Automated Guided Vehicle utilising thermal signatures for Human identification and tracking

H.R. Rabie

Submitted in fulfilment of the requirements for the degree:

MASTER OF ENGINEERING: ELECTRICAL

in the

Department of Electrical, Electronic and Computer Engineering

of the

Faculty of Engineering and Information Technology

at the

Central University of Technology, Free State

Supervisor: **Dr NJ Luwes** (DTech: Electrical Engineering)

Co-supervisor: **Prof. HJ Vermaak** (PhD: Electrical Engineering)

December 2016

Declaration

I, Hendrik Reitz Rabie, hereby declare that this research project which has been submitted to the Central University of Technology, Free State for the degree of Master in Electrical Engineering, is my own independent work; complies with the Code of Academic Integrity, as well as other relevant policies, procedures, rules and regulations of the Central University of Technology, Free State, and has not been submitted before by any person in fulfilment (or partial fulfilment) of the requirements for the attainment of any qualification.

Student Signature: Date: 01 Dec 2016

Acknowledgements

The Author would like to acknowledge the following individuals and institute, whom without the completion of this dissertation would not have been possible:

- The author’s family, for their love and patient support.
- Dr N.J Luwes, Prof. HJ Vermaak, the CUT and RGEMS, for the provision of guidance, granting the opportunity to undertake the project, monitory assistance and the work experience gained during the duration of the project.
- Friends and various fellow research students within the RGEMS research group.

Abstract

Industry requires the development of sophisticated autonomous guided vehicles (AGV) with sensory and software capabilities to allow a vision-based awareness of surrounding objects. To achieve this, a closely integrated control system for the AGV together with machine vision capabilities needs to be developed to efficiently and reliably detect objects of interest. Industry application of AGVs require detection of humans and to support that requirement thermal imaging cameras offer a broad set of advantages.

The aim of the study is to develop an AGV that uses a thermal imaging camera to detect a human in its environment. To achieve this, a literature study was done to determine the best type of components that should be used, reveal design issues and what characteristics the system must adhere to. LabVIEW was used to simulate AGV movement and operation together with the control system, develop machine vision capable of background noise filtering and verify the machine vision identification and tracking processes. Based on simulated results, the physical system was built and small modifications made to accommodate real world variables. The results indicate that a vision-based approach to detect, track and identify a person on a mobile robot in real time is achievable. It was found that LabVIEW is an excellent tool and platform for building the integrated system and expedites design and implementation. A key implication of this study is to show the versatility of thermal imaging as a method to extract a person from its background independently from current light conditions and in situations where full-colour cameras will fail.

Table of Contents

List of Figures

List of Tables

List of Equations

Acronyms and Abbreviations

AGV	Autonomous Guided Vehicle
CMYK	Cyan Magenta Yellow Key
CAN	Controller Area Network
COM	Communication Port
CAD	Computer-Aided Drafting
DC	Direct Current
DAQ	Data acquisition
FPS	Frames Per Second
FPGA	Field-Programmable Gate Array
HSV	Hue Saturation Value
HSL	Hue Saturation Lightness
IP	Internet Protocol
JPEG	Joint Photographic Expert Group
LabVIEW	Laboratory Virtual Instrument Engineering Workbench
PWM	Pulse Width Modulation
PC	Personal Computer
PUI	Point Under Investigation
PDF	Probability Density Function
QR	Quick Response
RAM	Random Access Memory
ROI	Region of Interest
RGB	Red Green Blue
RTE	Run Time Engine
USB	Universal Serial Bus
VI	Virtual Instrument
Wi-Fi	Wireless Fidelity

CHAPTER 1: INTRODUCTION

As robots become more essential in work environments, it is important for machines to react to human operators, with whom it shares a workspace. Different industries are adopting Automated Guided Vehicles (AGV) for unique roles in the manufacturing, automotive, search and rescue and military applications. As these technologies progress and the need to incorporate AGVs into more complicated scenarios like self-driving cars or delivery vehicles, the vision-based technology that drive AGVs needs to be robust.

Vision-based detection and tracking of humans using mobile robots that fulfil tasks in cooperation with humans are becoming increasingly important. Recent systems detect, track and identify humans based on data extracted from the person's face; these methods can locate a human based on skin colour face detection. These approaches assume that people are close to the camera and facing the robot. Numerous methods use face detection and a combination of sonar and laser scanners to locate a human [1].

In this study, AGVs utilising thermal signatures for Human identification and tracking is investigated. Thermal imaging shows promise as it would ease feature extraction by temperature differentiation. To evaluate thermal identification and tracking as an option for AGV's, an AGV needs to be built. Figure 1-1 below shows the computer-aided design of the concept AGV:

Figure 1-1 Concept of proposed AGV system

The CAD concept above in Figure 1-1 shows it would consist of two motors, one per wheel driven from an intelligent speed control system powered by two batteries. The rear wheels can move independently and use a front caster wheel for directional movement. As stated, the proposed AGV in Figure 1-1 should have the ability to detect, extract and track people. It needs to accomplish this by using a thermal imaging camera and vision-based methods to identify and track a person. This is in comparison with other systems that use colour-based vision, which often encounter light exposure and background interference problems. Results will evaluate a thermal image solution to overcome shortcomings of other systems.

1.1 Problem statement

Image processing with conventional full-colour cameras struggles with extraction or, in other words, to differentiate between a person and background. The reasons are that full-colour cameras depend heavily on favourable lighting conditions. They rely on the fact that the background will not change and that the illumination of the scene stays unchanged. Thermal imaging detects only temperature variations. As people are at a different temperature to their surroundings (normally warmer), thermal imaging should be better at extraction. This would entail that humans could be detected independently from current light conditions and situations where no skin colour is visible.

1.2 Hypothesis

An AGV equipped with a thermal imaging camera would be able to detect, with the aid of image processing, a human thermal signature. These thermal signatures could be utilised for tracking by applying image processing information in the AGV control.

When a person is extracted from the background, taken by a thermal imaging camera and real-time image processing is applied, then tracking based on thermal signature or body features will result in a viable tracking method.

1.3 Specific Objectives

The objectives of the study are to:

- Develop and evaluate Automated Guided Vehicle with payload carrying capabilities for platform testing.
- Develop an image processing system to extract humans from a background utilising thermal signature.
- Develop an image processing system to track the extracted humans.
- Develop an application to utilise image data for AGV control.
- Develop an evaluation method to evaluate the application of a vision-based thermal imaging camera on an AGV.

1.4 Research Methodology

As stated, the proposed hypotheses would be best evaluated using a real-world AGV with thermal image processing abilities. This proposed AGV system as explained in Figure 1-1 would consist of a chassis and drivetrain and will be used as a platform base for testing. The physical design will make use of an electric drivetrain and intelligent motor drive system. The power for the motor drive will be controlled via a high current controller getting its instructions from a data acquisition (DAQ) unit that is connected to a computer where the image processing and control is done. Thus, the AGV would be controlled using computer-based software. The software choice will be LabVIEW from National Instruments. LabVIEW allows for easy direct hardware interface and rapid prototyping of software designs before the final physical system is implemented.

A thermal imaging camera is mounted to the AGV platform to obtain the best thermal imaging videos. This camera will be used to capture different participants at different exposures. Using the captured video, an offline version of the real-time image processing software would be used to analyse and model the thermal signatures. Identification via signatures will focus on reducing background noise and extracting relevant data from the person's thermal signature. Extracted data could be used in tracking applications.

An identified shortcoming of the particular thermal camera is that the exposure of the thermal camera is not able to be set meaning the camera auto-corrects or adjusts its base value. Thus, a workaround system capable of setting the exposure need to be developed.

Conclusions drawn from this would be utilised in the identification and follow feature of AGVs. Techniques will be investigated for visual-base thermal imaging person identification and tracking. Identification will be investigated in tandem with tracking to allow the system to operate as a unit. Identification will use size as the primary method to single out a person. Extracted location data from the identification process will help the tracking process to operate. Tracking entails knowing the location of the individual even if he/she has moved from their starting position.

The system is divided into three parts. Firstly, the running gear that handles the movement of the AGV, secondly the thermal camera that acquire thermal images and lastly the processing part. The processing part can then be subdivided into two parts, one where all images are processed and then a second where data is quantified from the images to make decisions and to produce AGV running gear control signals. A computer will run the image processing software and control the motor controller via DAQ interface.

1.5 Outline of the Dissertation

Chapter 1 is an introduction to the dissertation which presents background, problem statement, objectives, methodology, hypothesis, delimitation of the study, as well as the research outputs.

Chapter 2 provides an overview of the use of thermal imaging in different conditions. The primary focus is the use of a thermal camera to assist an AGV in identifying a human target. A review of current image processing as well as hardware and software used in development is included.

Chapter 3 provides a comprehensive overview of the AGV system and explains the reason for using such a system and also the benefits. It covers the development of the software capable of analysing the thermal images from the thermal camera and discusses the detection of a human target and then also the tracking of a human. The chapter is concluded by a discussion of the complete system and component overview.

Chapter 4 discusses results obtained from thermal image analysis. The accuracy of identification is compared to target distance from the camera.

Chapter 5 presents the conclusions and suggests future areas of research.

CHAPTER 2: LITERATURE REVIEW

To develop an Automated Guided Vehicle utilising thermal signatures for human identification and tracking, one need to be able to understand the concepts of AGVs, machine vision, and thermal cameras. This chapter will show concepts of AGVs, the uses of AGVs and guidance methods for them. It will go further and explain theory and thermal camera methods that acquire thermal images, and image processing used to analyse images from mentioned camera. This chapter also provides justification for the choice of software used.

2.1 Automated Guided Vehicle

As stated, an evaluation AGV needs to be designed and built. An AGV is often regarded as a robot that is mobile and moves around on wheels and is used in an industrial application to move equipment and parts around the facility. The industries that most often use AGVs are the manufacturing and warehouse storage facilities. An AGV makes use of sensors to navigate itself through a facility autonomously to achieve certain tasks [2].

The ability of AGVs to transport has always given them the greatest flexibility in industry. AGVs are capable of transporting extremely heavy loads, but can be equally suited for accurate and nimble operations.

Other applications include:

- Warehouse storage and retrieval operations where stock can be stored vertically and then retrieved on request by and operator, these AGVs are known as (AS/RV) or automated forklifts.
- Motor vehicle manufacturing where the assembly line requires the use of vehicle chassis transport. AGV can also deliver parts to individual assembly stations.
- Exploring dangerous environments; and
- Search and rescue.

An AGV needs a method to navigate – examples include but are not limited to magnetic strips placed under the floor or optical sensors to do path planning. Different methods are used in different environments. One example is where a slot is cut into the floor, and a wire is placed below the surface. This slot needs to be cut along the path the AGV has to move. The wire then transmits radio signals that are picked up by the AGV [3].

This method holds many advantages, but also has some disadvantages. Once laid out and installed, the wire is set in place and cannot be moved – meaning changes cannot be made to the factory floor layout. Other options for navigation include guide tape, laser target navigation, natural features navigation and geo-guidance. All use existing features to navigate their environment. Vision-Guided AGVs have the advantage that it can be installed with no modifications to the environment or infrastructure, making them a better scalable option for a company. They operate by using cameras or sensors to record features along the route, allowing the AGV to replay the path by using the recorded features to navigate. The AGV will be able to adapt to a change in environment and continue to work. To help in localisation, AGVs could use Evidence Grid technology, an application of probabilistic volumetric sensing, which was invented and initially developed by Dr Moravec at Carnegie Mellon University [4].

The Evidence Grid technology uses probabilities of occupancy to compensate for random events picked up in sensors for each point in space to compensate for the uncertainty in the performance of sensors and in the environment. The primary function of the Evidence Grid technology is to do localisation because if the AGV knows its location in the environment, it can navigate in it. The end result of probabilistic volumetric sensing is to build a 3D map that can be stored and helps the AGV navigate the environment in the future. In conclusion, AGVs are widely used to transport material and can be used in numerous environments [4].

2.1.1 Data Acquisition

A Data Acquisition unit from National Instruments is used, which is a USB powered module capable of linking a standard PC to electrical components for manipulation or monitoring. It includes analogue and digital inputs and outputs. The DAQ could be used to interface with the motor drive controller. A DAQ is capable of interfacing electrical equipment with software running on a computer, allowing control of existing equipment [5].

2.1.2 Tablet Computer

A tablet is a mobile computer that has a display built into it as well as all other components that make it a computer. This single unit has all the necessary circuitry and batteries needed for operation in a mobile package. Tablets are also equipped with sensors including cameras, microphone, accelerometer, and touchscreen. The software that runs on a tablet can range from mobile operating systems to desktop operating systems like Windows or Linux.

2.1.3 Movement control for AGV

The speed of a DC motor is controlled using Pulse Width Modulation (PWM). The concept of PWM can be explained using a circuit setup that is shown in Figure 2-1. In this figure, there is a battery connected to a motor and a switch that can enable or disable the motor. If the switch is closed, the motor will rotate in accordance with the amount of DC power it receives from the batteries, and when the switch is open, the motor will not receive any power and stop rotating. Now suppose the switch is opened and closed rapidly. This will cause the motor to rotate and stop rotating in accordance with the time the switch is open and close. At the time between opening and closing, the switch is decreased the motor will not have sufficient time to come to a complete stop and thus, will continue to rotate at a speed proportionate to the time the switch is closed. This is the principle of controlling the speed of a DC motor using PWM signal [6].

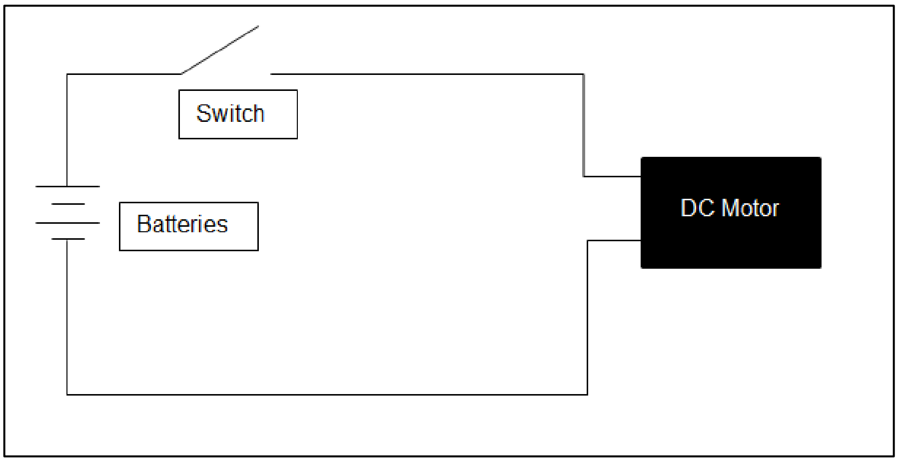


Figure 2-1 Controlling a DC motor

PWM is often used in conjunction with an H-Bridge, which is an electronic circuit that allows the voltage to be applied across a load in either direction. It allows the DC motor to run forward and backwards. Together they form the basis of the electric wheelchair controller that moves the motors.

There are different techniques that can be explored in the field of AGV movement and optically assisted robotic movement. The image that is received from the camera is seen as a two-dimensional flat image that has a fixed resolution; this resolution has two values, x and y. These values represent x and y coordinates that can be used in software to identify the position of an object.

Figure 2-2 is a picture of a chess board that has an imaginary resolution of (x=100 *X* y=100). If the black chess piece was to be represented as a coordinate, it would be roughly (x=45 *X* y=45).

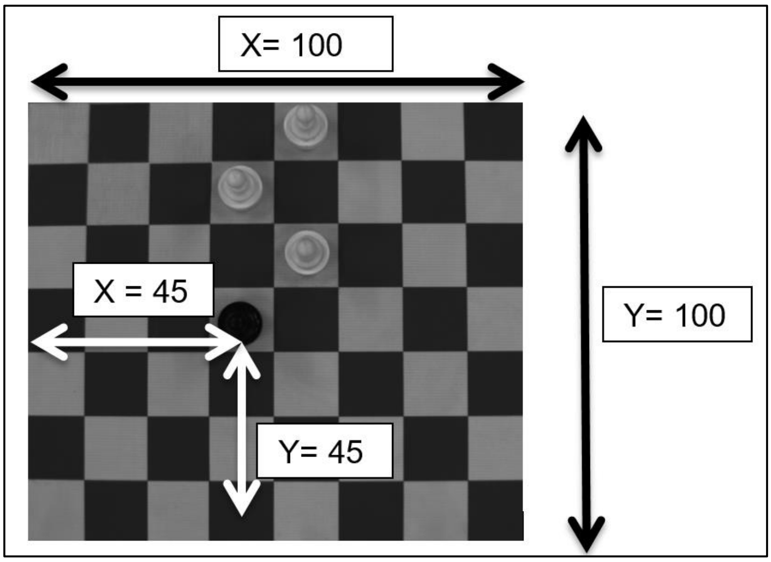


Figure 2-2 Chess Board

By knowing the resolution, it is possible to build a coordinate system much like a map to represent objects as numbers on the image coming from the camera. The data coming from the image processing part of the software gives these objects as coordinates it is then the job of the second part of the software that controls AGV movement to interpret the data and make decisions based on that. The AGV takes the coordinates it received from the image processing stage to make decisions regarding its movement. Figure 2-2 explains how this process works. The system will be designed to accommodate the necessary changes that might be implemented at a later stage.

2.1.4 CompactRIO

Made by National Instruments, it is a reconfigurable embedded control and acquisition system. It is a dedicated system used by industrial fields. Used in harsh conditions, it performs dedicated operations and can be fitted with a number of expansion slots to conform to its application. It runs the NI real-time OS with large and powerful field-programmable gate arrays (FPGAs). Used to do processing or vision analysis, this system is extremely versatile and because it can be expanded by adding additional hardware slots making it scalable and able to adapt to additional demands [7].

2.1.5 Wireless Communication

Wi-Fi is a microwave technology that operates in the 2.4GHz and 5GHz range. It is mainly used to connect mobile devices to the internet and to exchange data in a network. It can also be set up to work as direct communication from computer to computer. Wi-Fi is an 802.11 networking IEEE technology. The main advantage of Wi-Fi is that it is compatible with almost every operating system, desktop or mobile. Controller Area Networks (CAN) are used in industrial automation to allow intelligent machines to share data and communicate vital information to one another. Global industrial networks make it possible to program machinery from anywhere, even while they are in operation [8].

2.2 Digital Image Processing

Referring to image processing is to refer to the manipulation of an image by a computer. With modern computers, it is possible to manipulate multi-dimensional signals with advanced parallel computers. The goal of manipulation of an image is usually categorised into three divisions.

- Image processing (image in ⟶ image out)
- Image analysis (image in ⟶ measurement out)
- Image understanding (image in ⟶ high-level description out)

More advanced image processing uses ROI (regions of interest) to divide an existing image into regions for processing, helping to focus the processing to a particular area that holds the desired information instead of wasting processing power on redundant areas of an image. Computer vision is considered high-level image processing when a processor is to be intended to analyse the content of the image or video [9].

There are three basic processing techniques that cover the spectrum of how image processing can be used; namely image enhancement, image restoration and image compression.

Image enhancement refers to the altering of an image to allow some key feature to stand out; for example, sharpening of the image boundaries or contrast correction to highlight features for graphical display. This form of image processing will generally be used first so as to enhance the effects of the rest of the image processing system. This processing does not alter the inherent information content in data [10].

Image restoration is used by building a filter in the software to reduce the effect of image degradation. The effectiveness of image restoration depends on the filter that is being designed. The extent and accuracy of the knowledge of the degradation process help to design an effective filter and is key in handling the effect of degradation. In image restoration, more extraction or accentuation of image features are used to reduce degradation [11].

Image compression is used to minimise the number of bits that represent an image and is used in applications where it is necessary to transmit a large amount of data over a distance [10].

2.2.1 Grayscale

A grayscale image is the simplest of images to work with. The spatial coordinates are represented by x and y and their respected intensity values. Grayscale images can be represented as surface graphics with the z-axis representing the intensity of light. Figure 2-3 shows that the brighter areas in the image represent higher z-axis values. Intensity is represented by depth, which is the range of intensities that can be represented per pixel. For a bit depth of x, the image is said to have a depth of 2^x meaning that each pixel can have an intensity value of 2^x levels.

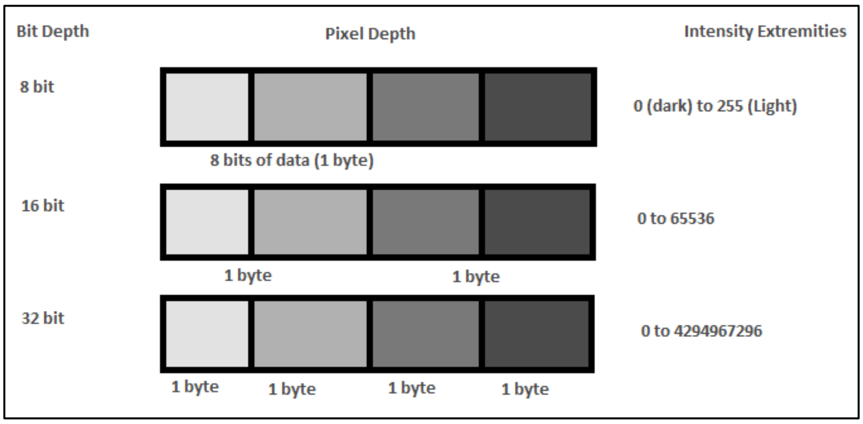


Figure 2-3 Shade Intensity Chart

Different bit depths exist depending on the desired resolution. When performing image analysis and features tracking an 8-bit pixel, depth is enough. The requirement for subtle changes and detail analysis will need more bit depth. The drawback of using more bits is that it requires more memory from the computer both Random Access Memory (RAM) and hard drive storage [9].

Equation 1 [9] explains how to calculate the amount of memory needed when analysing and image with provided quality.

$$\text{MemoryRequired} = \text{Resolution}_x \times \text{Resolution}_y \times \text{BitDepth}_z \qquad (1)$$

2.2.2 JPEG Image Format

The most common file format used in the LabVIEW Vision Toolkit is Joint Photographic Experts Group (JPEG). In order to achieve a high quality but also aggressive ratio of compression without sacrificing colour changes or tone reduction, JPEG format is used. The JPEG compression technique analyses images, removes data that is difficult for the human eye to distinguish, and stores the resulting data as a 24-bit colour image. JPEG conversion is definable making it versatile in the sense it can be adjusted for different applications. The compression level can be adjusted to 15 and is still very difficult to distinguish from the source even at high magnification [9].

2.2.3 Resolution

Resolution is defined as the smallest feature that a camera can distinguish on an object. A very good example of resolution is a barcode. Take Figure 2-4 for example; the QR barcode consists of black and white square blocks that have spaces between them. These spaces and patterns hold valuable data and if the resolution on the camera is not sufficient, valuable information might be lost. The easiest method to resolve this problem is to acquire a camera with a very high resolution. However, there are plenty of reasons not to choose that option: more processing power is needed the higher the resolution, more memory is needed to store images in higher resolution and processing is slower and much more expensive. The difficult part is to reach a compromise between resolution and processing power at your disposal [9].

Figure 2-4 Matrix Barcode

To calculate the necessary resolution needed there are three things to consider. The type of camera lens, the smallest features that need to be visible, and the maximum resolution capable for the imaging sensor. It is important to consider the lens used, because this will greatly determine the resolution. For instance, using a larger magnification lens effectively increases the resolution but decreases the field of view. To determine the sensor resolution required in analysing a feature of two pixels in size, Equation 2 [9] can be used to calculate the resolution needed.

$$\text{Sensor Resolution}_{1D} = \left\lceil 2\left(\frac{\text{FieldOFView}}{\text{SmallestFeature}}\right)\right\rceil \quad (2)$$

2.2.4 Contrast and Modulation

Contrast goes hand in hand with resolution, for whereas resolution resolves the spatial differences in an image, contrast resolves intensity differences. Contrast is a dimensionless quantity that is the difference between the lightest and darkest features of an image. Equation 3 [9] suggests that when the difference between the brightest and darkest contrast is high, the image has a large range of contrast. The opposite is true for a flat image with little contrast.

$$\text{Contrast} = \frac{I_{\text{Brightest}} - I_{\text{Darkest}}}{I_{\text{Brightest}} + I_{\text{Darkest}}} \quad (3)$$

Where $I_{\text{Brightest}} > I_{\text{Darkest}}, I_{\text{Darkest}} \neq 0$

2.2.5 Brightness

Intensity differs from brightness in the sense that intensity of a light source depends on the total amount of light emitted and the solid angle from which it is emitted. Thus, intensity can be measured because it is a physical property, whereas brightness is a psycho-visual concept and may be described as the sensation to light intensity. The perceived brightness can be defined as the contrast observed. The location of a bright spot depends not only on the brightness, size and duration, but it greatly depends on the contrast between the spot and the background. When the contrast is greater than a threshold, called just noticeable difference, then the spot can be detected. Also, this greatly depends on the average brightness of the surroundings. This dependence is also commonly known as brightness adaptation [12].

2.2.6 Camera Exposure

Camera exposure is the amount of light per unit area that reaches the electronic image sensor. The camera's shutter speed, lens aperture and available light all determine the amount of light that will reach the image sensor. Lux is used to measure exposure and is calculated from exposure value and scene luminance. Exposure is used to control the effects of underexposure and overexposure; both occurrences impact the clarity of the image and conceal detail. Overexposure is loss of highlight detail, and underexposure is loss of shadow detail; both occurrences are brought under control using manual exposure, automatic exposure or exposure compensation. Manual exposure is when all exposure settings are controlled manually by a person; automatic exposure is done by an onboard sensor in the camera; and exposure compensation is done by taking a reading from an external sensor to help adjust camera exposure settings [13].

2.2.7 Thresholding

Thresholding enables the selection of a range of pixel intensity values in a grayscale and colour image. Thresholding can often be seen as a method of compression or simplifying a complicated image. It forms part of segmentation, which is the identification of areas of an image that appear uniform to an observer and subdivides the image into regions of uniform appearance. Equation 4 is used to select ranges in pixel intensity values [9, 13].

$$I_{New} \begin{cases} 0 & I_{Old} < Range_{Low} \\ I_{Old} & Range_{Low} \leq I_{Old} \leq Range_{High} \\ 255 & I_{Old} > Range_{High} \end{cases} \quad (4)$$

Three methods of detection can be used during thresholding: Bright objects, Dark objects and Grey objects detection. First, consider how thresholding works on the pixel intensity scale. Table 1 represents an image that is in grayscale format that has been put in numeric table form.

Table 1 Grayscale table represented by intensity values

122	122	111	90	120	122	133	110	101	108
130	155	165	150	205	150	220	210	240	245
136	176	150	210	206	180	200	80	88	240
126	145	245	240	250	200	230	220	230	230
135	126	210	170	175	205	23	221	106	215
133	220	215	200	200	200	240	235	245	250

Figure 2-5 is the grayscale image used to build Table 1. This illustration represents a very small part of an image, each block representing an individual pixel on the image.

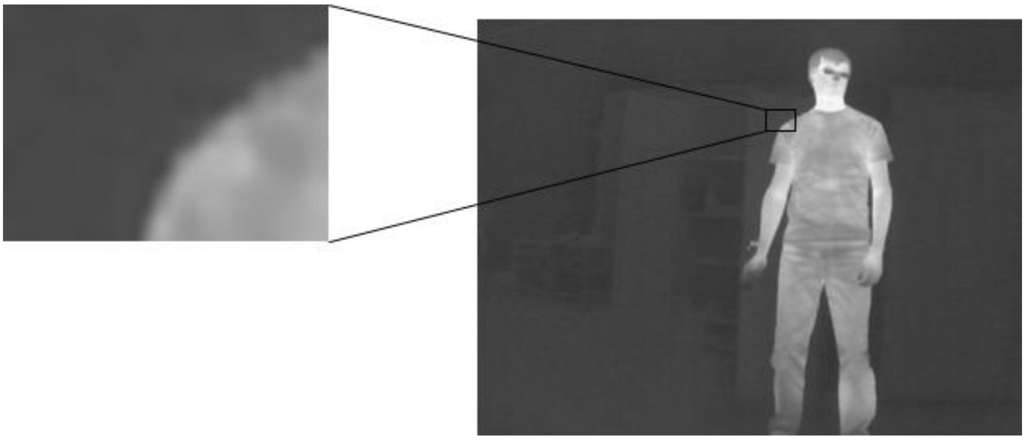


Figure 2-5 Grayscale Image

After thresholding, every pixel in the image gets a different intensity; these individual pixels get represented with only two values 0 and 255 meaning 0 equals to black and 255 equals to white. This transformed image is known as a binary image. Designing a filter that screens for intensity values of 200 and up it means all pixels with values between 200 and 255 change to white and all pixels between 199 and 0 change to black; Table 2 shows these changes clearly.

Table 2 Threshold Image intensity values

0	0	0	0	0	0	0	0	0	0
0	0	0	0	255	0	255	255	255	255
0	0	0	255	255	0	255	0	0	255
0	0	255	255	255	255	255	255	255	255
0	0	255	0	0	255	203	255	0	255
0	255	255	255	255	255	255	255	245	255

Once a grayscale image goes through the threshold process, it becomes a binary image and consists of only two colours; Figure 2-6 illustrates what a binary image looks like.

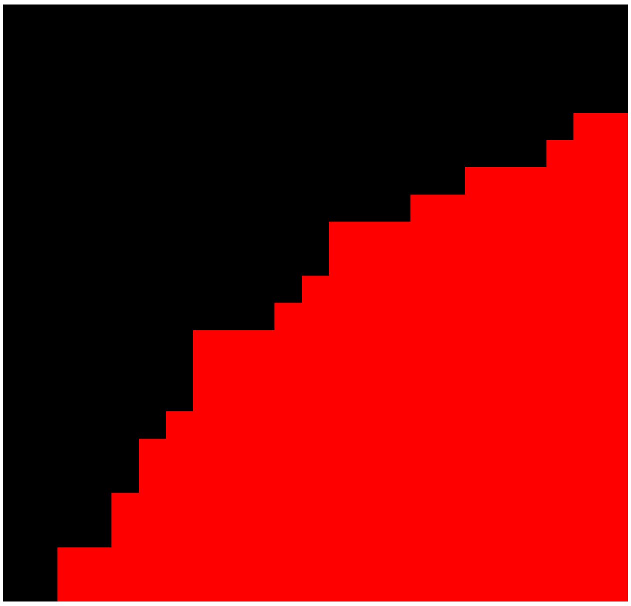


Figure 2-6 Image that has been thresholded

2.2.8 Frames per Second

Frames per second (FPS) is a term used to measure a number of frames displayed every second. In the field of image processing and for this project, the rule of FPS is the same as resolution. The higher it is, the more processing is needed. A balance has to be kept to manage the amount of processing available, and the amount of FPS needed. A low FPS result in loss of information and excessive FPS overtask the processor, and the system could lag.

2.2.9 Erosion and Dilation

Erosion and dilation are morphological operators and are commonly used in binary images. With binary images, the objects have a value of 1 and the background a value of 0. Erosion works in a field of 3x3 squares called the structuring element or kernel. The kernel is used to define the erosion type being used. Figure 2-7 is an example of binary erosion with the 3x3 square structure element starting at the centre. In Figure 2-7 and following diagrams, foreground pixels are represented by 1's and background pixels by 0's [14].

1	1	1
1	1	1
1	1	1

Figure 2-7 3x3 Kernel [14]

The binary image that is being eroded is represented as a matrix in Figure 2-8. Simple erosion occurs when the structuring element is applied to each pixel or point under investigation (PUI). If the pixel is 0, then it retains its value; if it is 1 and all of its cardinal neighbours are also 1, it is set to 1. If one of its cardinal neighbours were 0, it would have been set to 0 [9].

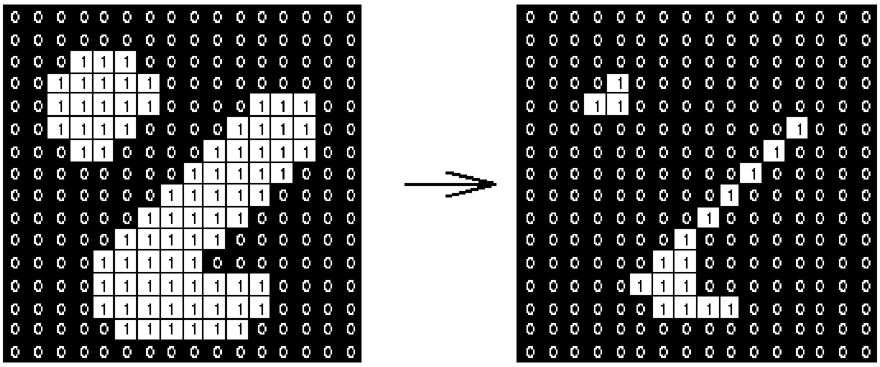


Figure 2-8 The Effects of erosion to a binary image [14]

Figure 2-9 shows the effect of erosion and how it could be used in image processing. Dilation works in the same way erosion works, only it has the opposite effect – instead of eroding it dilates the pixels and fills them in.

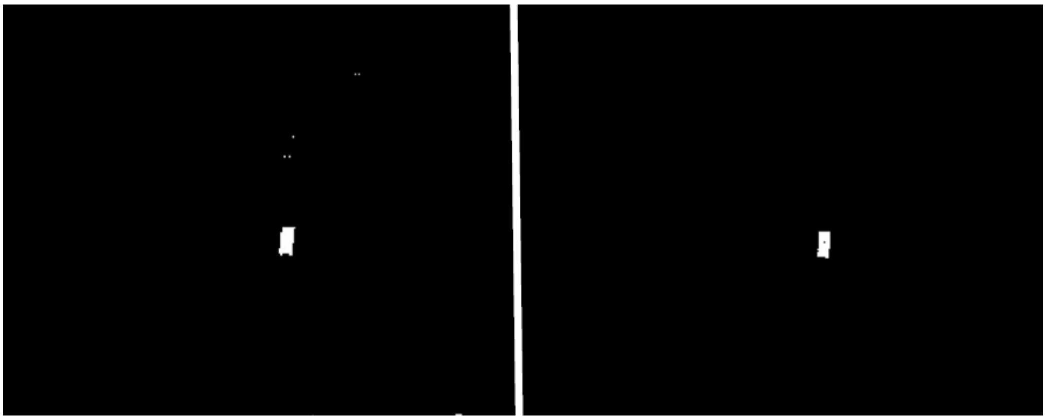


Figure 2-9 Side by side comparison of an image without and with erosion

2.2.10 Colour Space

An abstract mathematical model can be used to describe a colour model; colours can be represented as tuples of numbers that are commonly three or four values of colour components. Examples of these colour spaces are RGB, which is known to consist of channels: red, green and blue, or CMYK – that is cyan, magenta, yellow and black. Computer monitors use the RGB colour space to display all of its complicated colours; using a combination of these three colours it is possible to create different ones by combining the three primary colours. Red, green and blue can be seen as the X, Y and Z axes.

Another way of making the same colour is to use their hue (X axis), saturation (Y axis) and their brightness value (Z axis). This is also known as HSV colour space. Almost all colour spaces can be represented in this three-dimensional manner, and some have more dimensions and others fewer dimensions. This three-axis approach can be used to form different colours within their respected colour space [15].

Implementing the RGB colour model can differ greatly depending on the capabilities of the system. The most common implementation is the 24-bit implementation with 8-bits per channel and 256 discrete levels of colour per channel. The colour space using the 24-bit RGB model is limited to a range of 256x256x256 ≈ 16.7 million colours. Figure 2-10 shows how the RGB colour space is used to form different colours and how the HSV colour space is used in conjunction with it [9, 15].

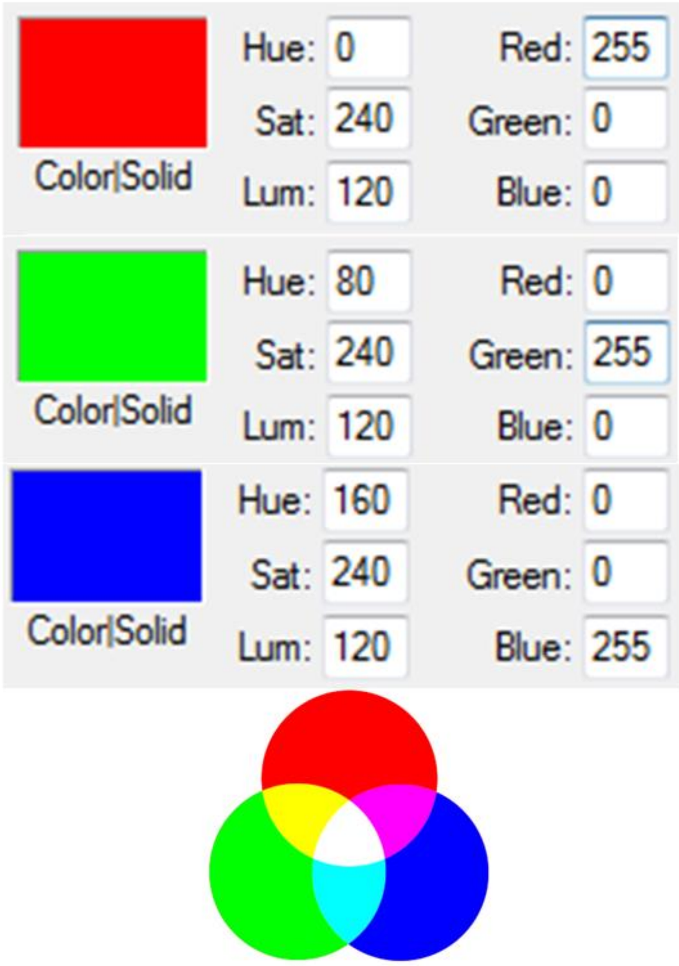


Figure 2-10 RGB colour space

2.2.11 Tracking of Thermal Targets

Full-colour cameras rely heavily on colour to track a human target making it susceptible to light changes. Thermal cameras return a grayscale image that shows only the heated parts of the environment; thus, the method of tracking a human target greatly varies from conventional cameras. A method called motion optical flow analysis can be used. It is widely used for mobile robot navigation in the sense that it distinguishes between a stationary target and a moving target. Optical flow is the pattern of apparent motion of objects, surfaces, and edges in a visual scene caused by the relative motion between an observer (an eye or a camera) and the scene [16].

The problem optical flow solves is not only the motion of the observer and the objects in the scene, but also the structure of the object and the environment it is in. An example of this is a human walking on the sidewalk and a car passing him/her on the street. For a human, it is easy to know that he/she is in motion and moving forward on the sidewalk and the car is moving in the opposite direction in the street. Since awareness of movement and the generation of mental maps of structures and environment are critical components of animal (and human) vision, the conversion of this innate ability to a computer capability is similarly crucial in the field of machine vision [17].

A subtraction base approach to tracking subjects can also be used, but the main disadvantage is that the capturing device has to be stationary for this method to work. This method works by taking two pictures of the same area and subtracting them from one another. This method can be applied only when the platform and the camera is completely still. Motion energy tracking is a method that calculates the temporal derivative of an image and thresholding at a suitable level to filter out noise. The temporal derivative is estimated by image subtraction and utilises Equation 5 [18].

$$\frac{df(x,y,t)}{dt} \approx \frac{f(x,y,t) - f(x,y,t-\delta t)}{\delta t} \tag{5}$$

2.2.12 Centroid Tracking

Centre mass or Centroid-based tracking refers to the concept of finding the centre of the energy of an image or part of an image. Centroid tracking is an easy and fast implementation of a basic equation. Through this equation, it is possible to analyse all pixels in the image or part of an image and determine the centre point of energy. Centre mass tracking works by analysing the pixel value of 8bit pixels and determining the energy of each pixel based on the pixel value of 0-255 where 0 is black and has no energy, and 255 is white and has the most energy. Figure 3-26 shows that once the centroid function is applied to an image with relatively scattered particles, it will align on the largest and densest object. The Centroid virtual instrument (VI) finds the centre of energy of an image through Equation 6 and 7 [19].

$$X = \frac{\mathrm{Sum}(\mathrm{X} * \mathrm{Pixel\ Values})}{\mathrm{Sum}(\mathrm{Pixel\ Values})} \tag{6}$$

$$Y = \frac{\mathrm{Sum}(\mathrm{Y} * \mathrm{Pixel\ Values})}{\mathrm{Sum}(\mathrm{Pixel\ Values})} \tag{7}$$

The Centre Mass (X, Y) coordinates parameter is a capable method of tracking because it can locate an object or pixel in an image with accurate (x, y) coordinates with respect to origin (0, 0) located at the upper-left corner of the image.

This can also be viewed as the centre of gravity of a particle. Where the particle composed of N pixels and Pi is defined as the point G and the centre of gravity in Equation 8 [20].

$$\overline{OG} = \frac{1}{N}\sum_{i=1}^{i=N} \overline{OP_i} \tag{8}$$

Centre of Mass for Coordinate X is calculated in Equation 9 [20].

$$X_G = \frac{1}{N}\sum_{i=1}^{i=N} X_i \tag{9}$$

X_G gives us an average location of horizontal segments in a particle.

Centre of Mass for Coordinate Y is calculated in Equation 10 [20].

$$Y_G = \frac{1}{N}\sum_{i=1}^{i=N} Y_i \qquad (10)$$

Y_G Gives an average location of vertical segments in a particle.

This method is affected by noise and requires different image analysis techniques to reduce artefacts. Spatial information combined with image subtraction is widely used to produce good real-time results. This method has plenty of disadvantages, but for its sheer simplicity has to be considered [18].

2.2.13 Mean Shift Tracking

Mean shift is used in statistics to seek the value that appears most often in a set of data. It is used in cluster analysis in computer vision and image processing. Mean shift is a tool for finding modes in a set of data samples representing an underlying probability density function [21].

Clustering is used to reach a point of convergence were the modes had been found in a set of data. It utilises a scanning area known as a kernel to move iteratively through the set of data to the point of convergence, also known as the hill climbing algorithm the mean shift vector always points towards the direction of the maximum increase in density. Many different types of kernel exist and impact the way mean shift works. Clustering forms the base of tracking, which forms a confidence map of consecutive frames based on the colour histogram; mean shift is then used to find the peak of a confidence map near the object's old position. The confidence map is a probability density function on the new image, assigning each pixel of the new image a probability, which is the probability of the pixel colour occurring in the object in the previous image [21].

The function of Mean shift should be seen as a tool. Mean shift is an analysis technique for locating the maxima of a density function; it is used to analyse a cluster of objects in computer vision. Suppose a cluster of points is given as red dots – as seen in Figure 2-11. Mean shift vector will estimate the exact location of the mean of the data by determining the shift vector from the initial mean [21].

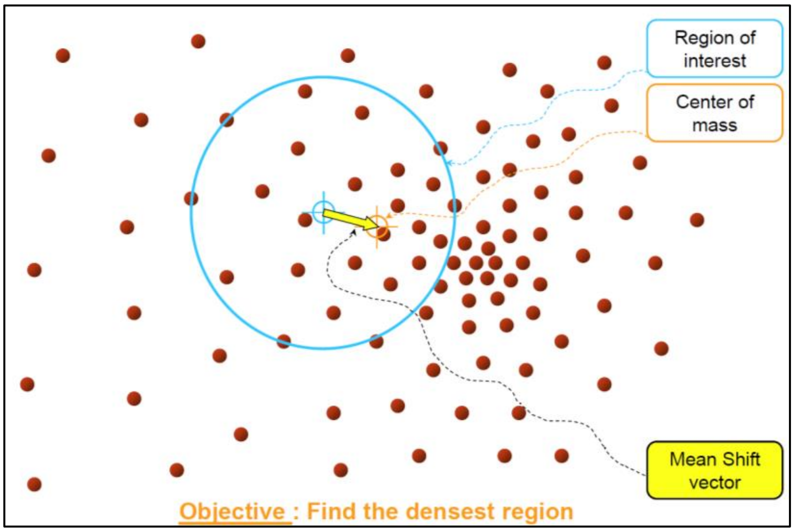


Figure 2-11 Mean-Shift Theory [21]

The first step in using mean shift is to define a region of interest (ROI). The ROI is used to focus the area of analysis of the mean shift algorithm. Figure 2-12 shows the mean shift vector, which will always point towards the direction of the maximum increase in density.

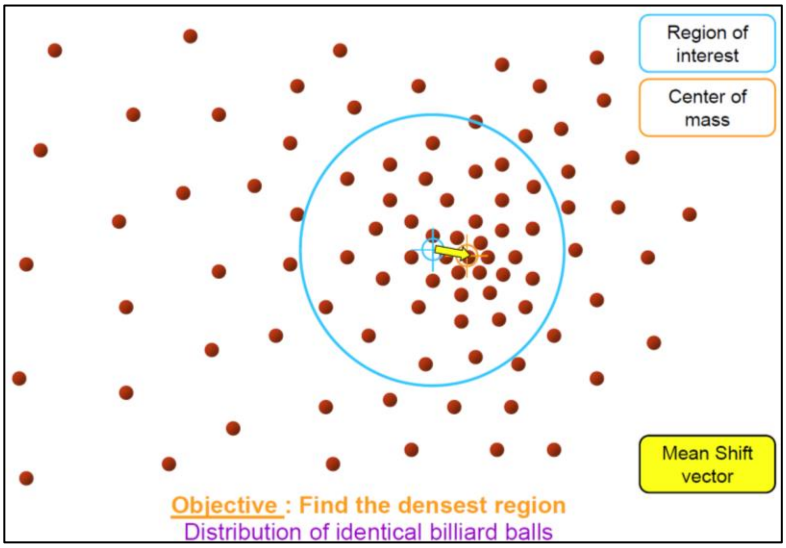


Figure 2-12 Mean-Shift Theory [21]

Properties of Mean Shift state that a vector has the direction of the gradient of the density estimate and that it is computed iteratively as seen in Figure 2-11 to Figure 2-13 to obtain the maximum density in the distribution [21].

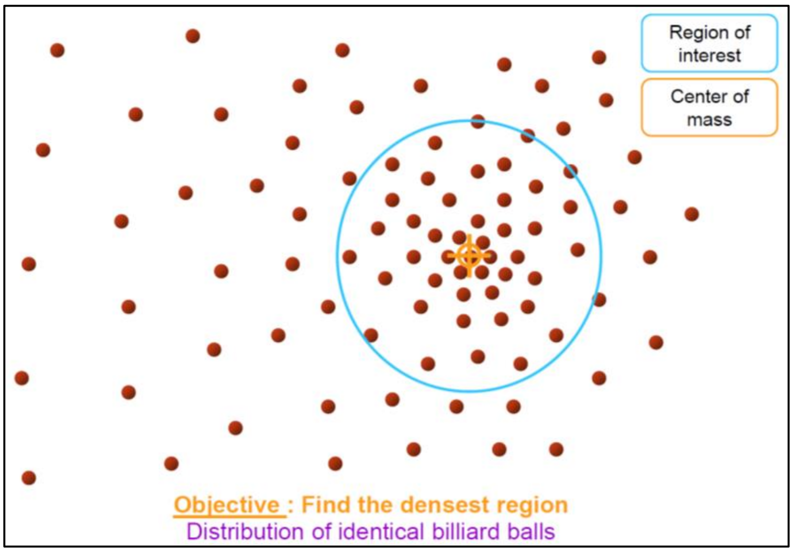


Figure 2-13 Mean-Shift Theory [21]

In a set of data points, the approximate location of the mean is defined as y_0. The desired result is to estimate the exact location of the mean or find y_1. First, the shift vector has to be determined from the first mean, and this is done iteratively until the mean shift vector is zero.

Equation 11 [22] is the Mean Shift vector of equal weighted points. The Mean shift vector always points towards the direction of the maximum increase in the density

$$\boldsymbol{M_h}(\boldsymbol{y}) = \left[\frac{\mathbf{1}}{\boldsymbol{n_x}} \sum_{\boldsymbol{i=1}}^{\boldsymbol{n_x}} \boldsymbol{X_i} \right] - \boldsymbol{y_0} \tag{11}$$

where $\boldsymbol{y_0}$ Is the initial estimate of the mean shift vector.

Equation 12 [22] is used to provide a more accurate vector. Weights w_i for each data point X_i is calculated by introducing a kernel h

$$M_h(y_0) = \left[\frac{\sum_{i=1}^{n_x} w_i(y_0) X_i}{\sum_{i=1}^{n_x} w_i(y_0)} \right] - y_0 \tag{12}$$

where y_0 is the initial mean location and n_x the number of points in the kernel. Equation 11 used a uniform kernel, which means each data point had equal weight. Examples of different kernel types include Gaussian and Epanechnikov [21].

Mean shift is a tool that is excellent at finding modes of the peaks in distributions. This can then be used for the probability density function (PDF). Examples of distributions are Gaussian distributions or bell shape curves and uniform distributions, which have equal uniform probabilities [21].

What Mean Shift is good at is to analyse a non-parametric distribution. Uniform, Gaussian and Bose-Einstein distributions are parametric distributions because they can be written as a formula; they have an analytical expression that is defined by a distribution with parameters. To determine the gradient of the distribution, Mean Shift can be used to ascertain the mode of the distribution were the mode will be similar to the gradient of the distribution as seen in Figure 2-14 [21].

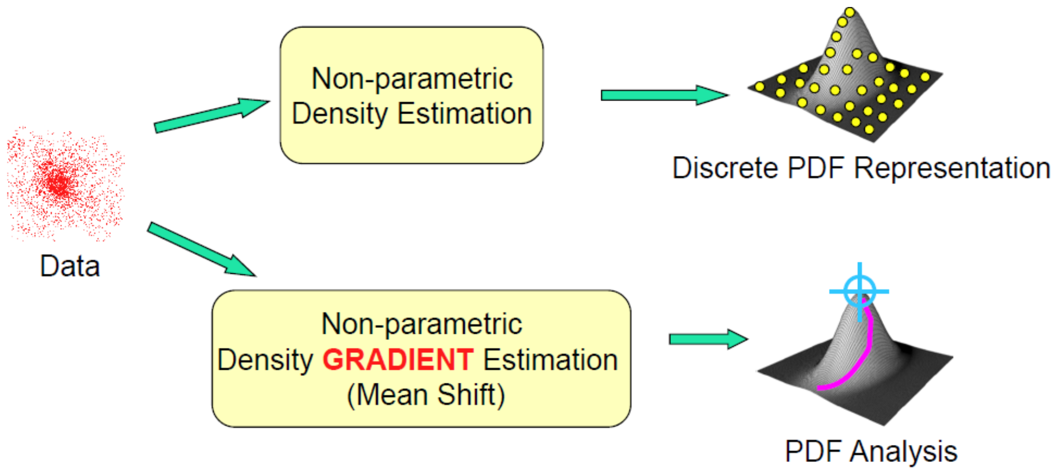


Figure 2-14 Density Gradient Estimation [21]

Figure 2-15 shows how to move from a 2D distribution where each point has an (x;y) location with weight, to PDF where density is used to plot the probability of each point that consists of an intensity value. Figure 2-15 shows two graphs: the 2D right side graph indicates data points and the 3D graph on the left side represents the data point, but with weight or probability. The data points are sampled from an underlying PDF.

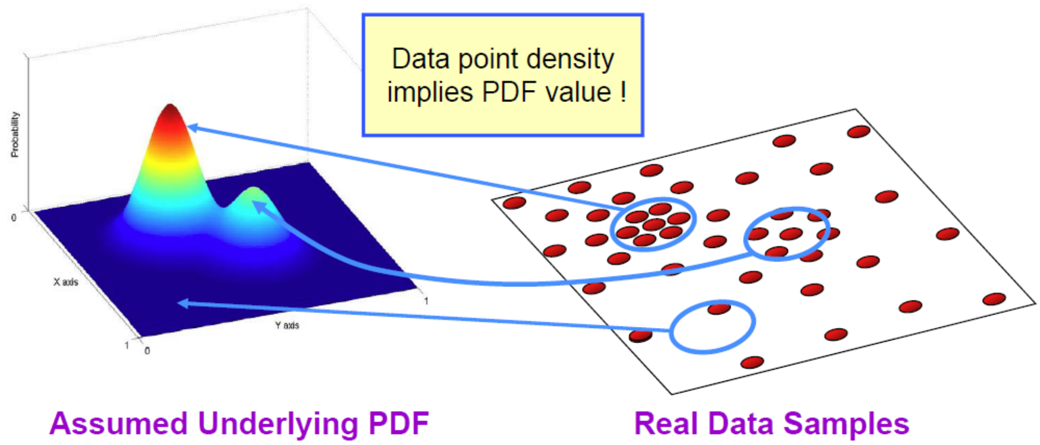


Figure 2-15 Distribution to PDF [21]

Equation 13 [22] is used to calculate the kernel density estimate.

$$\text{PDF}(\text{X}) = \sum_i C_i e^{-\frac{(X-\mu_i)^2}{2\sigma_i^2}} \qquad (13)$$

The density estimate uses all data points – meaning each point contributes to calculating the PDF. The PDF is computed based on the number of points μ_i or the number of samples and distance of each point to X. Figure 2-15 represents this in an assumed underlying PDF. Equation 13 [22] is a Gaussian function that is modified to sum up all data points that must be analysed.

The weight value of each point depends on the kernel that is used to calculate it. Normal, Uniform and Epanechnikov are some of the kernels that can be used to calculate the weight of each data point. Mean shift is the gradient of the probability density function in the non-parametric form.

Mean shift uses Equation 14 [22] to shift through the data points and find the gradient of the distribution. Each data point X_i is assigned a weight by multiplying by g_i and then summed up and divided by the weights.

$$\nabla \text{P}(\text{X}) = \frac{c}{n}\sum_{i=1}^{n} \nabla k_i = \frac{c}{n}\left[\sum_{i=1}^{n} g_i\right] . \left[\frac{\sum_{i=1}^{n} X_i\, g_i}{\sum_{i=1}^{n} g_i} - X\right] \qquad (14)$$

Equation 15 [22] identifies that Mean shift is the gradient of the PDF in the non-parametric form.

$$\nabla \text{P}(\text{X}) = \frac{c}{n}\sum_{i=1}^{n} g_i \times m(X)$$

$$\text{m}(\text{X}) = \frac{\nabla \text{P}(\text{X})}{\frac{c}{n}\sum_{i=1}^{n} g_i} \qquad (15)$$

$$\text{g}(\text{X}) = k'(\text{X})$$

Figure 2-16 defines how the mean shift equation is used in a two-dimensional image.

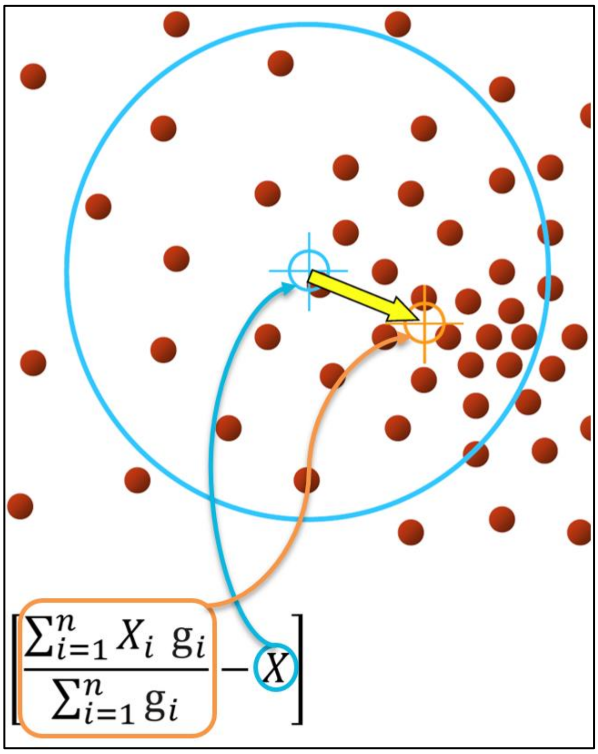


Figure 2-16 Computing Mean Shift [21]

2.2.14 Image acquisition

Image acquisition is the process of capturing the first image for analysis. Image processing does not only consist of image analysis, but also image acquisition from a source. Performing image acquisition in image processing is always the first step in the workflow sequence because, without an image, no processing is possible. The image that is acquired is completely unprocessed and is the result of hardware capturing it; this is known as a baseline. The way the image is acquired can impact image analysis. One of the ultimate goals of this process is to have a source of input that operates within such controlled and measured guidelines that the same image can, if necessary, be nearly perfectly reproduced under the same conditions. The actual hardware device can be anything from a desktop scanner to a massive optical telescope. One of the forms of image acquisition in image processing is known as real-time image acquisition. Real-time involves retrieving images from a source that is automatically capturing images.

2.3 Thermal imaging

As seen in Figure 2-17, infrared light is part of the thermal radiation spectrum. Infrared light has longer wavelengths than that of visible light, and thus a detector is needed to capture it. There are two methods of using infrared light to enhance vision capabilities in low light situations. The first is active infrared that sends out a beam of infrared light and gathers the reflected waves; the second is passive infrared, which collects infrared radiation from surrounding objects and amplifies it to form a picture [23].

2.4 Thermal camera

Thermal imaging cameras can form an image using infrared radiation. Unlike Infrared cameras, thermal cameras do not need an active beam of light to enhance visual capabilities. Thermal imaging is capable of capturing long wave infrared thermal radiation, and this allows for the visual representation of infrared radiation. Thermal cameras, therefore, use passive infrared light to see. A standard infrared camera can clearly see a human in the dark, and for this reason it is perfectly suited for identifying a human. These cameras are capable of seeing in the 7 μm – 14 μm infrared spectrum as explained by Figure 2-17 [23].

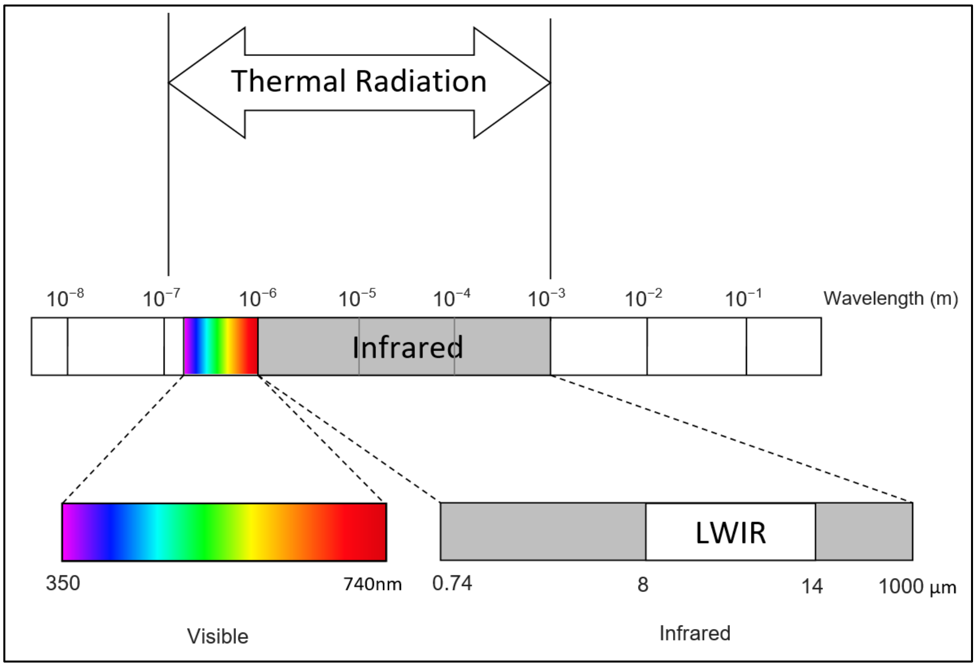


Figure 2-17 Thermal radiation spectrum

Different infrared cameras use different detectors to see infrared light. A thermal camera is an uncooled microbolometer camera when it is self-cooled and does not require any additional cooling. Previous high-resolution thermal sensors required extreme cooling methods and made them only usable as a stationary camera. The microbolometer is a type of bolometer that is used in detecting infrared radiation with wavelengths between 7.5 and 14 µm. Infrared radiation strikes the detector material and heats it, thus changing its electrical resistance. These changes are measured and processed into temperatures which are used to form an image [24].

Thermal camera lenses are different from lenses of normal cameras. Glass does not transmit infrared radiation well, and so the lenses of some thermal imaging camera are made of germanium. This material is an excellent transmitter of infrared radiation. However, it is a very expensive material. The AXIS Q1921 has a lower quality zinc sulphide lens [25].

2.5 LabVIEW

LabVIEW is a programming language like any other, except that the code is graphical. Its compiler directly translates the graphical code into machine code where the Run-Time Engine (RTE) executes the machine code. LabVIEW does not directly communicate with the desired hardware that must be controlled or monitored. Instead, it goes through an application programming interface (API). LabVIEW interacts with hardware drivers that do the lower level work and interacts with Data Acquisition (DAQ) hardware using DAQmx drivers [26].

LabVIEW programs are called virtual instruments, or VIs, because their appearance and operation often imitate physical instruments, such as oscilloscopes and multimeters. LabVIEW contains a comprehensive set of tools for acquiring, analysing, displaying, and storing data, as well as tools for troubleshooting the code written. When a VI is created two windows, the front panel or user interface and the block diagram or rear panel is used to write functional graphical code. LabVIEW VIs, are often referred to as programs throughout the document.

National Instruments vision products allow the ability to program functions. LabVIEW supplies the fundamental components needed for such projects. It offers various resources – not just in the field of machine vision, but also in the application of robotics. LabVIEW offers the ability to easily develop a system to quantify images and to analyse that data. It allows the flexibility to address the needs of this project through testing, measurement and industrial automation in vision application.

The analysed images produce data that requires further processing and decision-making. LabVIEW offers a diverse and comprehensive software platform that can do both image processing and logic decision-making. The vision acquisition express toolbox helps to streamline the process of acquiring and analysing images. Vision/Vision Express toolbox is located in LabVIEW and is one of the fastest and easiest ways to configure all of the camera's characteristics. The IMAQ vision is a library of LabVIEW VIs that can be used to aid the design and implementation of machine vision and image analysis in scientific applications. Vision Assistant is another module that enables prototyping an application without having to do any programming. Included in these modules are vision controls using different type of images and image processing operations. Another module called Vision Builder uses an interactive interface for prototyping, benchmarking and deploying applications. Using some of these modules during the development of the project will make it possible to build a robust program in a short period of time.

CHAPTER 3: Methodology

This chapter not only explains how the thermal imaging camera is used to monitor a person, but also how the software works that controls the AGV. Thermal calibration is used to improve image quality and form part of a stand-alone system.

3.1 AGV

The AGV consist of running gear connected to a motor drive, and a DAQ unit connecting the drive to a computer running a control program. The computer sends and receives signals through the DAQ unit and a thermal imaging camera attached to the computer. The computer aided design (CAD) model in Figure 3-1 will act as a blueprint and will be improved upon as the AGV construction continues.

3.1.1 Running Gear

The drive of the AGV consists of two 25amp motors, two 12V 50Ah batteries in series and a PG intelligent speed control system. Figure 3-1 is a (CAD) model that is a conceptual representation of the AGV. Through this design, it is possible to make changes in the software before the final structure is built and tested. This method of design helps to streamline the final builds and makes part placement easier.

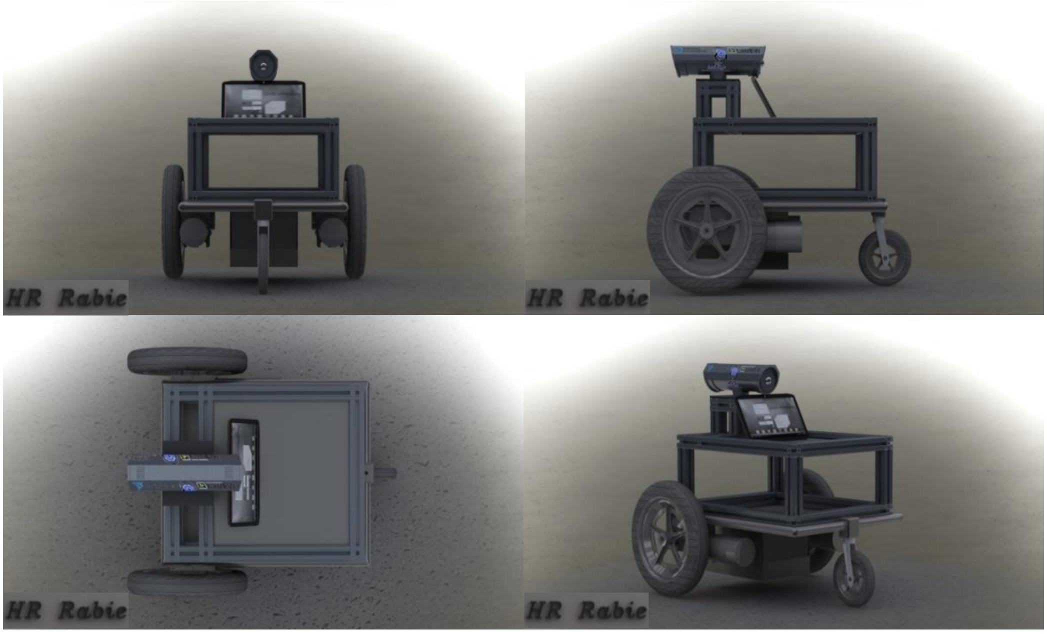


Figure 3-1 Rapid prototype CAD model of the AGV

Figure 3-1 shows the part placement and the structure of the AGV. The thermal camera is mounted at the highest point of the structure and the slate computer just beneath it. Mounting the camera higher should deliver a better viewing angle and the tablet computers display can be used to show information to the operator. Attaching the battery pack at the lowest point on the AGV lowers the centre of gravity and allows room for additional hardware and storage. The AGVs wheels consist of a differential drive with two driving wheels and one caster wheel in front. The AGV will carry all the components of the system including the thermal camera and the computer that will be doing the processing. The platform will have the capability to move around in all four major directions.

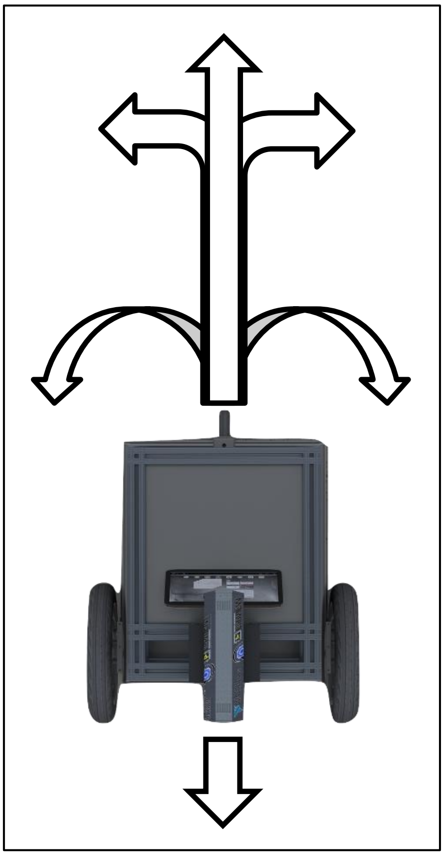


Figure 3-2 AGV Movement Capabilities

The tracking software is also responsible for the AGV movements and makes its decision based on the images it has analysed from the thermal camera. The manoeuvring capabilities of the AGV is similar to an electric wheelchair because of component equivalence. The advantages of this blueprint allow for less time spent on the design of the AGV and more time devoted to its movement and thermal imaging system. Designing and building a control system and then interfacing it to a camera would have taken up much time. To keep all focus on machine vision and human tracking, an electric wheelchair was used as a base for the running gear. The AGV is capable of making sharp turns, and because it uses a differential drive system, it can perform a pivot turn. The cast wheel helps the AGV to achieve a better turn radius and navigate tight spaces.

3.1.2 Modifications made to the intelligent speed controller

Initial modifications have been tested to allow the PG drive intelligent speed controller unit to interface with a computer through a data acquisition unit. The PG drive has an intelligent speed controller that is used to control the movement of the motors. The intelligent speed controller connects both motors and battery pack. Separating the controller as seen in Figure 3-3 reveals that the controller consists of two parts.

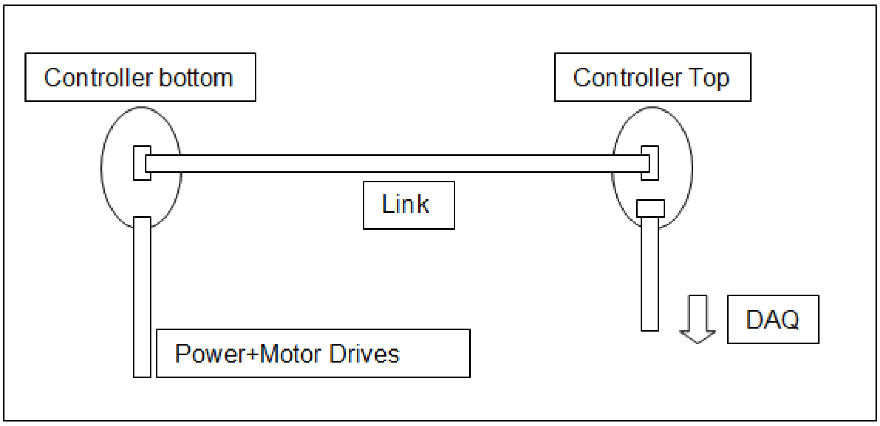


Figure 3-3 Controller separate parts

The bottom part regulates power and charges the batteries, and it also does pulse width modulation to move the motors. The second part is the joystick together with speed control and power buttons. Joining the two parts is an eight wire cable that relays user input from the joystick to the bottom part where it is processed and used to control the movement of the wheels. Figure 3-4 shows the dividing line where the controller is separated into two parts to allow modification.

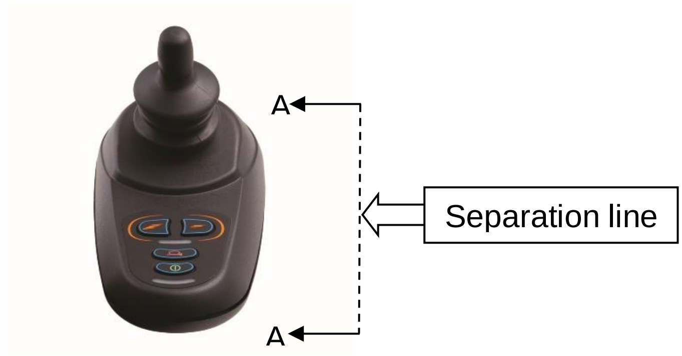


Figure 3-4 PG controller drive

The modification takes part between the eight wires linking the joystick to the base of the controller. Normally the link would be straight through – meaning pin one on the base will go to pin one on the joystick, as seen in Figure 3-5.

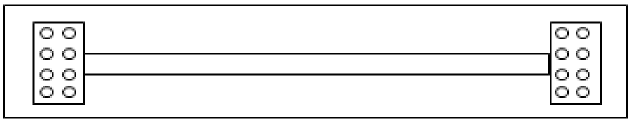


Figure 3-5 Unmodified link controller wiring

Figure 3-6 shows the modifications made to the connection linking the base and joystick; notice that some wires are unmodified. Safety features are built into the controller to prevent motor movement if the controller is damaged.

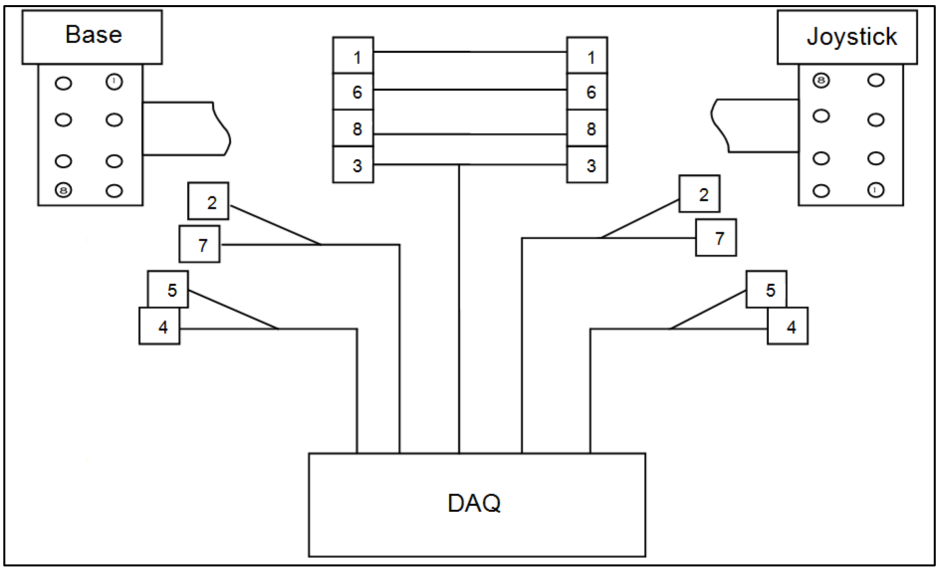


Figure 3-6 Rewired electric wheelchair controller

Interfacing via the controller reduced the time needed to design a motor drive and allowed fine control over the motors. The PG controller doubles as a safety system by preventing voltage spikes that can cause unwanted wheel movement. The controller constantly monitors the voltages received from the DAQ system and prevents voltages that exceed the input limit. Figure 3-7 shows the completed AGV with the thermal camera mounted to the rear of the frame with calibration heater, movement control system and computer. The placement of the PG controller drive allows for access to the joystick and speed buttons.

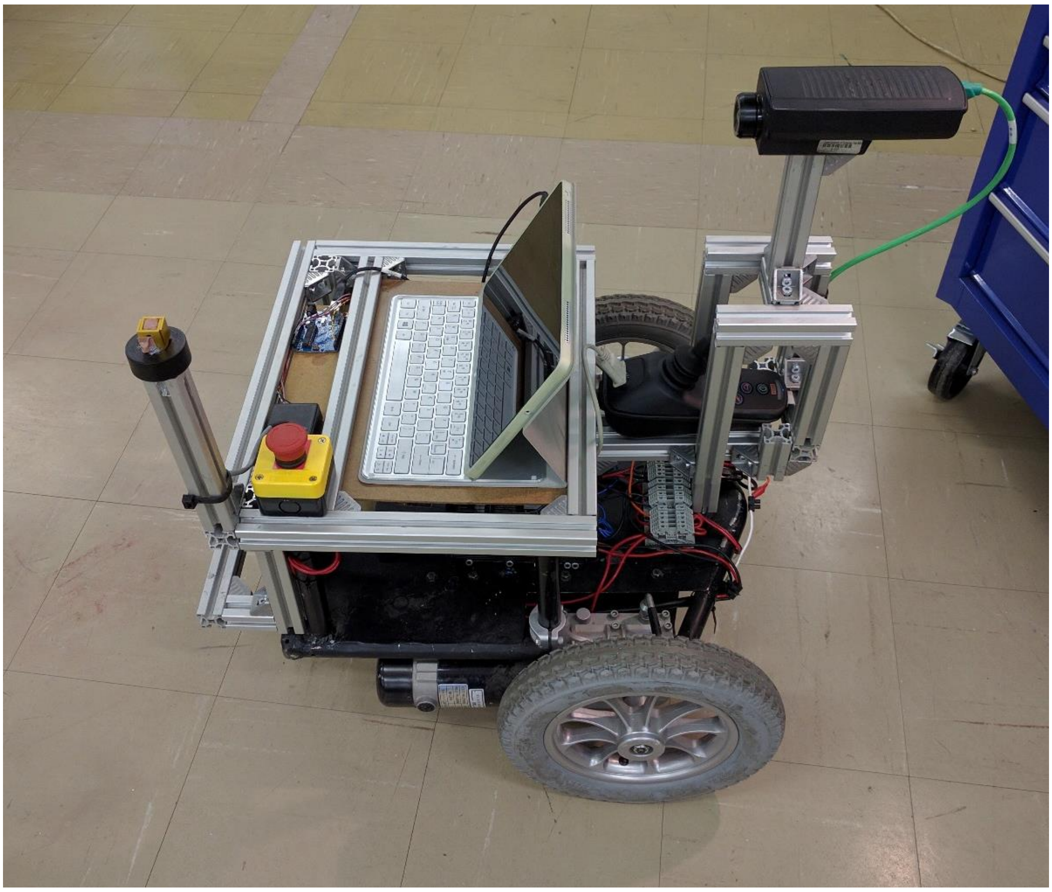


Figure 3-7 Is the completed AGV with all the particular systems mounted together

Figure 3-8 shows that the calibration point is within the camera's field of view without obstructing the visibility. The slate computer's screen faces the operator and is visible to anybody working with the AGV. Behind the ultrasonic sensor, the DAQ unit is visible and connects the PG drive to the slate computer.

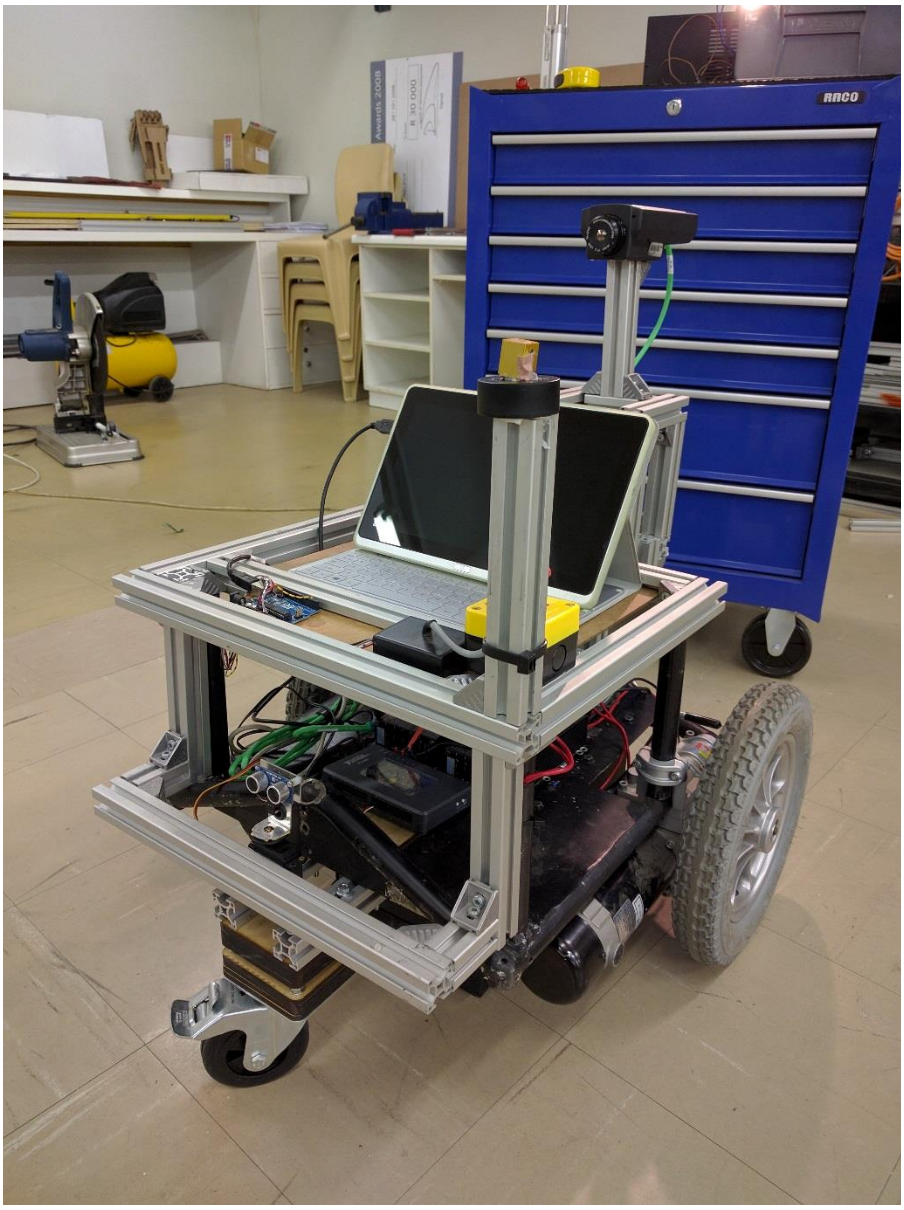


Figure 3-8 Frontal view of the AGV where the ultrasonic sensor is visible

3.2 System Hierarchy

Figure 3-9 shows the main system components and the interconnecting structure. USB and ethernet constitute the medium of communication to the computer. The thermal camera uses ethernet to communicate because its intended manufactured goal was to be used as a security camera. The software running on the PC is capable of detecting the camera and receiving images from the camera. The DAQ uses USB to communicate and is used to interface with the motor drive to control the AGV. The thermal calibration consists of hardware and software and uses an Arduino microcontroller to perform similar operations to the DAQ unit, as described in section 3.1.2.

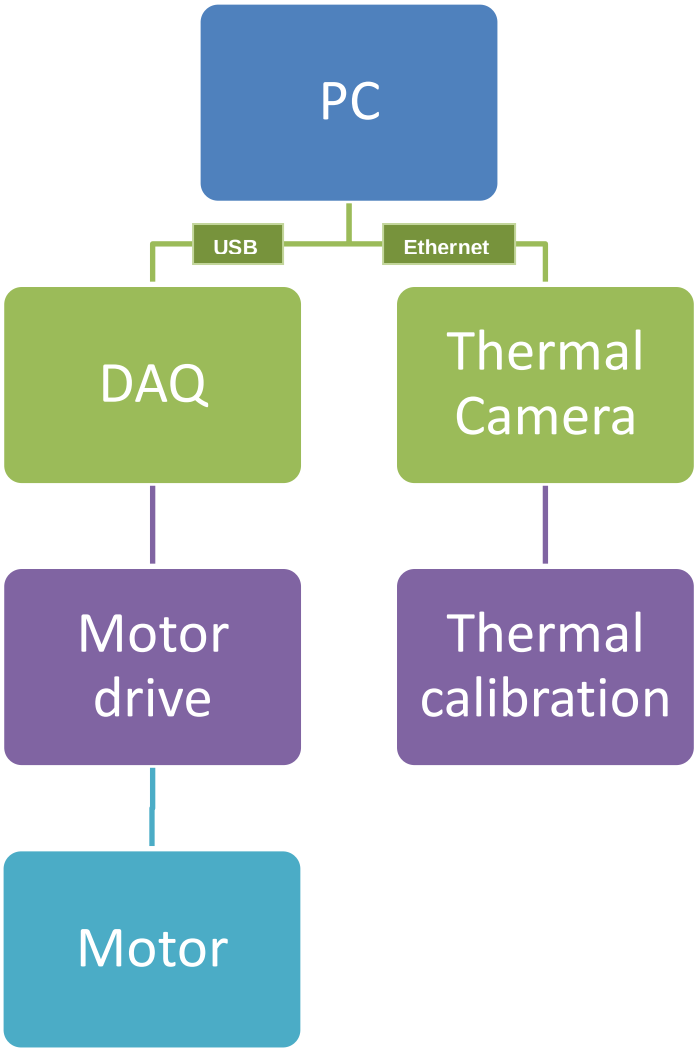


Figure 3-9 Complete hierarchy system breakdown of all individual components

3.2.1 System Operation

Figure 3-10 shows the process flow of the complete system. To start, the system LabVIEW has to be running on the AGV's computer. Two separate programs are running at the same time; one is the thermal calibration that controls a set temperature on an external heater. The second program analyses images from the thermal imaging camera and detects and tracks a human target. The movement part shares the data from the machine vision part of the program, and it is this coordination that is used to move the AGV.

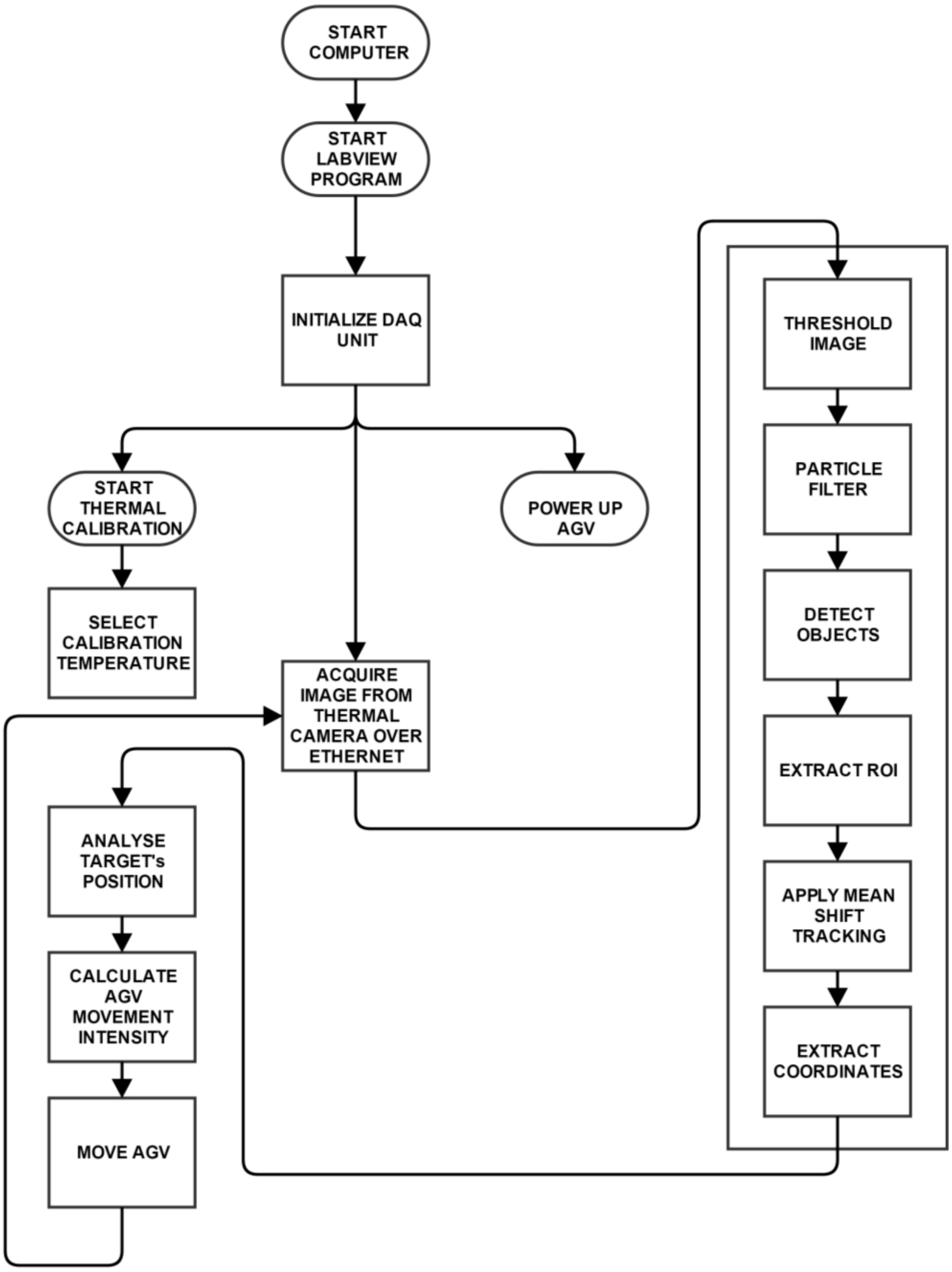


Figure 3-10 System Flow Chart Explaining Software and Hardware Processes of operation for AGV

Figure 3-10 shows an initial systems diagram that highlights key systems in the project. For the system to function optimally, certain procedures have to be followed at start-up. The slate computer is a fundamental component of the system that has to be started. The software running on the computer drives outputs that are connected to the running gear of the AGV. To have a successful start-up, the LabVIEW software needs to be initialised first. Once the software is active, the thermal camera can be started and the AGV running gear as well. LabVIEW is responsible for enabling drivers which communicate with the DAQ and IP thermal imaging camera. Software that was written using LabVIEW also enables the thermal calibration system that assists the thermal imaging camera.

3.3 Image acquisition

Cameras are widely used in the manufacturing and assembly industry to do quality checks and product inspection. Mounting cameras to moving robots present unique challenges in the form of changing light conditions and object reshaping based on perspective change. In this section, it is discussed how images are received from the camera and then imported into the software to be able to do image analyses on them.

3.3.1 Thermal camera

There is increasing interest in thermal cameras in the robotics and machine vision community. Using thermal cameras can significantly enhance the ability of a robot or AGV to identify and track a human. Thermal cameras work by capturing warm objects that emit energy in the infrared region of the electromagnetic spectrum. It is this infrared energy within the field of view that is optically captured and focused, as with visible light, and presented as an image. The proposed thermal camera is a network camera that connects to the computer via an IP protocol. The camera needs a constant power source which it receives via the AGV's onboard batteries. The images captured by the thermal camera are in Joint Photographic Experts Group format or JPEG. This format is an irreversible compression, and the digital images are captured at a rate of nine images per second. The images are received by the computer and will be analysed by software. The results of the analyses are used to detect and track a human.

3.3.2 Thermal camera software interface

The user interface for the thermal camera is accessed using a browser. The web interface controls the camera settings and also gives the user the ability to view a live stream from the camera. This interface is used to do the camera setup and is capable of adjusting the camera's gamma, shutter speed, contrast and calibration zone. These settings are adapted to achieve the best quality images for the experiment. Adjustable camera settings remain the same for the duration of the experiment. When capturing raw data, it is important not to make an adjustment to these setting as it will influence final results.

3.3.3 File Format

The format that the video information is transmitted in can be changed in the web interface. Some of the formats include a stream of images, which is called MJPEG. Each video format profile setting can be adjusted to optimise the quality of the video. The resolution and frame rate is the most important parameter influencing the quality of the video. The video streaming settings also allow compression over the network. This would be experimentally set up to achieve the best possible quality frame rate and image stream.

3.3.4 Thermal imaging colour palette

Under the video settings, there are different video pallets to choose from. Video pallets allow the video information displayed to be interpreted differently. Figure 3-11 explains this better by taking two of the pallets; one is White-hot, the other is called Black-hot. White-hot means all warm objects are displayed as white and in Black-hot all warm objects are displayed as black. Experiments would be done to select the optimal palette that is best suited for the environment the camera is going to be used in.

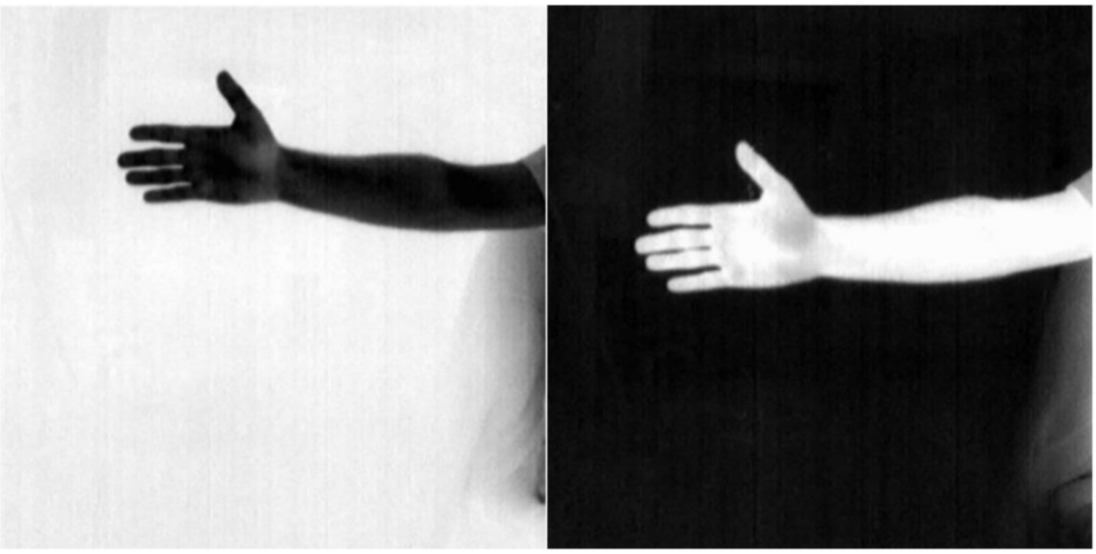


Figure 3-11 Thermal camera image pallets (Black hot: Left; White hot: Right)

3.3.5 Additional camera settings

Additional settings include access control and password setup. Because the camera uses internet protocol to communicate, it has an address, and it is thus important to secure it using a password. The setup of the IP address can also be done through this interface which includes the ability to setup a proxy. The camera web interface is used to control the camera settings directly. Initial setup is done to obtain the best possible image quality and clarity for the capturing environment. Figure 3-12 shows all the submenus available in the web interface. Within the camera settings, it is possible to control the brightness of the image; this helps to highlight details in the darkest parts of the image. The exposure setting was set to exposure zone. This setting determines a reference zone to adjust the camera's exposure constantly. Exposure settings are discussed in detail in section 3.4.

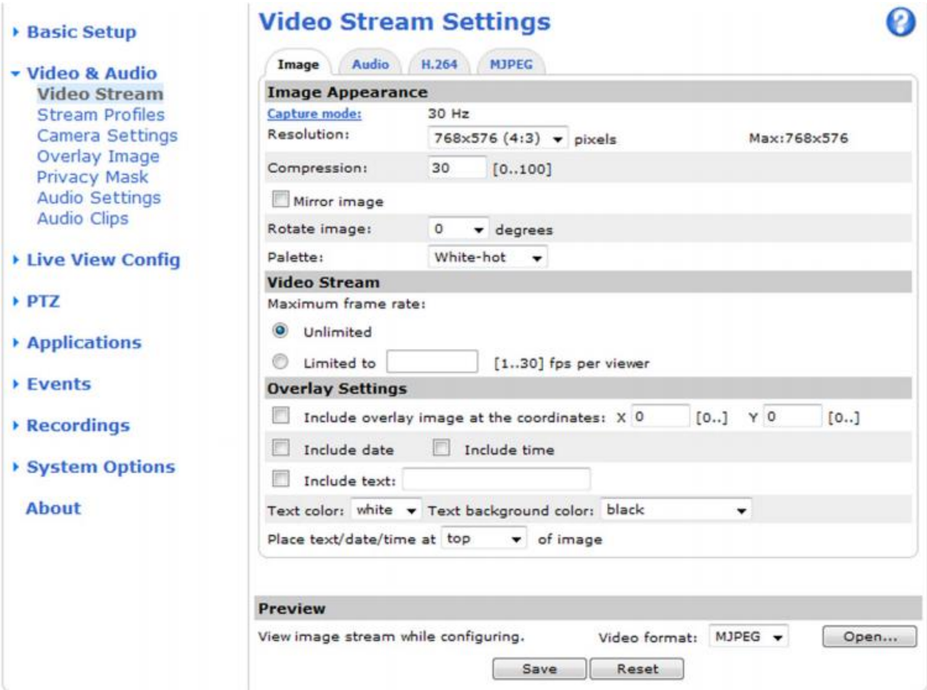


Figure 3-12 AXIS Thermal imaging camera web interface showing video settings [27]

3.3.6 Software and Drivers

Laboratory Virtual Instrument Engineering Workbench (LabVIEW) is a development environment for a visual programming language from National Instruments. LabVIEW is a system design platform used for data acquisition, instrument control, and industrial automation. Software required to complete this product include LabVIEW and two vision add-ons including vision acquisition software and vision development module. LabVIEW vision modules are required for drivers' support and software extensions used for identification and tracking.

3.3.7 Vision express

After initialization of the software is done, images can be retrieved from the camera. National Instruments' vision acquisition express toolbox helps to streamline the process of acquiring and analysing images. In Figure 3-13 it is shown that the Vision Acquisition toolbox block is used and is one of the fastest and easiest ways to configure all of the characteristics of the camera.

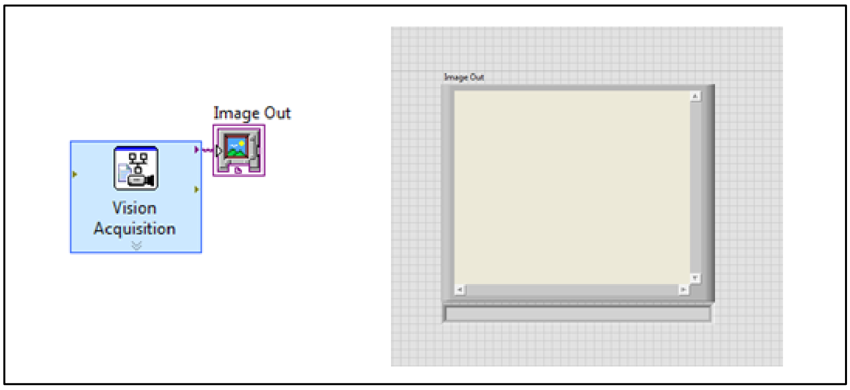


Figure 3-13 Video acquisition using IMAQ Vision Acquisition Express

The toolbox includes an initial camera setup. Connected cameras show up either as directly connected to the computer or the network and become available to acquire as a source. Once the acquisition source is selected, it will move on to select acquisition type. Acquisition type tells the program how the pictures must be acquired and consists of four modes. Single acquisition with processing, finite acquisition with inline processing, finite acquisition with post processing, and continuous acquisition with inline processing.

Single acquisition is used to acquire a single image. Finite acquisition with inline processing is used for acquiring a fixed number of images. When an image is acquired, it will be available for image processing. Finite acquisition with post processing is used for acquiring a fixed number of images, but all images must first be acquired before they can be processed. Continuous acquisition was employed in this project, to do real-time image acquisition and processing from the camera.

The next step is to configure the acquisition settings and change the attributes of the camera that include size, gamma, aperture, shutter speed and brightness. The last step is to select controllers and indicators that can change parameters; this allows the program to implement changes to the attributes of the camera during run time without having to do the setup procedure again.

3.4 Thermal Calibration

Thermal cameras give us the ability to see into the invisible part of the infrared spectrum. Although people can see infrared light up to 1050 nm, this is a small part of the broad infrared spectrum. Thermal imaging allows us to see long wavelength infrared radiation. Only specialised camera equipment can see the wavelengths of 8-15 micrometre namely thermal cameras [24].

3.4.1 Thermal Camera limitations

Thermal cameras are specialised equipment commonly used in military, search and rescue, police and security applications. It is also a very expensive piece of specialised equipment; the reason thermal cameras are so expensive has much to do with their lens material, as stated in section 2.4. The camera used is an Axis Q series thermal imaging camera. It should be noted that different countries have export control regulations that regulate which sensor or model of camera, at a specific resolution and frame rate, can be utilised in that country.

Microbolometer sensors used in thermal imaging cameras can be imported into South Africa and must meet specific criteria. The camera that will be used within the project was not designed for laboratory use, because it is not an ideal camera for machine vision and lacks adjustment features that industrial cameras have.

When working with machine vision, it is often desired to keep the lighting condition as constant as possible. Any changes in lighting conditions may disrupt results. That is why auto exposure is often turned off when doing image analysis. When working with the Q series thermal imaging camera, it was discovered that it could not switch off auto exposure. Automatic exposure control works by adjusting the shutter and lens aperture to limit the amount of light that reaches the camera’s light sensor [28].

Thermal imaging cameras function similarly to regular digital cameras that capture visible light, with the exception that thermal cameras capture infrared radiation. If a sharp light were to be shined at a regular camera, the camera would compensate by underexposing the image to make up for the excess light. It will also over expose if the light reaching the camera is insufficient, as seen in Figure 3-14. Thermal cameras use a similar method to regulate the exposure, but instead of visible light, it sees heat. Where in a regular camera light could come from reflections of natural or artificial sources, a thermal camera sees any object giving off heat – essentially turning every object producing heat into a light source [28].

Figure 3-14 Over and under exposure [31]

To not be able to turn auto exposure off is a problem when using thermal imaging in machine vision, because every object gives off a different amount of heat radiation. If auto exposure is left on, the camera will adjust its exposure to best focus on the warmest object in the frame. This camera feature can in most cases be switched off, but unfortunately not always. This makes image analysis difficult if the exposure is constantly changing.

3.4.1.1 External Camera Modification

A built-in feature in the camera's software allows the user to specify a part of the image the camera uses to calibrate its exposure. This calibration point can be small, and the reference position can be adjusted in the camera’s web software. An external thermal calibration surface had to be built in front of the camera so as to allow calibration of the exposure and to keep the exposure constant, as seen in Figure 3-15.

Figure 3-15 Heating element with a person is visible during thermal calibration

3.4.1.1.1 Hardware

The hardware solution consists of five components; the system is used to control the heat on an aluminium surface. Two MCP 9700E sensors are used to measure the temperature on the aluminium surface and an Arduino Uno microcontroller (Arduino) reads the voltages from the sensors. A 12-ohm resistor is used as a heating element to produce heat on demand from the Arduino. The heat sinks are used to regulate the heat more efficiently, and a computer running LabVIEW is necessary to process all sensor data.

To be able to control the heat on the calibration surface, the temperature has to be monitored. The two linear active thermistors are analogue temperature sensors that convert temperature to analogue voltage. Linear active thermistors do not require an additional signal conditioning circuit; this streamlines implementation and lowers circuit design time. The main advantage is that the voltage output pin can be directly connected to the analogue pin on the Arduino. These thermistors are immune to the effect of parasitic capacitance and can drive large capacitive loads. This allowed the sensors to be located away from the Arduino and gave some design flexibility.

Two sensors were used to increase the accuracy of the temperature. The calibration surface is one side of a square cube that consists of aluminium. The use of aluminium ensures even heat dispersal and more consistent temperatures. One sensor is placed on the calibration surface and the second sensor is positioned on the inside of the cube close to the resistor. This allows for better temperature control and an accurate calibration temperature on the surface, which is the part that the camera uses for calibration. The two thermal sensors are also run through a low pass filter to better manage interference. Some interference was noted when the resistive heating element was switched on.

3.4.1.1.2 Arduino Uno

The Arduino Uno microcontroller is used as a data acquisition device. The advantage of using this device is it has a fast analogue read speed which ensures high-speed data sampling. The Arduino can also interface with LabVIEW running on a computer – meaning data acquisition and processing takes place on the same software. The device is extremely power efficient, which makes it suitable for use on an AGV. The Arduino can be seen in Figure 3-16 connected to the computer.

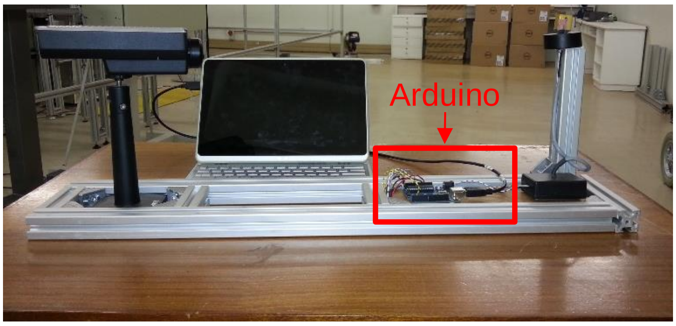


Figure 3-16 Thermal calibration distance testing and raw video data capture

The main purpose of the resistor is to heat up the calibration surface. It is more than capable of heating the aluminium surface to sufficient temperatures, and because the temperature is measured on the surface of the aluminium, constant feedback is received about calibration temperature, as seen in Figure 3-17. The Arduino is used to heat the resistor based on data it receives from the thermal sensors mounted in and around the aluminium calibration surface.

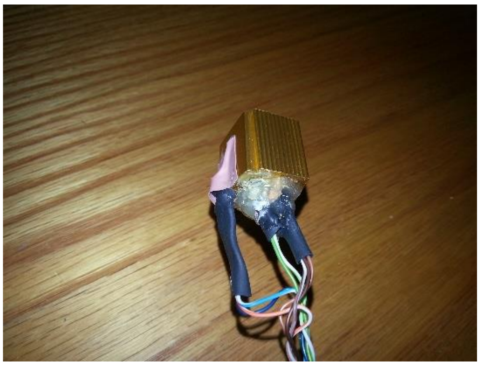


Figure 3-17 Thermal Calibration Heater

The heat sink is a 3x3x3 cm aluminium cube that is used to dissipate the heat produced by the resistor evenly across the calibration surface. It also better regulates the temperature and reduces the time the resistor is actively heating. The heat sink also disperses the heat and increases the size of the calibration surface.

The Arduino is connected to the computer via USB. The computer then executes software that monitors the voltages on both the thermal sensors. Based on the temperatures that the sensors are reading the resistor is switched on until the desired temperature is achieved. LabVIEW is used to convert the voltages received from the sensors to ^{o}C and then decide when to switch the resistor on or off.

3.4.1.1.3 Software

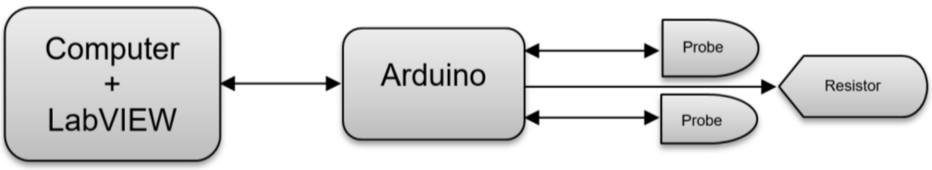


Figure 3-18 Calibration Hardware and Software Overview

The purpose of this LabVIEW program is to establish communication with the Arduino. The Arduino platform is used to read and write values to sensors and electronics connected to the Arduino. The connection has to be initialized to the Arduino to specify the baud rate, board type, bytes per packet and connection type. These parameters change with different types of Arduino boards. The thermal calibration VI also opens the serial connection port. It then sends a sync packet to the Arduino and waits for a response, which is an identical sync packet.

Once communication is established, the Arduino's analogue and digital pins can be controlled; before entering the main loop of the program, the digital pin responsible for switching the heating resistor has to be specified as output. The main loop of the program handles three processes:

- Analogue voltages from the two thermal probes are read.
- These voltages have to be converted to degrees Celsius and put through a smoothing filter.
- The heating resistor has to be switched on at the correct time to heat the calibration surface.

Figure 3-19 explains the hardware and software process to control and achieve the external calibration of the thermal imaging camera; it explains the logical flow that is followed to achieve the desired output in the program.

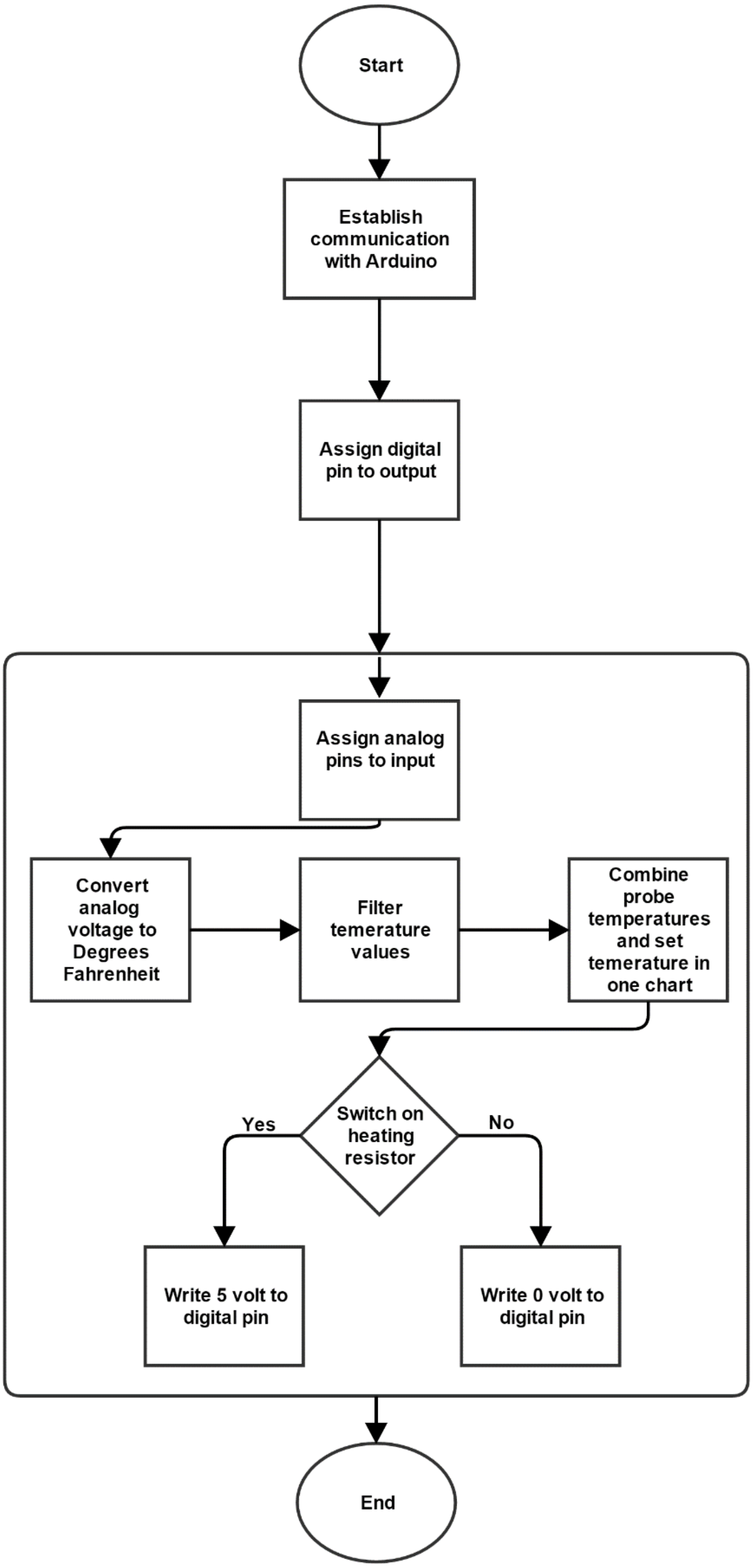


Figure 3-19 Flow Process of the Calibration Software used together with thermal camera

Figure 3-20 is developed in LabVIEW and follows a flow from left to right. All VIs outside of the while loop (box) is executed once, and everything inside of the while loop is repeated until stopped.

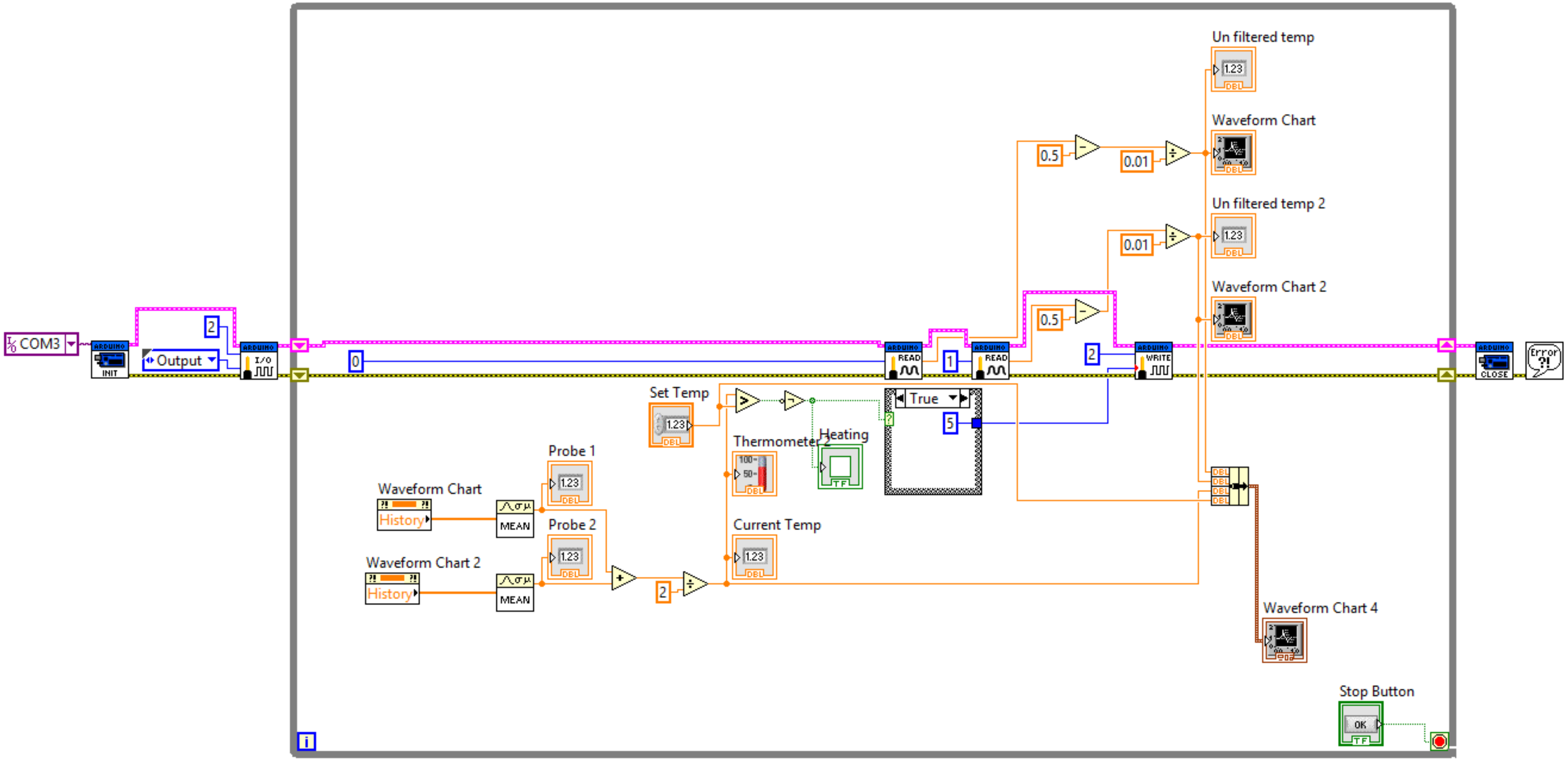


Figure 3-20 LabVIEW Code used to Control Temperature Calibration Software

Virtual Instrument Software Architecture (VISA) provides the programming interface between the hardware and development environments such as LabVIEW. First, the program finds the VISA resource names of all COM ports on the host machine. It finds the VISA resource name of all serial ports and then filters through these to find which of those COM ports are active. After the VISA resource is known, a session is opened to the desired COM port.

Figure 3-20 shows how the program uses COM port three to communicate with the Arduino and designate it as a VISA resource. The configuration of the given resource is done by setting baud rates, stop bits, data bits, and parity. When the setup is done, and the connection is syncing with the Arduino, writes and reads operations can be performed with the device. Pin two on the Arduino is configured as a digital output to allow the program to switch on the heating element. Figure 3-20 shows both pin zero and one set up as analogue read pins to allow them to read the voltages from the thermistors.

The thermistors voltages are converted to degrees Celsius in the program by deriving information from the thermistor's datasheet as seen in Equation 16 [29].

$$x = \frac{(y - 0.5)}{0.01} \tag{16}$$

Where y is the thermistor's read voltage value.

Because x is being plotted on a chart, it is first filtered by a mean calculation. The temperature is buffered 50 times, and the mean is calculated and used to represent a more accurate live temperature from the thermistor. Because the sample rate of the voltage is high, no delay in changing temperature is experienced.

The chart in Figure 3-21 is a screen capture of four values: a set temperature, two unfiltered temperature from the individual thermistors and the filtered real-time temperature, which is used to calibrate the infrared camera.

Based on the combined temperature from both thermistors, the heating resistor is switched on at the set temperature.

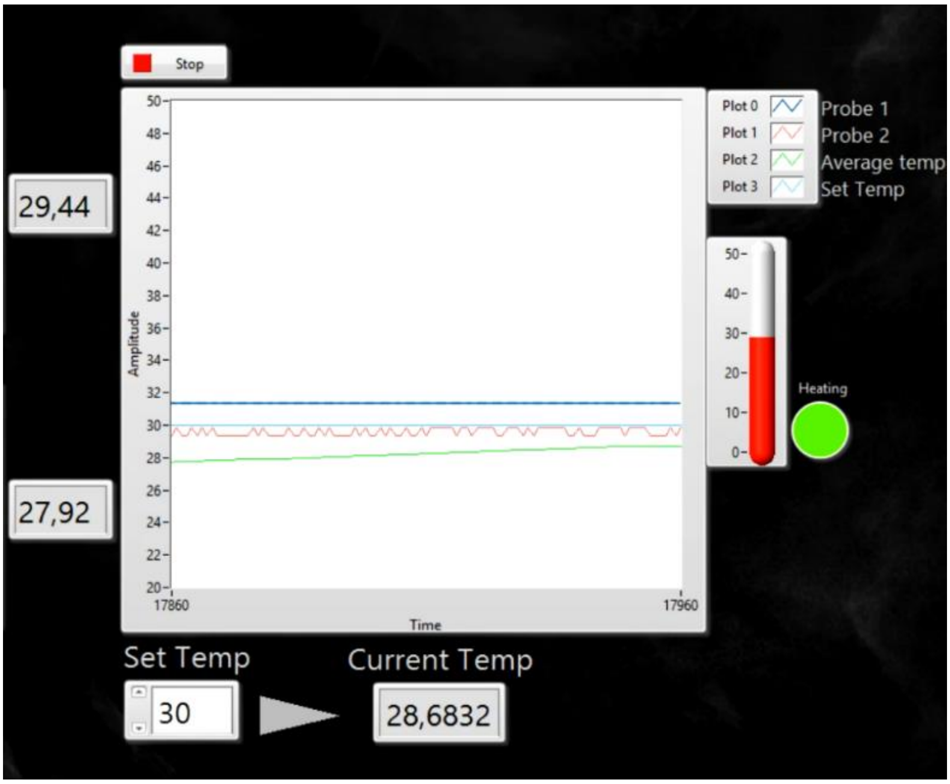


Figure 3-21 Front interface of thermal calibration program

After exiting the main loop, the active connection to the Arduino is closed down. This is important because leaving the connection open and exiting the program can leave the Arduino running and it might keep the heating surface on.

3.5 Identification

All the methods that were explored originated from the idea that a human possesses unique measurable features. These features are what makes humans unique and distinguishes us from all other objects. The first approach to the problem of identification was to measure the length and width of the target using a software calliper approach. The thought is that an average human will have a certain length at a fixed distance from the camera. The subject was asked to give an open stance at each meter mark so as to make it possible to measure their separate limbs. This approach showed potential and was considered a viable method for identification of a human subject. It was outperformed, however, by a second technique which uses template matching.

Measuring something in an image requires us to make it visible to the computer first. Template or image matching is a technique that teaches the computer to find a small part of an image which matches a template image. This method is processor expensive and has unpredictable results. The simplest method and thus the one used for identification is the thresholding of colour.

Colour thresholding works by taking an image from the camera and teaching the computer what colour to look for. With the thermal imaging camera, we receive grayscale images from the camera that are ready for segmentation. The output of thresholding is a binary image which consists of only black and white pixels. This gives us a measurable image which consists of 1's and 0's.

3.5.1 Colour plain extraction

This function is only required when working with post processing of videos. The reason for this is that the video received from the thermal camera is a grayscale image which does not require colour plain extraction to be thresholded. Simply extracting a single colour will turn the RGB image into a grayscale image. This will, however, have no effect on the image as it is already without colour. If the images were recorded using a full-colour camera, the extraction of a single colour from the image would have changed the image significantly. The main reason for using plain extractions is because of the way the videos were recorded and saved. Recording directly from the screen resulted in an image format that represented an RGB image and not a grayscale image. Once the images are in grayscale, it is possible to apply thresholding to them. Figure 3-22 gives an overview of the process behind thresholding an image.

Figure 3-22 Identification process

3.5.2 Thresholding

The thermal camera images go through a process called thresholding first to eliminate background noise; this is anything other than human body heat.

Once we have a grayscale image, it can be thresholded. Thresholding changes a grayscale image into a binary image, i.e. the range of pixel values goes from [0, 255] to [0, 1]. This is visually represented by a black (for pixel value 0) and red (for pixel value 1) image in NI-Vision. Since most machine vision operations expect grayscale images, certain settings need to be changed to make these operations work.

Thresholding allows the removal of unwanted temperature from the image. Because thermal images represent temperature in colour, we could extract valuable information from the images about human body temperature.

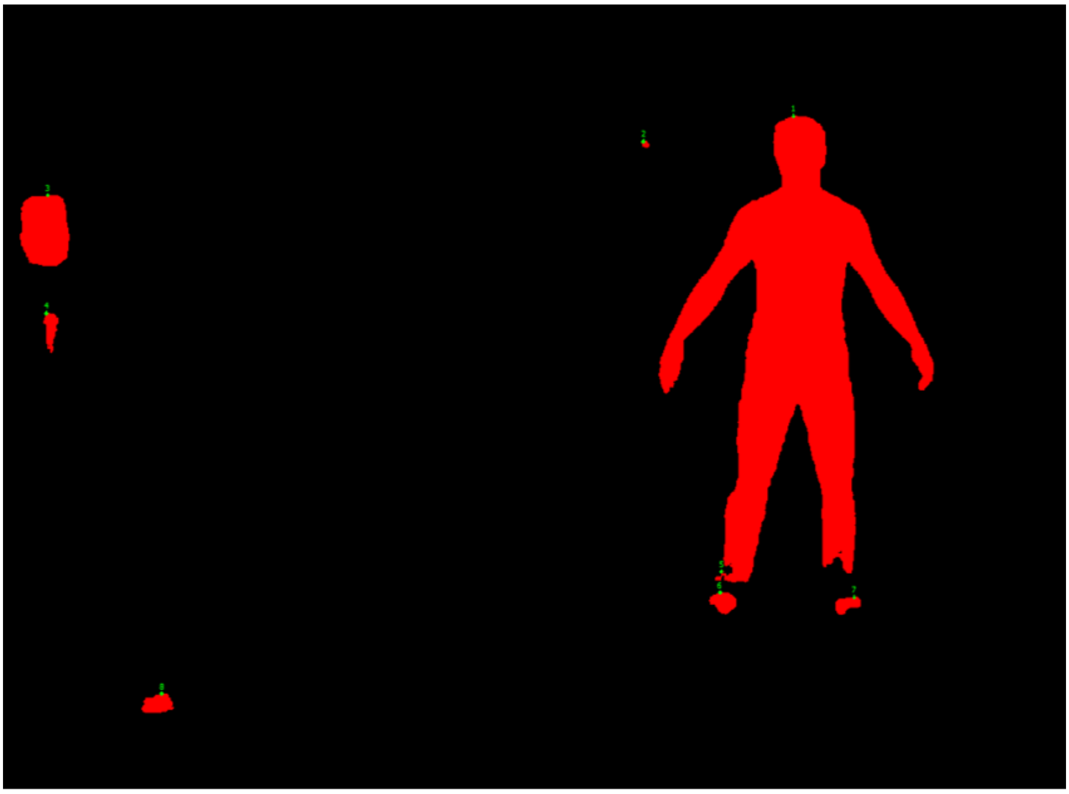
3.5.3 Particle filter

A particle filter helps to differentiate between human and artificial heat. It works on the principle of heat signature size. Human body heat will occupy the majority of the image; this assumption is true only when working with a single human in a thermally clean environment; therefore a human will occupy a certain number of pixels on the image; anything smaller than that can be discarded.

The particle filter acts as a software filter where the calibration surface which was discussed in section 3.4 serves as a hardware filter. The purpose of filters is to reduce the number of objects that need to be analysed by the software.

A high object count in the analysis of an image means that more processing power is required to distinguish non-human objects from human objects. This inevitably leads to a reduction in performance and reliability. Objects or particles in an image can either refer to any physical or light-based objects in the field of view of the camera.

LabVIEW has a function that allows filtering of particles in a binary image by first measuring the area of each particle in the frame. Each particle receives an area value which makes it possible to do filtering. In software, it is specified which particles to remove and which ones to keep. Figure 3-23 shows the breakdown of how many particles were detected and each particle's area size.



Results ...	1	2	3	4	5	6	7	8
Area	80793.00000	70.00000	5314.00000	626.00000	68.00000	740.00000	549.00000	845.00000

Figure 3-23 Particle filter results

With that information, a function in LabVIEW was used to filter out small, nonessential particles as seen in Figure 3-24.

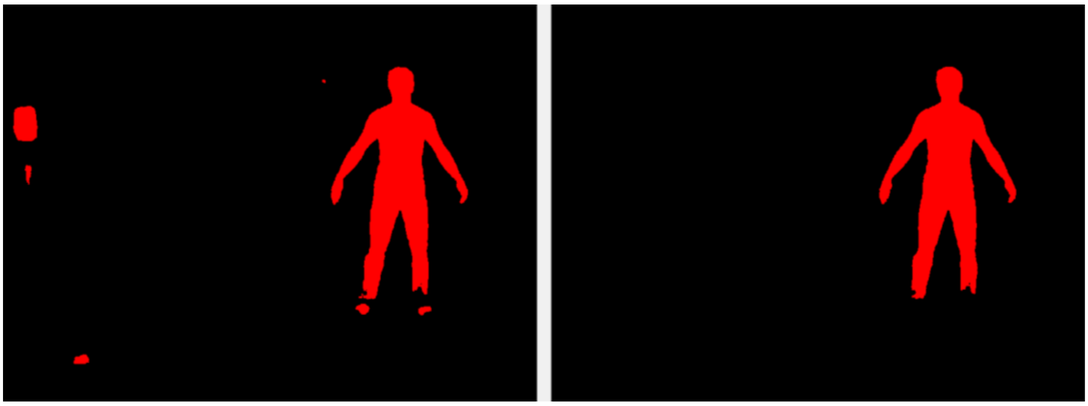


Figure 3-24 Particle filtering comparison

3.5.4 Centroid Tracking

LabVIEW is equipped with a VI called Centroid, which functions as a subroutine to find exact coordinates of a particle within the frame of the camera. The Centroid virtual instrument is being used in tandem with the AGV movement control and gives exact coordinates of the object being tracked to allow for decision-making when moving the AGV's orientation. Figure 3-25 is the connector panel of the Centroid VI subroutine and will be used to represent the VI in the block diagram of the LabVIEW program.

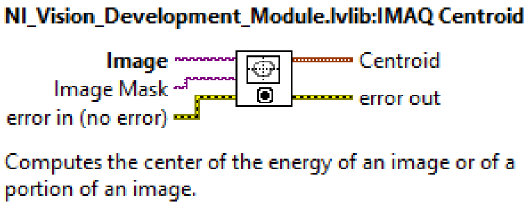


Figure 3-25 Centroid VI subroutine

Figure 3-26 is a representation of how LabVIEW centroid VI is used to find the centre mass of a particle. Even though only the person's coordinates are shown, it is capable of tracking every distinguishable object in the frame.

Figure 3-26 Centroid applied to image

Figure 3-27 shows where the Centroid VI is used in a vision analyses loop. Note the cluster called centroid2 that consist of two elements, namely the (x,y) coordinates which are used in section 3.6.2 for orientation of the AGV. The specific theory of operation is explained in section 2.2.12.

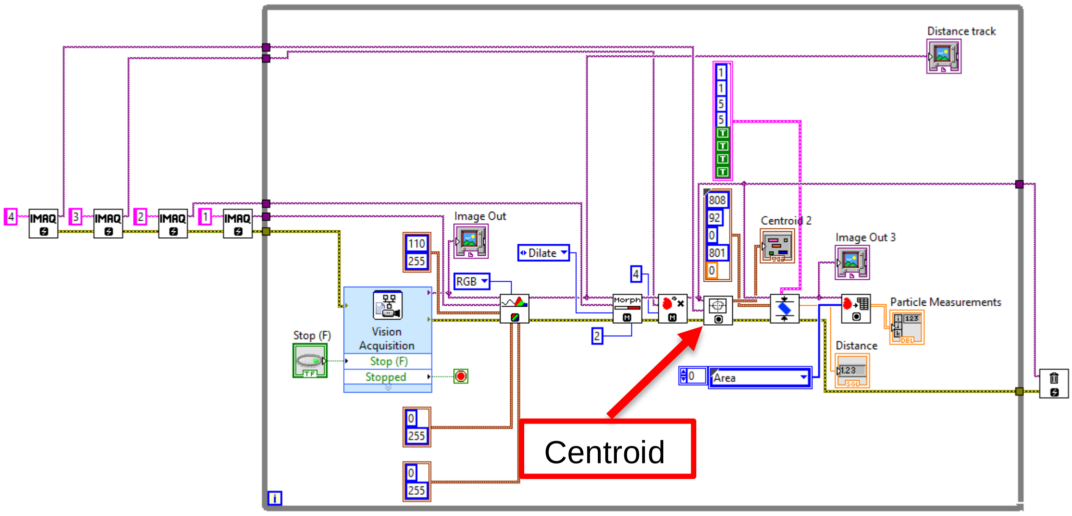


Figure 3-27 The Centroid VI is used to extract object coordinates from real time images

Figure 3-27 represents an application of the centroid function into a program; it is often used with other image processing techniques to allow the program to track an object.

3.5.5 Obstacle Detection

An obstacle detection system is implemented to allow forward-facing detection of obstacles. Using an ultrasonic sensor mounted to a servo motor the servo motor is constantly moving in a 160-degree arc, which allows the sensor to scan the immediate forward terrain. The ultrasonic sensor consists of two units, a sender and receiver; this allows the sensor to send out a burst of inaudible sound and then receive it again. The sensor uses echo location which works similarly to modern vehicle parking sensors. The distance that obstacles can be detected varies depending on shape and material. The average effective distance is two metres. The obstacle detection serves the purpose of detecting immediate collisions and is not used to for track targets.

Figure 3-28 is the front panel of the obstacle detection program and shows the program during operation. On the left side, the servo position can be seen together with a speed selector to adjust the rate that the sensor is moving from left to right. The two centre charts display every ping that is received back from the sensor. The charts show an obstacle to the front because the lines are short, meaning the returning sound reached the sensor faster. The gauge at the bottom right indicates the position of the sensor and is red if an obstacle is too close.

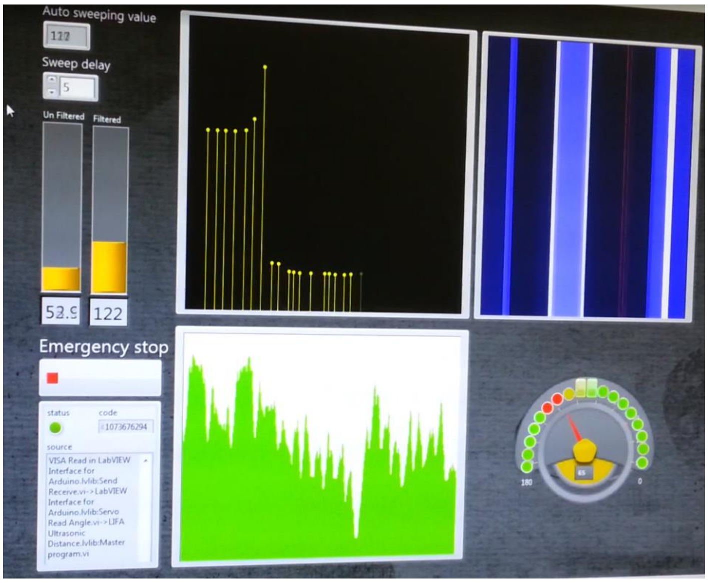


Figure 3-28 Ultrasonic Obstacle Detection GUI

Figure 3-29 is a prototype built before the sensor is mounted to the AGV. It shows the ultrasonic ping sensor mounted to a servo motor; both are connected to an Arduino. The servo motor is powered from a different source.

Figure 3-29 Panning Ultrasonic Sensor Mounted to the Servo Motor

3.6 Tracking

After the identification process is done, two values are returned to be used in tracking. The location of the subject and the onscreen size. The location is returned using the centre mass tracking process. Exact coordinates are returned to what is the centre mass of the person to be tracked. Figure 3-26 shows that the coordinates are close to the centre point of the individual's body. The accuracy of the coordinates depends on how much noise or false positives make it through the filter process. The particle filter gives the size of the person. Knowing the size of the objects and coordinates greatly increases the effectiveness of the tracking software.

3.6.1 Mean shift object tracking process

The first frame is used to select the target or model that should be tracked. The initial target select stage requires the placement of an ROI around the target.

LabVIEW uses VI's to apply an algorithm that tracks the ROI defined by the user through iteratively updating the location of the object as explained in section 2.2.13.

Mean shift convergence is used in the next frame to compare the current histogram to spatial data for the best target match candidate by maximising the similarity function. In the mean shift algorithm, the object centre moves from current location to a new location, as was shown in the section 2.2.13. The kernel is moved until the convergence of the similarity function; then the location of the object is updated.

Part of the second step is to choose the feature space by extracting the gradient and colour. Then the PDF of the ROI is calculated by obtaining the histogram of the initial frame.

Figure 3-30 shows that the basic idea is to move from frame to frame and compare the successive histograms of the ROI in the image.

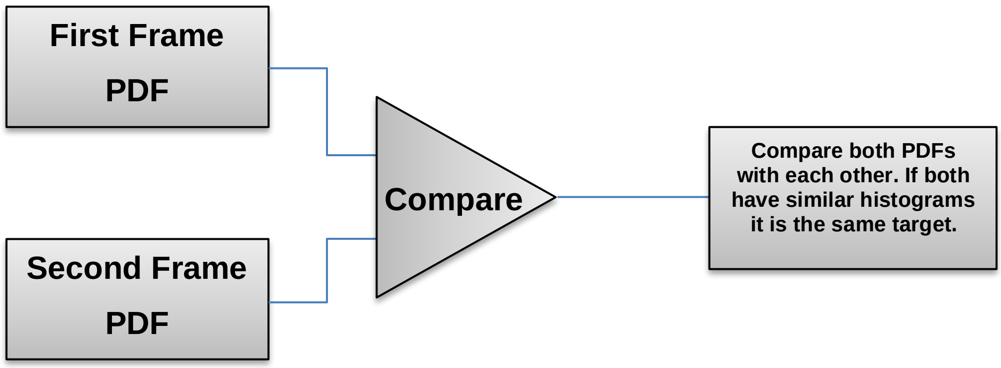


Figure 3-30 Mean Shift process

The Bhattacharyya Coefficient does the comparison and similarity function. The process of tracking a single target is outlined in Figure 3-31 [21].

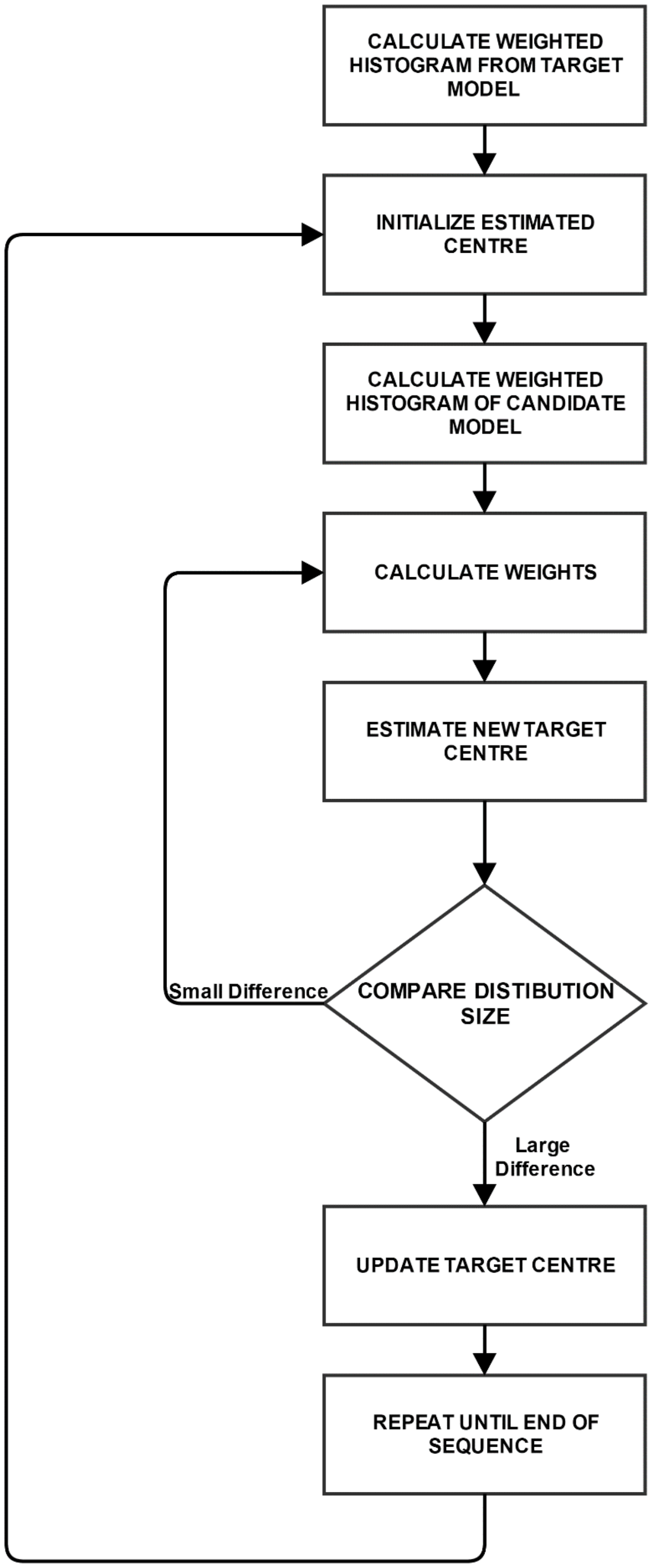


Figure 3-31 Flow diagram tracking single target [21]

3.6.2 Analyse calculated and movement control for AGV

The operation of deciding how the AGV should move is explained in Figure 3-32. Using the coordinates generated from the vision analysis program, it is possible to orient and move the AGV.

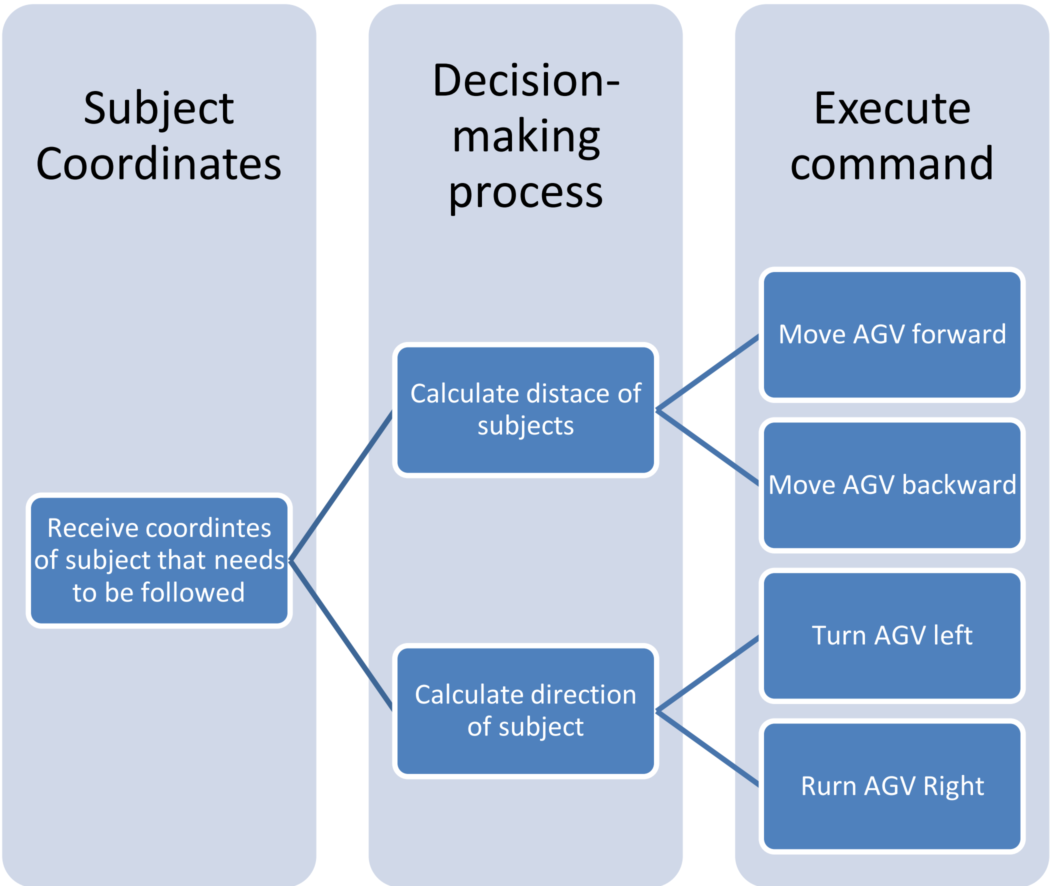


Figure 3-32 AGV Movement Process

Experimentation is done with two methods to determine the distance of the target to the AGV. The first method calculates the AGV's distance to target by measuring the objectives length. This approach assumes that the length of the target changes as distance changes. The image analysis measures the distance from the person's head to feet and sends the measurements to the movement VI, as seen in Figure 3-33.

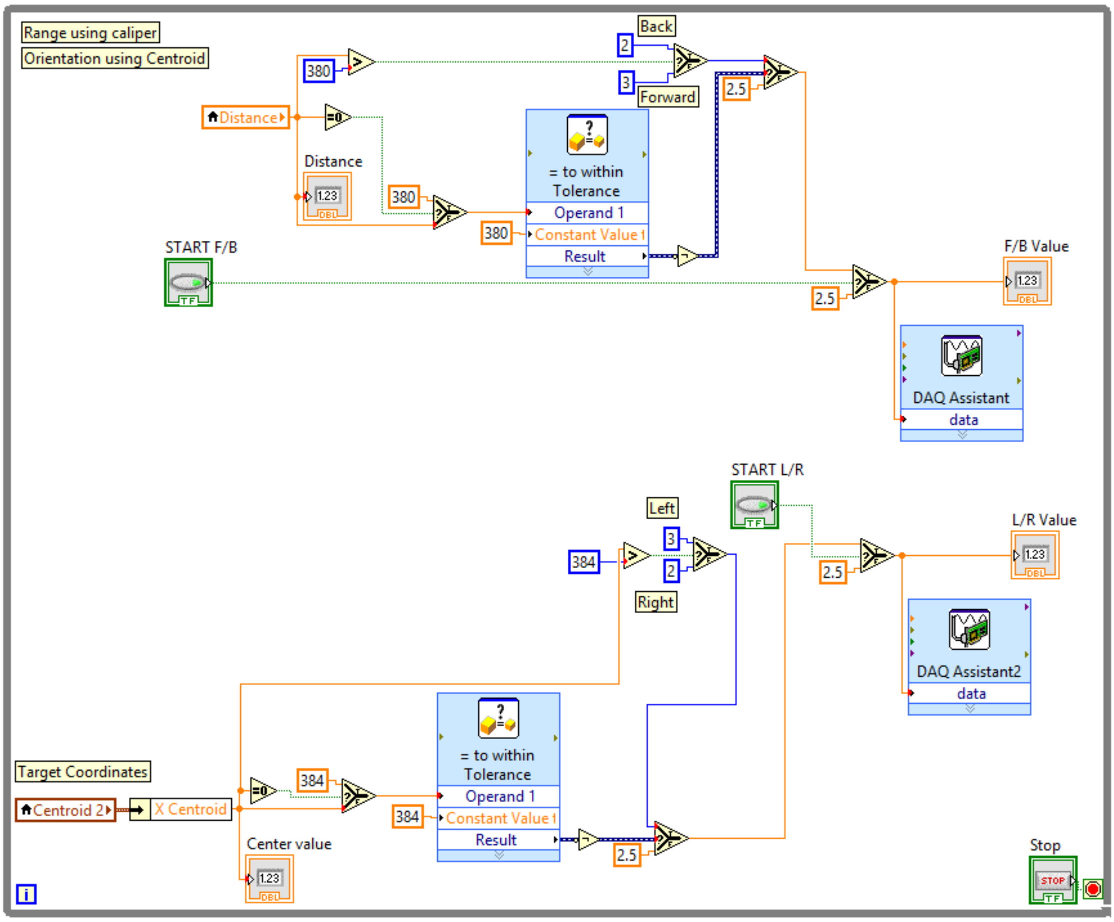


Figure 3-33 AGV movement control system using calliper and centroid

The second method seen in Figure 3-34 uses the target's area to calculate the distance of the target from the AGV. This approach assumes that the target's area changes as distance changes. A smaller area means the target is further, and a larger area means the target is close.

Both methods also use a Centroid VI to orient the AGV to the target. The centroid coordinates are received in the form of a cluster, which consists of two elements; the cluster is then separated to reveal (x,y) coordinates. The coordinates are passed through a process that decides whether the AGV should orientate left or right depending on the position of the target in the frame. The same logic is used when moving the AGV forwards or backwards.

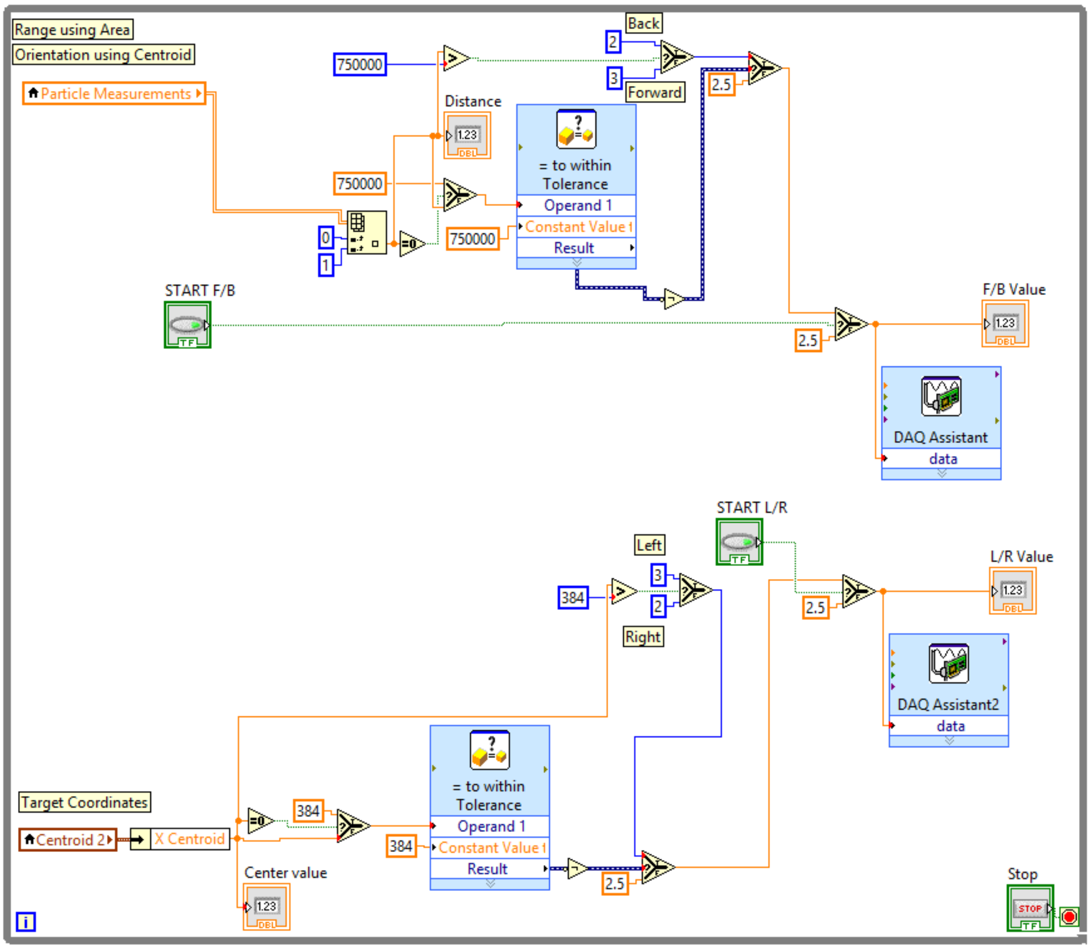


Figure 3-34 AGV movement control system using area and centroid.

Build into the process is a tolerance to deter the AGV from sporadic movement and overcompensating for a turn. Rather, the movement is elegant and appears smooth. Signals sent to the DAQ, which in turn controls the motors, consists of only three values; these values represent voltages that are sent to the DAQ and motor controller.

CHAPTER 4:Results

This chapter reports on a series of tests that were undertaken to verify the operation of the system and its subcomponents, by initially making use of a pre-recorded thermal camera video and afterwards verifying the results through physical testing.

A LabVIEW VI was developed to perform two operations. The first task is to analyse a real-time thermal imaging camera video and search for a human. Secondly, once a human subject is detected, it will be able to track that subject.

Pre-recorded videos of human subjects were analysed. In each video, only one person is present and is facing the camera at all times. The person starts from a position of two metres from the camera and moves to a distance of ten metres stopping on each metre mark. The subject is facing the camera so that a clear view of the person's face is visible. Distance is also introduced into the videos to allow for the development of a robust method of tracking. LabVIEW is used to develop a two-stage program that identifies and tracks a human target.

4.1 Person Identification

The expected outcome is that raw grayscale images from the thermal imaging camera can be processed and used to almost entirely remove background noise and only detect the heated objects in its field of view. This experimental program was developed in NI Vision Builder and was compiled with the intention of testing a concept.

Here follows the result of that program: Figure 4-1 shows the initial raw images that were received from the thermal camera. It can be seen that the person clearly stands out and that background objects are hardly visible.

Figure 4-1 Grayscale Image

Figure 4-2 is a thresholded image of Figure 4-1. It clearly shows how distinguishable the silhouette of a person is after thresholding. Additional image processing techniques should improve this binary image to remove additional background noise.

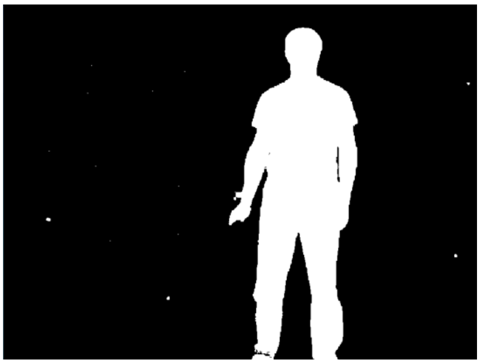


Figure 4-2 Binary Threshold Image

After thresholding the image, another technique is applied to eliminate small particles. In all cases, the heat was picked up from sources other than a human target. Some of these sources were lights, computers, reflections and electronic devices all significantly smaller than the human in the scene. Gray Morphology is applied in order to erode away small particles. Erosion works by removing particles in a binary image that are smaller than a certain size. Figure 4-3 shows the results of erosion applied to an image.

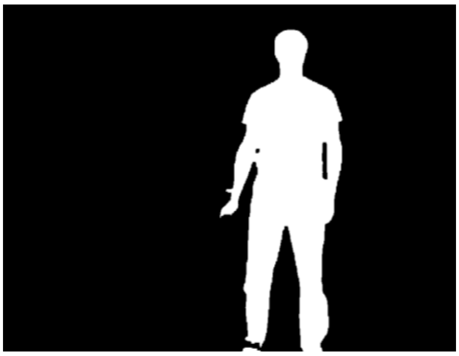


Figure 4-3 Eroded Image

The last step in the process is object detection, and this step focuses on detecting the remaining particles in the image. Roughly knowing how many pixels a human body occupies in the frame at a certain distance allows one to specify what to look for; anything smaller or bigger is ignored.

If all but the desired object could be filtered out, this method should always work. Unfortunately, small objects which are identified as noise will always be present after thresholding. Initially, the program looks for any heat signature. The idea is to make the program capable of distinguishing a human target from the rest of the environment. This first attempt at identification shows that although it is relatively straightforward to identify an object, it remains a challenge to identify a human target. The next step is to make the program more robust and capable of detecting a human target, without reducing performance and stability.

4.1.1 Thermal characteristics

To track a human target, it is necessary first to look at what makes a human unique. Thermal cameras detect heat only; this means that it detects humans very well, but it also detects computer monitors, server racks and lights equally proficiently; in fact, anything that gives off heat is visible.

Designing a program capable of detecting a human requires it to use human features as a reference. A human silhouette has obvious features that stand out; arms legs and head being the most prominent features that can be used to track a person. Either by measurement, object comparison or both can the program successfully track a human. The biggest problem so far has been scale. Once the target moves closer or further from the camera the size of the limbs changes, making it difficult to measure or compare.

Figure 4-4 represents some of the body features that present a measurable source to determine if the object is a human or not. Literature exists which explains how an elliptic contour measurement model is used to estimate the position of a person in the image. This method uses particle filters to estimate the state of the system at a given time based on current and past measurements. This elliptic measurement model is used to determine the position of the person and also the person's head [1].

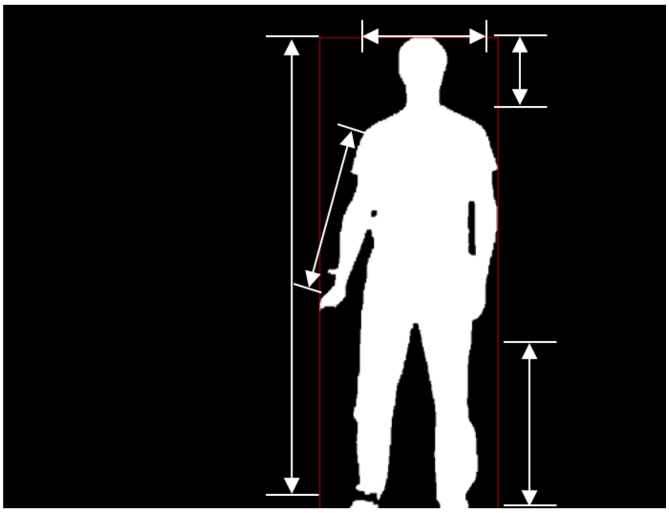


Figure 4-4 Example of Limb Measurement

A similar approach together with additional image processing techniques will enable a better method of identifying and tracking a person. A problem that persisted was the interference of background noise with the identification process. The problem was that the identification program was designed for a specific environment; if the environment is changed, the program is unable to adapt.

Rather than adapting the software, the focus was shifted to forcing the camera to adapt to changing environments. Figure 4-5 explores how an external device was used to significantly help with the stability of detection and also tracking. The simpler method of identification relies on the condition that the person that should be identified is the only person in the frame and that there should also be no significant background noise present.

The result of this initial program shows that it can better detect a person than conventional cameras. It can cut out background noise almost completely. Artificial lights do not influence the tracking program in a significant way, and that interaction between humans and robot can significantly improve with the use of thermal imaging cameras. The problem of identifying a human is still a major problem even with the addition of a camera capable of revealing infrared thermal radiation as light.

4.1.2 Setup of Identification Experiment

Figure 4-5 shows how the camera reacts to different calibration temperatures. The calibration surface is now used to adjust the exposure of the camera. Instead of relying on the camera's auto exposure function, the calibration surface is used to establish the most efficient temperature to use for the camera's exposure control. The first image shows that ambient temperature is noticeable if the calibration surface is at 25°C. Any object other than the human subject is classified as noise and needs to be eliminated or reduced.

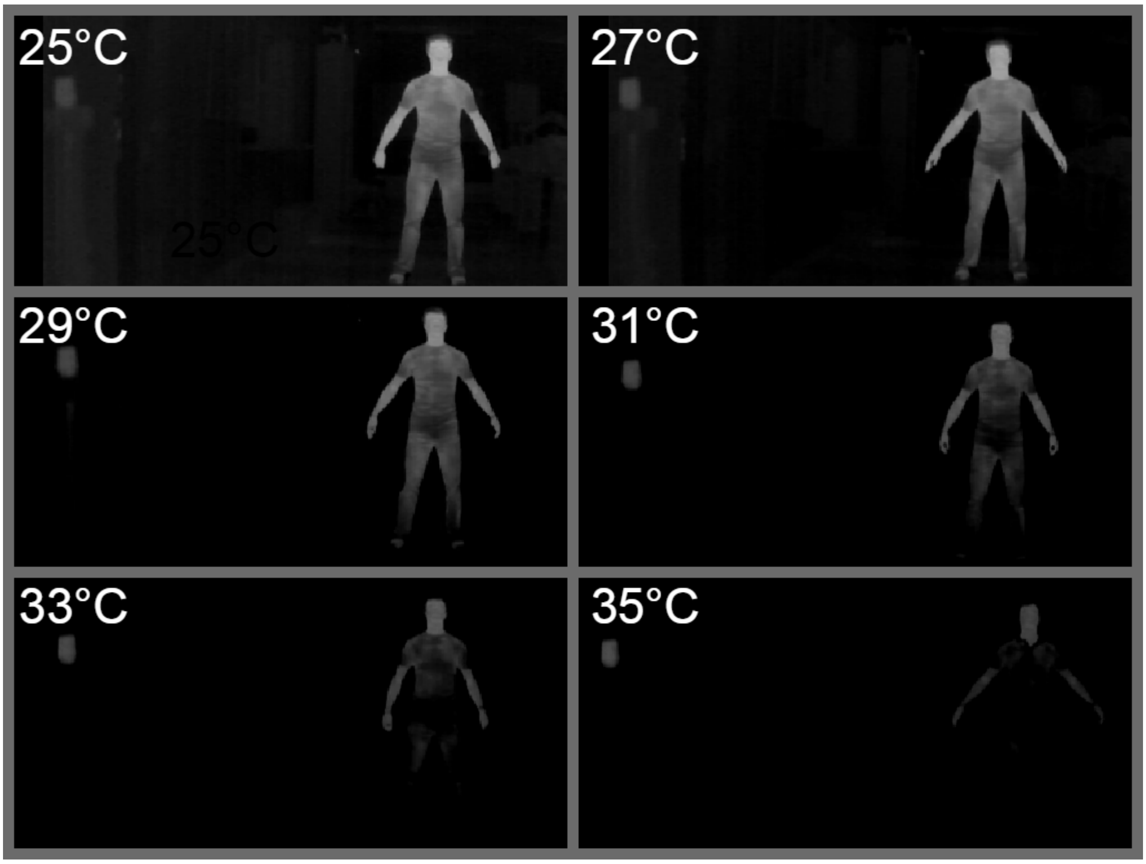


Figure 4-5 Image of how the camera reacts to different temperatures on the calibration point

Normal human body temperature will range from 37°C down to 20°C. This is because the average human body temperature has a maximum of 37°C and the skin and clothes (which are seen by the thermal camera) are influenced by external factors like ambient temperature [30].

The calibration surface allows us to eliminate some of the background noise by raising the surface temperature. If the target's temperature is between 27°C and 37°C the calibration surface will be adjusted to 26°C.

Adjusting the exposure with the calibration surface allows us to cut out all objects with temperatures below 26°C and only observe objects with temperatures higher than 26°C. All objects that have a temperature higher than the calibration set temperature will be visible. Background noise consists of mostly lower temperatures and are therefore reduced.

Unfortunately, nothing can be done to implement a ceiling temperature, and this problem has to be solved in the computer vision software. All results are captured with LabVIEW, Vision Assistant and NI Vision Builder. To keep all variables constant, the environment had a steady temperature of 22°C. The subjects always face the camera and move from three metres to 10 metres from the camera, stopping at every metre.

Five subjects were recorded in this way at different exposure values from 25°C to 35°C in 2°C intervals. The results consisted of 30 different videos to analyse. A subject was recorded moving an object on the floor in front of the camera. The video allows testing of less scripted movement and more natural human behaviour and seeing how the tracking program reacted to changing scenarios.

Figure 4-6 shows a diagram of the experiment and illustrates the camera on the left pointing to a person on the right; close to the camera the calibration surface can be seen visible in the camera's field of view.

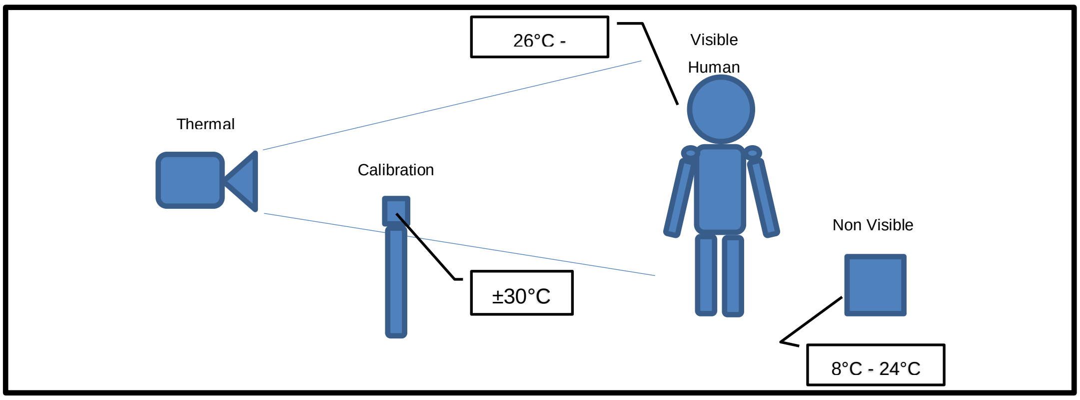


Figure 4-6 Representation of Thermal calibration setup overview

Figure 4-7 is a photo of the calibration system under development. Figure 4-7 shows the system before it was mounted to the AGV. The camera is slightly angled to allow the calibration surface to be visible in the field of view but not obstructing a human target.

Figure 4-7 Thermal calibration setup overview

4.1.3 Results on Identification of Human Target

To test the identification capability of the program, the best calibration temperature has to be selected. A human target is captured at the four metre mark at five different exposure settings, which is regulated with the calibration surface and is measured in temperature.

The temperature starts at 25°C and is incremented by 2°C until 35°C. This means six photos have to be analysed to determine what temperature will work best for human target identification. Figure 4-8 is a chart that represents the analysis of six photos taken with the thermal imaging camera. Each photo was taken at a different exposure value, which was set using the external calibration device. Figure 4-8 is a chart that shows how an increase in temperature decreases background noise and also target area. To successfully identify a human target, the background noise has to be reduced, but a suitable target area must be kept.

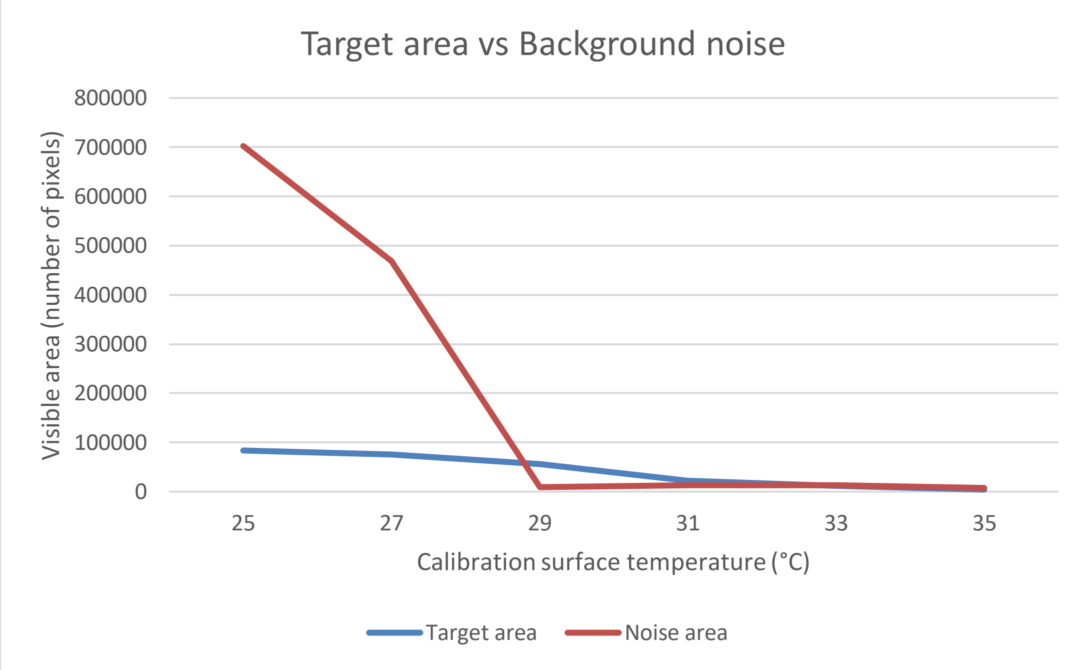


Figure 4-8 Human visibility versus background noise visibility

After analysis, the results are displayed in Figure 4-8 in chart form. The results show the optimal temperature for identification of a human target in an indoor environment. The optimal temperature will be at a point where the target retains a large area after thresholding, and the background noise is low enough to discern the target from the background noise. The chart in Figure 4-8 points to 29°C being the best possible temperature for identification. Figure 4-9, however, shows that while 29°C allows for clear human target identification, the tracking program prefers 25-27°C. Based on these results it is shown that the identification and tracking program use different computer vision approaches. Mean-shift tracking prefers a higher number of trackable pixels and therefore react better to the 25-27°C calibration zone.

4.2 Person Tracking

4.2.1 Setup of Mean Shift Tracking

Individually, four subjects were asked to face the thermal imaging camera – which was set at a fixed height of one metre from the ground. The camera was mounted on a table and not the AGV for this experiment. The subjects were then recorded walking from three to 10 metres facing the camera. All subjects were recorded at 25°C, 27°C and 29°C.

The tracking program was set to lock on to the complete body of the subject as shown in Figure 4-10. Data from 12 videos are shown in Figure 4-9 in chart form. Each colour represents a different subject, and each subject was recorded at three calibration temperatures. Each video ran for 670 program cycles, which gave 670 tracking scores per video. The program scored each program cycle out of a 1000 meaning a score of 1000 equals a 100% tracking score. For each tracking run a calibration temperature was set and a total of three temperatures used for each subject.

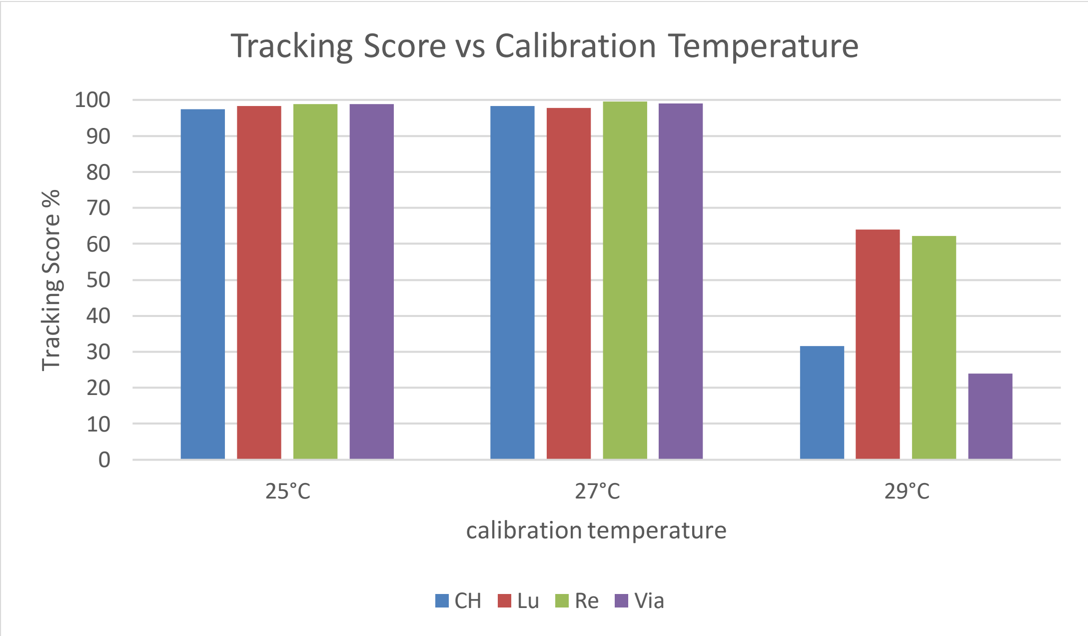


Figure 4-9 Tracking complete human target at different distances and calibration temperatures

Figure 4-10 and Figure 4-11 shows four images and two different tracking parameters. Figure 4-10 shows green square boxes that surround the part of the object that is going to be tracked and is set before every run.

Figure 4-10 The participant's face or complete body is selected for tracking

Figure 4-11 shows red square boxes indicating the parts of the participant that are being tracked with the mean shift program.

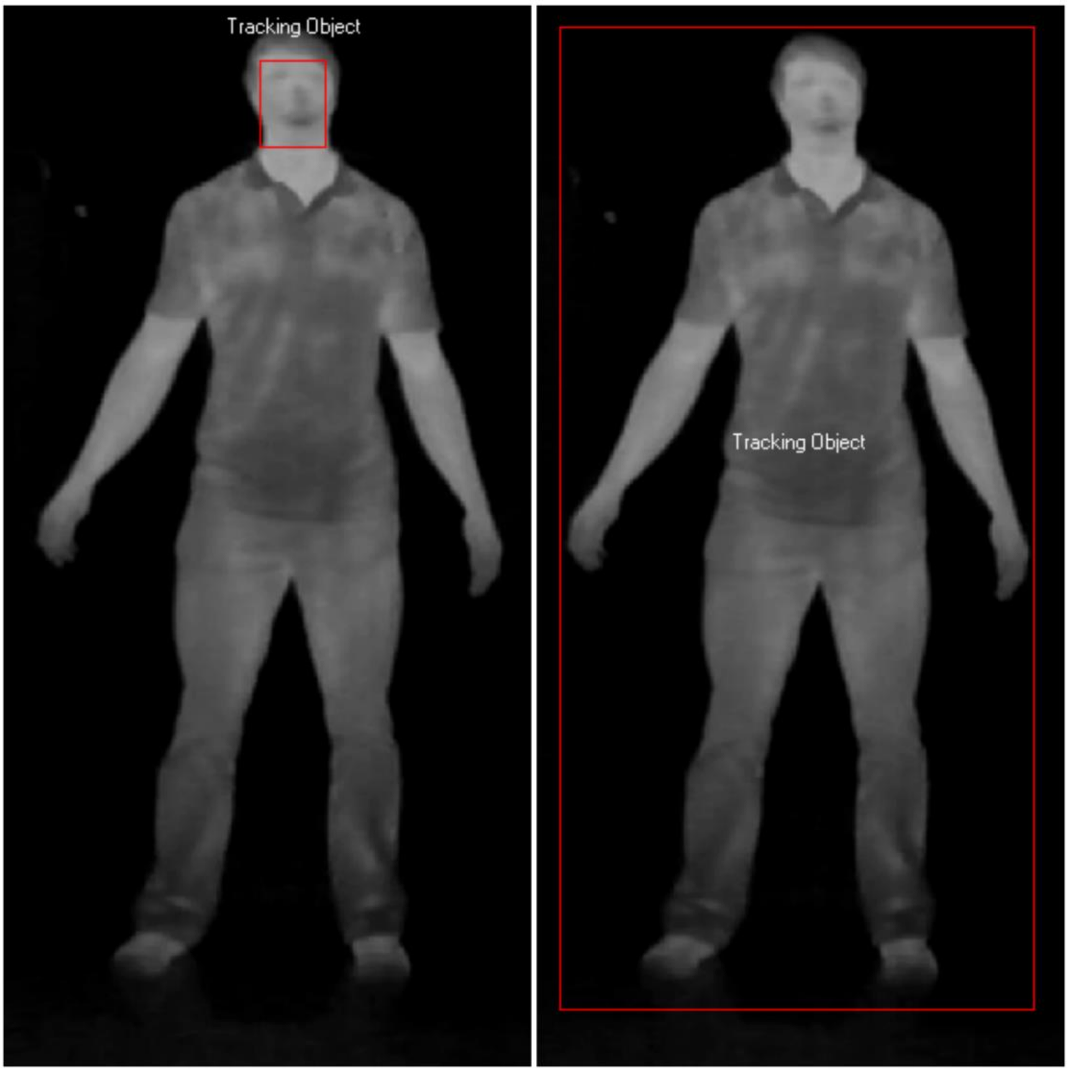


Figure 4-11 The participant's face or complete body is being tracked with the mean shift program

The results of Figure 4-9 show that Mean-shift tracking prefers 25-27°C temperatures on the calibration surface. The chart shows that the program had a +95% tracking record. Section 2.2.13 indicated that Mean Shift would prefer higher contrast image with more tracking points that give concise histograms.

4.2.2 Whole Body versus Face Tracking

A comparison between complete and partial body tracking was made at an exposure adjusted temperature of 25°C, as shown in Figure 4-12. The subject was recorded walking from the three metre mark to the 10 metre mark and asked to move a box out of the way at the nine metre mark. The subject is then also recorded walking back to the three metre mark.

Figure 4-12 Person recorded at 25°C in different stances performing work

This action simulates work that had to be done to move an object and allows analysing a subject performing a task to better test the tracking program in real-world conditions. The results are shown in Figure 4-13 in chart form. The y-axis shows the tracking score and measures the tracking certainty out of a 100. The x-axis is the program count and indicates the amount of program cycles for the analysed stretch of video.

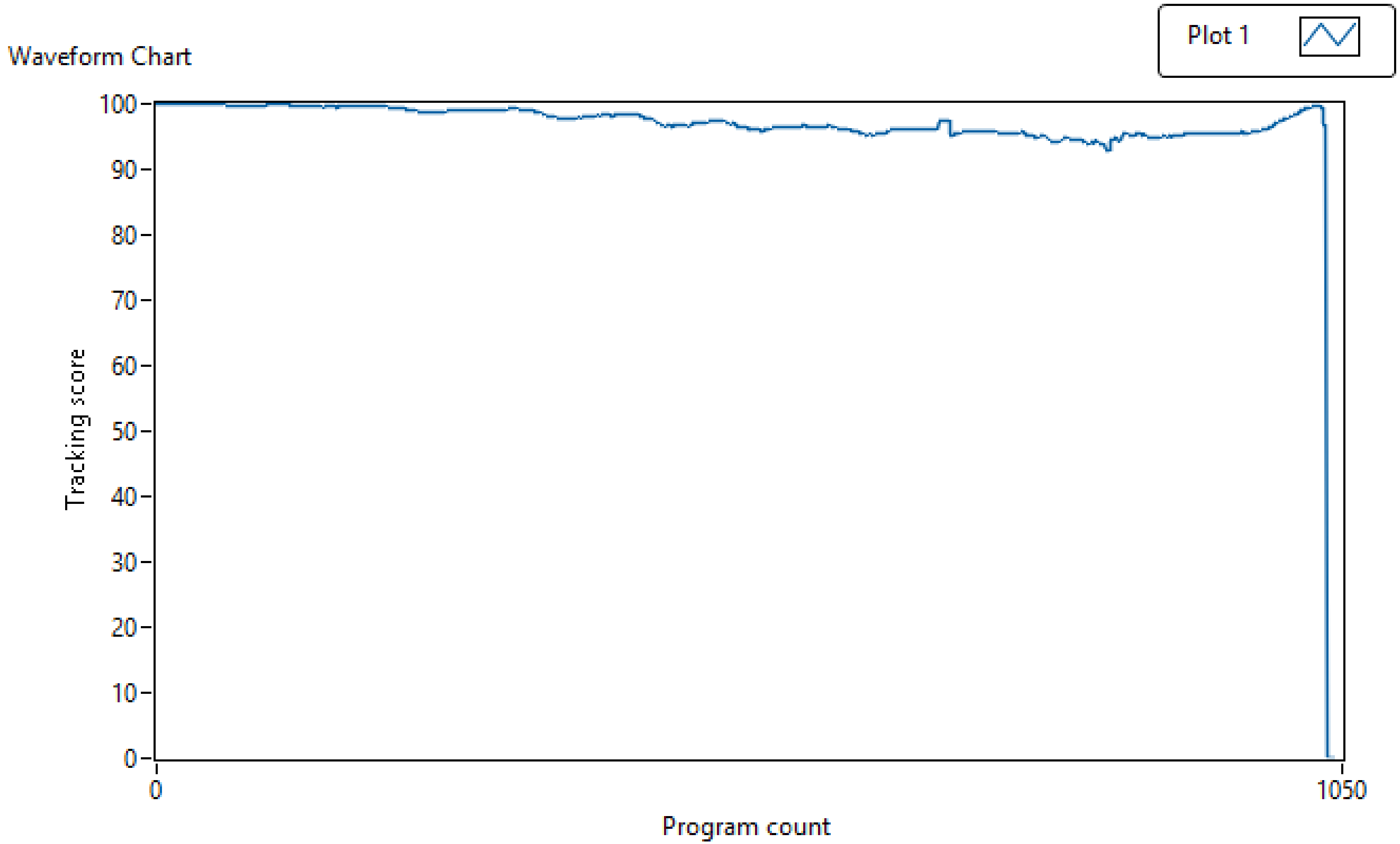


Figure 4-13 Complete body tracking at 25°C of person in different stances performing work

Figure 4-13 is a chart that plotted every program cycle's tracking score of Figure 4-12. It shows that even though the object that was being tracked moved away from the camera and changed stances along the way, the Mean Shift tracking method allowed for continuous tracking with a high tracking score. Figure 4-14 represents face tracking at 25°C and indicates the gradual loss in tracking certainty as the subject moves away from the camera.

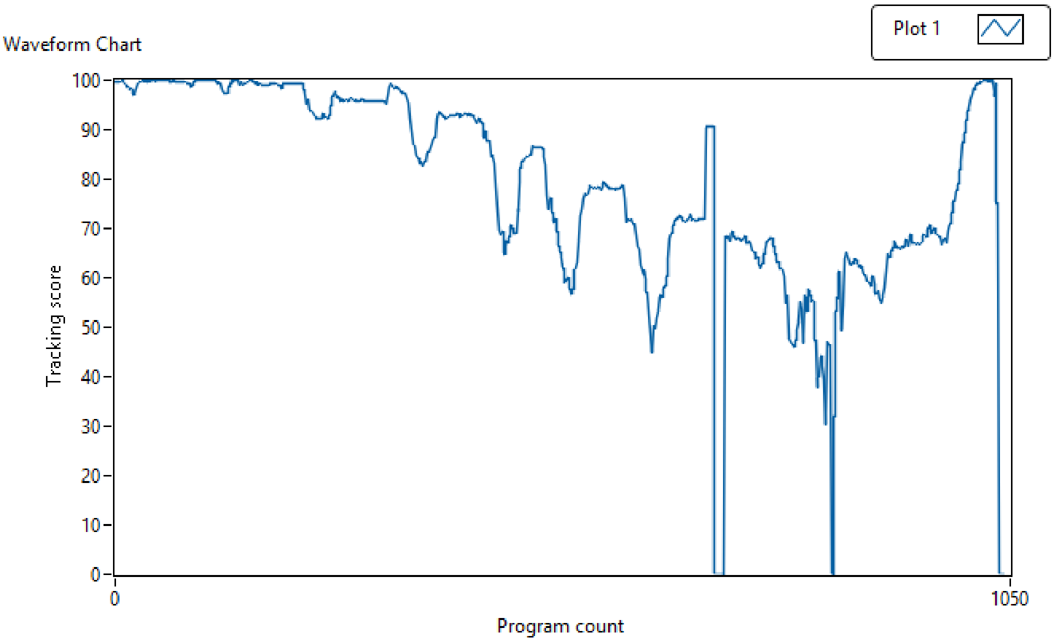


Figure 4-14 Face tracking at 25°C of Person in different stances performing work

Only tracking the subject's head allows the test to analyse what happens when the subject's face is no longer visible. Figure 4-14 shows sudden spikes where the program is unable to identify the object in the set tracking field. To compensate, the tracking program remembers the last known location of a good tracking field and waits for the subject's face to return and locks on to it again. If the participant were to move rapidly, or what can be considered fast for a camera capturing at nine frames per second, the program will lose its ability to track. An increase in the tracking score at the end of the run is due to the target moving closer to the camera at the end of the run.

In Figure 4-14 the tracking score starts high as is expected because of clear similarities between the first frames and better object's quality due to the subject's proximity to the camera. The score then decreases as the subject moves away from the camera and returns to high as the subjects are asked to move towards the camera again. When comparing this to complete body tracking in Figure 4-13, the tracking score remains constant and high for complete body tracking.

Figure 4-15 is a chart that shows tracking score over program cycles when tracking the face of a subject at a calibration temperature of 25°C. The participant moved from three metres to 10 metres.

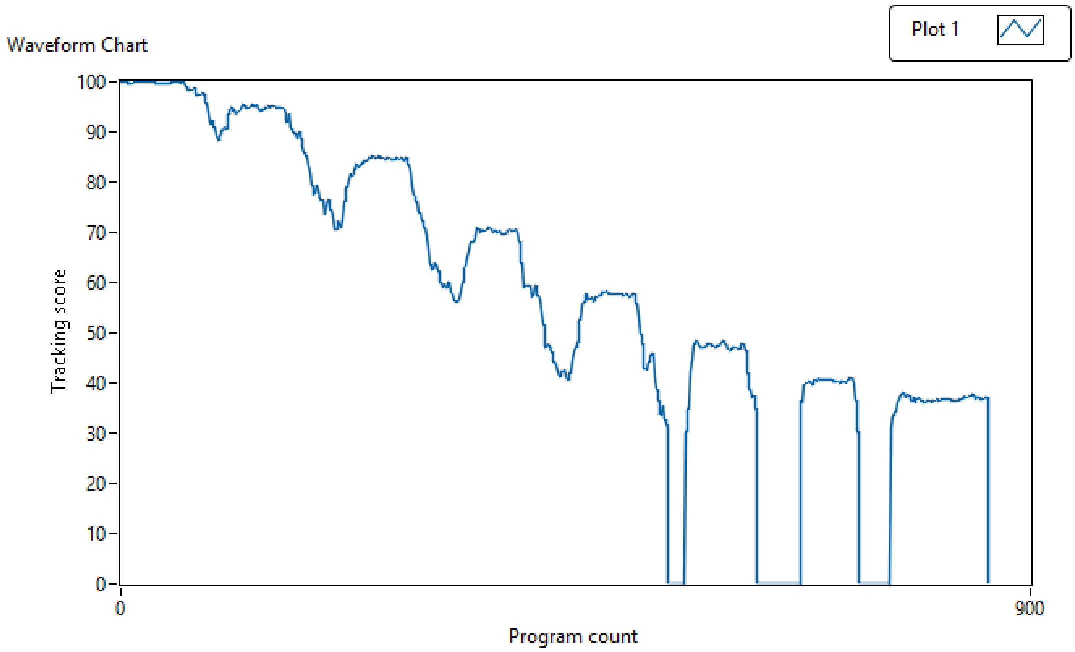


Figure 4-15 Face track at 25°C with increase in distance from camera

Compare this to Figure 4-16, which is the tracking of the same subject, but the complete body. The complete body tracking remains constant with face tracking varying based on distance and face visibility. Figure 4-15 shows sharp spikes in the tracking score where the participant turns his head away from the camera. The dips also clearly show a loss in tracking capability as distance increases. Both Figure 4-15 and Figure 4-16 show that the tracking stops abruptly in the middle of the run; this is not a loss in tracking, but rather the participant moved out of frame.

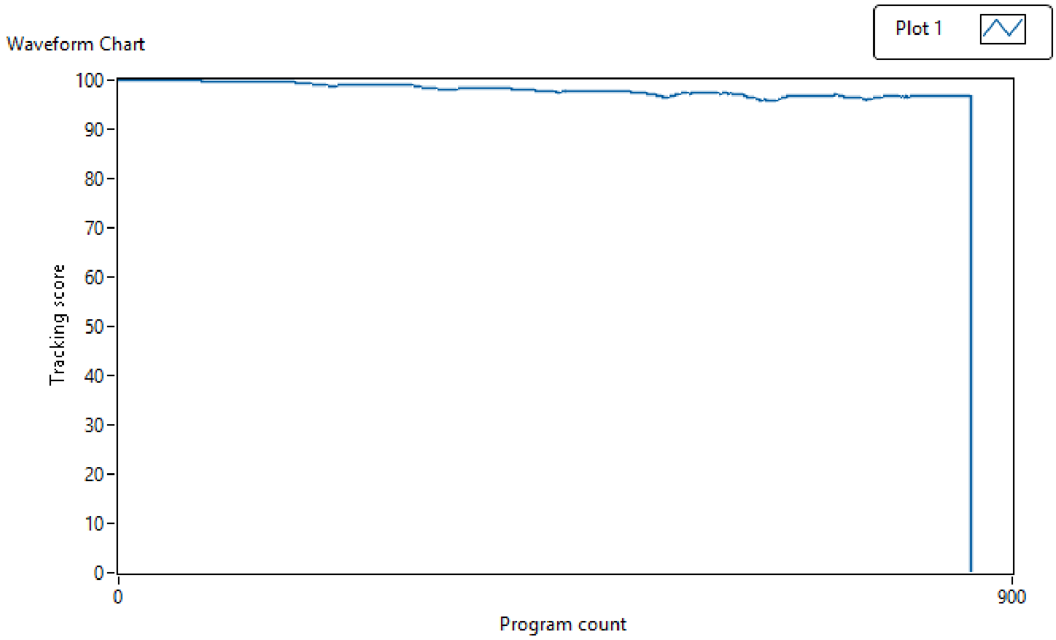


Figure 4-16 Complete body track at 25°C shows constant high tracking scores

Figure 4-16 indicates a very high tracking certainty and only shows a gradual decrease in the tracking score due to the increase in target distance from the camera.

4.2.3 Summary

To summarise, the results show the capability of the system to track a target over a distance by either complete or partial body tracking. In addition to human identification, the operation of the Mean Shift tracking was verified by using real-world video. Mean shift tracking can be run in real time because it is efficient and capable of tracking large or small objects.

Mean shift tracking is also not limited to thermal imaging videos and can work on full-colour video as well. The drawback in using this tracking system is that it does not scan the complete frame for the targeted object; this means that once target lock is lost, it is unlikely to regain it. This is, however, different for complete body tracking because the larger object means it is unlikely to lose the target. This method of tracking will, however, not work if the subject cannot be kept in the frame.

4.2.4 Mean shift Histograms

Figure 4-17 is a histogram that displays the number of pixels versus pixel value; the histogram shows how different areas of an image can pose a greater certainty of tracking based on pixel information.

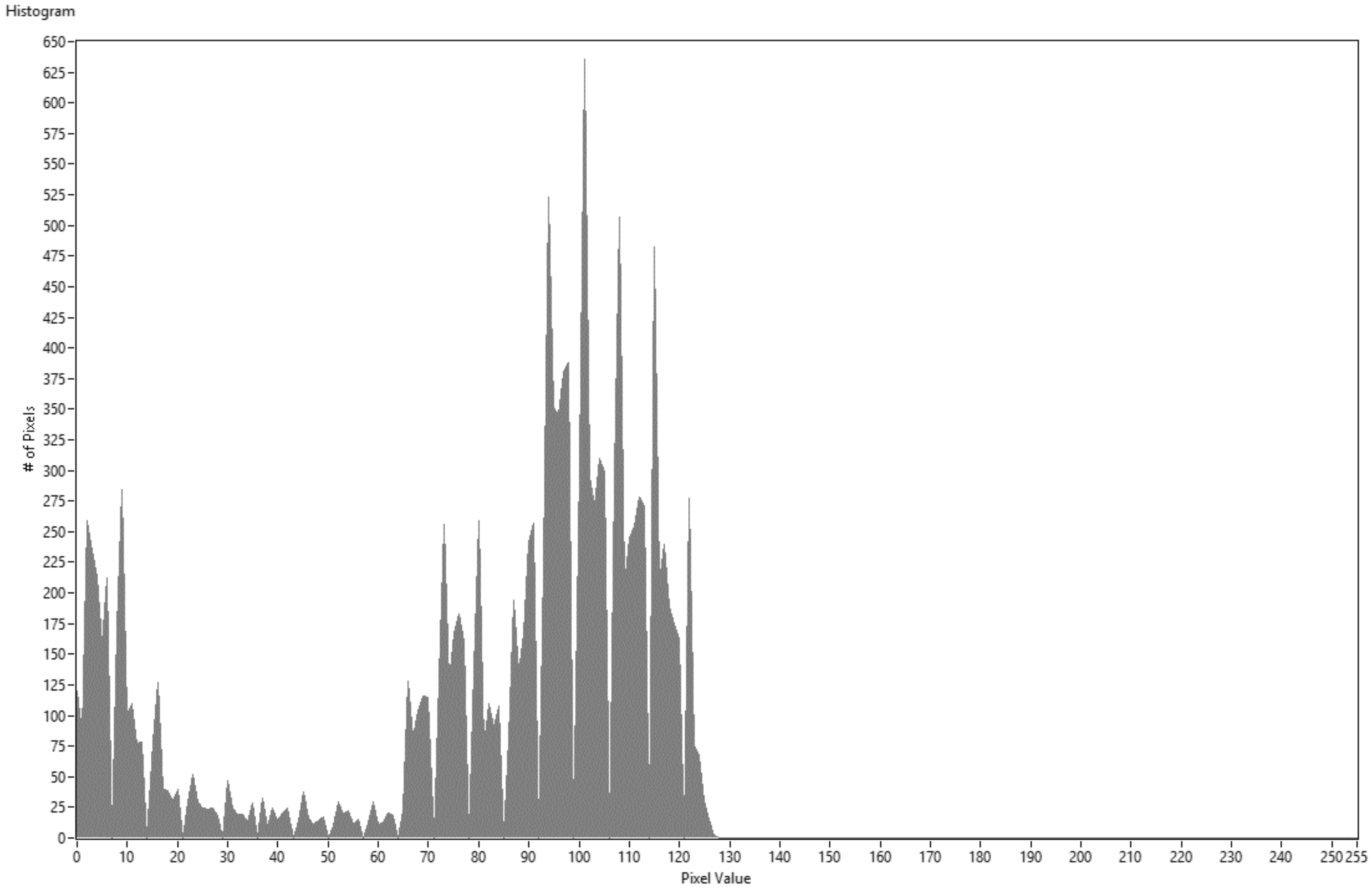


Figure 4-17 Histogram of the tracking area of a person's complete body at a calibration temperature of 25°C

Pixel intensity is clearly different depending on the tracking method. Complete body tracking in Figure 4-17 shows a higher intensity in the mid pixel values than face tracking in Figure 4-18, because the surface area is much larger, and the tracking certainty greater. This, however, does not rule out face tracking as a viable method, but rather as a less reliable form of tracking.

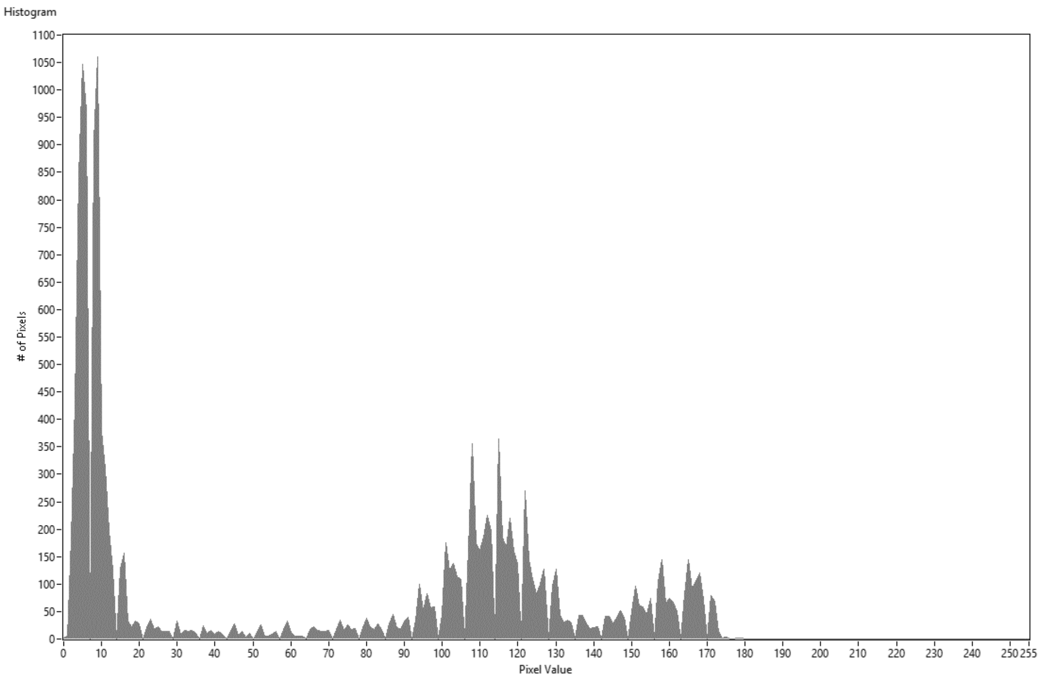


Figure 4-18 Histogram of the tracking area of a person's face at a calibration temperature of 25°C

Because Mean Shift tracking relies on comparing successive frames, having more information to compare helps produce better tracking results. To visually show how the histograms look that was described in section 3.6 which was used for Mean Shift tracking, two histogram extracts were taken from a frame in the same video but with different tracking methods.

CHAPTER 5:Contributions and Conclusion

This chapter summarises the project and glances over the research goals and objectives; it briefly touches on the contributions made and reveals future work to be done.

5.1 Summary

Chapter 1 introduced the project and placed it in perspective. It stated the problem, contained the hypothesis, research methodology and listed the objectives of the project.

Chapter 2 A literature study was conducted to gain preliminary knowledge on the various aspects of thermal imaging cameras. These aspects included characteristics of machine vision that require design and flexibility, typically considered hardware and software components and the use of LabVIEW software in this project.

Chapter 3 examines the methods used to generate results and solve the problem at hand; Here it is described how the AGV was constructed and used to carry the thermal imaging camera. It shows how software is used to accomplish identification and tracking in a vision-based system.

Chapter 4 shows how testing was done to achieve the objectives of the project. Test preparation, raw data gathering and analysis verifies that the system works as intended. The physical construction and real-time machine vision analyses are also shown in detail. Testing indicates that an AGV is capable of identification and tracking of a human utilising thermal signature.

5.2 Research Goals and Objectives

The main goals of the project were to develop an AGV equipped with a thermal imaging camera to track a human thermal signature. Also, this had to be achieved by creating and utilising image processing software.

This was met through firstly researching all the components that were needed to complete such a task in chapter 2. Afterwards, the physical system was developed and combined with software as shown in chapter 3 that enabled the system to be tested and generate results as specified in chapter 4.

5.3 Contributions

The project delivered the following contributions:

5.3.1 Efficient image processing

An efficient and compact method of identification and tracking of a human was developed as part of the main program. It consists of various image processing techniques, where Mean Shift Tracking was the main method of tracking a target. In addition to Mean Shift Tracking, software functions are developed to be modular so that they can be easily added or removed from the system software, or be enabled or disabled in a particular operational state. Identification and tracking are separated to allow rapid program changes to accommodate different cameras.

5.3.2 Single system

The AGV includes all components and their various subcomponents. The camera calibration system was assembled separately and mounted to the AGV once testing was completed. The rolling chassis with motors and drive were tested individually and then joined with the camera and calibration system. Obstacle detection with ultrasonic sensor and servo motor was mounted at the end of the AGV's assembly.

A Compact computer (slate computer) processes all sensor data and controls all the modules using a DAC unit. The system runs off battery power and can move and make decisions without external assistance. A single compact AGV capable of moving, identifying and tracking humans with frontal collision detection was developed.

5.3.3 Modular system

A modular approach was taken with hardware and software construction. Every module was tested and completed on its own, once all modules were assembled the AGV operated as a single system. Modular systems meant that once the tracking software was completed, it could be implemented immediately onto the AGV. To improve the AGV's tracking software, a version of the software was created to allow testing of tracking in an offline mode. This allows for modification to the software without having to reprogram the entire AGV; this can also be seen as a form of simulation before implementation.

5.4 Future work

5.4.1 Operator feedback

The system currently only displays information on the mobile computer attached to the AGV. However, methods can be implemented to allow remote operation and monitor of the AGV. Other methods can also notify the operator if the tracking of the subjects is not possible via sound notifications.

5.4.2 Upgrade camera

Consideration to further improve the system by upgrading the thermal imaging camera could increase performance. A better camera will allow for a higher frame rate and increase resolution. This could increase tracking performance and subject identification at a range. It might also enable the removal of external exposure control and reduce system complexity. Presently, the system will not operate as efficiently outdoors because heat affects the exposure of the camera severely. This might not be the case with a better-suited camera.

5.4.3 Better depth sense

The AGV is equipped with a sweeping ultrasonic sensor that covers only the front for obstacle detection. The current build is unable to determine an accurate distance of targets at range, and this remains a problem where the

AGV is required to follow a person. Stereoscopic vision or laser radar might be implemented in the future to allow for better awareness and localisation.

5.4.4 Obstacle avoidance

The current method of detecting frontal objects at ranges of three metres satisfies the needs of the project, but different methods of detecting surrounding obstacles and also avoiding them are worth looking into. Finding the best method and also building it into the vision software will significantly improve the capabilities of the project.

5.5 Conclusion

The aim of this study was to design an AGV capable of utilising thermal signatures to identify and track a human. The AGV should also be able to move based on tracking information it receives from the image processing unit. To accomplish this, a study was done on AGVs, and how they are used in industry to accomplish similar tasks; in addition, different vision-based systems were considered for human identification. Digital image processing and thermal imaging cameras were investigated to assess different types and methods of tracking a human.

An AGV was first designed as a CAD model to determine a solid and logical build; this approach saved vital time and allowed for better system integration of every component. LabVIEW offered a rapid platform for designing the AGV's control system and doubled as a machine vison platform.

A thermal imaging camera was selected to allow the AGV to see based on factors that clearly differentiates a human from a non-heated backdrop. Literature also supported the benefits of thermal signature as a means of detecting a person.

Exposure problems are common with machine vision applications, and the thermal imaging camera that was used did not allow manual exposure control, and that created difficulties. A solution in the form of camera exposure control was explored. Integrated with the vision-based tracking system, the external controlled heating element adjusts the exposure of the camera. Together with the integrated Axis camera software (which designates an exposure zone) the background noise was reduced significantly. Results obtained show that an optimal exposure can significantly benefit the machine vision process and reduce complicated software filters that impact performance.

The external camera exposure system together with Mean Shift tracking offers consistent and robust tracking results when using thermal imaging. Equipping an AGV with thermal imaging capabilities allows it to track a human's thermal signature and a method of Mean Shift image processing offers a viable solution to tracking. Mean Shift tracking is, however, not limited to thermal imaging as tracking source and will deliver excellent results when using conventional cameras.

Future research could allow the system to function in a variety of environments. Better hardware could enable more applications for the AGV and improve vision-based tracking using thermal imaging cameras. AGVs are being introduced into factories and need the capacity to work together and identify humans.

References

[1] Andr´e Treptow, Grzegorz Cielniak and Tom Duckett, "Active People Recognition using Thermal and Grey Images," *International Conference on Intelligent Robots and Systems,* no. 2153-0858, pp. 2103 - 2108, 2005.

[2] B. Trebilcock, "What is an AGV," *The Big Picture,* pp. 1-2, June 2011.

[3] M. P. Groover, Automation, Production Systems, and Computer-integrated Manufacturing, 3rd ed., Prentice Hall, 2008, pp. 285-286.

[4] H. P. M. Martin C. Martin, "Robot Evidence Grids," pp. 2-4,7-9,10-14,17-20, March 1996.

[5] N. Instuments, "NI," [Online]. Available: http://www.ni.com/pdf/products/us/20043762301101dlr.pdf. [Accessed May 2014].

[6] D. J. P. Steven F. Barrett, Microcontroller Programming and Interfacing Texas Instruments MSP430, Part 2 ed., Morgan & Claypool Publishers, 2011, pp. 176-179.

[7] National Instuments, "NI," [Online]. Available: http://www.ni.com/compactrio/whatis/#controller_chassis. [Accessed 28 November 2016].

[8] C. R. A. Valentinos K. Kongezos, "Wireless communication between A.G.V.'s and the industrial network CAN," *IEEE International Conference on Robotics and Automation,* pp. 434-437, 2002.

[9] C. G. Relf, Image Acquisition and Processing with LabVIEW, CRC PRESS, 2003, pp. 2,7,30,32,85-88,117-122.

[10] C. P. Maria Petrou, Image Processing: The Fundamentals, 2nd ed., John Wiley & Sons, 2010, pp. 377-378,767.

[11] H. He and L. P.Kondi, "Superresolution Color Image Reconstruction," in *Color Image Processing: Methods and Applications*, illustrated ed., K. N. P. Rastislav Lukac, Ed., CRC Press, 2006, p. 499.

[12] D. D. M. Bhabatosh Chanda, Digital Image Processing and Analysis, 2nd ed., PHI Learning Pvt. Ltd, 2004, p. 50.

[13] H. Frey, *Machine Vision,* University of Applied Sciences Ulm Departmnet of Computer Science, 2005, p. P46 chapter 5 Segmentation.

[14] Dr. Robert Fisher, Simon Perkins, Ashley Walker and Erik Wolfart, "Image Processing Lerning Resources," [Online]. Available: http://homepages.inf.ed.ac.uk/rbf/HIPR2/erode.htm. [Accessed 28 November 2016].

[15] T. B. Moeslund, Introduction to Video and Image Processing: Building Real Systems and Applications, Springer Science & Business Media, 2012, pp. 27-30.

[16] J. A. Amar Mitiche, Computer Vision Analysis of Image Motion by Variational Methods, Springer Science & Business Media, 2013, pp. 41-45.

[17] C. M. B. Dana Harry Ballard, Computer vision, Psychology Press, 1982, pp. 102,199.

[18] A. B. Don Murray, "Motion Tracking with an Active Camera," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 16, no. 5, pp. 449-459, May 1994.

[19] National Instuments, "NI," [Online]. Available: http://digital.ni.com/public.nsf/allkb/4F4947441856350B86256B3E0074FD39. [Accessed 06 Nov 2016].

[20] IMAQ Vision Concepts Manual, "NI," [Online]. Available: http://www.ni.com/pdf/manuals/322916a.pdf. [Accessed 06 November 2016].

[21] D. Comaniciu, V. Ramesh and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 24, no. 5, pp. 603-619, May 2002.

[22] D. Comaniciu, V. Ramesh and P. Meer, "Real-Time Tracking of Non-Rigid Objects using Mean Shift," *Computer Vision and Pattern Recognition,* pp. 673-678, 2000.

[23] K.-P. M. Michael Vollmer, Infrared Thermal Imaging: Fundamentals, Research and Applications, John Wiley & Sons, 2011, pp. 500-507,157-158.

[24] T. L. Williams, Thermal Imaging Cameras: Characteristics and Performance, CRC Press, 2009, pp. 4-5.

[25] FLIR Systems, "Flir," [Online]. Available: http://www.flirmedia.com/MMC/CVS/Appl_Stories/AS_0015_EN.pdf. [Accessed 19 November 2016].

[26] P. Douglas Stamps, LabVIEW for Data Acquisition, University of Evansville: SDC publications, 2012, p. 6.

[27] A. Communications, "www.axis.com," July 2012. [Online]. Available: http://www.axis.com/files/manuals/um_q192x_48466_en_1207.pdf. [Accessed 30 January 2016].

[28] R. Hirsch, Exploring Color Photography Fifth Edition: From Film to Pixels, 5th ed., Taylor & Francis, 2013, pp. 35-42.

[29] Microchip, "Microchip," [Online]. Available: http://ww1.microchip.com/downloads/en/DeviceDoc/21942e.pdf. [Accessed 20 November 2016].

[30] M. Steven B. Heymsfield, Human Body Composition, 2nd ed., vol. 918, P. Timothy G. Lohman, ZiMian Wang, Phd and Scott B. Going, PhD, Eds., Human Kinetics, 2005, p. 168.

[31] Avigilon, "Avigilon," [Online]. Available: http://avigilon.com/white-papers/adaptive-ir/. [Accessed 11 October 2015].