we construct the MVQ index as a measure of the method's performance. With the ability to tune the enhancement operations for either perceptual quality or MVQ, a filter bank architecture is created to achieve image enhancement of distorted pictures for face detection. It is demonstrated how MVQ may be utilized to efficiently manage both comparative recall and accuracy efficiency that can be attained on improved pictures. Focusing on MVQ rather than subjective qualities can boost the accuracy of some face detection systems. The image pre - processing system is evaluated on a dataset built using the IDEAL-LIVE Distorted Face Database, and the MVQ is produced for three distinct face identification techniques. It was demonstrated that MVQ could be set to balance the recollection and accuracy that are attained when perceived quality only reaches a particular level just on PR curve from the points of view of image enhancement for enhancing face detection. By directly reflecting the relative gain in accuracy and recall brought on by image augmentation, the MVQ index achieves this. The conclusion is that the machine vision performance indexes created are more applicable to picture augmentation for face detection than perceptual quality indices in this situation. It also demonstrates that, when compared to photos upgraded using perceived quality measures, the quality improvement on images enhanced owing to MVQ was superior to the DPM-based face detection system. The performance loss in unrestricted face detection caused by obstruction, attitude, coverings, and eyewear, which differs from picture quality degradation processes, is a limitation of our study.

9. Min Yong Kim, Shirazi Muhammad Ayaz, Jaechan Park, Young Jun Roh (Min Yong Kim, Shirazi Muhammad Ayaz, Jaechan Park and Young Jun Roh 2014) this work, Adaptive 3D sensing system based on variable magnification using stereo vision and structured light provides examination results on the develpment and construction of a 3D sensing system using structured light and dual vision, this study examines the development and construction of a 3D sensing system with configurable magnification. To gather high-resolution data from a far-off object, the suggested system comprises of a stereo camera with an illuminating projector and zoom lenses.

Along with the creation of a zoom lens calibration technique for a stereo camera based on linear and non-linear models, a calibration target appropriate for zoom lens calibration is also developed. The reliable zoom lens control scheme is designed to provide high-quality photos and is built on accurate automation hardware and image processing methods. Additionally, stereo matching of the coded pictures is done in order to get the results of the reconstruction using a linear equation approach, and normalized cross correlation (NCC), which is resilient to intensity offsets and contrast shifts, is employed for this. According to the experimental findings of this study, the suggested 3D sensing system's integration of NCC with structured light and epipolar geometry further produces accurate 3D reconstruction results. Using optimum fitting, the intricate link between each camera's parameters and variable zoom is discovered. To confirm the calibration of a zoom lens, linear interpolation is utilized to estimate calibration data for a given zoom level. The 3D reconstruction with variable zoom demonstrates that a more precise 3D sensing system is produced by using lenses with increased concentration. The suggested 3D sensing system with adjustable zoom is designed and implemented in this study. It is built on structured light and stereo vision. To collect high-resolution data from 3D objects, this system consists of two digital cameras with optical zoom. The creation of the stereo calibration technique employing linear and non-linear models with variable zoom is part of this study. A linear regression technique was utilized for stereo matching the coded images to get the restoration findings because standardized comparator is resilient to intensity offsets and contrast alterations.
10. H. Golnabi, A. Asadpour (H. Golnabi and A. Asadpour 2007). This study, Design and application of industrial machine vision systems provides the function and significance of vision systems for machines in industrial uses are discussed in this study. It is discussed how to first comprehend the vision in terms of a general idea. A general machine vision model is described and system design technique is explained. The systems and sub-systems that make up such a computer rely, of course, on the types of applications and tasks that must be performed. In general, expected functions from a vision machine include exploiting and imposing the environmental factors of a scene, taking pictures, analyzing those pictures, identifying specific items and traits in each picture, and starting follow-up actions to accept or reject the related objects. A vision system will have virtually finished the task at hand after it has performed all these steps. Here, it is detailed how each system and its related subsystems work in order to produce high-quality photographs. When the machine is operating in a scene with constraints, the first phase is picture capture, followed by pre-processing, segmentation, feature extraction, classification, inspection, and ultimately actuation, which is an interaction with the scene under investigation. Industrial image vision applications are thoroughly detailed towards the conclusion of this research. These applications play a significant role in robotic navigation and control, process control, components recognition, and automatic visual inspection (AVI). Manufacturing visionary advances that might lead to enhancements in product quality, dependability, and enabling technologies for a new production method are discussed. The main components of a machine vision system's design and uses are also discussed. These factors may be roughly categorized into six areas, including scene restrictions, picture capture, image pre-processing, image processing, machine vision justification, and

systematic considerations. Here, each component of such processes is detailed in detail, along with the necessary conditions for an ideal design.

11. Tianhai Wang, Bin Chen, Zhenqian Zhang, Han Li, Man Zhang (Tianhai Wang, Bin Chen, Zhenqian Zhang, Han Li and Man Zhang 2022) this article Applications of machine vision in agricultural robot navigation examined machine vision navigational uses in agricultural robots. Additional criteria for the autonomous navigation of agricultural robots are present in many smart agriculture jobs. Machine vision is frequently employed in agricultural robot navigation because it provides priceless visual data at minimal hardware costs. The benefits, drawbacks, and functions of various vision sensors, as well as machine vision techniques, in agricultural robot navigation, were first discussed. Then, the machine vision for agricultural robot navigation's state of development was examined. This article examined machine vision steering applications in agricultural robots from 2017 to 2021. According to this study, the monocular camera offers the greatest number of potential uses for agricultural robot navigation. However, many studies select depth sensors when dealing with complicated surroundings, such as curving routes and rough terrain, since depth information can assist robots effectively comprehend complex situations. The development of the machine vision navigation system for agricultural robots faces several obstacles. This includes travelling under curving routes and rough terrain, avoiding obstacles, and turning around headlands. The robustness of agricultural robot navigation based on machine vision has to be further improved. Contrarily, there are less open-access datasets for machine vision in agricultural areas, which restricts the involvement of researchers from many sectors and the impartial assessment of various navigation techniques. Presently, agricultural robot navigation is currently using a lot of new techniques, tools, and platforms, which can assist in resolving issues. The multi-sensor fusion approach, the advanced machine vision algorithm, and the inter cooperation technique will be the main areas of attention for machine vision development in agriculture mobile robots in subsequent times.

12. Rana Qarooni, Jonathan Prunty, Markus Bindemann, Ron Jenkins (Rana Qarooni, Jonathan Prunty, Markus Bindemann and Ron Jenkins 2022) this study dicussed the Capacity limits in face detection. The first step in face processing, such as the extraction of identification or semantic data, is face detection. Those subsequent procedures seem to have strong capacity constraints, but it's not apparent where the bottleneck is. It is unclear if the bottleneck happens before or after face detection in particular. Here, we provide a brand-new test for evaluating face detection's capability constraints. In four behavioral tests, we measured observers' capacity to distinguish between two different display styles in order to recognize several faces. The elements on fixed displays were all of the same type (all faces or all non-faces). Faces and non-faces were merged in mixed displays. Importantly, a "fixed" response necessitates processing of all items. The finding that many faces may be recognized simultaneously suggests that the bottleneck in face processing occurs after the detection stage rather than before or simultaneously with it. Other elements of the visual landscape may affect whether numerous faces are really perceived simultaneously in a given scenario.

13. Erik Hjelmas, Boon Kee Low (Erik Hjelmas and Boon Kee Low 2001) this work was A Survey on face detection. İt focuses on thorough and critical analysis of face

identification algorithms is presented in this study. Face detection is an essential first step in face recognition systems that locates and extracts the face area from the backdrop. Furthermore, it has several uses in fields including intelligent human-computer interactions, video analysis, crowd monitoring, and content-based picture retrieval. The face identification issue did not, however, start to draw a lot of attention from academics until lately. Face identification in computer vision is a challenging topic since the human face is a dynamic entity with significance rapidly increases fluctuation. There have been many different ways put forth, from straightforward edge-based algorithms to composite high-level systems leveraging cutting-edge pattern recognition techniques. The algorithms covered in this study are categorized as either feature-based or image-based, and their specific techniques and effectiveness are reviewed. We do not offer a thorough comparative review due to the absence of standardized testing, however comparisons are provided when findings are given on widely used databases.

14. Mike Burton, Markus Bindemass (Mike Burton and Markus Bindemass 2009) this study, The role of view in human face detection pprovides a number of fresh perspectives on human facial recognition. The first addition is that diverse facial perspectives do not all have similar detection. Face view in tasks involving visual discrimination and identification has been the subject of a significant amount of study. The results of the current study demonstrate that vision clearly influences one of the most basic activities involving faces, namely the process by which faces are found in our visual world. In Experiment 1, faces could be discovered with comparable speed and accuracy in frontal and mid-profile views, but profile views had slower relative speed and accuracy. The study's second contribution is an analysis of the view effect's theoretical underpinnings. The fact that both frontal and mid-profile faces performed similarly in Experiment 1 implies that this task's left-right (2D) symmetry is not the only factor influencing detection accuracy. Mid-profile faces were asymmetrical, but only one eye was visible in profile view, but both eyes were visible in frontal and mid-profile views. Thus, Experiment 2 looked at whether a pair of eyes contributes significantly to the structure of face identification. The results of Experiment 3 then demonstrated that information in a face's upper half is still what drives detection disparities between face perspectives. The same detection pattern as in the previous trials was seen when the top halves of faces were incorporated into the visual sceneries, with comparable and higher performance for frontal and mid-profile views. However, detection was equivalent for all of the face views when only the bottom halves of the faces were shown. This reveals unequivocally that information in the upper half of a face, rather than only the visual cue of a pair of eyes, is what drives variations in detection accuracy between face perspectives (Experiment 2). The top and bottom regions of faces were recognized differently from one another in terms of speed and accuracy. All of the lower face conditions showed significantly longer response times and more identification mistakes than the higher frontal and upper mid-profile conditions. This demonstrates that the information that underlies the detection advantage for the top portions of frontal and mid-profile faces is significantly diminished, if not completely absent, in their equivalents in profile.

15. Markus Bindemann, Mike Burton, Rob Jenkins (Markus Bindemann, Mike Burton and Rob Jenkins 2005) this research on the Capacity limits for face processing, was purposefully carried to determine if reactions to a face target may be influenced

by a distractor face that is shown simultaneously in an interference paradigm. Participants in Experiment 1 ignored a distractor name or face and made a sex determination in response to faces or forenames. In a semantic categorization task, subsequent experiments reproduced similar design with famous faces and famous names (Experiment 2) and famous faces and pictures of country flags (Experiment 3). Through its congruency effects on target RTs, distractor processing in these studies was evaluated (e.g. same vs. different sex in Experiment 1). The FACE-face condition was unable to generate a consistent consistency impact in any of the three studies. This surprising lack of distraction technique distraction, though, was not caused by the fact that face targets are immune to interference. In fact, nonface distractions consistently hampered the categorization of faces in Experiments 2, 1, and 3. Similar to the FACE-face condition, single face distractors in all three experiments produced significant congruency effects with the nonface targets, indicating that the absence of distraction technique interference in this condition is not due to the fact that face diversions cannot affect aim categorization. Furthermore, all of the nonface comparisons generated trustworthy in between interference, in contrast to the FACE-face condition. Four distractions were used in the fourth experiment to enhance the amount of unimportant information. No congruency impacts between face distractors and face targets were seen, as they weren't in the earlier tests. Nevertheless, in this instance, the numerous face distractors also did not affect the categorization of nonface items (images of flags). The discovery of limited capacity processing for faces is thought to advance our understanding of the characteristics of visual attention and serve as a springboard for further research on face processing.

16. Haozhe Xie, Hongxun Yao, Shangchen Zhou, Shengping Zhang, Xiaojun Tong, Wenxiu Sun (Haozhe Xie, Hongxun Yao, Shangchen Zhou, Shengping Zhang, Xiaojun Tong and Wenxiu Sun 2021), this work Towards 3D object reconstruction from stereo images framework makes inferences about the 3D structure of the object by taking into consideration feature correspondences and bidirectional discrepancies between the two views. The 3D form of an item may be recreated from a pair of stereo photos using a novel deep learning framework proposed in this study. The StereoShapeNet dataset, which contains 1,052,976 pairs of stereo pictures generated from ShapeNet as well as the accompanying bidirectional depth and disparity maps, is also offered. The 3D form of an item may be recovered from a pair of stereo photographs using the unique framework presented in this research. The suggested technique analyzes bidirectional differences and feature correspondences between the two perspectives to reason about the 3D model. To the best of our knowledge, our work is the first to investigate 3D reconstruction using deep learning from stereo photos. We also create a massive synthetic dataset called StereoShapeNet that comprises 1M stereo picture pairings produced from ShapeNet together with the matching bidirectional depth and disparity maps in order to complement our work and encourage more research in this new path. On StereoShapeNet, both quantitative and qualitative evaluations of 3D volumes and point clouds demonstrate that the recommended method works better than cutting-edge techniques. Impressive results have been obtained when inferring the full 3D shape of an item from an RGB image; nonetheless, current approaches mostly depend on selecting the most comparable 3D model from the training set to address the issue. These approaches have limited generalizability and might produce subpar reconstructions of hidden items.

Nowadays, stereo cameras are widely used in new technologies like dual-lens smartphones and robotics. This makes it possible to employ stereo pictures' two-view nature to study the 3D structure and hence enhance reconstruction performance.

17. Abdulla Mohamed, Chenguang Yang, Angelo Cangelosi (Abdulla Mohamed, Chenguang Yang and Angelo Cangelosi 2016) in this study, Stereo Vision based Object Tracking Control for a Movable Robot Head, a visual tracking control system was created and put on a Baxter robot with a specially constructed moveable head. The stereo vision system from a Bumblebee camera was utilized by the tracking control method, made possible by the MATLAB computer vision toolbox. This project aims to make it possible for a robot to move its head in a way that is comparable to how a person would to concentrate their vision on a moving object. Using image processing techniques, an item is identified, and its coordinates are estimated. In order to give the robot mobility that is similar to that of a person, a fuzzy logic approach is used to regulate the moving head. Any other mobile robot head platform can use the technique established in this study. The usefulness and efficiency of the suggested strategy have been examined in depth through extensive experimental experiments. In order to imitate the movements of a human head, we have created a stereo vision-based object tracking control system on a mobile robot head with two degrees of freedom. The tracking control method was created utilizing fuzzy logic theory, which resulted in a tracking system that tracked more accurately and used less computing. The ability to measure distance alone using cameras is made possible by stereo vision. These two characteristics will improve an independent robot's capacity to function in an unfamiliar environment. The outcome of the experiment demonstrates how the moving head might be effective for locating nearby objects and tracking moving objects. Additional work will be done on the platform and the Baxter robot to improve their capacity for decision-making and complete automation.

18. Mahmoud Badawy, Hisham Khalifa, Hesham Arafat (Mahmoud Badawy, Hisham Khalifa and Hesham Arafat 2019) this study provides and highlights New approach to enhancing the performance of of cloud-based vision system of mobile robots. Mobile robots need a common computing environment, high compute capacity, and authentic performance. While cloud computing provides computational capacity, network slowness might have a negative impact on real-time performance. This study's primary goal is to enable a mobile robot vision system to successfully meet real-time restrictions by employing cloud computing. Along with a data flow mechanism set up on both the mobile robot and cloud server sides, a human cloud mobile robot architecture is suggested. A real-time image clustering method for mobile robots and a modified expanding neural gas technique for cloud servers are the two algorithms that are suggested. According to the experimental findings, the overall reaction time is improved by 25% to 45%, depending on the connection bandwidth and picture resolution. Additionally, superior performance compared to other cutting-edge methods is shown in terms of information quantity, path planning time, and reliability. Planar-constrained points and non-planar-constrained points are two categories for which 3D points in the current study were divided using MGNG. In order to accommodate more regular geometric surfaces, such as cones, ellipses, cylinders, or even incomplete geometric forms, the suggested MGNG method needs be expanded.

19. Mahmoud Abdelaal, Ramy M.A. Farag, Mohamed S. Saad, Ahmed Bahgat, Hassan M. Emara, Ayman El-Dessouki (Mahmoud Abdelaal, Ramy M.A. Farag, Mohamed S. Saad, Ahmed Bahgat, Hassan M. Emara and Ayman El-Dessouki 2021) this research, Uncalibrated stereo vision with deep learning for 6-DOF pose estimation for a robot arm system, a brand-new technique for object posture estimation with six degrees of freedom is proposed for pick-and-place robot arm applications. Since a stereo vision system doesn't need to be calibrated, it is based on the usage of that technology. Four corner points of the item are found using both cameras. From the four corner points that were detected in each picture from both cameras, a deep neural network (DNN) is trained to estimate the object's 6 DOF posture. An inexpensive vision system with a specially designed configuration is employed for the stereo vision. The robot is programmed to automatically gather data in a preset workspace before the DNN's training phase. The shared field of vision of the stereo vision system and the spatial viability of the robot arm are what constitute this workspace. Images of a 2D marker linked to the gripper on the robot arm are represented by the data gathered. To make it easier to find the four corner locations, data is collected using a 2D marker. Without determining either the intrinsic or extrinsic characteristics of the stereo vision system, the suggested technique is successful in predicting the six degrees of freedom posture of the object. A comparison of various activation functions and DNN-related optimizers leads to the best design for the proposed DNN. The performance of the suggested uncalibrated DNN-based technique is contrasted with that of the established calibration-based method. The rotational matrix connecting the robot coordinates to the stereo vision coordinates is generated using two methods in the calibration-based technique. Singular Value Decomposition (SVD) is used in the first strategy, and Hyper Particle Scouts Optimization (PSO), a recently suggested variation of particle swarm optimization (PSO), is used in the second strategy (HPSO). HPSO performs better than PSO and the genetic algorithm, two other metaheuristic optimization techniques (GA).

20. Mohidul Alam Laskar, Amlan Jyoti Das, Anjan Kumar Talukdar, Kandarpa Kumar Sarma (Mohidul Alam Laskar, Amlan Jyoti Das, Anjan Kumar Talukdar and Kandarpa Kumar Sarma 2015) this study Stereo Vision-based Hand Gesture Recognition under 3D Environment provides an unique method for detecting hand gestures that move toward and away from the camera in the forward and backward directions, accordingly. Our method, which relies on stereo vision, recognizes gestures using a conditional random field (CRF) as a classifier and a centroid movement and strength change feature based on a disparity map. Due to its many uses, including sign language detection, human computer interaction, gaming, virtual reality, etc., human hand gestures as a natural method of communicating with computers are becoming an emerging topic of research. Given that the z-axis is not taken into account, hand gesture detection in a 2D environment has several drawbacks. Therefore, the study of hand gesture detection in 3D environments is expanding. To obtain the corrected pictures, stereo calibration and rectification are performed, and correspondence results in the depth map. For Arabic numbers (0–9), we tried our suggested strategy, and it performed well, with an average identification rate of 88%.

21. Fathi, H., Brilakis, I. (Fathi, H., and Brilakis, I. 2016) this paper Multistep Explicit Stereo Camera Calibration Approach to Improve Euclidean Accuracy of Large-Scale 3D Reconstruction, discusses the intrinsic and extrinsic camera characteristics identified during camera calibration have a significant impact on the spatial accuracy of point clouds produced by stereo image-based 3D reconstruction techniques. When utilized for large-scale situations like mapping civil infrastructure, the current camera calibration techniques introduce a considerable amount of inaccuracy because of inaccurate estimations of camera characteristics. As a result, 3D point locations are less definite, which increases the likelihood that the entire reconstruction process would fail. This research suggests an innovative method to deal with this issue. It postulates that the calibration accuracy may be improved by videotaping a series of successive calibrations along a predetermined path. This is accomplished by performing independent estimations for a few predetermined distance values using typical camera calibration techniques. Multiple sets of camera parameters are included in the outcome, which is subsequently employed in the Structure from Motion procedure to increase the reconstruction's Euclidean accuracy. The suggested technique has been tested on infrastructure sceneries, and experimental studies show that the spatial accuracy of 3D points has increased by more than 25%.

22. In 2020, Mayur Surve, Priya Joshi, Sujata Jamadar and Minakshi Vharkate created a model that captures the real-time pictures from cameras. Then it employs several face detection and facial recognition methods. Additionally, they developed a GUI (graphical user interface) that, with a single click, can capture photos, build datasets, and include datasets. To identify the face in the picture, they applied the Haar cascade technique.

23. In 2019, Jenif D Souza W S et al. introduced a system that uses image processing approaches for facial recognition. To compare with the staved catalog, the improved photo is used. Four modules were used to carry out the first procedure: image capture, group photo cleavage, face detection, comparison, and recognition.

24. In 2019, Nandhini R, Duraimurugan N et al. developed a method that can rapidly and correctly recognize student faces in pictures or videos that are subsequently captured using a camera. Numerous technologies and techniques have been introduced to improve how Deep Learning is used to recognize faces.

25. Using facial detection, Omar Abdul Rhman Salim et al. suggested a method for fully implanting a student attendance system. The Raspberry Pi, which runs the Raspbian operating system, is used in the process. The Raspberry Pi is connected to a camera and a 5-inch screen. The Raspberry Pi will get the image that was recorded by the camera. It develops the LBPs and is internally coded to manage face recognition. The door will be opened, the attendance will be positively recorded, and it will be saved if the face in the input image, or image captured, matches the training dataset image.

26. In 2017, Sujay Patole et al. built a device that combines methods for justification examination. It is based on feature extraction and uses Voila-Jones for face detection. PCA is used to identify faces. The database initially contains images of people that are used to identify the subject of a photograph. The outcome is good and it can detect how people's looks have changed over time.

1.6 OBJECTIVES.

Making the computer recognize and identify the direction of human sight angle.

The main objective is to have a robot that can give directions to human. Those directions shown in a display behind the robot. The robot wants to make sure the user pay attention to what is displayed on the screen. If the robot detects that the user is not paying attention to the screen, it makes an alert on the screen to make sure and ensure the user focuses on the desired location on the screen.

A system that helps a user with navigation round a map, for example, a tourist in a new city trying to navigate and find his way round the new city, he find himself in the front of the machine system (the robot), the map of the city is displayed on the screen behind the robot. The robot captures the user’s face and also detects the user’s eyes, the detecteion and guide is being given to the user by the robot based on the user’s eyes movement. İn a bid to give directional guide to the user, the robot detects the user is not following the guide on the map, a caution sign is popped up to call and catch the user’s attention to the right point.

For the algorithms to work accurately, large data sets comprising millions of both positive and negative pictures must be used for training. With practice, the algorithms' ability to locate faces in a picture and identify them grows. A system of computer-human interaction with a focus on user-computer interfaces, it entails designing and utilizing computer technology.

CHAPTER TWO

FACE AND EYES DETECTION

2.1 FACE DETECTION METHODOLOGIES.

Face detection, commonly referred to as facial detection, is a computer program that searches through digital images to find and recognize human faces. Face detection technology may be utilized in a multitude of fields, including security, identification, enforcement agencies, entertainment, and personal safety, to communicative and collaborative monitoring and tracking of people. Automatically identifying human faces in digital photos improves content-based video indexing systems. (C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and Wavelet packet analysis", *IEEE Trans Multimedia*, vol. 1, no. 3, pp. 264-277, 1999).
Face identification has progressed, resulting in continuous performance gains, from fundamental computer vision techniques through advancements in machine learning (ML) to more complicated artificial neural networks (ANN) and associated technologies. It now forms the basis for several important applications, including face tracking, face analysis, and facial recognition. The accuracy of face detection greatly depends on the application's capacity to carry out consecutive operations.

In order to identify age, gender, and emotions through facial expressions, face analysis usually focuses on certain areas of a photograph or video. Face detection helps with this process. A facial recognition system, which mathematically maps a person's facial characteristics and saves the data as a faceprint, uses face detection data to identify which components of an image or video is essential to generate a faceprint. When a new faceprint is discovered, it may be compared to faceprints that are already in storage to determine if there is a match.

A new strategy for automated face recognition (AFR) is presented after studying the discrimination capacity of distinct human facial characteristics. The first section of the study focuses on the linear discriminant analysis (LDA) of various elements of human faces in both the spatial and wavelet domains. This study provides for an impartial assessment of the importance of visual information in various areas (features) of the face for identifying the human subject. Face LDA also offers us with a short number of features that include the most important information for categorization. The characteristics are derived using scatter matrices using eigenvector analysis with the goal of maximizing between-class variations and decreasing within-class variations. As a consequence, an efficient projection-based feature extraction and classification strategy for AFR has been developed. Each projection generates a decision axis with varying degrees of discriminating power or dependability. To produce more trustworthy recognition results, soft judgments based on each projection are integrated, and probabilistic or evidential techniques to multisource data analysis are applied. The use of very-low-dimensional feature vectors results in outstanding classification accuracy for a medium-sized database of human faces. Furthermore, the strategy utilized is broad and may be used to a wide range of additional image-recognition tasks. (K. Etemad and R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images", *J. Optical Soc. Am.*, vol. 14, pp. 1724-1733, 1997).

Computers are getting smarter all the time thanks to the quick rise in computational power and the availability of cutting-edge sensor, analytical, and rendering equipment and technologies. The ability of a computer to interact with people naturally by watching people through cameras, hearing interacting with others via microphones, understanding their inputs, and being amiable in return way better has been demonstrated in numerous research projects and commercial applications. One of the key techniques that makes human-computer interaction (HCI) so seamless is face detection. All facial analysis algorithms, including those for gender/age recognition, head posture tracking, facial emotion tracking/recognition, face alignment, face modeling, face relighting, face recognition, and face authentication, among many more, start with face detection. Computers can now clearly grasp facial expressions, and once they do, they can genuinely comprehend what individuals are thinking and intending. Check to see whether there are any faces in a digital picture is the main objective of face detection. For humans, this can seem like a simple task, but for computers, it's incredibly difficult, making it one of the most widely researched academic areas in recent years. Numerous scale, position, orientation (in-plane rotation), posture (out-of-plane rotation), facial expression, illumination, occlusions, etc. differences all come into play. contribute to the challenge of face detection. In the literature, there are numerous reports when detecting faces and there lots of different reports linked and associated with and to face detection. Face detection has evolved dramatically during the last 10 years.

The most crucial visual information sources for human-computer interaction are human faces. Face recognition, however, is quite challenging since it is affected by variations in lighting and positions of the human face. One of the key elements that impacts a lot of image and machine vision applications' success or failure is scene lighting. (H. Liu, W. Gao, J. Miao and J. Li, "A novel method to compensate variety of illumination in face detection", *Joint Conference on Information Sciences*, pp. 692-695, 2002).

Figure 2.1: Typical Faces. (http:// link.springer.com/article/10.1007/s10462-018-9650-2/figures/1)

Figure 2.2: Face detection. (http:// link.springer.com/article/10.1007/s10462-018-9650-2/figures/2)

As shown in figure 2.1, shows the image of sample faces of 3 asian men. While figure 2.2 shows square boxes on their faces to indicate face detection. The 3 asian men's faces was detected.

It provided an introduction to the idea behind the human face detection method as well as how to use DSP to create and improve it. The AdaBoost algorithm was used to find human faces. The original picture data was first optimized using techniques such image scaling, median filtering, histogram equalization, and edge detection. Following the detection, it assessed and adjusted the findings to raise the detection's level of precision. To fulfill the need for real-time, it then employed many optimization techniques. The results of the experiments demonstrate that this DSP system can recognise faces in 720576 high-resolution photos in real time while maintaining excellent accuracy and a low miss rate in a challenging setting. (L. Kong, J. Peng, Z. Xia, Y. Lian and S. Fan, "Realization and optimization of face detection algorithm based on DSP", *Proc. of 2010 International Conference of Information Science and Management Engineering (ISME)*, vol. 1, pp. 246-249, 2010).

2.1.1 CHALLENGES WITH FACE DETECTION.

One of the trickiest challenges in automatic facial detection is pattern recognition. Face photographs frequently have a plain backdrop in many practical applications (such as personal identification utilizing photo-IDs). In this study, we provide a novel approach based on shape data. For photographs with a plain backdrop, our technique can be highly effective. Histogram equalization is used to improve the input picture before edge detection using a multiple-scale filter. Then, a technique based on an energy function is used to connect the extracted edges. Finally, the face contour is retrieved using the associated edges' direction data. (J. Wang and T. Tan, "A new face detection method based on shape information", *Pattern Recognition Letters*, vol. 21, pp. 463-471, 1999).
Face detection challenges are what lower the accuracy and detection rate of face detection. These difficulties include complex backgrounds, an excessive number of faces in the photographs, unusual expressions, illuminations, low resolution, face occlusion, skin color, distance and direction, and so forth.

- Strange expressions

Face detection can be challenging when a picture shows a human face with peculiar expressions.

- Occlusive face.

Face occlusion is when an object obscures the face. It might be a hand, hair, scarf, cap, or any more item, etc. Additionally, the rate of facial detection is reduced.

- Illuminations.

There might be inconsistent light impacts in the photograph. There could be very high illumination in certain areas and very low illumination in others of the image.

- Complicated background.

There are more items in the image when the background is complex, which lowers the speed and precision of facial detection.

- The image has several faces.

This indicates that there are too many human faces in the photograph, making it difficult to recognize faces.

- Low resolution.

The The image's resolution can be quite poor, which makes face detection almost impossible.

- Skin tone.

Skin tone varies according on geography. Chinese people have a distinct skin tone than Africans, who in turn have a different skin tone than Americans, and so forth. Changes in skin tone make it difficult to recognize faces.

- Distance.

Large space between the camera and a person's face could make it harder to distinguish their faces in pictures.

- Orientation.

An angled face position is known as face orientation. Additionally, it lowers facial detection's accuracy and detection rate.

2.1.2 APPLICATION OF FACE DETECTION SYSTEM.

Applications for face detection discover people's faces in larger photographs, which commonly contain non-facial objects like buildings, landscapes, and other human body parts like the feet or hands, using algorithms and machine learning (ML). The eyes are one of the easiest features to recognize in a face, thus algorithms for identifying faces frequently start their search there. After that, the computer may make an effort to recognize the iris, mouth, nose, and nostrils. Once the algorithm concludes that it has discovered a facial area, it performs further checks to ensure that it has truly located a face. In order to help assure accuracy, the algorithms need to be trained on big data sets comprising hundreds of thousands of both positive and negative photographs. Training improves an algorithm's ability to locate faces in a picture and identify them.

The most crucial feature in identifying anybody is their face. Face recognition aids in the authentication of every person's identification utilizing his unique personal traits since it serves as a distinct identity for everyone. The entire process of authenticating any face data is divided into two phases. The first phase involves quickly detecting faces, with the exception of situations where the object is placed quite far away. The second phase then begins, during which the faces are identified as belonging to specific people. The entire procedure is then repeated, assisting in the development of a face recognition model, one of the most carefully examined biometric technologies. (Gurlove Singh and Amit Kumar Goel, "Face Detection and Recognition System using Digital Image Processing", *IEEE International Conference on Innovative Mechanisms for Industry Applications*, 2020).

Field and areas where face detection system is been applied or could be applicable:

1. Categorization of gender Images of people can reveal gender information.
2. Control of documents and access with a facial identification system, control over document access can be established.
3. Human-computer interaction system With a focus on user-computer interfaces, it covers the development and use of computer technology.
4. Biometric identification. It is a mechanism for recording attendance using a person's face, finger, or other features.
5. Photography The focusing feature on certain modern digital cameras uses facial detection. For highlighting particular areas of interest in photo slideshows, face detection is very helpful.
6. Extraction of facial features. Images can be used to extract facial features including the nose, eyes, mouth, skin tone, etc.
7. Face identification. Face recognition software may be used using a digital picture or video frame system can identify or confirm a person. Selected face a database of face attributes and picture characteristics can be compared as one method of doing this. Usually, security systems employ it.
8. Marketing. Marketers are becoming more interested in face detection. A webcam that is built into a television can recognize any passing faces. The approach next establishes the age range, gender, and race of the face. After this data is gathered, a selection of advertising that are tailored to the detected race, gender, and age can be played.

2.1.3 TECHNIQUES OF FACE DETECTION.

Face detection algorithms might be based on appearance, features, knowledge, or templates. Each has advantages and disadvantages.:

1. In knowledge-based or rule-based techniques, a face is described using rules. This technique has a challenge since it is hard to establish precise restrictions..
2. For feature invariant techniques, which rely on characteristics like a person's eyes or nose to detect a face, noise and light might be detrimental.

3. In order to detect faces, template-matching approaches function by connecting pictures to previously saved common face patterns or features. Unfortunately, these methods disregard variations in size, shape, and location.

Appearance-based approaches use machine learning and statistical analysis to identify the pertinent features of face photos. This technique, which is divided into smaller steps, is also applied to feature extraction for face recognition.

A useful geometrical face model and an efficient facial feature detection approach are proposed. Based on the fact that human faces are constructed in the same geometrical configuration, the proposed approach can accurately detect facial features, especially the eyes, even when the images have complex backgrounds. The average computation time for one image of size 512 × 340 is less than 5 s by a SUN-Sparc 20 workstation. Experimental results demonstrate that the proposed approach can efficiently detect human facial features and satisfactorily deal with the problems caused by bad lighting condition, skew face orientation, and even facial expression. (Jeng, S.H., Liao, H.M. Liu, Y.T., Chern, M.Y., An efficient approach for facial feature detection using geometrical face model. In: Proceedings of the ICPR 1996, pp. 426–430).

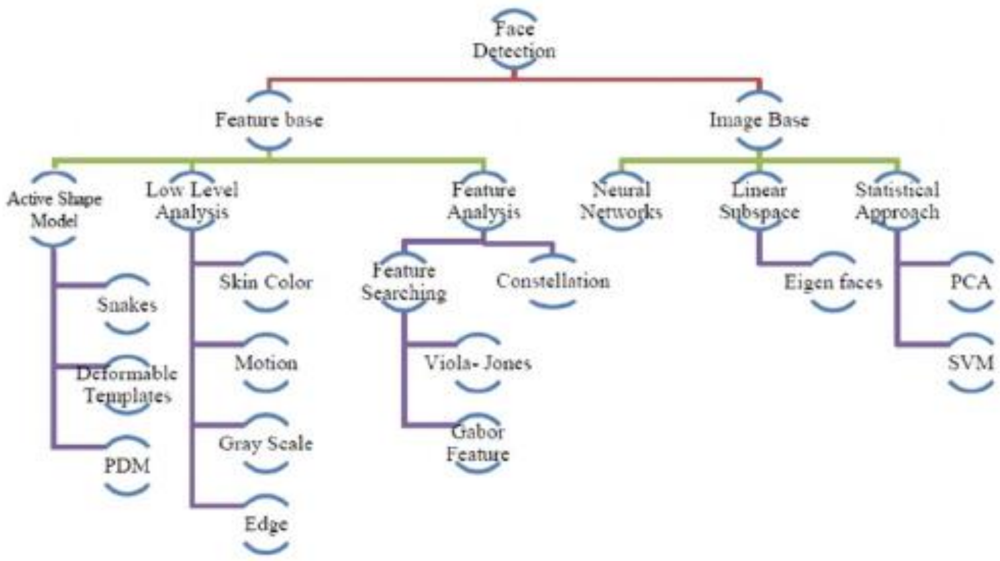


Figure 2.3: Various methods of Face Detection. *(http:// link.springer.com/article/10.1007/s10462-018-9650-2/figures/4).*

A computer program called face detection can locate and measure human faces in digital images. Any other things in the digital image, such as bodies, trees, and buildings, are ignored in favor of the facial traits. It might be viewed as a particular instance of finding and measuring each object in a photo that belongs to a specific class is the work of object-class detection. A more broad illustration of face detection is face localization. The goal of face localization is to determine the positions and measurements of a number of known faces (usually one). The two primary techniques for identifying face parts in a particular digital photo include feature-based and image-based techniques. The goal of the feature-based technique is to extract image features and compare them to the facial feature information. The image-based technique seeks the ideal alignment between the training and test images.

The following are techniques used in face detection:

1. Taking away the backdrop. Removing the background can aid in revealing the face borders, for instance, if an image has a static, pre-defined backdrop or a basic, monochromatic one.
2. Skin tone can occasionally be used to recognize faces in color photographs, however this may not always be the case always be the case depending on the subject's complexion.
3. Another method for locating faces is via motion. Users of this technology must figure out the moving area because a face is virtually always moving in real-time video. The potential for misinterpretation with other moving things in the background is a disadvantage of this approach.
4. A thorough face detection method can the result of combining the methods mentioned atop earlier.

There are several features, such as attitude, expression, location, and orientation, as well as differences in skin tone and pixel values, the presence of spectacles or facial hair, camera gain, lighting conditions, and picture quality, it can be challenging to identify faces in photographs. Deep learning breakthroughs made possible in recent years have improved face detection, which has the advantage of outperforming more conventional computer vision techniques.

It was a big leap in face recognition approach when computer vision researchers Paul Viola and Michael Jones developed a system that recognise faces in real time with excellent accuracy. The training of a model to identify what is and is not a face is the basis of the Viola-Jones framework. In order to compare features from new photographs with those that were previously stored at various stages, the model extracts certain characteristics after training and saves them in a file. Processes can proceed if the image under review successfully completes each step of the feature comparison because a face has been found.

The development of a system to recognize faces in real time with excellent accuracy by computer vision researchers Paul Viola and Michael Jones represented a significant progress in face detection technology. The Viola-Jones approach is based on training a model to distinguish what is and is not a face. Following training, the model extracts particular characteristics, which are then recorded in a file, allowing features from new photographs to be compared with features previously saved at various stages. A face has been detected, and procedures can proceed if the image under review completes each level of the feature comparison effectively.

The Viola-Jones system, while still commonly used for facial recognition in real-time applications, has certain disadvantages. The framework, for example, may not operate if a face is obscured by a mask or scarf, or if a face is not properly oriented, in which case the algorithm may be unable to find it.

Other approaches have been developed, such as region-based convolutional neural network (R-CNN) and Single Shot Detector (SSD) to aid in process improvement and help alleviate the shortcomings of the Viola-Jones architecture.

2.1.4 ADVANTAGES AND DISADVANTAGES OF FACE DETECTION SYSTEMS.

It is generally agreed that facial recognition plays a significant role in surveillance systems since it doesn't require the consent of the target. Uniqueness and acceptability are the real benefits of face-based identification over other biometrics. Face identification is a challenging topic in computer vision because the human face is a dynamic entity with a high degree of visual variability. Accuracy and speed of identification are major problems in this discipline (F. Ahmad, A. Najam and Z. Ahmed, "Image-based Face Detection and Recognition: State of the Art", *IJCSI International Journal of Computer Science Issues*, vol. 9, no. 6, pp. 169-172, 2013).

Advantages of Face Detection

Face detection is an essential part of facial imaging techniques like facial recognition and face analysis, and it provides users with a variety of advantages, including as:

1. Enhanced security: Face detection improves surveillance efforts and aids in the capture of terrorists and criminals. Since there is nothing for hackers to take or alter, like passwords, personal security is also increased.
2. Easy integration: Facial recognition and face detection technology are easy to profer solutions and straightforward to install and work with the majority of security software.
3. Identification authentication: Inefficient and usually wrong, identification used to be done manually by a human. Automating the identification process with face detection increases accuracy and reduces processing time.

Disadvantages of Face Detection

Face detection provides several important benefits for customers, but it also has a number of disadvantages, including:

1. Enormous load of data storage: Face detection ML technology necessitates strong data storage that not all users may have access to.
2. Detection is prone to error: Compared to manual identification techniques, face detection is more accurate, but it is also more prone to mistakes because of alterations in appearance or different camera perspectives.
3. A possible invasion of privacy: Face detection's ability to help the government track down criminals has many benefits, but the same monitoring might also let the government watch over private citizens. Strict regulations must be put in place to guarantee that technology is utilized fairly and in compliance with human privacy rights.

2.1.5 USES OF FACE DETECTION SYSTEM

1. Despite the fact that face detection is a component of face recognition, not all face detection methods or facial recognition systems use facial recognition.
2. Facial motion capture, which uses cameras or laser scanners to electronically transform a person's facial gestures into a digital database, is another use for face

detection. This database may be used to produce lifelike computer animation for movies, video games, or avatars.

3. Face detection may be used to determine the number of people that have entered a space or to automatically focus cameras. The technique also has marketing uses, such showing customized adverts in response to a recognized face.
4. A software that uses emotional inference can, for example, assist autistic persons understand the feelings of those around them, is another usage for face detection. The software uses cutting-edge image processing to "read" the emotions displayed on a person's face.

2.2 EYE DETECTION METHODOLOGIES.

A number of applications involving human computer interaction use eye detection as a core technology. For example, it can be used to simplify operation of human-machine interfaces or to identify drowsy driving detecting systems to prevent accidents. For face detection, face recognition, and facial expression analysis. All of which will be used extensively in the near future, eye detection is particularly crucial.

A key component of human-machine interactions is accurate eye detection. In this research, a novel developed gadget for the new real-time and efficient eye detection technique is suggested. It is based on various physical properties. The Red-eye effect and the 850-property noted in the trials are two physical properties used in this gadget. Pretreating, matching using SURF (Speeded Up Robust Features), determining the difference, and obtaining the points of the eyeballs are the four processes of the detecting technique. Following these four processes, the locations of the eyes are determined, and it is also possible to determine if an eye is present in the image or not. The matching parameters are periodically recorded and changed to speed up the detection (Quan, D., Zhao, Y., Li, J., & Zheng, H. H. An Effective Method and Device for Eye Detection with Physical Properties. In *Proceedings of the 2nd international conference on computer science and electronics engineering*. 2013).

The three categories of currently used eye detection techniques include the classic image-based passive approaches, mixed approaches, and IR-based active approaches.

Model-based techniques, template-based approaches, and feature-based approaches are examples of classical approaches. Despite minor differences, all of them use the distinctions between the form and appearance of the eyes and the rest of the face to detect eyes. They all use the face box as their input and presumptively do face detection correctly. In actual applications, the performance isn't promising.

The hybrid techniques make use of a near-infrared camera that first takes photos of the bright pupil. Next, use the conventional image-based methods to find the eyes. The active approaches use images captured by the near infrared sensors to identify eyes. They do this by utilizing the redeye effect to determine the position of the pupils, which makes it simple to identify the eyes. The equipment that the active methods that are currently in use may be divided into two categories.

Antonio Haro, et al. have presented one form that uses just one CCD camera and two identical IR led groups that are located on and off the camera axis, respectively, to provide two images. As these two groups of lights are turned on and off alternately, they might see pictures with pupils of varying brightness. So, it was possible to accurately determine the position of the eyes by subtracting the dark pupil image from the brilliant pupil image.

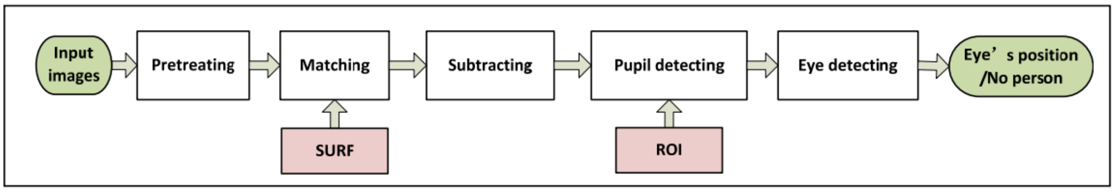


Figure 2.4: Overview of eye detection. *(http:// www.researchgate.net/publication/266652013_An_Effective_Method_and_Device_for_Eye_Detection_with_Physical_Properties).*

Relevant activities like iris and face identification, gaze tracking, behavior and expression analysis, or the detection of driver weariness rely on an earlier prediction of the person's eye location. This estimate needs to be done in high-resolution, distortion-free pictures at a high frame rate, depending on the application. For instance, this is the situation when we must take iris photographs of a person who is moving. The iris must have a radius of at least 70 pixels (Daugman J, "How iris recognition works", *IEEE Transactions on Circuits and Systems for Video Technology*, 11 (3) (2004), pp. 21-30).

2.2.1 EYE DETECTION PROCESS.

The basic premise of this study, eyeballs may be quickly recognised and detected in photographs because, regardless of the iris', hue is usually darker than the sclera's. This makes it reasonably simple to identify the iris' edge as a collection of dots that are arranged around a circle. In order to use PGAs to solve the issue with feature selection for eye recognition, the pattern area must first be mapped into a representation that can be used for genetic search. The most basic representation of an image base considers each feature as a binary gene, because the main goal is to represent the area feasible within the original feature list's subdivisions. A limited version of the original feature list represented as a fixed-length binary string. is subsequently used to represent each particular chromosome. This method is used to first carve out the pupil and the margin of the eye. and the position of the eyes is also more precisely recognized.

The individual's chromosome is set as the initial former array for the extraction of the eye area, which is composed of location data, eye outline size, and pupil diameter. Additionally, a person's fitness is determined by an evaluation function that considers three aspects of the eyes (the distance between the pupil and the white of the eye, the pupil's color, shape, and size, and the edge of the eye). The chromosome data of the person with the best fit at the time of the end of evolution is used to extract the ocular region.

The pupil is presumed to be at the centers of the eyes when the pupil is expressed as a circle for detection purposes and the contours of the eye as an oval. The X and Y coordinates, the circle's radius, the pupil's size, and the oval's shape, which indicates the form of the eyes, are used to define a person's chromosomal. The radius of a circle and the major and minor axes of an oval are each ten bits in the first former array, making a total of 36 bits. Random initialization is utilized in PGA.

The defined chromosome is then utilized to determine whether it is appropriate for the eye using the evaluation function. Due to the fact that this approach places a strong emphasis on ocular features, the following three are used as an assessment function: "The pupil and eye white are present in eyes," "The contour of the eye may be roughly compared to the oval, and the pupil is almost round in shape and black in color."

Eye detection is a critical component in the development of compression techniques as well as face identification and recognition, human computer interface design, and driver behavior analysis. The gaze can be identified by finding where the eyes are situated.

One of the most significant human-computer interfaces today, eye detection and monitoring has been proven to be helpful in a variety of applications. The process of measuring an eye's gaze point or eye movement in relation to the head is known as eye-gaze tracking. Eye tracking and detection analyzes the positions and movements of the eyes.

The objective of eye detection is to pinpoint the location of the eyes in a given face image. Simply put, in eye detection, the locations of both eyes are found or each eye is localized separately. A rectangle is typically used to show the eye areas as a result of the treatment.

The objective of this category, on the other hand, is to provide specific information, such as the contour of the visible eyeball region, the iris-pupil circle, the location of the pupil in the visible eye area, and the state of the eye (blinking or not blinking). This kind of work is particularly challenging in the field of computer vision since it depends heavily on changing environmental variables to successfully detect or track minute elements in real-time.

There has been a great deal of work done in the field of eye detection, including eye pupil movement detection, eye feature extraction, eye state recognition, and eye gaze identification utilizing various algorithms in still images as well as in video sequences for real-time applications.

An improved method of eye detection has been roposed based on a cascade Faster R-CNN with Gabor filters and a nave Bayes model (FRCNN-GNB), which locates this same locations of both the eyes with increased precision even in images that only encompass partial faces. This method addresses the issues with existing eye detection systems (Nsaid A.K., Ali S.H.M., Jassim K.N., Nseaf A.K., Sulaiman R., Al-Qaraghuli., Wahdan O., Nayan N.A. "Cascade faster R-NN with gabor filters and naïve Bayes for enhanced eye detection". *IEEE Access*, 2021, pp. 15708-15719).

Eye/gaze detection can be done in a variety of ways; some require additional hardware support, while others only require a basic webcam to complete the work. Below, many approaches are described:

1. **Electrooculography (EOG)**: A method for calculating the retina's resting potential. The Clinical Electro-oculogram (EOG) is an electrophysiological test of the function of the outer retina and retinal pigment epithelium (RPE) in which the change in electrical potential between the cornea and the ocular fundus is measured during sequential periods of dark and light adaption (Brown, M., Marmor, M. and Vaegan, ISCEV Standard for Clinical Electro-oculography (EOG), *Documenta Ophthalmologica*, 113:3205-212. 2006).

2. **Infra-Red Oculography**: The amount of light reflected back to a fixed detector when a fixed light source is pointed at the eye will change depending on where the eye is located. A variety of eye trackers that are sold commercially use this approach.

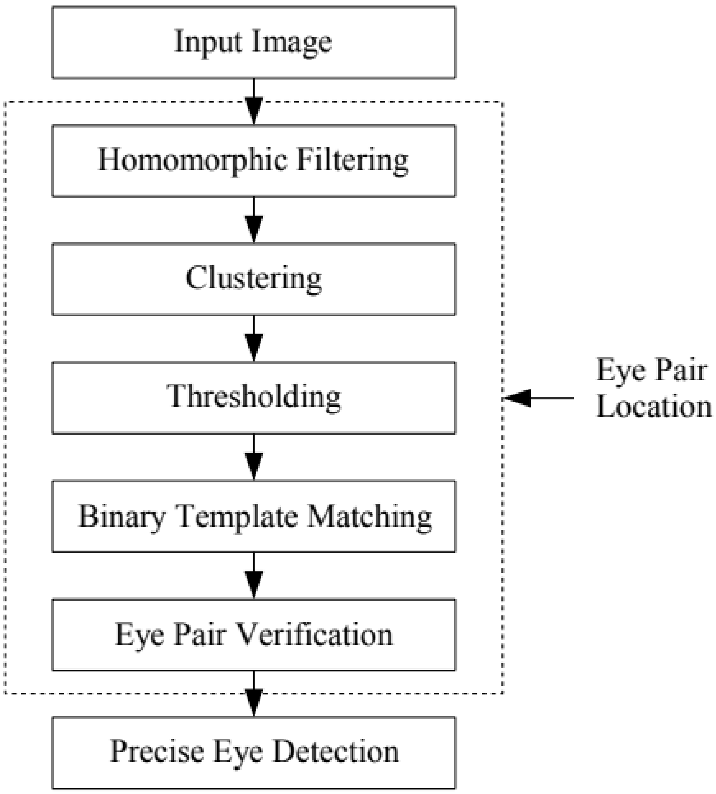


Figure 2.5: Eye Detection Method. *(http:// citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.400.677&rep=rep1&type=pdf page56)***.**

An eye pair is located before two eyes are precisely detected. The structure of the eye area is thought to be a reliable and consistent characteristic that can be used to separate an eye pair from other patterns. This reliable indication is used by the suggested strategy to identify eye pairs. Fig. 2.7 depicts the method's fundamental flowchart. The structural picture is first created by binarizing the augmented face image. The binary image that contains the human face's structural features is referred to as a "structure image." After that, eye pair candidates in the image are found using a binary eye pair template. Then, after being rescaled to a given size, each eye pair candidate is sent to an SVM classifier for verification in order to identify true eye pairs. Finally, based on the verification results, eye pairs are located.

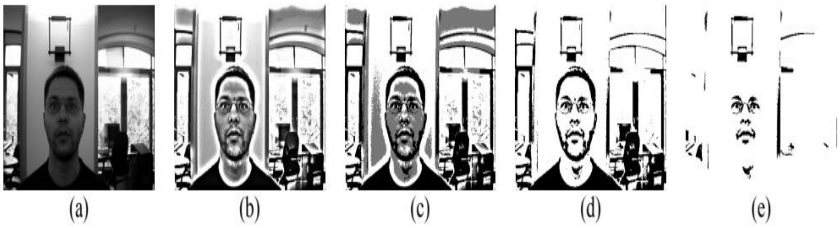


Figure 2.6: An example of preprocessing. *(*Wang and Yang: *http:// citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.400.677&rep=rep1&type=pdf page57).*

1. **Homomorphic filtering**: A generalized method for nonlinear picture enhancement and correction is homomorphic filtering. It simultaneously boosts contrast and balances brightness across a picture. Low-light image (LLI) improvement techniques have lately made significant advances, particularly with the application of deep learning methodologies. However, the majority of present algorithms are designed as stand-alone solutions and do not consider the influence of LLI augmentation on high-level computer vision tasks such as object categorization (Rayan Al, Sobbahi Joe Tekli. "Low-Light Homomorphic Filtering Network for integrating image enhancement and classification". *Signal Processing: Image Communication.* Vol 100. 2022).
2. **Clustering and Thresholding**: In a facial image with a grey level, the intensity around the eyes and other facial features is dark. Homomorphic filtering has been used to improve the image. Then, using the K-mean clustering technique, the grey-level image is divided into three clusters and the features of interest are separated from the skin and other pixels. The lightest grey level, which represents the image's light pixels, is set to 255. The middle level, which depicts the skin, is set to 128, and the darkest level, which depicts both the image's features and other dark pixels (such as the hair, beard, and some dark backdrop), is set to 0. The face picture processed by the clustering technique is displayed in Fig. 2.6(c). (A. Haro, M. Flickner, and I. Essa, “Detecting and tracking eyes by using their physiological properties, dynamics, and appearance,” in Proc. *IEEE Conf.. Computer Vision and Pattern Recognition*, vol. 1, pp. 163–168, 2000.)
   Following clustering, a threshold is set at 128 to remove all but the darkest pixels, including the eye pair structure. The face structure is then clearly visible in the binary image that has been created. The thresholding outcome is displayed in Fig. 2.6(d). The typical connected components labeling procedure eliminates the oversized black area, which is useless in the binary image, taking into consideration that the nonface area can affect the speed and outcomes of template matching. Finally, as shown in Fig. 2.6(e), the final feature image is obtained.
3. **Binary Template Matching**: A binary template matching is used to search for the two eyes in the feature image in order to identify the group of rows that contain the eyes (the eyes band). The problem is that we aren't looking for anything with a predetermined shape. For this reason, a binary template is used to represent two eyes in a very crude manner. All of the photos, which are noticeably different in size, were created using the same template, demonstrating the desired scale

independent property. The use of color information, template matching, neural networks, edge detection, and Hough transformations are just a few of the several approaches that have been used to recognize faces. This report's methodology follows a rejection scheme algorithm-like pattern. In the first stage, color thresholding and skin segmentation are used to reject areas of the picture that aren't faces. The last step is to apply binary image processing to define these zones more precisely. The Sirovich-Kirby technique of template matching is used to distinguish between faces and non-faces as well as to locate faces utilizing both training picture faces and eigenfaces (Elad, M., Rejection Based face Detection. *EE368: Digital Image Processing Lecture,* Stanford University, Stanford, CA. May 19, 2003).

4. **Eye Verifier**: A straightforward eye verifier is used in conjunction with the suggested approach to determine the successful face detection rate. Instead of using the face verifier, the eye verifier is utilized since the area below the eyes may be obscured and because the results may also be affected by the exaggerated face expression that appears below the eyes. Eye verifier is a means for accurate face color pictures for the identification of eyes. The circular filter applied to the binary facial picture initially searches for candidates for the eye area. After that, an eye-verifier receives the eye candidates. The eye-verifier employs a ternary template made from the iris, sclera, and skin of the eye. Then, ternary Hamming distance is used to match the template. The suggested technique performs well in terms of detection rate according to experimental findings using our own face database and the FERET database. The suggested method performed effectively with regard to the facial positions and eye directions (A. Senior, R.L. Hsu, M.A. Mottaleb, and A.K.Jain,. "Face Detection in Color Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002. 24(5): p. 696-706).
5. **Precise Eye Detection**: After finding the eye pair, the eye variance filter approach suggested by Feng and Yuen is used to accurately detect the eyes. On a human face, it is easy to see how the intensity of the gray changes where the eye's white, pupil, and eyelids are located. The second-order instant that represents the measurement of variation in gray intensity is called a domain's variance. The development of an eye variance filter is based on these observations. When the eye variance filter is applied to an eye region, the eye window exhibits a significant response, but the non-eye window exhibits a comparatively weak response. To precisely extract the two eyes in the ocular area is the main goal of applying the variance filter. The augmented image that was subjected to a homomorphic filter and extracted from the facial image is the image of the eye pair. By splitting the eye region into two sections, the left and right eyes can be distinguished individually. a real-time, non-contact eye-gaze monitoring system whose precision is based on an extremely accurate calculation of the pupil center. The pupil is kept in the center of the picture when the eye camera moves with the head. When a tracking error occurs, the eye may be located and the tracking process is swiftly restored using the picture from a camera with a broader field of vision. To create corneal glints, four infrared light sources are timed to the eye camera's shutter. The unique form of it has been used to optimize the image processing algorithms created for this system. The lighting power has been carefully constrained, and work has been done at extremely low levels (Perez A.,

Cordoba M.L., Garcia A., Mendez R., Munoz M.L., Pedraza J.L., and Sanchez F., "A precise eye-gaze detection and tracking system", *The 11-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2003*, 3.-7. February 2003, Plzen, p. 105-108). Following a calibration process, the pupil-glint vector is used to identify the line of sight. In order to prevent glint distortion brought on by variations in the curvature of the ocular globe, the glints are confirmed using the contour of the iris. By minimizing measurement error in the pupil-glint vector, the suggested methods pinpoint the pupil's center with sub-pixel accuracy.

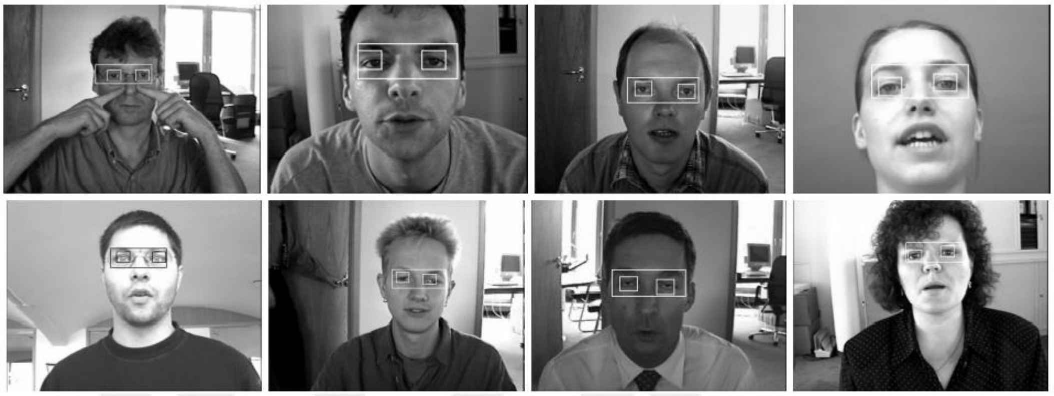


Figure 2.7: Eye Detection Results. (Wang and Yang: *http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.400.677&rep=rep1&type=pdf page61).*

2.3 HAARCASCADE.

A snapshot or live video may be used to identify a face using the Haarcascade acquisition technique. The research article "Fast Detection with the Advanced Casual Cascade," written by Viola and Jones in 2001, makes use of edge or line detection characteristics. The program generates lots of stunning pictures that don't show faces. negative pictures without any faces you may practice on. (Viola P., and Jones M., "Fast Detection with the Advanced Casual Cascade". 2001). A classifier is trained using the machine learning technique known as Haar Cascading using a large number of both positive and negative images. Michael Jones and Paul Viola advance the algorithm.

The OpenCV GitHub repository contains the model produced by this training.

https://github.com/opencv/opencv/tree/master/data/haarcascades.

The repository includes models that may be read using OpenCV techniques and are saved as XML files. These include models for detecting faces, eyes, the upper and lower bodies, license plate recognition, etc. Listed below are a few concepts from Viola and Jones' investigation.

The most popular machine learning (ML) and deep learning features are face detection and face recognition (DL). These characteristics are gradually gaining acceptance in industries like CCTV surveillance, mobile phone security (biometric locks), etc.

(Samiksha Malhotra, Vaibhav Aggarwal, Himanshu Mangal, Preeti nagrath and Rachna Jain. “Comparison between attendance systems implemented through haar cascade classifier and face recognition” librarynIOP *Conf. Ser.: Mater. Sci. Eng.*, 1022. 2021).

The classifiers used for object detection are cascade classifiers based on Haar features. This classifier uses a machine learning process that incorporates a cascade operation from the images to find objects in further images. Successful face recognition and facial expression recognition in images. The classifier is then presented with positive and negative images to complete the task. The traits are then extracted from the image. Every feature has a unique value that is obtained by deducting the total of the pixels in the white rectangle from the sum of the pixels in the black rectangle. In which it recognizes faces of various people in various settings. Integral pictures allow for the constant-time calculation of the Haar-like feature of any size.

2.3.1 FEATURES OF HAARCASCADE

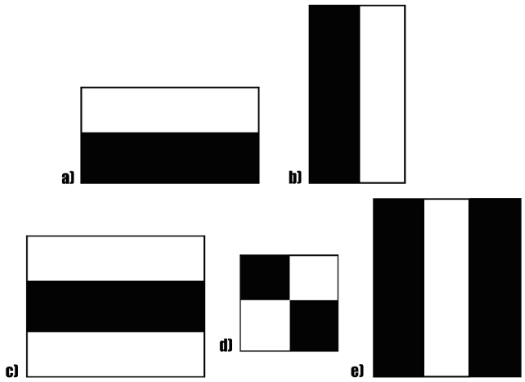


Figure 2.8: An example of the Haar characteristics utilized in Viola and Jones' original research paper. *(https:// ir.uitm.edu.my/id/eprint/60662/1/60662.pdf page72)*

The introduction of the aforementioned haar elements was the study's first original contribution. These characteristics of the picture make it simpler to identify the image's borders or lines or to choose regions where there is a sharp shift in pixel density.

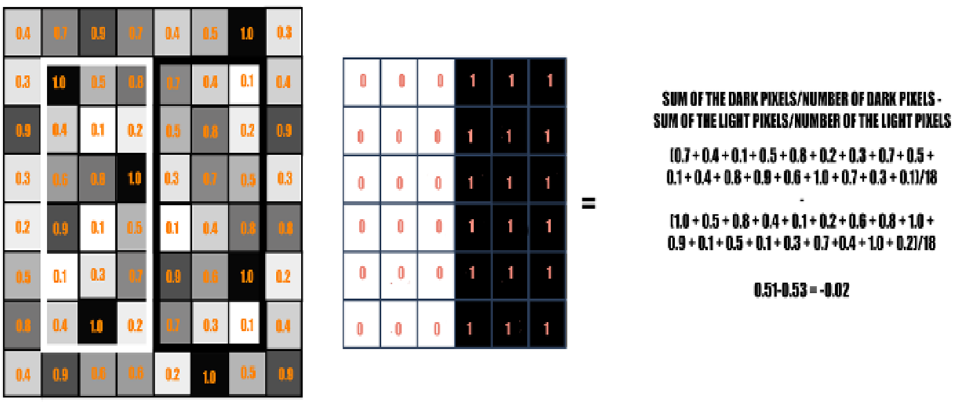


Figure 2.9: A depiction of 0.0 to 1.0 pixels may be seen in the rectangle on the left. A haar kernel serves as the focal point and is composed of just dark pixels on the right and only light pixels on the left. By calculating the difference between the average

pixel values in the simple region and the average number of pixels in the dark area, the haar is calculated. If the difference is near to 1, the haar element can produce a hem. *(http://towardsdatascience.com/face-detection-with-haar-cascade-727f68dafd08)*

The rectangular image section's estimated sample of the Haar value is displayed here. On the haar feature, the bright portions are 0 pixels in size, while the dark parts are 1 pixels. Each of these is in charge of locating a certain component in the image. Like an edge, a line, or any other structure where the intensity of the picture abruptly changes. For instance, the haar feature in the image above can develop a dry border with bright pixels on the left and dark pixels on the right.

The objective of this exercise is to calculate the total of all the pixels in the image that are located in the haar element's simple and dark areas, respectively. then identify where they vary. The quantity of haar will be closer to 1 if the image has a limit that divides the light pixels on the left from the dark pixels on the right. To put it another way, when the number of haar is near to 1, we say that a limit has been reached. There is no cap in the aforementioned case since there are many more haar than one.

This is only one example of a haar element that splits the straight edge. There are now some haar features that will see the boundaries of some of the references as well as any other picture structures. The haar feature must cut across the entire image to obtain the edge anyplace in the image.

2.3.2 METHODOLOGY

Face Detection

When it comes to technology Face detection is regarded as the challenging and widely used method. The main aim of face detection is to identify each face that appears in a picture. The implementation in this case makes use of OpenCV.

i. Taking in the supplied picture data.
ii. Creating grayscale versions of the input pictures.
iii. Using the LBP classifier and the Haar cascade.
iv. Assessing the accuracy and speed of each classifier.

a. the import of the necessary libraries
b. Using the pictures that the camera has taken.
c. The image is changed into a grayscale version in order to be processed by classif iers.
d. OpenCV will be used to load the image.
e. The image will by default be loaded in BGR color space.

Haarcascade Classifier

i. Using the builtin method *cv2.imread(img_path),* the input picture is loaded, with the image path being passed as a parameter.
ii. Grayscale mode conversion, followed by display.
iii. The Haar cascade classifier being loaded.

2.3.3 VIOLA JONES ALGORITHM

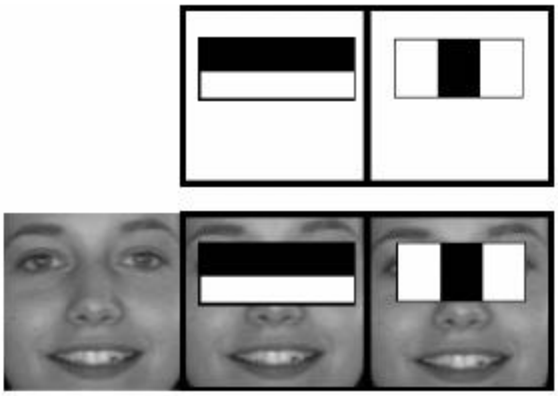


Figure 2.10: Utilization of Classifiers. *(http:// medium.datadriveninvestor.com/haar-cascade-classifiers-237c9193746b)*

Paul Viola and Michael Jones created the Viola-Jones Discovery Framework in 2001 as a machine learning discovery methodology. Although nearly anything may be trained to be detected by this frame, the issue of real-time face detection is essentially resolved.

This algorithm has four steps.

1. Feature Selection by Haar
2. Representation of an Integral Image
3. Adaboost Instruction
4. Cascade Classifier Design

1. **Haar Feature Selection**

Much quicker than the pixel method, items are classified by very basic properties like the ad-hoc domain information encoding feature. The term "Haar" refers to the feature's resemblance to haar filters. The difference in the total number of pixels in a rectangular region, which can be any size and shape within the original image, is an example of one of these properties. Here, both three and four rectangular features are employed.

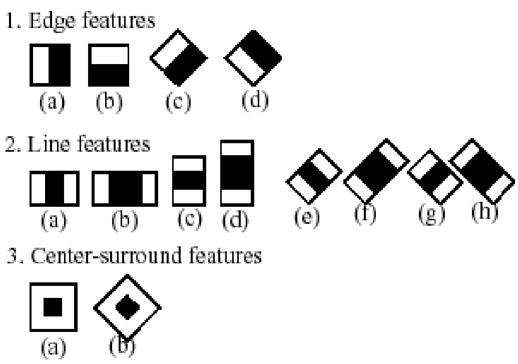


Figure 2.11: Haar Features. *(http://towardsdatascience.com/face-detection-for-beginners-e58e8f21aad9).*

2. Integral Image Representation

As the name suggests, a significant picture representation. Any point in the integral image has a value equal to the sum of all the pixels above and to the left of it. One ground over the image allows for accurate calculation of the combined image.

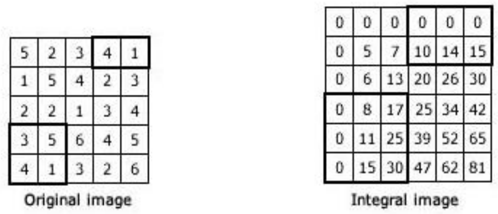


Figure 2.12: Chart showing integral image representation. *(http://towardsdatascience.com/face-detection-for-beginners-e58e8f21aad9).*

3. Adaboost Training

In a 24x24 pixel window, there may be some 162,336 features that are too expensive to test. Therefore, the class is trained using just advanced features using the AdaBoost method.

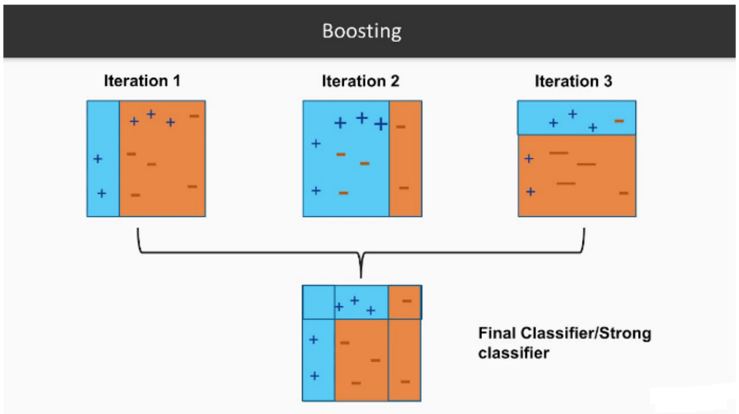


Figure 2.13: Representing blocks of Adaboost training. *(https:// ir.uitm.edu.my/id/eprint/60662/1/60662.pdf page73)*

4. Cascade Classifier Architecture

A combination of numerous categories presented in chronological sequence is referred to as a cascade classifier. It frequently decides whether to fulfill its mission or not. The cascade classifier's structure resembles a dying tree.

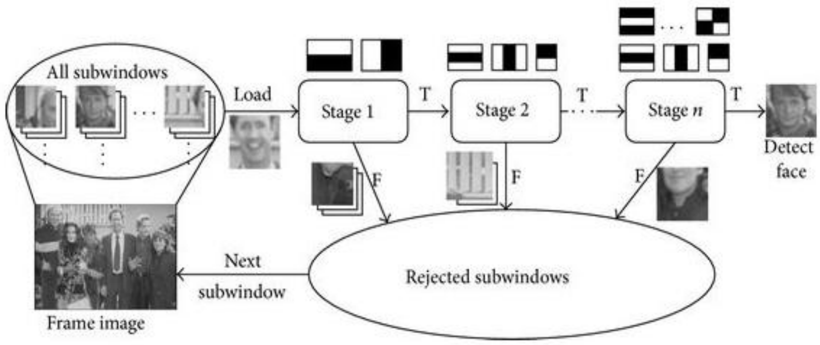


Figure 2.14: Design of Cascade Classifier.
*(https://miro.medium.com/max/1200/0*O4k602EVv7smbTFG)*

Application

Despite the development of in-depth learning (RCNN, YOLO, and so on), this approach is still utilized in many face and object identification systems since it is basic yet effective.

2.4 PROBLEM DEFINITION.

Detecting the face, being able to create a system that detects the face in the image, and detects the correct eyes in this face. After detecting the face and eye location, specify the center point for each eyes; both left eye and right eye, and then calculate the relative sight direction from the eyes location.

Angles of eye location are captured and calculated with respect to head movement, either right or left, up or down.

The key elements in this expereiment are the face and the two (2) eyes, the right and left eyes properly captured in their right and appropriate position.

The system (the machine) detects the face, locates the two eyes on the face, the aim of the experiment wouldn't be achieved if the eyes are not appropriately detected in the right posiotion, sizes and proportions. The system capturing the face and detecting the lip which is also on the face as eyes would be a wrong detection.

The face being captured and detected with the right and left eyes detected on the face in appropriate proportion and size is a correct and best detection for the system as accurate results would be gotten.

In larger photographs that commonly include non-facial objects like buildings, landscapes, and other human body parts like the feet or hands, applications for face identification employ algorithms and ML to find people's faces. The eyes are one of the most straightforward features of a face to recognize, and algorithms for identifying faces frequently start their search there. The iris, mouth, nose, and nostrils may then be identified by the computer. After determining that it has discovered a facial area, the algorithm performs additional checks to ensure that it has in fact located a face.

CHAPTER 3

SYSTEM DESIGN

3.1 SYSTEM FLOWCHART DIAGRAM

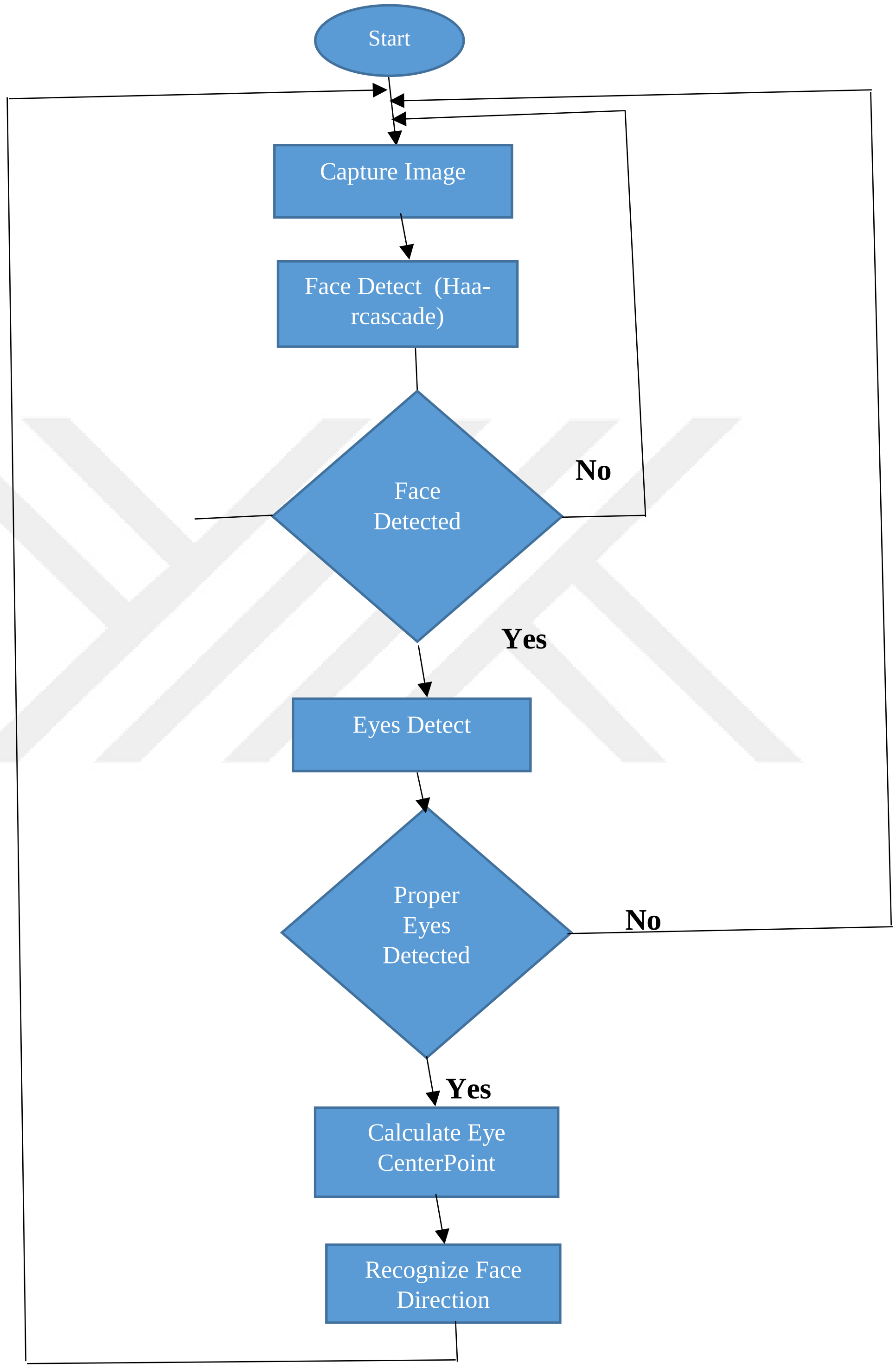


Figure 3.1: System Flowchart ***(****Drawn and Designed by Ahebhamen Samuel Osahon****).***

As shown in figure 3.1 above (the system flowchart), firstly, the system starts, then images are been captured by the capturing device, then the system tries to detect the face using the haarcascade algorithm, if a face is successfully detected, eyes are detected next in the face, if a face isn't detected, the system jumps back to start the process all over again. After the eyes have been detected, the system makes sure the eyes detected are rationally proportional being the proper eyes. If proper eyes are not detected, the system jumps back to start the process all over again. But if proper eyes are detected, center point is been calculated then the face direction is being recognized, which means the system process and transition is successful.

3.2 FACE DETECTION (USING HAARCASCADE)

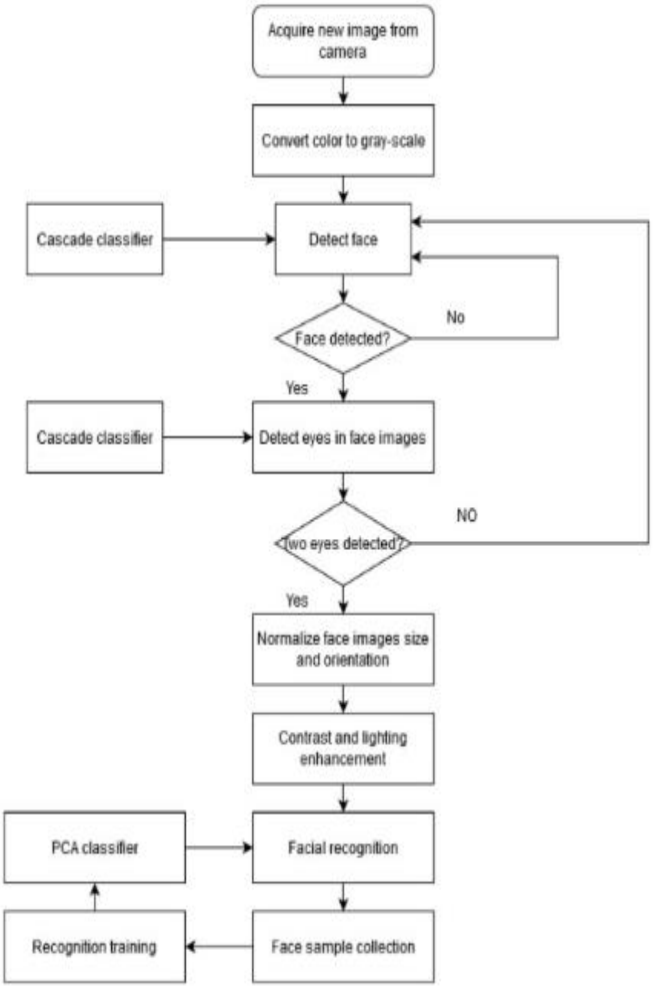


Figure 3.2: Haarcascade Flow chart. *(http:// www.researchgate.net/figure/Flowchart-for-real-time-face-detection-and-recognition_fig1_317012786)*

The figure above (fig. 3.2), displays the Haar cascade classifier's flowchart. The camera turns the image to grayscale once it has captured it. When a face is discovered, the cascade classifier checks to see whether both eyes are there. If both eyes are present, the classifier normalizes the size and orientation of the face picture. The picture is then subjected to face recognition processing, where it is compared to a library of face samples.(B.D. Parameshachari, "Logistics sine map (LSM) based partial image encryption". *National Computing Colleges Conference (NCCC), IEEE* March 2011, pp1-6).

Although people find it simple to identify someone by their face, vision-based automated systems find it difficult. It has been a focus of current research in a number of fields, including anthropometry, neural networks, statistics, pattern recognition, and computer vision are all related to image processing. Facial recognition and facial identification, automatic systems with vision can be utilized in a variety of commercial contexts, including interaction between humans and computers, surveillance, video games, and multi-media entertainment. Face recognition is a very accepted biometric because it is non-invasive and requires no direct physical touch between the user and the device. The four processes of face recognition using vision-based automated systems are Face detection, picture pre-processing, feature extraction, and matching are some of the techniques used.

Faces belong to a comparable because facial features such as the eyes, lips, nose, and chin generally share the same geometrical arrangement, detecting faces can be challenging. It is possible to pre-process the face image after it has been captured to correct for lighting differences.

The method of feature extraction involves assembling significant characteristics displayed on the face into a geometrical or vectorial representation. Three methods of feature extraction exist: holistic, feature-based, and hybrid. Examples of a holistic approach include fisher discriminant analysis, principal component analysis, and support vector machines. The geometric relationship of the face characteristics is the foundation of the feature-based approach. Applying an active shape model, gathered significant data from certain of the facial features. In the matching step, the characteristic vector can be compared to other classes (individuals) using statistical classifiers like Mahalanobis distance, Euclidian distance, Bayes classifier, and neural classifiers.

The ViolaJones object detection framework's contribution to the application of haar-features has enhanced the speed of face detection. Implementations of this framework, like OpenCV, offer many face classifiers developed by authors who used various training datasets. These classifiers' efficiency and dependability vary greatly, assessed certain classifiers' accuracy as well as their performance.

This study focuses on assessing facial classifiers in light of facial traits present in the discovered face. We propose a method that rates each facial characteristic identified in the located face with a range of values. Two separate face databases were used to assess ten face classifiers (FEI database and Yale face database).

1. Yale Face Database

The Yale Face Database features photos of 15 persons, with an average of 11 photographs taken under varied lighting conditions. The participants' faces show a variety of expressions (with glasses, sad, sleepy, surprised, wink). Each picture measures 320x243 pixels.

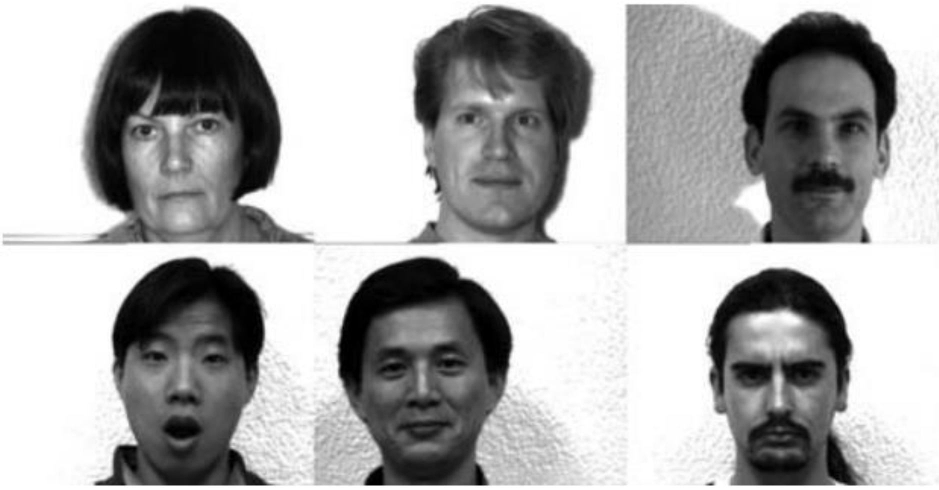


Figure 3.3: Yale Face Database Images.
(https://www.semanticscholar.org/paper/Evaluation-of-Haar-Cascade-Classifiers-Designed-for/5b90bf3ebad1583beebcae5f892db2add248bcad/figure/0)

2. **FEI Face Database**

With a total of 2800 images, the FEI face database is a Brazilian database, 14 photographs for each of 200 different people. The photos have neutral, smiling, and non-smiling emotions and are vibrant in various rotations. With regard to both happy and frowning expressions, we used two frontal photographs of each person out of a total of 400. Each image's original dimensions are 640x480 pixels.

Figure 3.4: FEI Face Database Images
(http://www.semanticscholar.org/paper/Evaluation-of-Haar-Cascade-Classifiers-Designed-for-Padilla-Filho/4de9506cc8abc617cebdc265b0e49efab50dd4e6/figure/1)

There are various techniques for detecting faces. The haar cascade is one of these.

A camera that is outfitted with high-tech capabilities is no longer a novel idea in this era of modern technology. The most significant advancement in the biometric system is its capacity to distinguish each component of the face as an item. Face detection is still actively being researched by locating a specific object within a digital image, then examining, comparing, and evaluating its pattern. As far as we are aware, a person's face is a genuine thing that reflects their self-identity and sets them apart from one another. To discover and index images and videos with backdrop, size, and position, face detection can be employed. It enables the system to identify a person's face in various lighting conditions.

3.3 EYE DETECTION

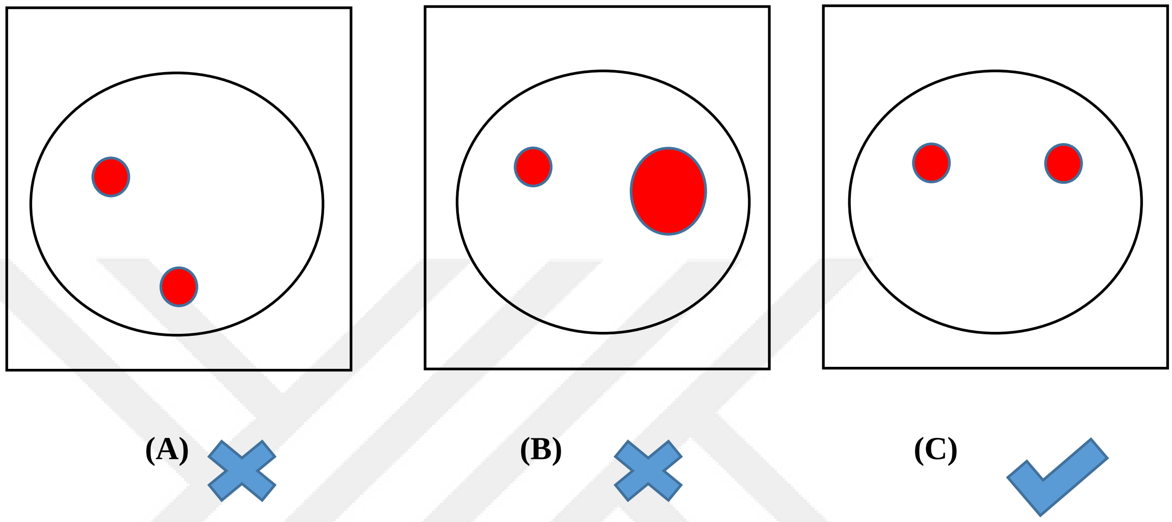


Figure 3.5 Pictorial Illustration of Detected faces with eyes placement. *(Drawn and Designed by Ahebhamen Samuel Osahon)***.**

As shown in the figure above (figure 3.5) **A, B, C** are illustration sample of 3 detected faces with the eyes detected. In A, the eyes has been detected wrongly because one eyes detected properly on the eye window region while the other has been detected outside the eye window region but on the lip region. In B, the eyes has also been detected wrongly because both eyes, though in the right eye window region, they are widely different in size and proportion. In C, the system detected the eyes correctly in the eye window region and the eyes are same sizes or slightly different in size proportion which makes the detection correct.

Using C, the system that will detect the eyes inside the detected face will get the face from the image and this face will detect the eyes, if the eyes are found in the way it is on the illustrated diagram C using the library called haar-cascade, the center point can be accurately calculated.

3.4 CALCULATE CENTERPOINT

To calculate the center point of the eye, add all the points gotten from both left camera and right camera, then X, Y, Z for both right eye and left eye on the face together and divide by 2, thereby getting the average which serves as centerpoint for eye, both left and right.

Below is the formular for calculating centerpoint:

R_Cam_R_Eye_X = Faces_R[i].Eye_R.X + Faces_R[i].Eye_R.Width/2;
R_Cam_R_Eye_Y = Faces_R[i].Eye_R.Y + Faces_R[i].Eye_R.Height/2;
R_Cam_L_Eye_X = Faces_R[i].Eye_L.X + Faces_R[i].Eye_L.Width/2;
R_Cam_L_Eye_Y = Faces_R[i].Eye_L.Y + Faces_R[i].Eye_L.Height/2;
L_Cam_R_Eye_X = Faces_L[i].Eye_R.X + Faces_L[i].Eye_R.Width/2;
L_Cam_R_Eye_Y = Faces_L[i].Eye_R.Y + Faces_L[i].Eye_R.Height/2;
L_Cam_L_Eye_X = Faces_L[i].Eye_L.X + Faces_L[i].Eye_L.Width/2;
L_Cam_L_Eye_Y = Faces_L[i].Eye_L.Y + Faces_L[i].Eye_L.Height/2;

To calculate the centerpoint of X of the right eye of the right camera, X of the right eye of the right face an integer, summed up with width of the right eye of the right face also an integer all divided by 2. R_Cam_R_Eye_X = Faces_R[i].Eye_R.X + Faces_R[i].Eye_R.Width/2

To calculate the centerpoint of Y of the right eye of the right camera, Y of the right eye of the right face an integer, summed up with height of the right eye of the right face also an integer all divided by 2. R_Cam_R_Eye_Y = Faces_R[i].Eye_R.Y + Faces_R[i].Eye_R.Height/2

To calculate the centerpoint of X of the left eye of the right camera, X of the left eye of the right face an integer, summed up with width of the left eye of the right face also an integer all divided by 2. R_Cam_L_Eye_X = Faces_R[i].Eye_L.X + Faces_R[i].Eye_L.Width/2

To calculate the centerpoint of Y of the left eye of the right camera, Y of the left eye of the right face an integer, summed up with height of the left eye of the right face also an integer all divided by 2. R_Cam_L_Eye_Y = Faces_R[i].Eye_L.Y + Faces_R[i].Eye_L.Height/2

To calculate the centerpoint of X of the right eye of the left camera, X of the right eye of the left face an integer, summed up with width of the right eye of the left face also an integer all divided by 2. L_Cam_R_Eye_X = Faces_L[i].Eye_R.X + Faces_L[i].Eye_R.Width/2

To calculate the centerpoint of Y of the right eye of the left camera, Y of the right eye of the left face an integer, summed up with height of the right eye of the left face also an integer all divided by 2. L_Cam_R_Eye_Y = Faces_L[i].Eye_R.Y + Faces_L[i].Eye_R.Height/2

To calculate the centerpoint of X of the left eye of the left camera, X of the left eye of the left face an integer, summed up with width of the left eye of the left face also an integer all divided by 2. L_Cam_L_Eye_X = Faces_L[i].Eye_L.X + Faces_L[i].Eye_L.Width/2

To calculate the centerpoint of Y of the left eye of the left camera, Y of the left eye of the left face an integer, summed up with height of the left eye of the left face also an integer all divided by 2. L_Cam_L_Eye_Y = Faces_L[i].Eye_L.Y + Faces_L[i].Eye_L.Height/2

3.5 RECOGNIZE FACE DIRECTION

After many trials, the reading was carried out by the use of training mechanism, the program was read for a long time during the training and many values were generated.

Getting the mean and standard deviation was used to see the proper location according to certain specific points.

The user is sitting in front of a robot with a big screen behind the robot. The robot is trying to detect which direction the user is looking at with respect to positions. The position detected by the robot is Position1-9.

Figure 3.6.1a POSITION 1

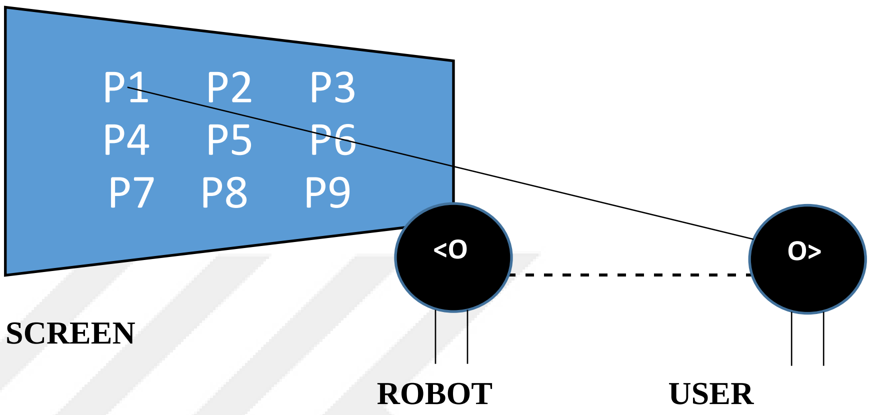


Figure 3.6.1b POSITION 1 Result

P1	P2	P3
P4	P5	P6
P7	P8	P9

Position 1 detected.

In the figures above (figure 3.6.1a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 1 (P1), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 1. The robot detected the user sitting in front of it is looking at position 1 and the result is been reflected on the result table indicator as shown in figure 3.6.1b.

Figure 3.6.2a POSITION2

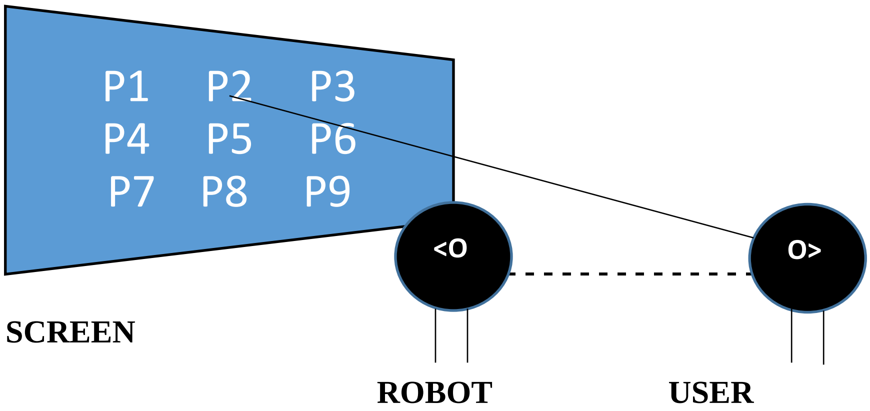


Figure 3.6.2b POSITION2 Result

P1	**P2**	P3
P4	P5	P6
P7	P8	P9

Position 2 detected.

In the figures above (figure 3.6.2a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 2 (P2), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 2. The robot detected the user sitting in front of it is looking at position 2 and the result is been reflected on the result table indicator as shown in figure 3.6.2b.

Figure 3.6.3a POSITION 3

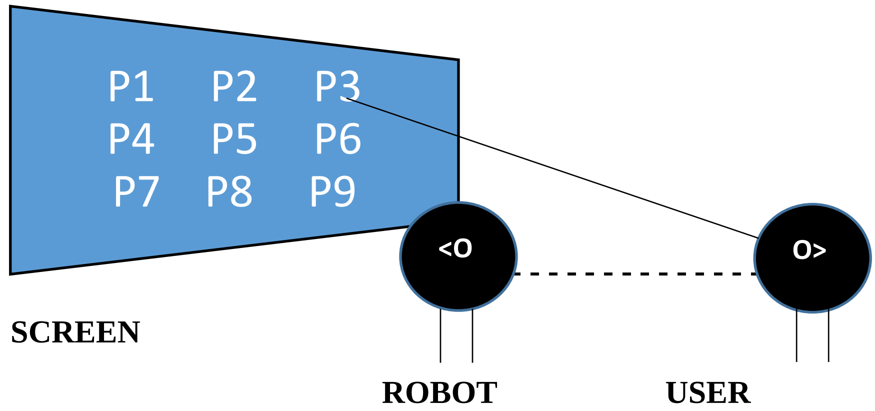


Figure 3.6.3b POSITION3 Result

P1	P2	**P3**
P4	P5	P6
P7	P8	P9

Position 3 detected.

In the figures above (figure 3.6.3a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 3 (P3), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 3. The robot detected the user sitting in front of it is looking at position 3 and the result is been reflected on the result table indicator as shown in figure 3.6.3b.

Figure 3.6.4a POSITION 4

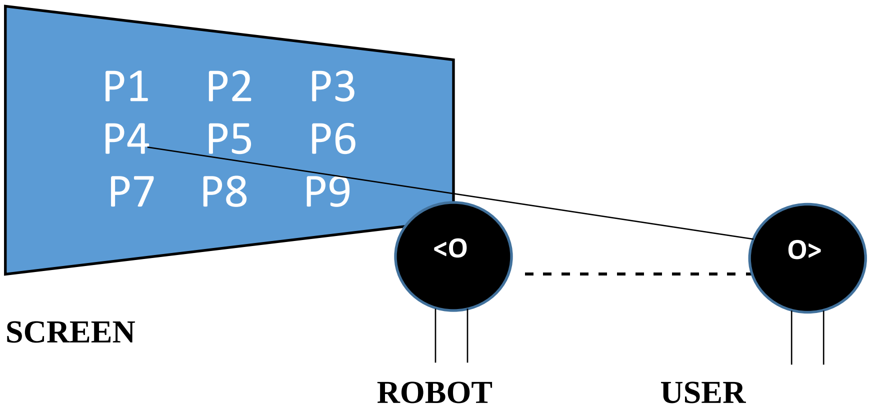


Figure 3.6.4b POSITION4 Result

P1	P2	P3
P4	P5	P6
P7	P8	P9

Position 4 detected.

In the figures above (figure 3.6.4a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 4 (P4), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 4. The robot detected the user sitting in front of it is looking at position 4 and the result is been reflected on the result table indicator as shown in figure 3.6.4b.

Figure 3.6.5a POSITION 5

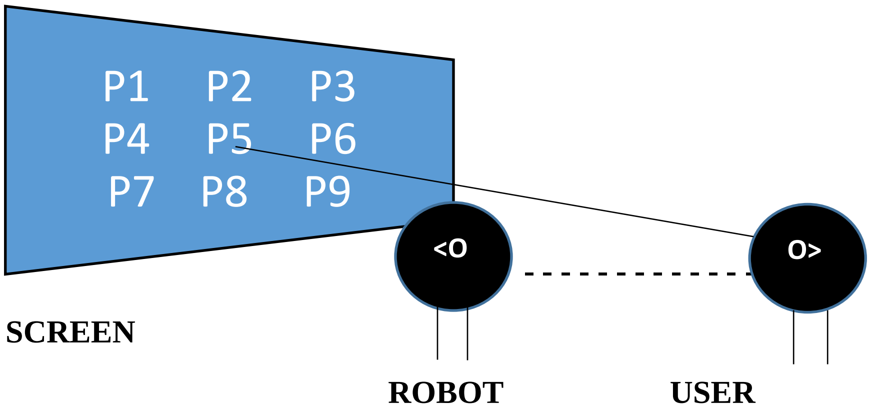


Figure 3.6.5b POSITION5 Result

P1	P2	P3
P4	**P5**	P6
P7	P8	P9

Position 5 detected.

In the figures above (figure 3.6.5a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 5 (P5), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 5. The robot detected the user sitting in front of it is looking at position 5 and the result is been reflected on the result table indicator as shown in figure 3.6.5b.

Figure 3.6.6a POSITION 6

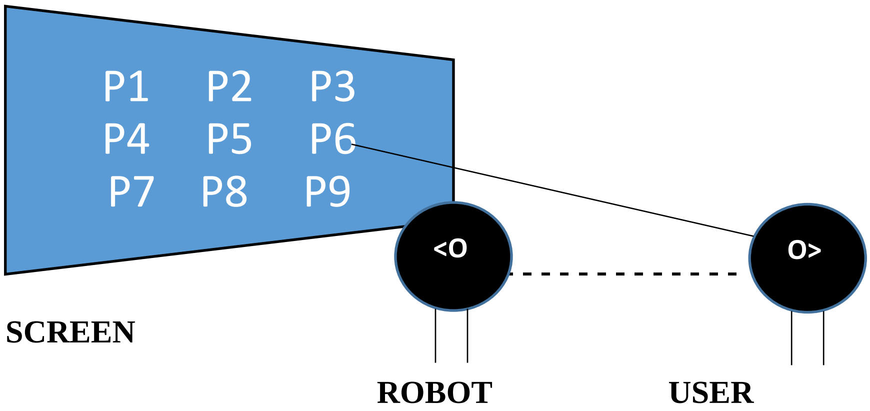


Figure 3.6.6b POSITION6 Result

P1	P2	P3
P4	P5	**P6**
P7	P8	P9

Position 6 detected.

In the figures above (figure 3.6.6a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 6 (P6), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 6. The robot detected the user sitting in front of it is looking at position 6 and the result is been reflected on the result table indicator as shown in figure 3.6.6b.

Figure 3.6.7a POSITION 7

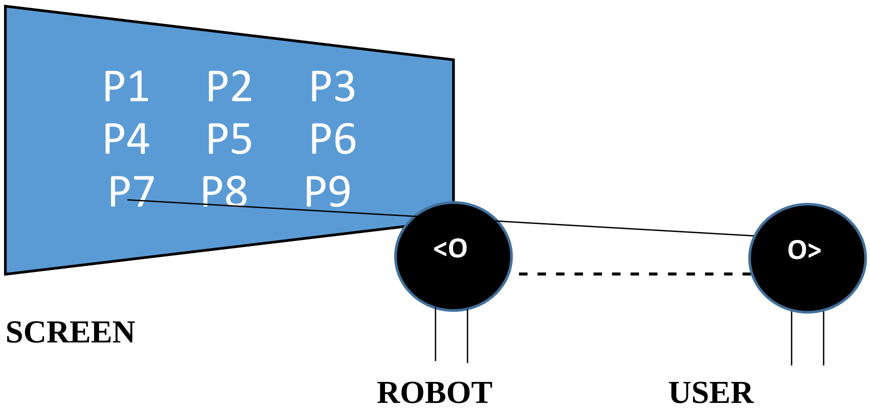


Figure 3.6.7b POSITION7 Result

P1	P2	P3
P4	P5	P6
P7	P8	P9

Position 7 detected.

In the figures above (figure 3.6.7a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 7 (P7), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 7. The robot detected the user sitting in front of it is looking at position 7 and the result is been reflected on the result table indicator as shown in figure 3.6.7b.

Figure 3.6.8a POSITION 8

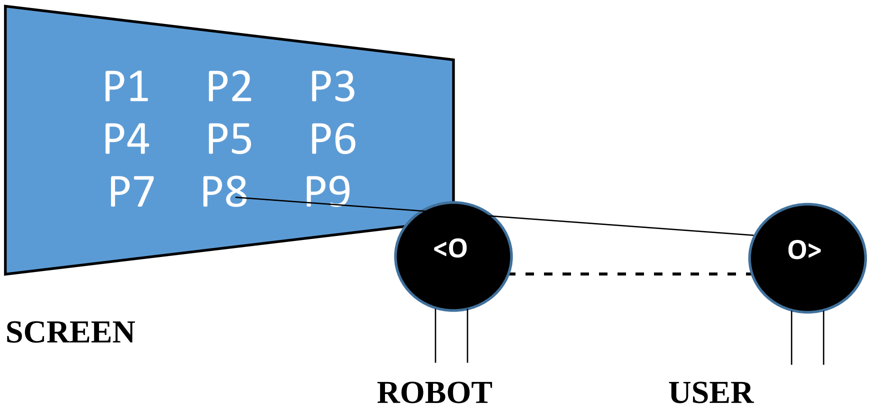


Figure 3.6.8b POSITION8 Result

P1	P2	P3
P4	P5	P6
P7	**P8**	P9

Position 8 detected.

In the figures above (figure 3.6.8a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 8 (P8), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 8. The robot detected the user sitting in front of it is looking at position 8 and the result is been reflected on the result table indicator as shown in figure 3.6.8b.

Figure 3.6.9a POSITION 9

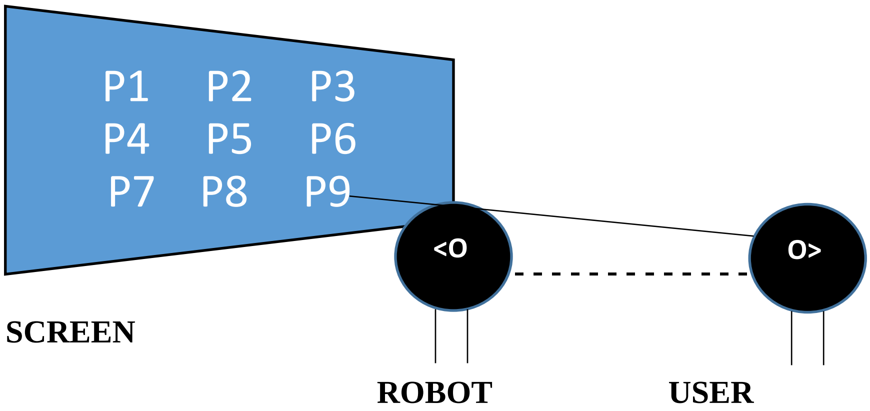


Figure 3.6.9b POSITION9Result

P1	P2	P3
P4	P5	P6
P7	P8	**P9**

Position 9 detected.

In the figures above (figure 3.6.9a), The user is sitting in front of a robot with a big screen behind the robot. The user looking at the direction of position 9 (P9), calculating the angle of eye wih repect to head movement, the robot detects the use is looking at position 9. The robot detected the user sitting in front of it is looking at position 8 and the result is been reflected on the result table indicator as shown in figure 3.6.9b.

CHAPTER 4

RESULTS

4.1 SYSTEM SETUP

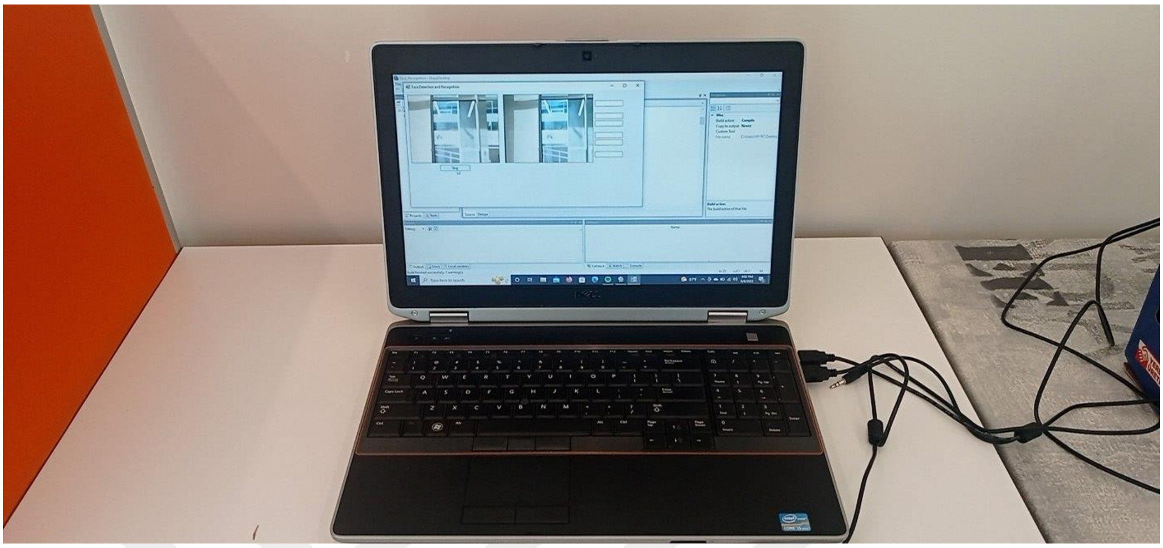


Figure 4.1: Computer System Dell Latitude E6520 x-64 Based PC running on OS Microsoft Windows 10 Pro.

As shown in figure 4.1, the image is a picture of the computer system, blow are the specifications and qualities of the computer system:

Processor: Intel Core i5-2540M 2 x 2.6 - 3.3 GHz.

Operating System: Microsoft Windows 10 Pro

Graphics adapter: Intel HD Graphics 3000

Memory: 4096 MB, DDR3

Display: 15.60 inch 16:9, 1366 x 768 pixel, PPCTF-156WH4, not gossy

Mainboard: Intel QM67

Storage: WDC Scorpio Blue WD2500BEVT-75A23T0, 250 GB, 5400 rpm

Soundcard: HD Audio

Connections: 1 Express Card 54mm, 4 USB 2.0, 1 Firewire, 1 VGA, 1 HDMI, 1 Kensington Lock, 1 eSata, 1 Docking Station Port, Audio Connections: line out/microphone combination, Card Reader: SD, 1 SmartCard, E-module bay II (DVD/30Wh battery)

Networking: Intel 82579LM Gigabit Network Connection (10/100/1000MBit/s), Dell Wireless 1501 Wireless-N Half-Mini Card (g/n = Wi-Fi 4)

Optical drive: TSSTcorp CDDVDW TS-U633J

Size: height x width x depth (in mm): 34.2 x 384 x 258 (= 1.35 x 15.12 x 10.16 in)

Battery: 60 Wh Lithium-Ion, 6 cell T54FJ 11.1V

Weight: 2.819 kg (= 99.44 oz / 6.21 pounds), Power Supply: 244 g (= 8.61 oz / 0.54 pounds)

Figure 4.2: 2 NUO cameras with focal length of 3.85mm megapixel 10x Digital zoom. Mounted on a flat prime base 10cm apart.

As seen above in figure 4.2, there are 2 cameras sitting on a base 10cm apart. Both camera having same specifications of focal length of 3.85mm megapixel 10x Digital zoom. The camera is the image capturing device. The cameras labelled and represented as LeftCamera and RightCamera.

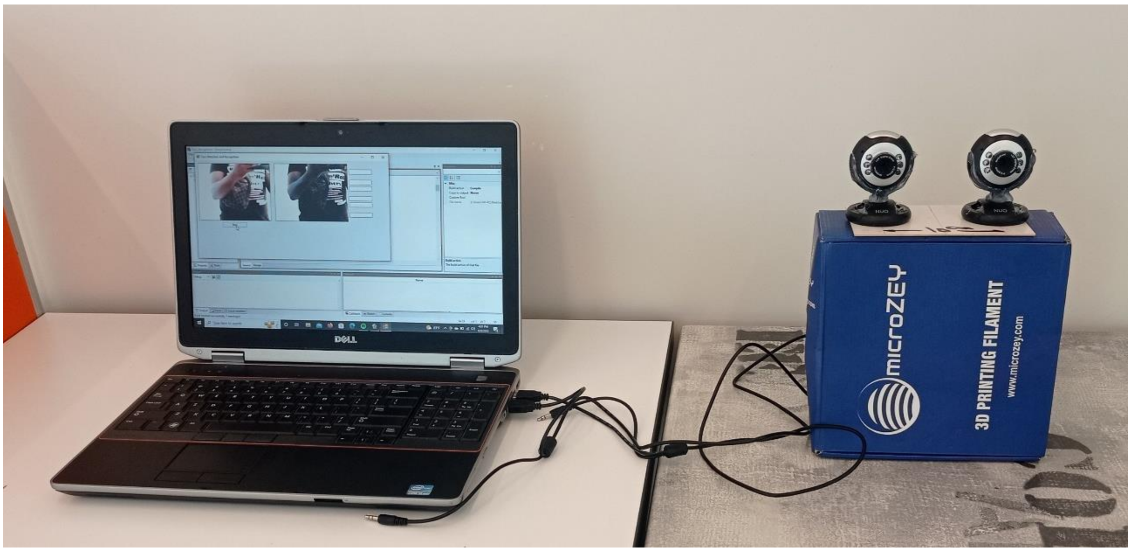


Figure 4.3: The Machine.

As shown in figure 4.3 above, the computer and cameras connected together to make up the machine. The image cameras being the capturing devices and the computer is the processing and display device.

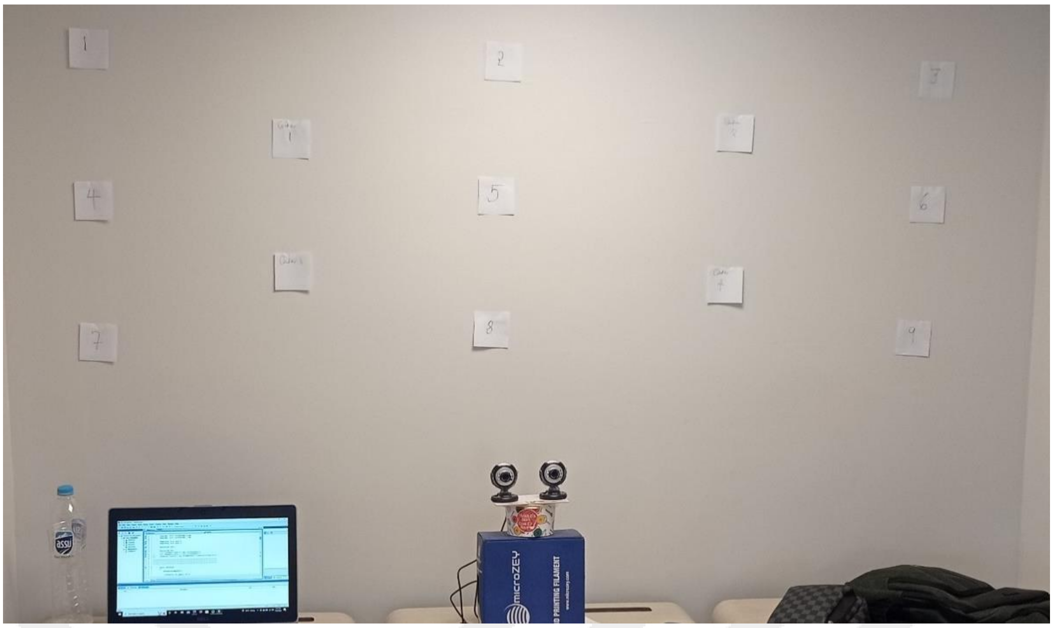


Figure 4.4 Point 1 to 9 during training. Points pinned on the wall representing positions.

As shown in figure 4.4 above, pinning of numbers on the wall, numbers representing positions numeriacally. This is a procedure during training. Point 1 to 9 are being pinned on the wall, with the midpoint of p1,p2,p4,p5 labelled as centerpoint1, the midpoint of p2,p3,p5,p6 labelled as centerpoint2, the midpoint of p4,p5,p7,p8 labelled as centerpoint3 and midpoint of p5,p6,p8,p9 labelled as centerpoint4.

Horizontal distance between points 70cm

Vertical distance between points 23cm

Each point (from 1 to 9) is being captured, calcuated and the data is been saved as the output of the procedure.

4.2 DATA ACQUISITION PROCEDURE

Sitting in front of the machine, and the machine detects the face and locate the eyes. The eyes are been kept at a particular point and positions (between point 1 to 9) at different times, the machine generates data figures for the points angles of those positions looked at with respect to placement, the machine records the results of data generated and saves/stores it in a text file for each position. This procedure and process is been run over and over again until it reaches the 9 points, and results are gotten and saved.

Position1

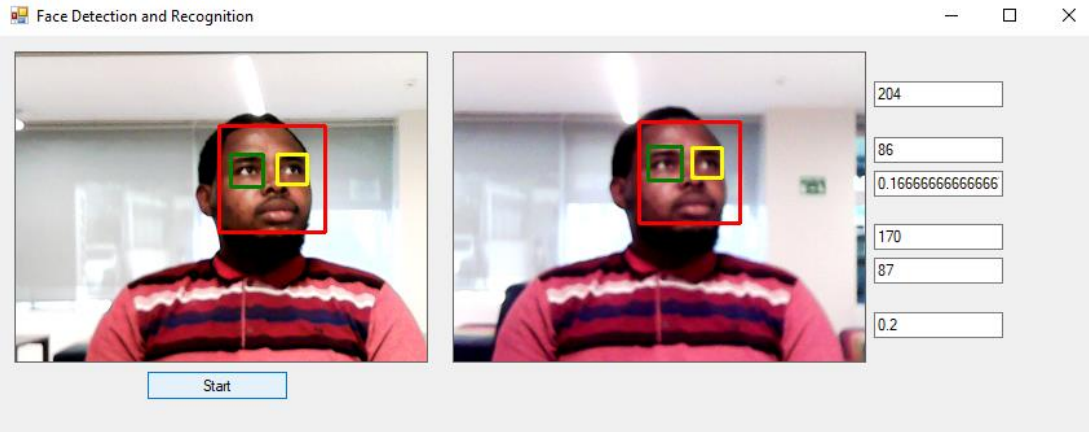


Figure 4.5.1 Training capturing for position 1 with results

In the figure above (figure 4.5.1), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 1 (P1), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 1, the result of the capture states;

X of Right eye = 204,

Y of Right eye = 86,

Z of Right eye = 0.1666,

X of Left eye = 170,

Y of Left eye = 87,

Z of left eye = 0.2.

Position2

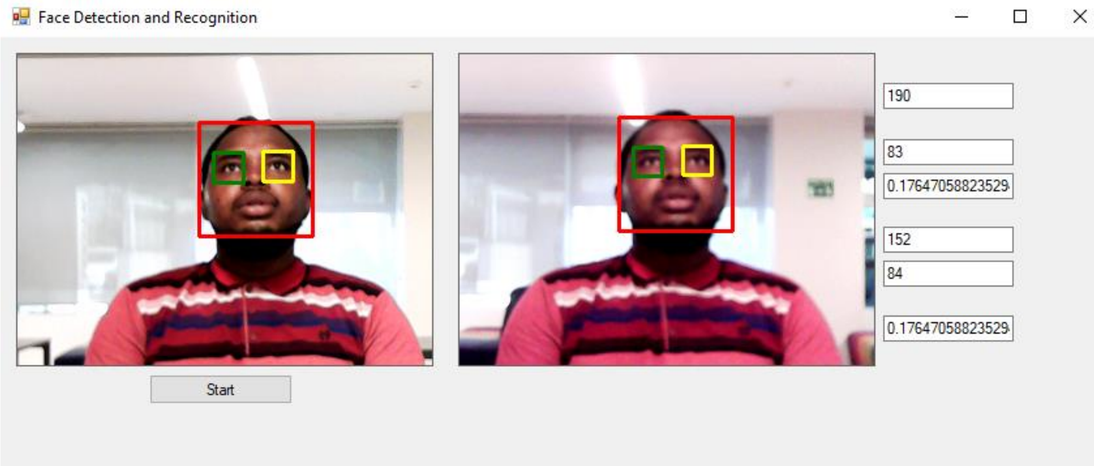


Figure 4.5.2 Training capturing for position 2 with results

In the figure above (figure 4.5.2), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 2 (P2), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 2, the result of the capture states;

X of Right eye = 190,

Y of Right eye = 83,

Z of Right eye = 0.1764,

X of Left eye = 152,

Y of Left eye = 84,

Z of left eye = 0.1764.

Position3

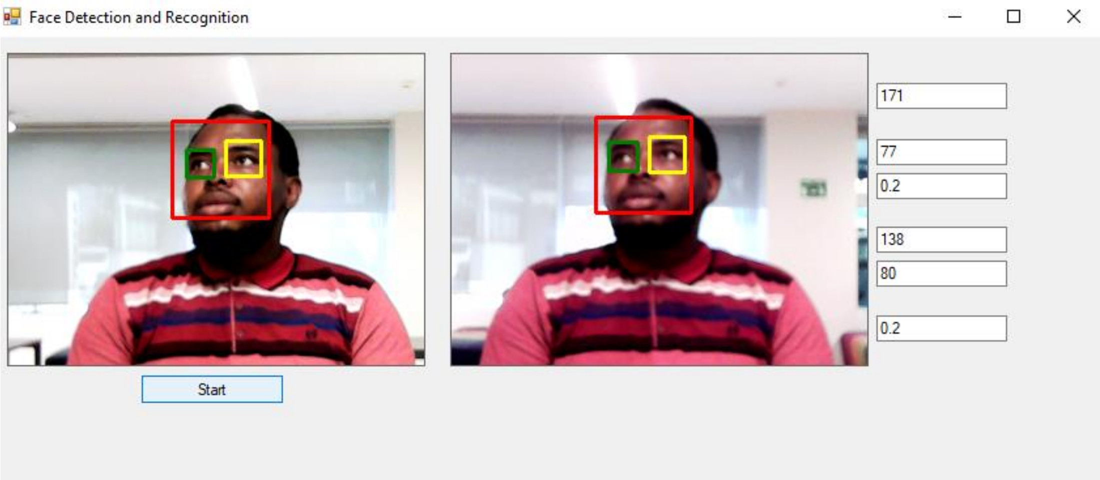


Figure 4.5.3 Training capturing for position 3 with results

In the figure above (figure 4.5.3), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 3 (P3), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 3, the result of the capture states;

X of Right eye = 171,

Y of Right eye = 77,

Z of Right eye = 0.2,

X of Left eye = 138,

Y of Left eye = 80,

Z of left eye = 0.2.

Position4

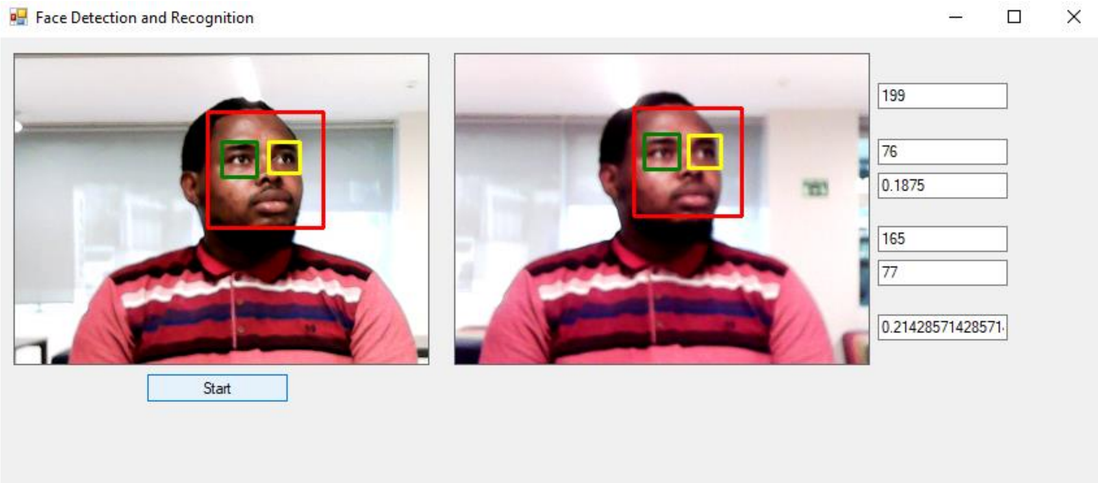


Figure 4.5.4 Training capturing for position 4 with results

In the figure above (figure 4.5.4), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 4 (P4), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 4, the result of the capture states;

X of Right eye = 199,

Y of Right eye = 76,

Z of Right eye = 0.1875,

X of Left eye = 165,

Y of Left eye = 77,

Z of left eye = 0.2142.

Position5

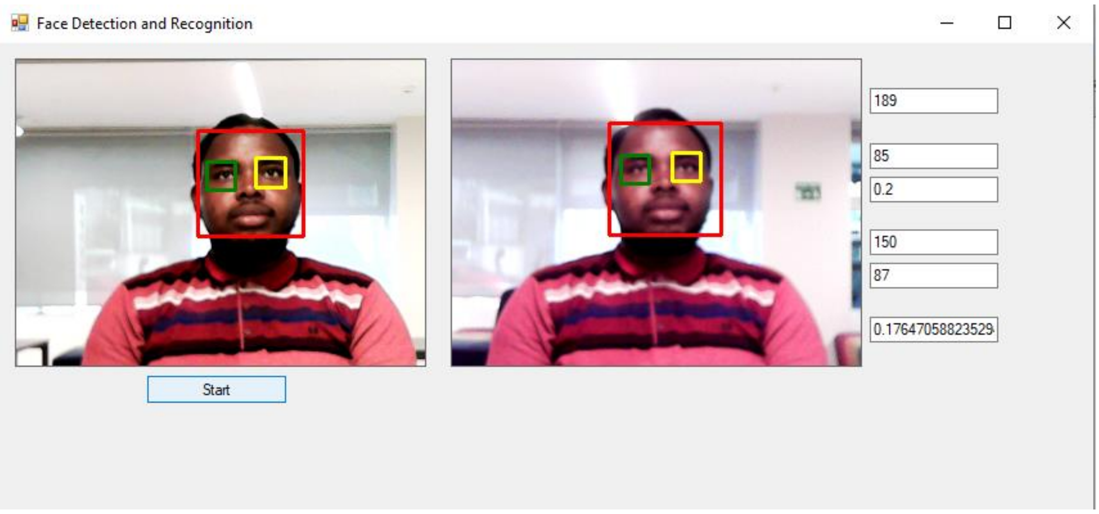


Figure 4.5.5 Training capturing for position 5 with results

In the figure above (figure 4.5.5), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 5 (P5), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 5, the result of the capture states;

X of Right eye = 189,

Y of Right eye = 85,

Z of Right eye = 0.2,

X of Left eye = 150,

Y of Left eye = 87,

Z of left eye = 0.1764.

Position6

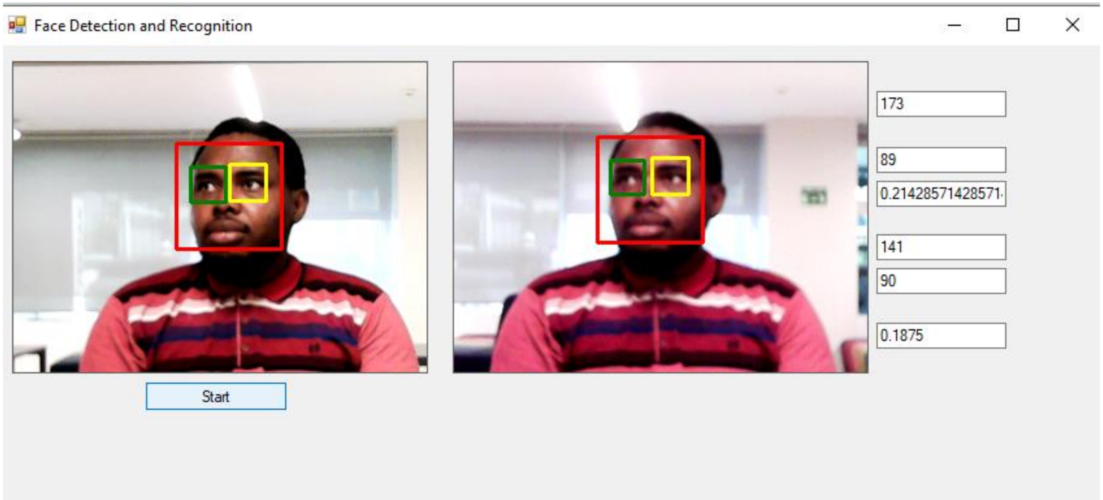


Figure 4.5.6 Training capturing for position 6 with results

In the figure above (figure 4.5.6), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 6 (P6), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 6, the result of the capture states;

X of Right eye = 173,

Y of Right eye = 89,

Z of Right eye = 0.2142,

X of Left eye = 138,

Y of Left eye = 90,

Z of left eye = 0.1875.

Position7

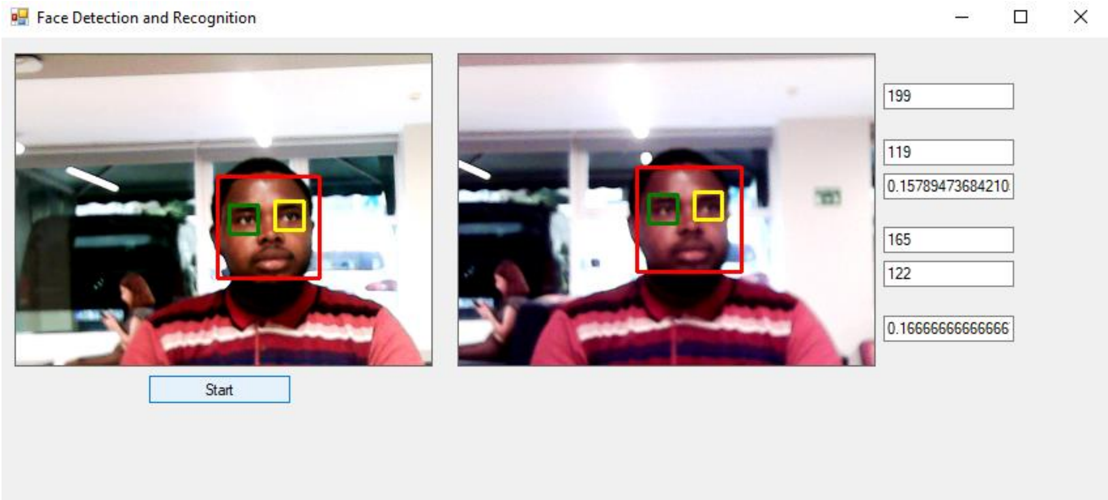


Figure 4.5.7 Training capturing for position 7 with results

In the figure above (figure 4.5.7), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 7 (P7), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 7, the result of the capture states;

X of Right eye = 199,

Y of Right eye = 119,

Z of Right eye = 0.1578,

X of Left eye = 165,

Y of Left eye = 122,

Z of left eye = 0.1666.

Position8

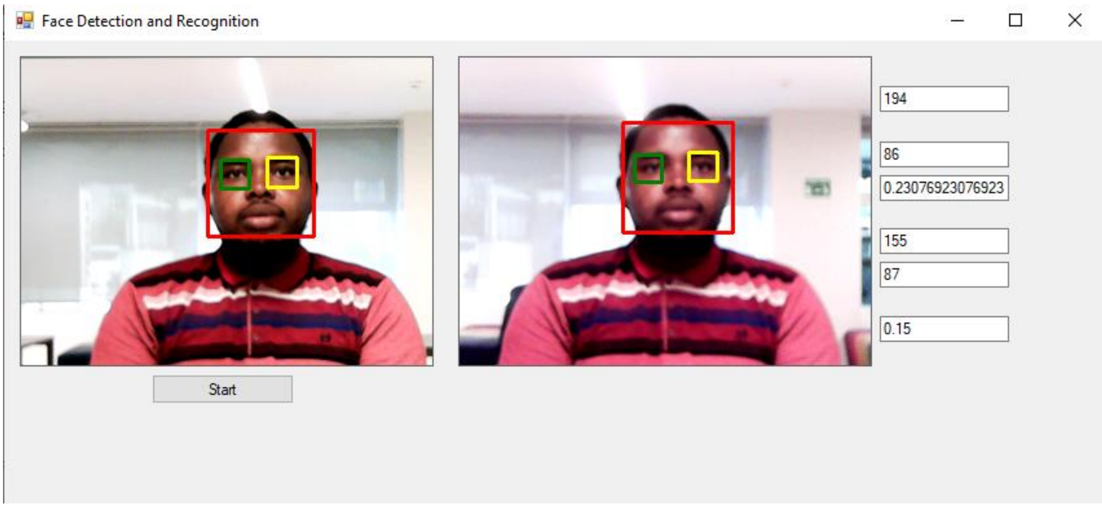


Figure 4.5.8 Training capturing for position 8 with results

In the figure above (figure 4.5.8), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 8 (P8), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 8, the result of the capture states;

X of Right eye = 194,

Y of Right eye = 86,

Z of Right eye = 0.2307,

X of Left eye = 155,

Y of Left eye = 87,

Z of left eye = 0.15.

Position9

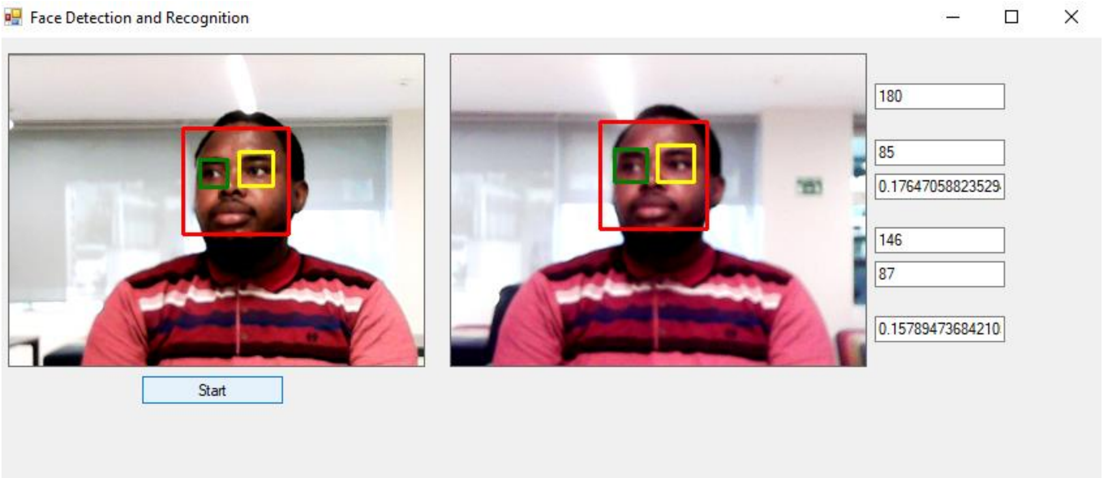


Figure 4.5.9 Training capturing for position 9 with results

In the figure above (figure 4.5.9), as shown in the figure, a training procedure is being processed, the user sitting in front of the machine looking at position 9 (P9), his face has been captured and both eyes (right and left) has been captured in the appropriate proportion and sizes. The face is been captured and indicated by a red colored squared shape, while the right eye is captured in a green small square shape and the left eye is captured in a yellow small square shape. Both eyes differentiated by colors are in appropriate proportion and right positon. And the results are been generated.

The result gotten are represented as XYZ for the right eye and XYZ for the left eyes as shown in the text boxes. The robot captures the face and eyes, and calculates the angles with respect to movement.

For position 9, the result of the capture states;

X of Right eye = 180,

Y of Right eye = 85,

Z of Right eye = 0.1764,

X of Left eye = 146,

Y of Left eye = 87,

Z of left eye = 0.1578.

4.3 DATA ANALYSIS (TRAINING)

Training is finding the mean and average of many readings for face specification.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Position1	160	50	0.25	123	61	0.3	Position2	170	55	0.2143	144	56	0.2143	Position3	180	53	0.2	148	55	0.2143
2		159	49	0.3	124	60	0.25		170	55	0.25	142	56	0.2143		181	54	0.2308	147	56	0.2727
3		160	50	0.2308	124	61	0.3333		169	56	0.25	146	54	0.25		180	53	0.2727	148	54	0.25
4		160	51	0.2308	124	61	0.3333		171	54	0.2727	144	56	0.2308		180	53	0.2308	147	53	0.2143
5		159	52	0.2308	123	61	0.3333		170	54	0.2727	145	56	0.25		180	53	0.2143	147	54	0.2308
6		160	49	0.25	125	60	0.3		170	56	0.25	145	55	0.25		180	53	0.2143	148	55	0.25
7		161	50	0.25	124	61	0.2727		169	54	0.2	144	55	0.2308		180	53	0.2143	146	54	0.2308
8		160	50	0.25	123	59	0.2727		170	55	0.3	145	55	0.25		179	53	0.2143	145	53	0.2143
9		161	50	0.2308	126	60	0.3		171	56	0.2	142	53	0.25		180	52	0.25	147	54	0.2143
10		160	49	0.2	124	56	0.3333		170	55	0.3	143	54	0.25		180	53	0.2143	147	52	0.25
11																					
12	Average	160	50		124	60			170	55		144	55			180	53		147	54	

Table 4.1

	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP
1	Position4	161	59	0.2308	125	61	0.2727	Position5	171	61	0.0588	152	57	0.3	Position6	182	59	0.25	147	57	0.272
2		161	62	0.25	126	59	0.25		171	61	0.0588	152	55	0.3		182	60	0.25	148	57	0.2
3		161	58	0.2308	125	60	0.3		172	60	0.25	139	60	0.25		181	60	0.25	147	58	0.2
4		161	60	0.2	126	61	0.2308		173	62	0.25	142	56	0.25		183	59	0.2	147	53	0.272
5		162	60	0.2308	126	59	0.2727		172	63	0.2308	145	56	0.25		184	60	0.2308	146	54	0.230
6		160	61	0.2308	130	61	0.3		173	61	0.2308	144	56	0.25		183	59	0.2	147	53	0.272
7		162	59	0.2727	129	60	0.3		171	60	0.2727	144	56	0.2308		182	59	0.3	146	56	0.230
8		160	60	0.1765	127	59	0.2727		172	61	0.25	145	57	0.25		180	59	0.25	146	57	0.230
9		161	61	0.25	129	60	0.2308		173	62	0.2308	143	56	0.25		181	58	0.25	147	57	0.2
10		161	60	0.2	127	60	0.2727		172	59	0.25	144	58	0.2308		182	57	0.25	147	58	0.
11																					
12		161	60		127	60			172	61		145	56.7			182	59		146.8	56	

Table 4.2

	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL	BM	BN
1	Position7	162	69	0.2308	128	59	0.25	Position8	171	71	0.2727	144	56	0.25	Position9	180	69	0.2727	145	61	0.3			
2		164	71	0.2308	130	60	0.25		171	74	0.2727	144	56	0.2143		181	69	0.3	144	61	0.2143			
3		160	68	0.2308	124	61	0.3333		170	71	0.25	143	56	0.25		180	70	0.1875	146	53	0.2143			
4		160	70	0.2143	122	62	0.25		172	70	0.2308	145	56	0.25		179	70	0.2143	147	52	0.25			
5		161	70	0.2	122	62	0.3333		170	70	0.2308	144	56	0.25		180	71	0.2308	147	53	0.25			
6		159	71	0.2727	122	61	0.3		171	71	0.2727	144	56	0.2308		180	70	0.2308	145	52	0.25			
7		161	70	0.2143	123	61	0.3333		171	69	0.25	143	55	0.25		182	71	0.2143	147	52	0.2308			
8		160	70	0.2143	123	61	0.2727		172	71	0.2308	144	55	0.25		180	69	0.2143	146	52	0.25			
9		162	71	0.2	122	62	0.25		171	72	0.25	145	56	0.2308		178	71	0.2143	146	52	0.2308			
10		161	70	0.25	124	61	0.2727		171	71	0.25	144	58	0.2308		180	70	0.2727	147	52	0.2308			
11																								
12		161	70		124	61			171	71		144	56			180	70		146	54				

Table 4.3

Average of X in position1 = 160	Average of X in position6 = 182
Average of Y in position 1 = 50	Average of Y in position6 = 59
Average of X in position2 = 170	Average of X in position7 = 161
Average of Y in position2 = 55	Average of Y in position7 = 70
Average of X in position3 = 180	Average of X in position8 = 171
Average of Y in position3 = 53	Average of Y in position8 = 71
Average of X in position4 = 161	Average of X in position9 = 180
Average of Y in position4 = 60	Average of Y in position9 = 70
Average of X in position5 = 172	
Average of Y in position5 = 61	

Table 4.4

As shown in the tables above (Tables 4.1,4.2,4.3) are showing figures generated from the training process of positions 1 to 9 captured and data were generated. While Table 4.4 shows the calcualed average of X and Y for the positions 1 to 9. The figure displayed in the tables are data generated from 10 different training results for 9 different positions (positions 1 to 9).

From Table 4.4 the data below is generated

(160,50) **Position1**	**(170,55)** **Position2**	**(180,53)** **Position3**
(161,60) **Position4**	**(172,61)** **Position5**	**(182,59)** **Position6**
(161,70) **Position7**	**(171,71)** **Position8**	**(180,70)** **Position9**

Table 4.5 As shown above in (Table 4.5), the table shows the mean values for positions 1 to 9 in tabular form.

4.4 VERIFICATION (TESTING)

Testing is trying to find and check the samples in other to specify the eyes.

Verification process is a procedure that has to be run and processed multiple times over and over again just like the training process. The user sits in front of the machine(robot), then, the face and eyes are being captured, angles are then calculated with respect to head movement. And results are being generated, and then, data is being saved and stored in a text file.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Position1	161	50	0.214286	134	57	0.3	Position2	170	55	0.214286	136	56	0.230769	Position3	180	54	0.214286	149	58	0.2
2		160	50	0.2	133	56	0.3		169	54	0.230769	136	58	0.272727		179	53	0.230769	149	58	0.2
3		160	50	0.230769	135	57	0.272727		170	56	0.214286	137	58	0.214286		180	52	0.25	147	57	0.23076
4		159	49	0.214286	133	60	0.3		171	55	0.214286	138	57	0.214286		180	53	0.25	148	57	0.23076
5		160	51	0.25	137	59	0.25		170	55	0.230769	139	57	0.25		181	53	0.25	148	58	0.23076
6																					
7	Average	160	50						170	55						180	53				
8																					

Table 4.6

	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ	AK	AL	AM	AN	AO	AP
1	Position4	161	60	0.2	134	65	0.230769	Position5	171	61	0.214286	145	63	0.230769	Position6	182	59	0.230769	147	64	0.25
2		162	59	0.230769	135	64	0.214286		172	62	0.214286	146	64	0.25		181	58	0.214286	146	63	0.230769
3		161	60	0.2	129	68	0.214286		173	60	0.2	146	64	0.230769		183	60	0.214286	146	63	0.214286
4		160	61	0.214286	132	68	0.3		171	61	0.214286	146	63	0.25		182	59	0.230769	146	63	0.25
5		161	60	0.25	132	68	0.3		173	61	0.214286	146	64	0.214286		182	59	0.230769	146	63	0.25
6																					
7		161	60						172	61						182	59				
8																					

Table 4.7

	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	BC	BD	BE	BF	BG	BH	BI	BJ	BK	BL	BM	BN	B
1	Position7	162	70	0.214234	136	66	0.428571	Position8	171	71	0.25	136	66	0.428571	Position9	181	70	0.25	151	68	0.25				
2		161	70	0.288235	133	66	0.272727		171	72	0.214286	136	67	0.25		179	70	0.214286	151	69	0.2				
3		159	69	0.2625	134	67	0.272727		172	70	0.230769	137	67	0.25		180	70	0.25	152	69	0.272727				
4		161	70	0.214286	135	67	0.333333		170	70	0.230769	137	67	0.214286		181	70	0.272727	152	69	0.25				
5		162	71	0.2	136	68	0.230769		171	72	0.25	135	68	0.25		179	70	0.230769	151	69	0.214286				
6																									
7		161	70						171	71						180	70								
8																									
9																									

Table 4.8

Average of X in position1 = 160	Average of X in position6 = 182
Average of Y in position1 = 50	Average of Y in position6 = 59
Average of X in position2 = 170	Average of X in position7 = 161
Average of Y in position2 = 55	Average of Y in position7 = 70
Average of X in position3 = 180	Average of X in position8 = 171
Average of Y in position13= 53	Average of Y in position8 = 71
Average of X in position4 = 161	Average of X in position9 = 180
Average of Y in position4 = 60	Average of Y in position9 = 70
Average of X in position5 = 172	
Average of Y in position5 = 61	

Table 4.9

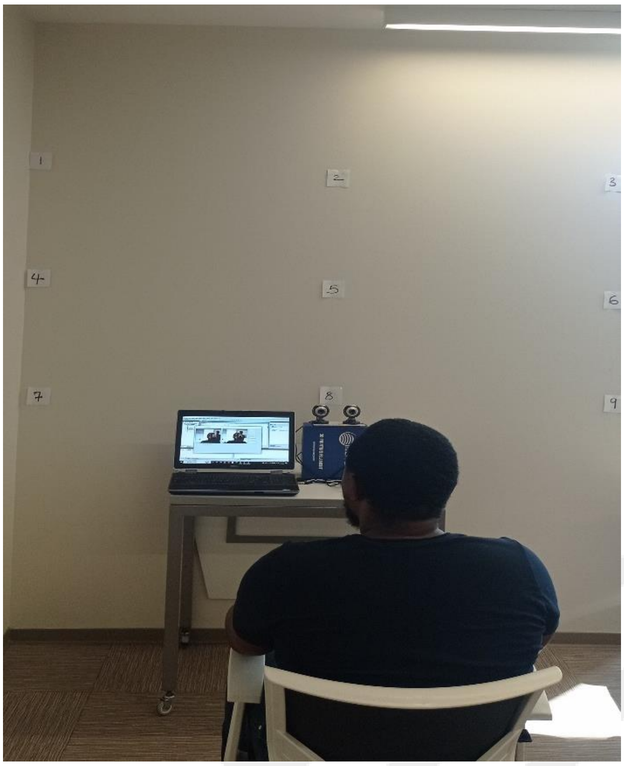
As shown in the tables above (Tables 4.6,4.7,4.8) are showing figures generated from the testing process of positions 1 to 9 captured and data were generated. While Table 4.9 shows the calcualed average of X and Y for the positions 1 to 9. The figure displayed in the tables are data generated from 5 different testing results for 9 different positions (positions 1 to 9).

From Table 4.9 the data below is generated

(160,50) Position1	**(170,55) Position2**	**(180,53) Position3**
(161,60) Position4	**(172,61) Position5**	**(182,59) Position6**
(161,70) Position7	**(171,71) Position8**	**(180,70) Position9**

Table 4.10 As shown above in (Table 4.10), the table shows the mean values for positions 1 to 9 in tabular form.

Position 1



P1	**P2**	**P3**
P4	**P5**	**P6**
P7	**P8**	**P9**

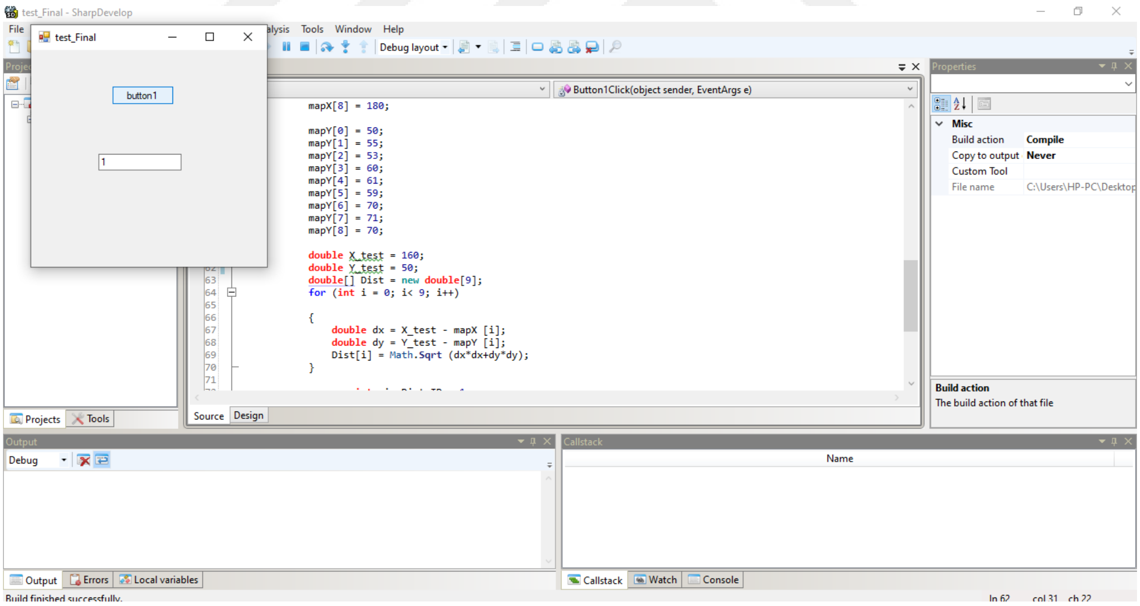
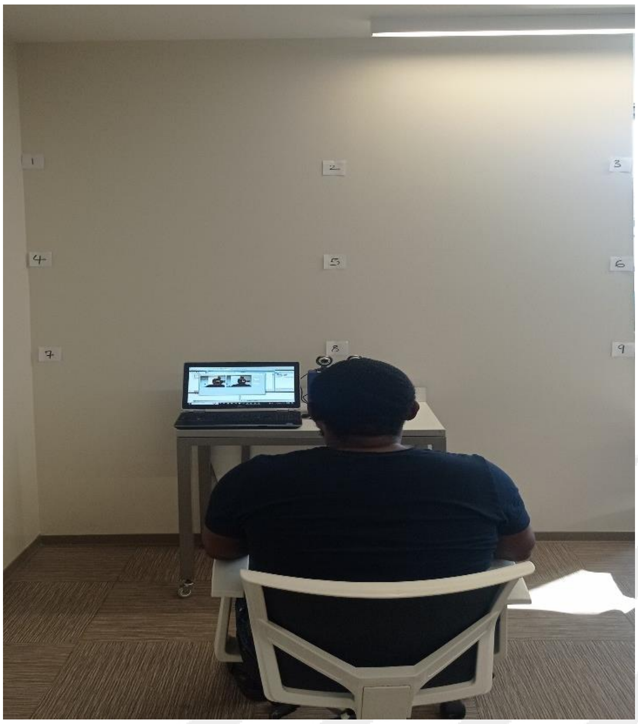


Figure 4.6.1 Verification procedure for position 1 with position indication image and result.

Verification process and procedure for position 1 (P1) as shown in the figure (figure 4.6.1) above.

The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 160,50 and 161,51 are been used as the parameters and the machine automatically detected the user is looking at position 1 because the output is 1.

Position 2



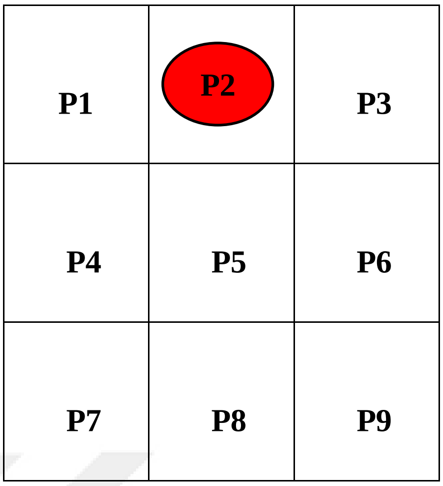
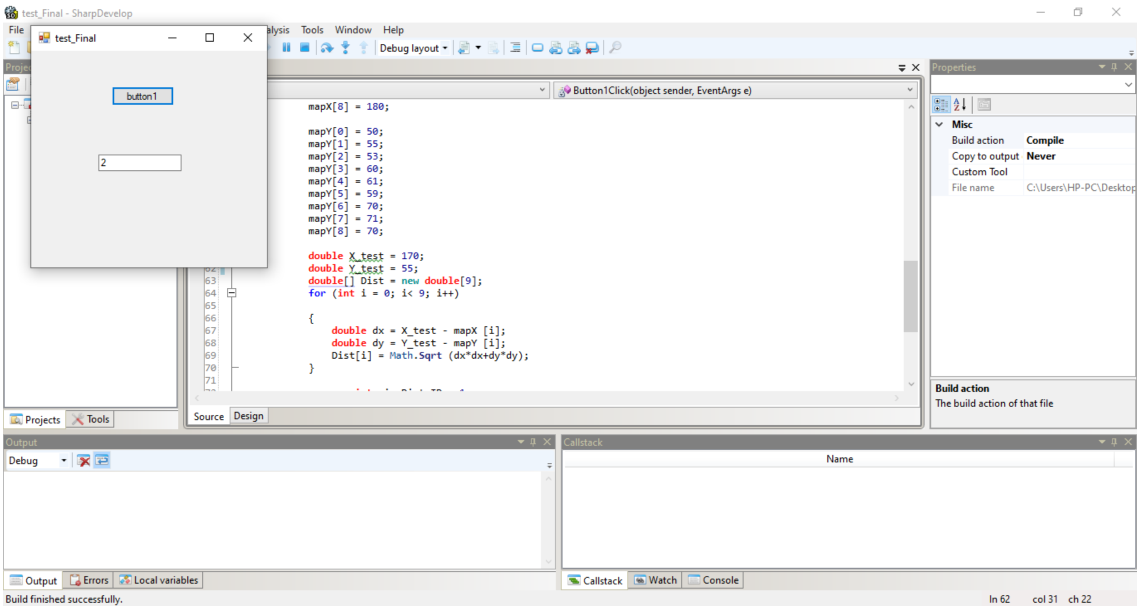


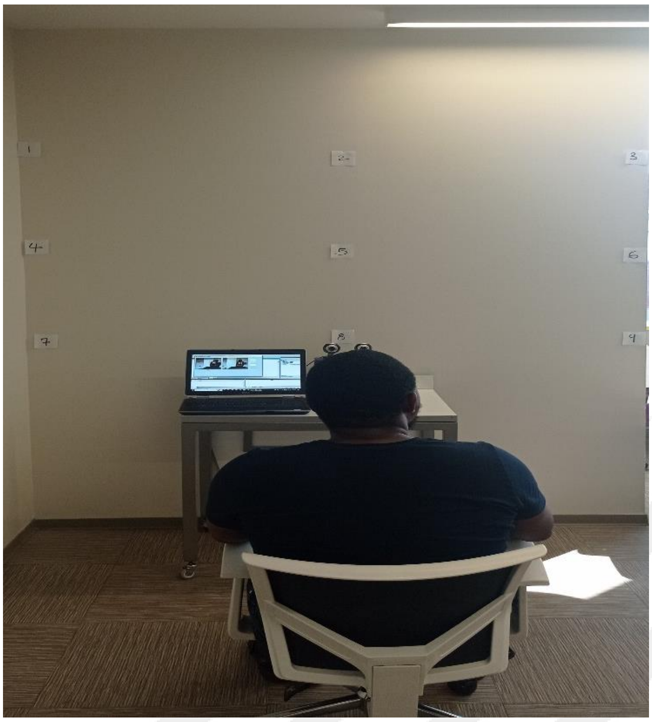
Figure 4.6.2 Verification procedure for position 2 with position indication image and result.

Verification process and procedure for position 2 (P2) as shown in the figure (figure 4.6.2) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 170,55 and 172,56 are been used as the parameters and the machine automatically detected the user is looking at position 2 because the output is 2.

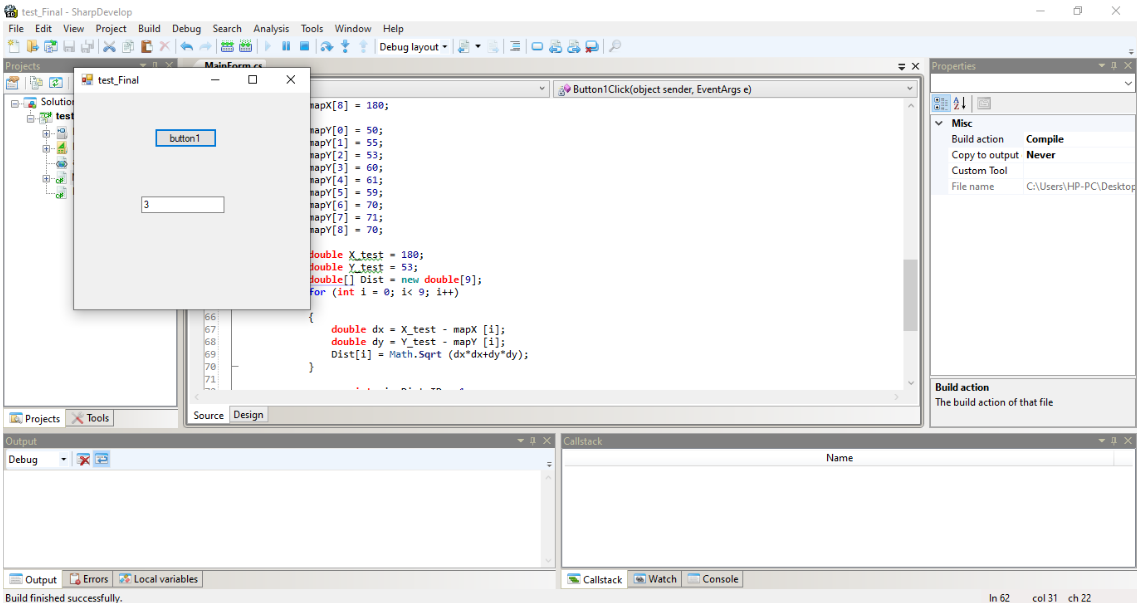
Position 3



P1	P2	P3
P4	P5	P6
P7	P8	P9

Figure 4.6.3 Verification procedure for position 1 with position indication image and result.

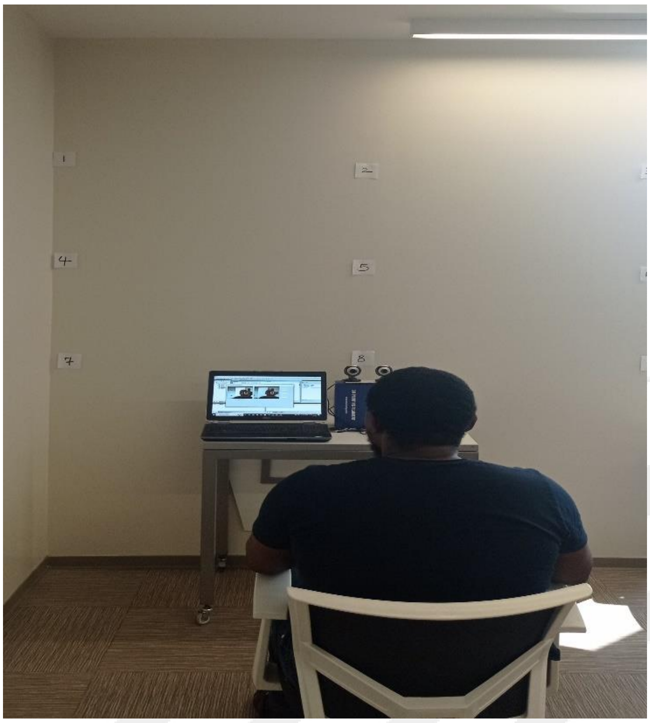
Verification process and procedure for position 3 (P3) as shown in the figure (figure 4.6.3) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are

entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 180,53 and 184,55 are been used as the parameters and the machine automatically detected the user is looking at position 3 because the output is 3.

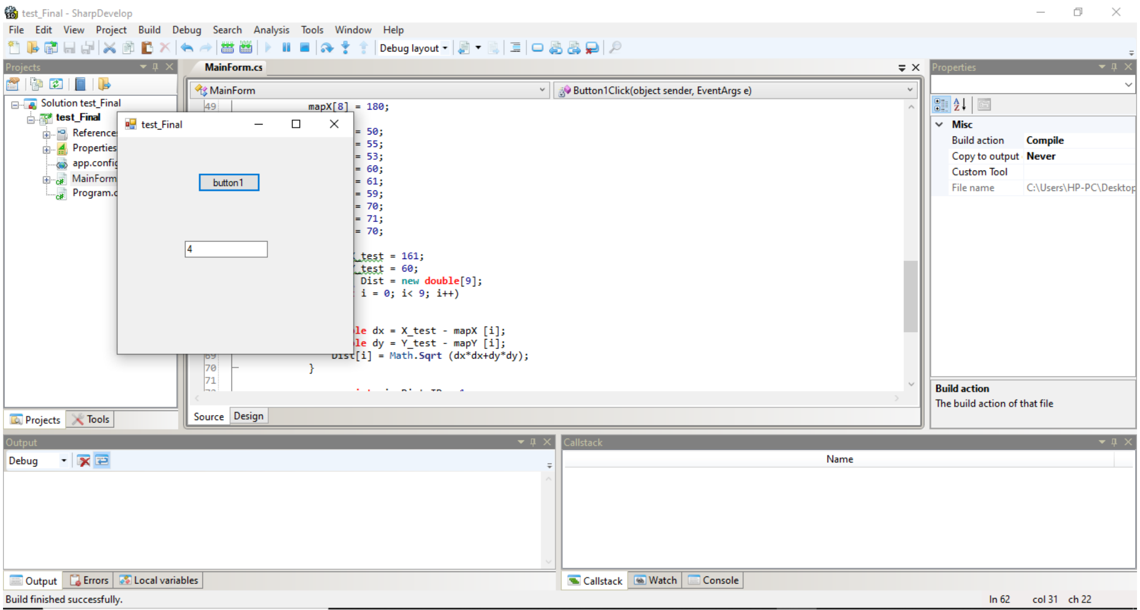
Position 4



P1	**P2**	**P3**
P4	**P5**	**P6**
P7	**P8**	**P9**

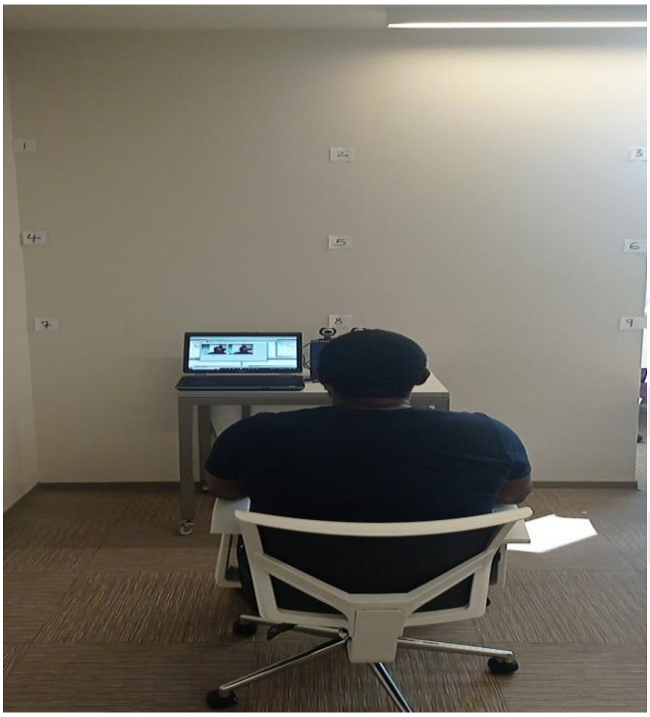
Figure 4.6.4 Verification procedure for position 4 with position indication image and result.

Verification process and procedure for position 4 (P4) as shown in the figure (figure 4.6.4) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 161,60 and 163,65 are been used as the parameters and the machine automatically detected the user is looking at position 4 because the output is 4.

Position 5



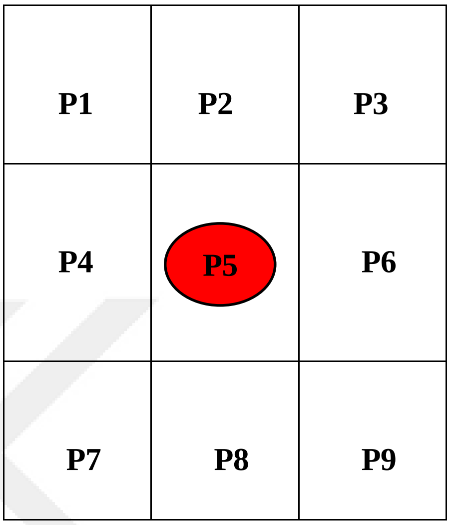
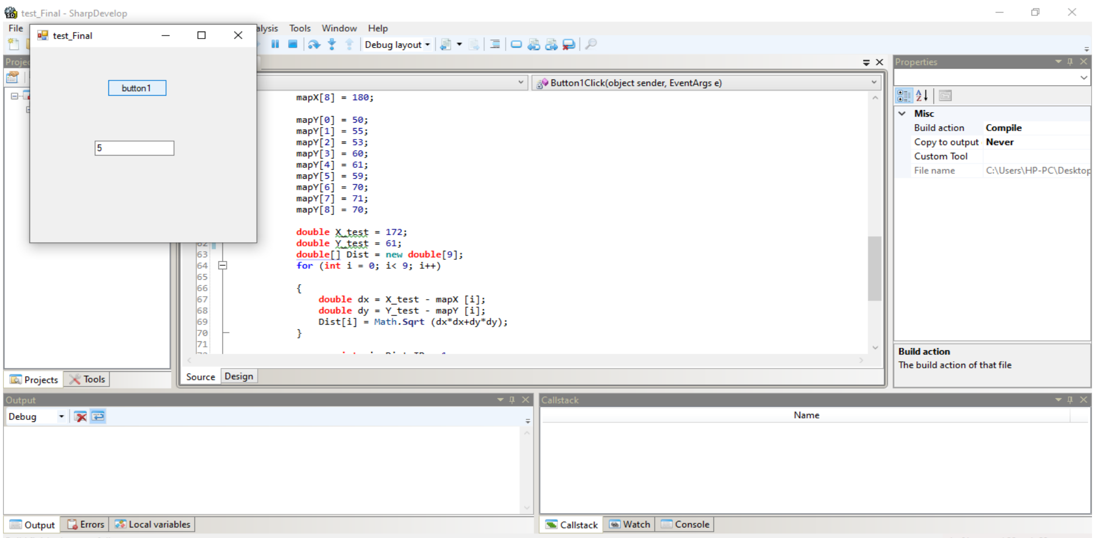


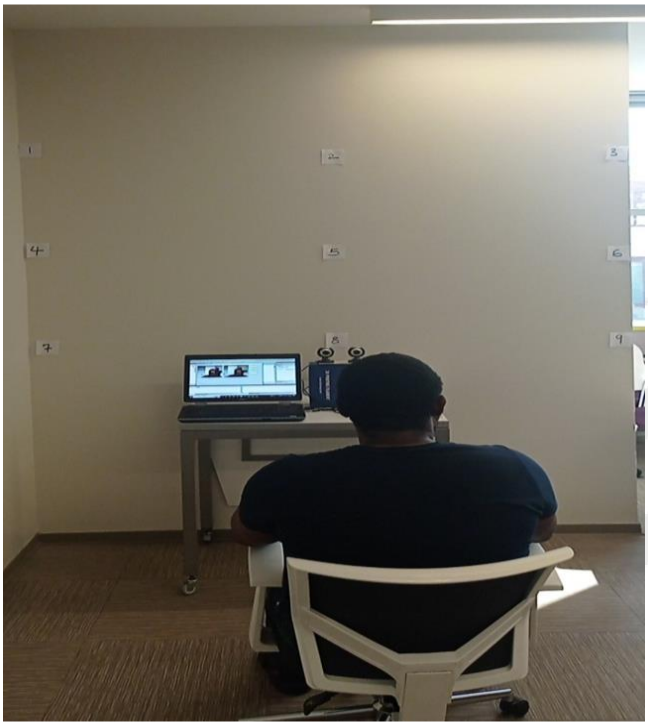
Figure 4.6.5 Verification procedure for position 5 with position indication image and result.

Verification process and procedure for position 5 (P5) as shown in the figure (figure 4.6.5) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 172,61 and 176,60 are been used as the parameters and the machine automatically detected the user is looking at position 5 because the output is 5.

Position 6



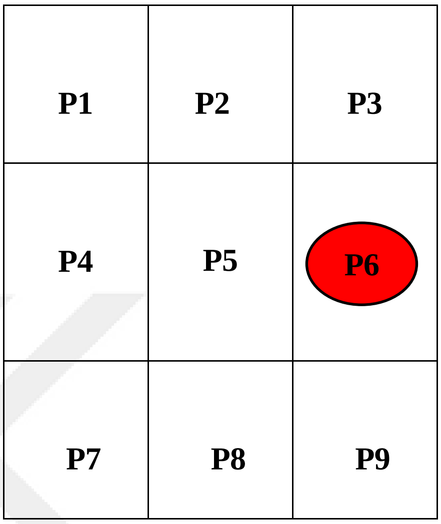
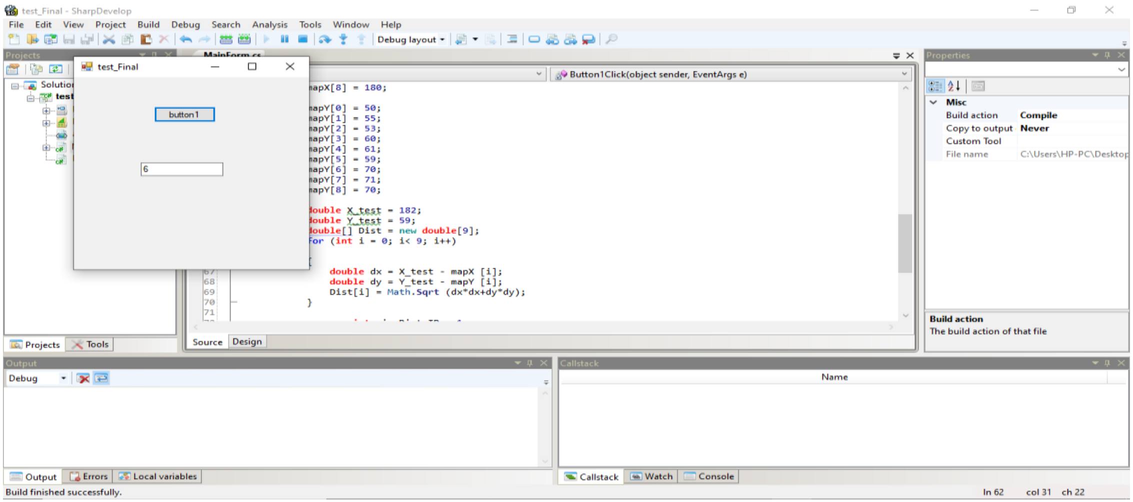


Figure 4.6.6 Verification procedure for position 6 with position indication image and result.

Verification process and procedure for position 6 (P6) as shown in the figure (figure 4.6.6) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are

entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 182,59 and 187,62 are been used as the parameters and the machine automatically detected the user is looking at position 6 because the output is 6.

Position 7

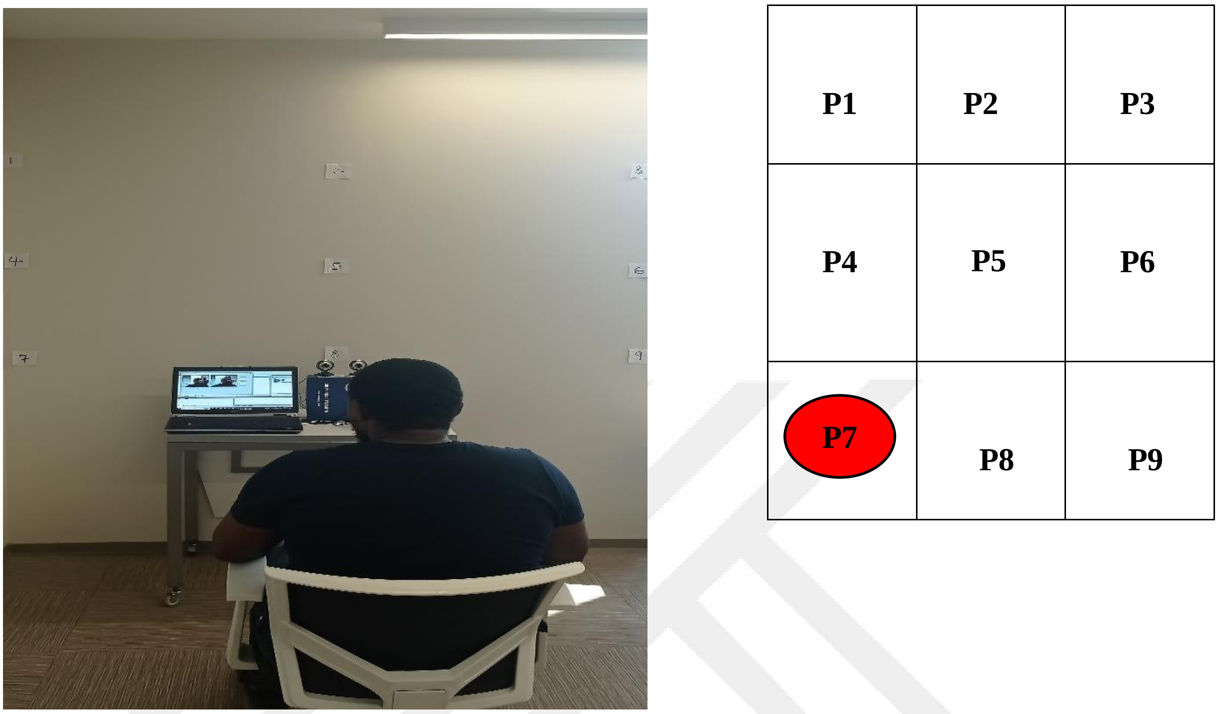
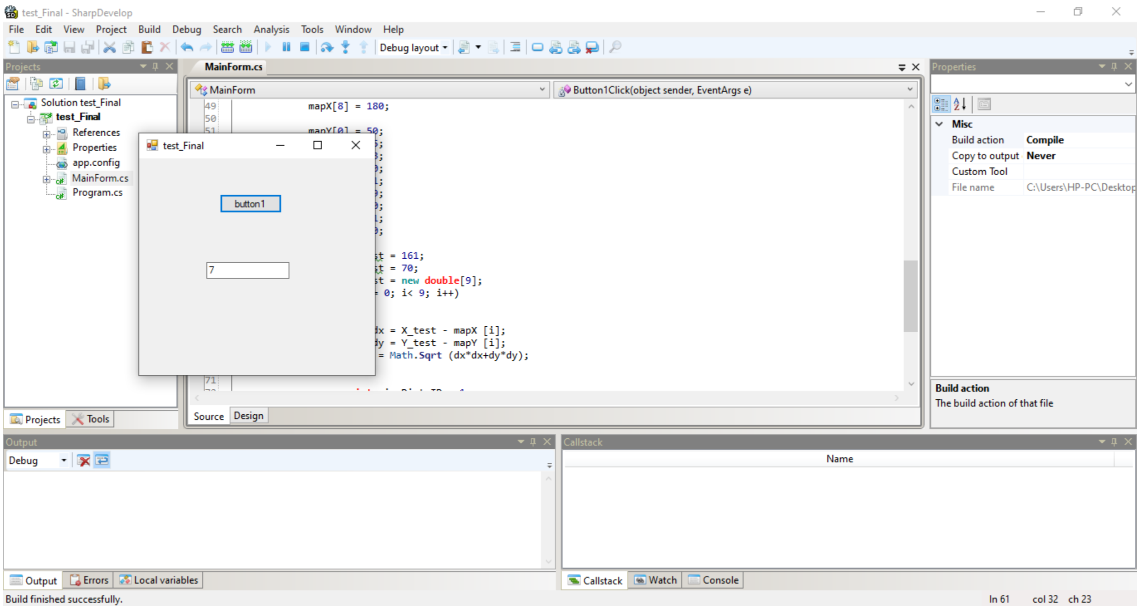


Figure 4.6.7 Verification procedure for position 7 with position indication image and result.

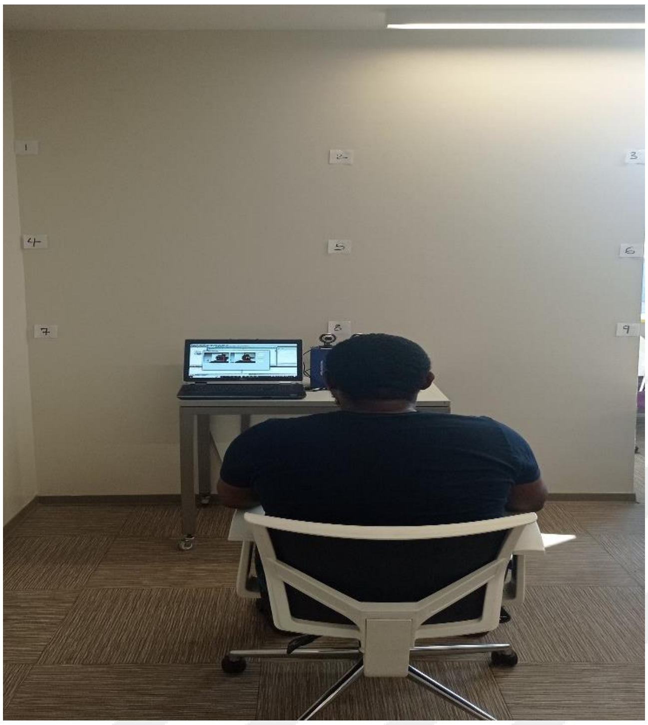
Verification process and procedure for position 7 (P7) as shown in the figure (figure 4.6.7) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are

entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 161,70 and 166,69 are been used as the parameters and the machine automatically detected the user is looking at position 7 because the output is 7.

Position 8



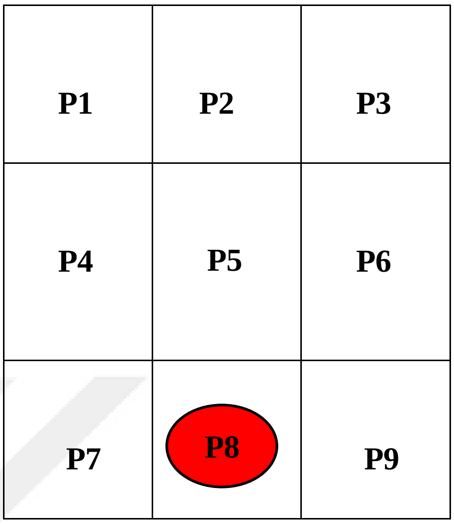
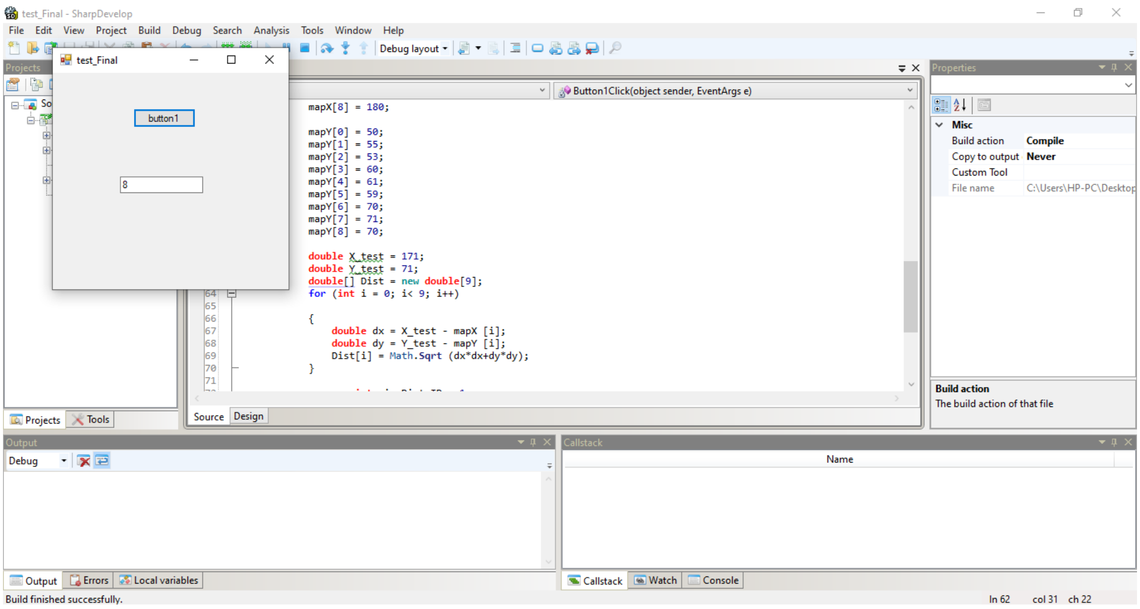


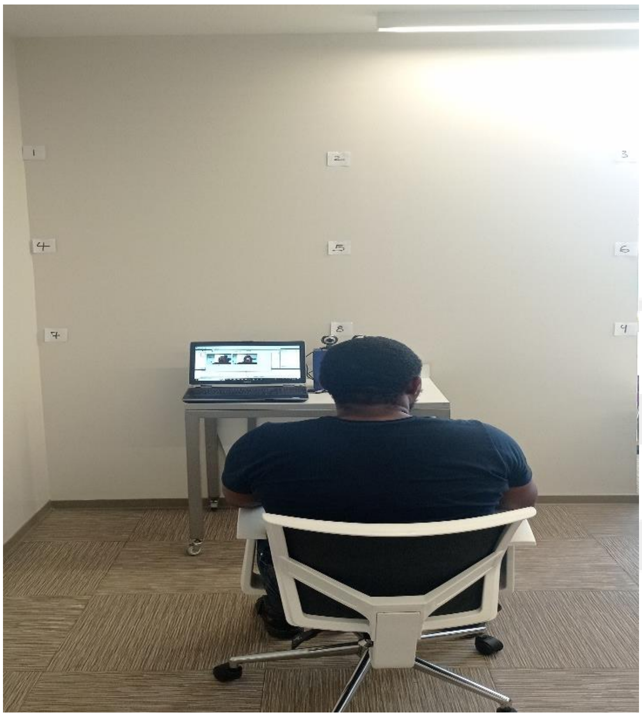
Figure 4.6.8 Verification procedure for position 8 with position indication image and result.

Verification process and procedure for position 8 (P8) as shown in the figure (figure 4.6.8) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 171,71 and 174,70 are been used as the parameters and the machine automatically detected the user is looking at position 8 because the output is 8.

Position 9



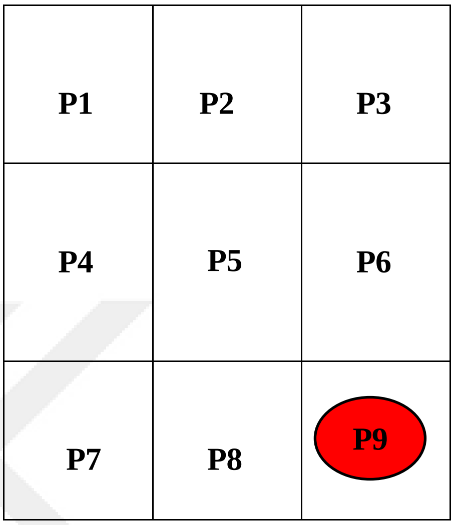
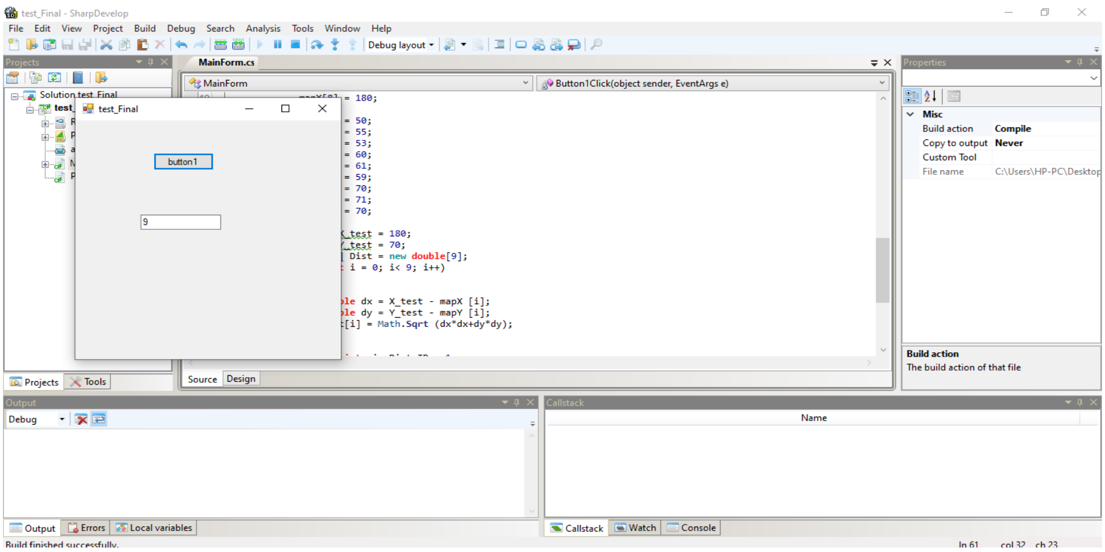


Figure 4.6.9 Verification procedure for position 9 with position indication image and result.

Verification process and procedure for position 9 (P9) as shown in the figure (figure 4.6.9) above.



The mean of the training generated figures are been saved and the mean of the testing generated figures are been verified simultaneously on the machine, when the figures are entered and checked, the machine automatically calculates the vector and tells the possible position the user is looking at from the figures generated and inputed. From the result above 180,70 and 178,74 are been used as the parameters and the machine automatically detected the user is looking at position 9 because the output is 9.

4.5 GRID MODEL.

1	**2**	**3**
4	**5**	**6**
7	**8**	**9**

Table 4.11: A 3-by-3 grid model

The table above is a 3-by-3 grid model of proving idea of a precise vector. More grids could be added in the future 5-by-5, 7-by-7, 9-by-9 are also model type of grid. But on this project, a 3-by-3 grid model because I have a low resolution camera, when higher resolution cameras are gotten in the future, other high grid model could be experimented. Low resolution cameras cannot give accurate eyes location, thereby cause limitations to accurate results.

The solution to this limitation is getting an higher resolution cameras. By getting those, the grids could be increased from 3-by-3 to 5-by-5, 7-by-7, 9-by-9 to get more precise vectors.

Grid numbers are meant to always be odd and not even because of the center-line.

1	**2**	**3**	**4**	**5**
6	**7**	**8**	**9**	**10**
11	**12**	**13**	**14**	**15**
16	**17**	**18**	**19**	**20**
21	**22**	**23**	**24**	**25**

Table 4.12: A 5-by-5 grid model

1	**2**	**3**	**4**	**5**	**6**	**7**
8	**9**	**10**	**11**	**12**	**13**	**14**
15	**16**	**17**	**18**	**19**	**20**	**21**
22	**23**	**24**	**25**	**26**	**27**	**28**
29	**30**	**31**	**32**	**33**	**34**	**34**
35	**36**	**37**	**38**	**39**	**40**	**41**
42	**43**	**44**	**45**	**46**	**47**	**48**

Table 4.13: A 7-by-7 grid model

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
10	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
19	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**
28	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**
37	**38**	**39**	**40**	**41**	**42**	**43**	**44**	**45**
46	**47**	**48**	**49**	**50**	**51**	**52**	**53**	**54**
55	**56**	**57**	**58**	**59**	**60**	**61**	**62**	**63**
64	**65**	**66**	**67**	**68**	**69**	**70**	**71**	**72**
73	**74**	**75**	**76**	**77**	**78**	**79**	**80**	**81**

Table 4.14: A 9-by-9 grid model

CHAPTER 5

CONCLUSION AND FUTURE WORKS

Face detection has received a lot of interest recently from researchers in the domains of biometrics, pattern recognition, and computer vision. Face recognition technology are needed in various security and forensic applications. As you can see, face detection technology is crucial to our daily lives. The most accurate biometric system is the face detection and recognition system. We have offered a survey of face detection methods in this post. It's great to see face detection technology being used to more products and applications in the real world. Face detection applications and difficulties were also mentioned, which inspired us to conduct face detection research. The most straightforward future line of action is to improve face identification when there are problems like facial occlusion and uneven illumination. Face recognition in the presence of occlusion and non-uniform lighting is the focus of current research in the subject of face detection and recognition. Face detection has received a lot of attention, and then not when there is a problem with occlusion and inconsistent illumination. If it does, it will greatly aid in the recognition of faces and facial expressions, among other things. Many businesses now offer facial biometrics on mobile phones for access purposes. It will be utilized in the future for things like payments, security, healthcare, advertising, and criminal identification.

REFERENCES

1. R. Anchini, G. Di Leo, C. Liguori and A. Paolillo, "Metrological characterization of a vision-based measurement system for the online inspection of automotive rubber profile", *IEEE Trans. Instrum. Meas.*, vol. 58, no. 1, pp. 4-13, Jan. 2009.
2. G. Di Leo and A. Paolillo, "Uncertainty evaluation of camera model parameters", *Proc. 2011 IEEE Conf. Instrum. Meas. Technol.*, pp. 1-6, 2011.
3. R.A. Brooks,, "A robust layered control system for a mobile robot", *IEEE Trans. Robotics and Automation*, vol. 2, pp. 14-23, Apr. 1986.
4. F. Ferrari,, E. Grosso,, M. Magrassi, and G. Sandini,, "A stereo vision systemfor real time obstacle avoidance in unknown environment", *Proc. Intl Workshop Intelligent Robots and Systems*, 1990-July.
5. J. Aloimonos,, I. Weiss, and A. Bandyopadhyay,, "Active vision", *Intl J. Computer Vision*, vol. 1, no. 4, pp. 333-356, 1988.
6. A. Laurentini. "The visual hull concept for silhouette-based image understanding". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. pp. 150–162, February 1994.
7. J. Aloimonos, "Visual shape computation," *IEEE Proc.*, vol. 76, no. 8, pp. 899-916, 1988.
8. P. J. Besl and R. C. Jain, "Three-dimensional object recognition," *Comput. Surveys*, vol. 17, pp. 75-145, 1985.
9. C. Chien and J. K. Aggarwal, "Model reconstruction and shape recognition from occluding contours," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 11, pp. 372-389, 1989.
10. R. Szeliski, *Computer Vision: Algorithms and Applications*, Springer, 2010.
11. B. H. Bodkin, *Real-time mobile stereo vision [M.S. thesis]*, University of Tennessee, 2012.
12. M. Y. Kim, S. M. Ayaz, J. Park, and Y. Roh, “Adaptive 3D sensing system based on variable magnification using stereo vision and structured light,” *Optics and Lasers in Engineering*, vol. 55, pp. 113–127, 2014.
13. H. Chen and W. Cui, *A Comparative analysis between ative structured light and multi-view stereo vision technique for 3D reconstruction of face model surface.* 2020.
14. Zhang Zhengyou. A flexible new technique for camera calibration*[J]*. *IEEE Transactions on Pattern Analysis and machine Intelligence*, 22 (11): 1330-1334. 2000.
15. K. Achour, M. Benkhelif. A new approach to 3D reconstruction without camera calibration*[J]*. *Pattern Recognition*, 12 (34): 2467-2476, 2001.
16. Wikipedia Free Encyclopedia, http://en.wikipedia.org/wiki/Computer_stereo_vision.
17. Sumi, Yasushi, et al. "3D object recognition in cluttered environments by segment-based stereo vision." *International Journal of Computer Vision* 46.1 (2002): 5-23.

18. Johannes Burge, Victor Rodriguez-Lopez, and Carlos Dorronsoro. "Monovision and the Misperception of Motion." *Science Direct Current Biology* Volume 29, Issue 15, 5 August 2019, Pages 2586-2592.
19. Google search, https://drtcarlson.com/eye-health/what-is-monovision/
20. A. J. Soini, "Machine Vision News", *proceedings of Machine Vision Seminar Finnish Society of Automation*, pp. 7-9, 26.-27 5.1998.
21. Google search, http://www.intel.com/content/www/us/en/manufacturing/what-is-machine-vision.html
22. Andy Wilson, "Keystone of machine vision systems design", *Vision Systems Design 200th anniversary issue*, 2013.
23. K. Etemad and R. Chellappa, "Discriminant Analysis for Recognition of Human Face Images", *J. Optical Soc. Am.*, vol. 14, pp. 1724-1733, 1997.
24. Georgiou T, Liu Y, Chen W, Lew M, "A survey of traditional and deep learning-based feature descriptors for high dimensional data in computer vision". *Int J Multimed Inform Retrieval* 9(3):135–170. 2020.
25. Utton R.S., Barto A.G. "Reinforcement learning: an introduction." *MIT Press, Cambridge*. 2018.
26. Zia I. Carrim, Nicholas M. Hickley, and Fiona Bishop, " Monovision: a refractive consideration in cataract surgery after vitrectomy", *International Ophthalmology* volume 32, pages623–625. 2012.
27. C. Garcia and G. Tziritas, "Face detection using quantized skin color regions merging and Wavelet packet analysis", *IEEE Trans Multimedia*, vol. 1, no. 3, pp. 264-277, 1999.
28. L. Kong, J. Peng, Z. Xia, Y. Lian and S. Fan, "Realization and optimization of face detection algorithm based on DSP", *Proc. of 2010 International Conference of Information Science and Management Engineering (ISME)*, vol. 1, pp. 246-249, 2010.
29. H. Liu, W. Gao, J. Miao and J. Li, "A novel method to compensate variety of illumination in face detection", *Joint Conference on Information Sciences*, pp. 692-695, 2002.
30. J. Wang and T. Tan, "A new face detection method based on shape information", *Pattern Recognition Letters*, vol. 21, pp. 463-471, 1999.
31. B. Miller, "Biometrics: Vital signs of Identity", *IEEE spectrum*, 1994.
32. Jeng, S.H., Liao, H.M. Liu, Y.T., Chern, M.Y., An efficient approach for facial feature detection using geometrical face model. In: *Proceedings of the ICPR* 1996, pp. 426–430
33. Gurlove Singh and Amit Kumar Goel, "Face Detection and Recognition System using Digital Image Processing", *IEEE International Conference on Innovative Mechanisms for Industry Applications*, 2020.
34. F. Ahmad, A. Najam and Z. Ahmed, "Image-based Face Detection and Recognition: State of the Art", *IJCSI International Journal of Computer Science Issues*, vol. 9, no. 6, pp. 169-172, 2013.
35. Daugman J, "How iris recognition works", *IEEE Transactions on Circuits and Systems for Video Technology*, 11 (3) (2004), pp. 21-30.
36. Quan, D., Zhao, Y., Li, J., & Zheng, H. H. An Effective Method and Device for Eye Detection with Physical Properties. In *Proceedings of the 2nd international conference on computer science and electronics engineering.* 2013.

37. Nsaid A.K., Ali S.H.M., Jassim K.N., Nseaf A.K., Sulaiman R., Al-Qaraghuli., Wahdan O., Nayan N.A. "Cascade faster R-NN with gabor filters and naïve Bayes for enhanced eye detection". *IEEE Access,* 2021, pp. 15708-15719.

38. Brown, M., Marmor, M. and Vaegan, ISCEV Standard for Clinical Electro-oculography (EOG) 2006, in: *Documenta Ophthalmologica,* 113:3(205—212)
39. Rayan Al and Sobbahi Joe Tekli. "Low-Light Homomorphic Filtering Network for integrating image enhancement and classification". *Signal Processing: Image Communication.* Vol 100. 2022.
40. Min Young Kim, Shirazi Muhammad Ayaz, Jaechan Park, and YoungJun Roh. "Adaptive 3D sensing system based on variable magnification using stereo vision and structured light", 2014.
41. Melvyn L. Smith, Lyndon N. Smith, Mark F. Hansen. "The quiet revolution in machine vision - a state-of-the-art survey paper, including historical review, perspectives, and future directions". 2021.
42. Pengcheng Wei, Xiangping Yu, Zhenpeng Di, Xiaojun Dai, Bo Wang, Yushan Zeng. "Design of Robot automatic navigation under computer intelligent algorithm and machine vision". 2022.
43. Tiejun Huang, Yajing Zheng, Zhaofei Yu, Rui Chen, Yuan Li, Ruiqin Xiong, Lei Ma, Junwei Zhao, Siwei Dong, Lin Zhu, Jianing Li, Shanshan Jia, Yihua Fu, Boxin Shi, Si Wu, Yonghong Tian. "1000x Faster Camera and Machine Vision with Ordinary Devices". 2022.
44. Safouane El Ghazouli, Alain Vissiere, Loius-Ferdinand Lafon, Mohammed-Lamjed Bouazizi, Hichem Nouira. "Optimized calibration of machine vision system for close range photogrammetry based on machine learning". 2022.
45. Chenglin Yu, Hailong Pei. "Face recognition framework based on effectiveness computing and adversarial neural network and its implementation in machine vision for social robots". 2021.
46. . Rajiv Soundararajan, Soma Biswas. "Machine vision quality assessment for robust face detection". 2018.
47. Min Yong Kim, Shirazi Muhammad Ayaz, Jaechan Park, Young Jun Roh. "Adaptive 3D sensing system based on variable magnification using stereo vision and structured light". 2014.
48. H. Golnabi, A. Asadpour. "Design and application of industrial machine vision systems". 2007.
49. Tianhai Wang, Bin Chen, Zhenqian Zhang, Han Li, Man Zhang "Applications of machine vision in agricultural robot navigation". 2022.
50. Rana Qarooni, Jonathan Prunty, Markus Bindemann, Ron Jenkins. "Capacity limits in face detection". 2022.
51. Erik Hjelmas, Boon Kee Low. "Face Detection: A Survey". 2001.
52. Mike Burton, Markus Bindemass. "The role of view in human face detection". 2009.
53. Markus Bindemann, Mike Burton, Rob Jenkins. "Capacity limits for face processing". 2005.
54. Haozhe Xie, Hongxun Yao, Shangchen Zhou, Shengping Zhang, Xiaojun Tong, Wenxiu Sun. "Towards 3D object reconstruction from stereo images". 2021.

55. Abdulla Mohamed, Chenguang Yang, Angelo Cangelosi. "Stereo Vision based Object Tracking Control for a Movable Robot Head" 2016.
56. Mahmoud Badawy, Hisham Khalifa, Hesham Arafat. "New approach to enhancing the performance of of cloud-based vision system of mobile robots". 2019.
57. Mahmoud Abdelaal, Ramy M.A. Farag, Mohamed S. Saad, Ahmed Bahgat, Hassan M. Emara, Ayman El-Dessouki. "Uncalibrated stereo vision with deep learning for 6-DOF pose estimation for a robot arm system". 2021.
58. Mohidul Alam Laskar, Amlan Jyoti Das, Anjan Kumar Talukdar, Kandarpa Kumar Sarma. "Stereo Vision-based Hand Gesture Recognition under 3D Environment" 2015.
59. A. Haro, M. Flickner, and I. Essa, "Detecting and tracking eyes by using their physiological properties, dynamics, and appearance". in Proc. *IEEE Conf.. Computer Vision and Pattern Recognition*, vol. 1, pp. 163–168, 2000.
60. Elad, M., Rejection Based face Detection. *EE368: Digital Image Processing Lecture,* Stanford University, Stanford, CA. May 19, 2003.
61. A. Senior, R.L. Hsu, M.A. Mottaleb, and A.K.Jain. "Face Detection in Color Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002. 24(5): p. 696-706.
62. Perez A., Cordoba M.L., Garcia A., Mendez R., Munoz M.L., Pedraza J.L., and Sanchez F., "A precise eye-gaze detection and tracking system", *The 11-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2003,* 3.-7. February 2003, Plzen, p. 105-108.
63. Paul Viola and Michael Jones, "Fast Detection with the Advanced Casual Cascade". 2001.
64. Samiksha Malhotra, Vaibhav Aggarwal, Himanshu Mangal, Preeti nagrath and Rachna Jain. "Comparison between attendance systems implemented through haar cascade classifier and face recognition" *librarynIOP Conf. Ser.: Mater. Sci. Eng.*, 1022 (2021), Article 012045.
65. B.D. Parameshachari, "Logistics sine map (LSM) based partial image encryption". *National Computing Colleges Conference (NCCC), IEEE* March 2011, pp1-6.
66. Mayur Surve, Priya Joshi, Sujata Jamadar, and Minakshi Vharkate. "Automatic Attendance System Using Face Recognition Technique*", International Journal of Recent Technology and Engineering.* 2021.
67. D. Jenif, W. S. Souza, S. Jothi, A. Chandrasekar. "Automated attendance marking and management system by facial recognition using histogram". 5th International Conference on Advanced Computing and communication Systems (ICACCS), IEEE (2019).
68. R. Nandhini, N. Duraimurugan, S.P. Chokkalingam. "Face recognition based attendance system". *Int. J. Eng. Adv. Technol. (IJEAT),* 8 (3S) (February2019) ISSN: 224-8958.
69. Omar Abdul Rhman Salim, Rashidah Funke Olanrewaju, Wasiu Adebayo Balogun. "Class attendance management using face recognition". *7th Intenational Conference on Computer and Communication Engineering* (ICCCE), IEEE (2018).

70. Sujay Patole and Yatin Vispute. "Face recognition-based automatic system", *Int. J. Innov. Res. Sci., Eng. Technol.*, 6 (August 2017).