

Delft University of Technology
Master's Thesis in Embedded and Networked Systems

Indoor Localization Using Thermal Sensors

Kavya M Managundi



Indoor Localization Using Thermal Sensors

Master's Thesis in Embedded and Networked Systems

Embedded Software Section
Faculty of Electrical Engineering, Mathematics and Computer Science
Delft University of Technology
Mekelweg 4, 2628 CD Delft, The Netherlands

Kavya M Managundi
K.M.Managundi@student.tudelft.nl

22nd May 2019

Author

Kavya M Managundi (K.M.Managundi@student.tudelft.nl)

Title

Indoor Localization Using Thermal Sensors

MSc presentation

29th May 2019

Graduation Committee

Dr. ir. Alessandro Bozzon	Delft University of Technology
Dr. Venkatesha Prasad	Delft University of Technology
Dr. ir. Vijay S Rao	Cognizant Technology Solutions
Ir. Sujay Narayana	Delft University of Technology

Abstract

Locating people inside buildings is still an unsolved problem. There is a lot of research going on in this field and many different solutions using different techniques have been proposed. However, there is no widely accepted indoor localization solution like how GPS is for outdoor localization due to less accuracy, higher hardware requirement, cost etc,. We introduce a system that locates people indoors more accurately.

Preface

With the advancements in IOT devices, the medical care and location based services in assisted living for elderly are improving. However this improvement is at the cost of elderly privacy issues or inconvenience. As a result such services are avoided by the old people even though they are helpful. This problem motivated me to come up with a better system that can monitor people indoors without bothering them.

This work wouldn't have been possible without the help of my supervisors VP Sir, Vijay and Sujay. First of all I would like to thank VP Sir for giving me an opportunity to work on this topic. I would like to express my sincere gratitude to Sujay for his continuous help and support, for his motivation, enthusiasm, and also for tolerating all my repeated doubts and solving them every time. I can't thank Vijay enough for his patience, for hours of brain storming sessions, for constantly trying keep me on track and trying to get the best out of me. I would also like to thank fellow master students for all the laughs, ideas, coffee breaks and for helping me with experiments. Delft now feels like a home away from home and it's all because of friends, so I would like to extend my gratitude to them as well. Lastly, I would like to thank my Family not for one reason but for everything.

Kavya M Managundi

Delft, The Netherlands
22nd May 2019

Contents

Preface	v
1 Introduction	1
1.1 Problem statement and Challenges	3
1.2 Contribution	4
1.3 Report overview	5
2 Background and Related Work	7
2.1 Working principle of sensors	7
2.2 Literature survey	8
2.2.1 Detection	12
2.2.2 Tracking	13
2.2.3 Localization	14
3 The Hardware Platform	17
3.1 Hardware Platform	17
3.1.1 Selecting suitable sensors	18
3.1.2 Efficient processing of thermopile data	19
3.1.3 Variable gain setup for PIR sensor	23
3.1.4 Software	23
4 Localization Algorithm	25
4.1 Features for machine learning	25
4.1.1 Observations from thermopile array sensor	26
4.1.2 Observation from PIR sensor	28
4.1.3 Feature extraction	29
4.2 Localization and tracking steps	30
4.2.1 Human detection and horizontal location estimate	30
4.2.2 Distance estimate	33
4.2.3 Tracking	35
4.2.4 Height classification	36

5 Results	39
5.1 Performance Evaluation	39
5.1.1 Experimental Setup	39
5.1.2 Localization and tracking accuracy	40
5.1.3 Impact of obstacles	44
5.1.4 Height classification accuracy	44
5.1.5 Raw data compensation accuracy	45
5.1.6 Occupancy detection	46
6 Conclusions and Future Work	49
6.1 Conclusions	49
6.2 Future Work	49

Chapter 1

Introduction

With the miniaturization of computing and communicating devices we are witnessing abundant growth of innovative Location Based Services (LBS) both for indoor and outdoor environments. For an outdoor environment GPS has firmly established its standards in terms of precision and accuracy to provide user location. On the other hand in spite of a lot of research carried and techniques proposed to estimate location of a target indoors, there is still no indoor positioning technique or system that is widely accepted. Several reasons for such a set back being user acceptance in terms of convince and privacy, hardware requirement, ease of deployment, power consumption and cost of the system. The technologies employed today for presence detection and localization can be categorized as shown in Figure 1.1 into device-based and device-free, whether or not the user carries an identifiable device. Each of these categories are further classified into active and passive technologies, depending on whether an active or passive elements are employed.

In case of device based localization [19, 41], target has to carry devices such as smartphones, wearables, and tags etc., These systems provide an accuracy of up to centimeter level, however it is very inconvenient to carry a device everywhere all the time. Hence device-free systems and solutions are also very much in demand for its simplicity and convenience. In active device-free systems like Wifi[17], Radar [3] and Lidar [15], a transmitter is constantly emitting waves or pulses of light and changes in these waves due to the presence of target are measure to find the location of target. It is evident that these systems are power hungry and cannot be employed for domestic services. However, Passive device free systems are low power systems, because they use data like changes in temperature or pressure due to the presence of target to estimate the location.

From figure 1.1 we can see that passive device-free system has in general four different techniques. Camera, is a vision based passive localization technique which is used to identify and track multiple people, it is very popular in 3d games. But, camera is dependent on illumination and most importantly it is

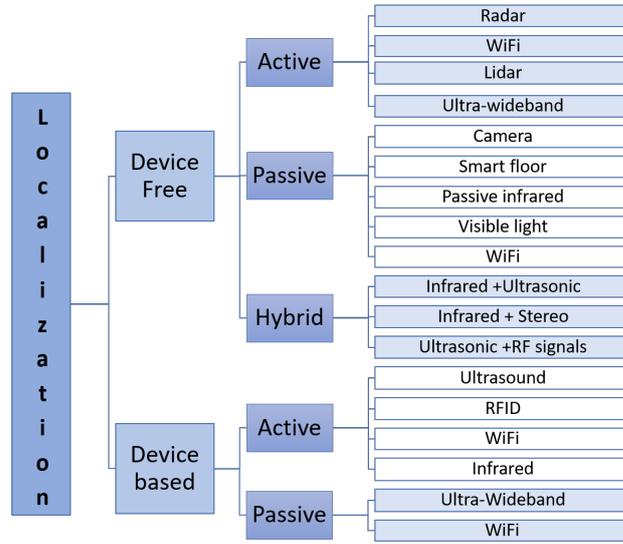


Figure 1.1: Categories of localization systems.

obtrusive. Smart floors [4] or also known as pressure sensitive tiles rely on foot pressure of the person to detect the presence or measure their location. The deployment of these smart floors is not trivial and since the whole of the tracking area should be covered with these sensors, it is very expensive. Visible lights, like conventional LEDs which are readily available can also be used to estimate indoor location of a target, with slight variations in the hardware. The accuracy of visible light communication (VLC) based localization systems [27] depend on how close the lights are placed and the median error increases as the distance between these lights increases. All these limitations are hindering the acceptance of such systems by the public. Hence there is a need for a system which is convenient, unobtrusive, inexpensive, low power consuming and are easy to install. Since humans emit thermal radiations which can be captured by passive infrared sensors and these sensors are less power consuming, non intrusive and also cost efficient they are perfect to be practically installed and used in indoor environments. Therefore, in this work, we focus only on thermopile and PIR based systems.

In literature, we can find several techniques, algorithms and systems that have been developed using infrared sensors like thermopile arrays and/or PIRs for localization [34, 32, 1]. However, they have one or more of the following issues:

- **Inflexible.** Many systems/algorithms promise accuracy in tens of centimeters in their *specified deployment scenario*, which is a key point to be noted, even when the deployment area is indoors. Some works create zones that may differ from room to room; some would require multiple towers placed at pre-specified angles and so on. Furthermore, with thermopile arrays, many works choose to deploy their module on the ceiling in order to avoid quasi-stationary

background noise due to lights or hot objects in the Field of View (FoV) of the sensor. These solutions do not work when deployed in non-ideal conditions. These make the systems inflexible.

- **Unscalable.** Many proposed systems require lot of sensors to guarantee localization with sufficient accuracy over a larger coverage area. Most of the work consider the sensor has a fixed range and cannot be dynamically tuned. Moreover, the deployment is not easy and it should be properly planned depending on the rooms.
- **Low accuracy.** Existing PIR based localization techniques fail when the signals get weaker and cannot accurately distinguish between a person walking slow at a farther distance and a person walking fast at a closer distance to the PIR.

These issues leave a gap in the literature and mandate newer systemic methods to address them, thus leads to our problem statement.

1.1 Problem statement and Challenges

Develop and implement localization algorithm using passive infrared sensors, to achieve high accuracy, in high coverage area, with reduced infrastructure, cost and power consumption.



Figure 1.2: Our proposed hardware platform that houses a thermopile sensor and a PIR sensor.

In this work, we propose and build a new system, consisting of both hardware and software modules. The hardware platform, as shown in Figure 1.2, consists of a Melexis 32×24 thermopile array and a PIR sensor. We use the analog waveform from the PIR sensor instead of simple ON/OFF signal. The novelty of our system is the gain of the PIR sensor can be dynamically controlled through software. By combining both the sensors into one platform, we will be able to eliminate noise in the background more effectively, eliminate the need for multiple PIR sensors, and bring in more flexibility for deployment. Our platform

can be used for many applications – to compute the height of a person, localize, track the movement, recognize some activity (such as fall detection), and count the number of people (upto five people with a reasonable accuracy). Since the breadth of applications is huge, we focus only on *computing the height and localization of a person* as localization is the basis for tracking and better activity recognition and better height estimation is necessary for fall detection and also gives an idea of whom you are tracking.

Challenges.

1. To increase the ease of deployment, we plan to deploy only one system for the entire room. Thus, it is significantly difficult to achieve a better localization accuracy in three dimensions even with our fusion of sensed data from thermopiles and PIR.
2. With a PIR sensor, it is extremely difficult to discern the fast moving person at a nearer distance from the sensor, from the slow moving person at farther distance from the sensor. This is because the sum total of incident heat energy falling on PIR is the same.
3. Low-power and real-time computation – thermopile requires heavy floating point computation for temperature measurements. This is compounded because total number of pixels to be processed is higher. The challenge is to directly use the thermopile output.
4. To use the low power computing platform the sampling rate needs to be low. However, the location of the person should be computed when the person is present at that spot rather than when (s)he is gone. We observed this in [34] that negates the whole purpose.
5. Thermopile array sensor is extremely sensitive even to slight disturbances in ambient heat. Thus removing the background hot objects and ambient heat noise is mandatory for effective localization and height estimation.

1.2 Contribution

Apart from the design novelty of our detection system, we propose many innovative solutions. Specifically our major contributions are as follows:

1. We use machine learning based classifier on peak to peak voltage Vs gain curve to find the distance of the target from sensor, unlike in the literature where only a single peak to peak output value is used, hence, our algorithm is agnostic to the clothing worn (e.g., jackets, full-clothing, and semi-clad).
2. This is the one of the first works to achieve joint height classification and localization using two different types of IR sensors in a miniaturized low power and low form factor and high FoV system that can be easily mounted.

3. The novelty is in the use of variable gain at PIR sensor to avail the spatial diversity gain that helps in avoiding multiple PIR sensors (like in [34]) without complex deployment.
4. We use raw values from thermopiles unlike other works. This enables the usage of low power MCU as processing is less stressful.
5. Since the gain control is done in software, our system works in real-time to localize moving persons with the accuracy of less than 35cm 80% of the times.

1.3 Report overview

The rest of the report is organized as follows: Chapter 2 describes the relevant literature and state of the art. Chapter 3 outlines the development of our hardware platform followed by its characterization. Subsequently, we describe our algorithm in Chapter 4, which is evaluated in Chapter 5. Finally, we conclude and discuss the future work in Chapter 6.

Chapter 2

Background and Related Work

2.1 Working principle of sensors

A thermopile sensor consists of many thermocouples that are connected in series on a silicon substrate to form hot and cold junctions attached to very thin IR detectors as shown in Figure 2.1. The incident infrared energy on this material makes thermocouples to generate a potential proportional to the incoming energy because of the *Seebeck effect*. Hence, thermopile sensors are used for non-contact temperature measurement. A thermopile array contains multiple thermopile sensors arranged in the form of a grid, such as MLX90640 array sensor having 768 pixels.

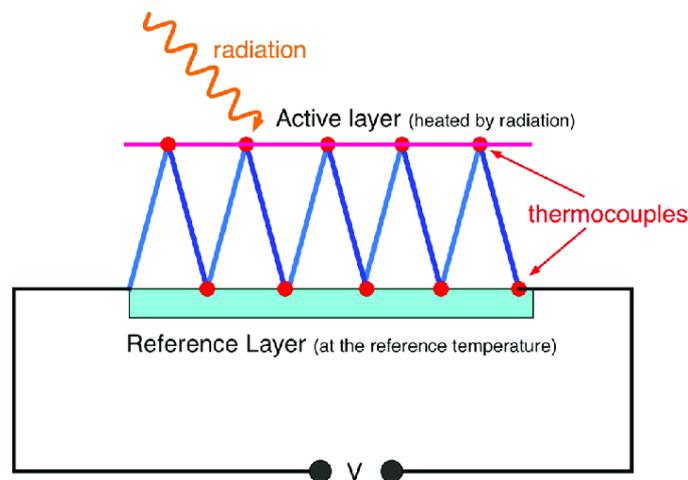


Figure 2.1: Working of thermopile infrared sensor [39]

Unlike thermopiles, PIR sensors work on the theory of pyroelectric effect. The sensor is made up of two halves as shown in Figure 2.2 and are wired up so that they cancel each other out. If one half sees more or less IR radiation than the other, the output will swing high or low [?]. This generates a voltage proportional

to the differential change in the generated charge on these elements when IR energy is incident on them. When the sensor is idle, both elements detect the same amount of incident energy from the ambient temperature. When a warm body (e.g., human) passes across in front of the sensor, it first intercepts one of the elements of the PIR sensor, causing a positive differential change between the two elements. When the warm body leaves the sensing area, the sensor reports a negative differential change.

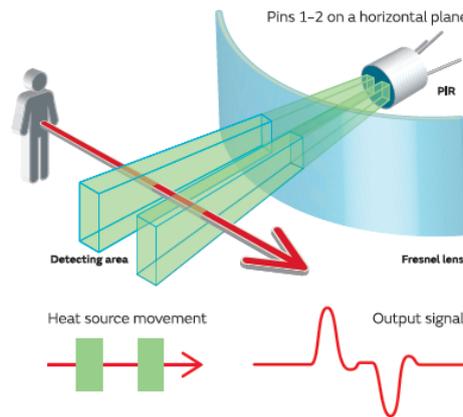


Figure 2.2: Working of PIR sensor [9].

While thermopile can detect both stationary and moving warm objects, PIR can detect only when the objects move. Hence, we use a fusion of these sensors to identify humans from the rest of the background warm objects. Additionally, there is a flexibility of adjusting the detection range of the PIR using a variable gain scheme that is not possible with the commercially available thermopile sensor arrays.

2.2 Literature survey

Localization has been a very active area of research for decades now. Algorithms and systems using both active and passive techniques, with device-based and device-free localization have been proposed. Our work targets the usage of passive and device-free type of technique. In this section the works related to device-free passive localization using low-power thermal sensors are mentioned. The state of the art techniques that are relevant and comparable to our work is listed in Table 2.1.

Narayana et al. presented a novel sensor tower containing four collocated PIR sensors that can perform height classification and localization of moving objects [34]. Sensor tower was designed with two spot lens at the top and two multiple lens PIR sensors at the bottom. In order to amplify the output of multiple

Table 2.1: State of the art that use thermopiles and PIR sensors for human detection, tracking, and localization. Our work outperforms the rest in terms of the coverage area, objective, accuracy, and deployment.

Work	Type of sensors	Coverage area	Objective	No. of people	Accuracy	Notes on deployment
Narayana et al. [34]	8 PIRs	8m x 8m	Localization + Height classification	1	30 cm	Two sensor towers, having 4 sensors each, are placed at 90° (IPSN '15)
B. Mukhopadhyay et al. [32]	4 PIRs	7 m x 7.5 m	Localization	1	65 cm	4 PIRs placed on the edge of the experimental arena (WCNC '18)
W. Chen et al. [7]	2 Thermopiles	2.35 m x 3 m	Fall detection + Localization	1	13.39 cm ¹	Two sensors attached to wall at 30° (ICENAS '15)
C. Basu et al. [1]	1 Thermopile	2.5 m x 2.5 m	Occupancy detection + Tracking	5 Occupancy, 1 tracking	80%	Ceiling mounted (arXiv '15)
Z. Chen et al. [8]	1 Thermopile	2.4 m x 1.2 m	Tracking	1	19 cm ²	Thermopile placed on a motor across the wall (IEEE Sensors '18)
J. Kemper et al. [22]	4 Thermopiles	4.9 m x 6.2 m	Localization + Tracking	2	25 cm	4 sensors placed in 4 corners (WPNIC '10)
M. N. Hock [35]	5 Thermopiles	4.6 m x 2.7 m	Human detection	1	50 cm	2 sensors on one side and 3 on adjacent side (IPSN '13)
C. Shih [40]	4 Thermopiles	3 m x 4.25 m	Tracking	1	87%	Ceiling mounted (RACS'17)
This work	1 Thermopile + 1 PIR	9 m x 8 m	Localization + Tracking + Height classification	1	12.5 cm ³	Single platform mounted on ceiling or on the wall

*^{1,2} Proper comparison is not possible with the available information, this value is presumably the best case. *³ Best case is 12.5 cm. 50% < 22 cm and 80% < 35 cm.

lens PIR sensor, a two stage amplifier is used and the gain of each stage is adjusted using a potentiometer placed at the feedback of each stage. Frequency and amplitude of the analog signal are the two features that have been made use to characterize the movement by object like speed, direction and distance. Making use of the sensor placements in the tower the objects were classified based on width and height of the object. The authors estimated the range between humans and the sensor tower by fixing different gains for different PIR sensors to form various detection zones. Two such towers placed spatially apart at 90° localizes the moving warm object.

Mukhopadhyay et al. improved the work in [34] by reducing the number of PIR sensors to one on each tower but by placing four such PIR sensor systems to form a square inside which localization can be performed [32]. Their algorithm exploits the peak to peak value of analog output of PIR sensor and its relation with distance to propose the distance estimation model. Distance of a person from every node is estimated either by using a hyperbolic function based model (in which the relationship between distance and peak-to-peak output voltage of PIR sensor is hyperbolic in nature) or by using piece wise linear model. This estimated distance from every sensor node is used in multi-lateration and Support Vector Regression (SVR) based techniques are used to compute the location co-ordinates of a person.

Chen et al. make use of two 16×4 thermopile array sensors, placed at 30° , and 3.3 m away to capture 3D image with which tracking elderly and fall detection is done [7]. Human detection is done initially by subtracting averaged 3 background frames. However, to update the background over the period of time, mean and variance of frame is calculated and compared with a fixed threshold value. Once the human is detected, the location of the human is obtained using the angle of arrival (AOA) from each sensor. With AOA of both the sensors, the position is estimated by calculating the angle of interception further to reduce the position error two quadratic regression models to describe the correction in x and y direction.

Occupancy detection and tracking of people within an FoV of $2.5 \text{ m} \times 2.5 \text{ m}$ is reported in [1]. The main issue that is addressed in this work is the wrong estimate of number of people who are in close proximity, when connected component labelling technique is directly applied on the background subtracted image. Since the head and chest region of human body emits more thermal radiation than the other parts of body, the local excitation peaks carry more information along with the connected component features. It is assumed that the temperature or intensity recorded at the sensor array is proportional to the height of the person and hence a sliding window of size 4 pixels is used to count people of different height accurately. Support Vector Machine(SVM) classification is used to count the number of people in the FOV. Results show that a maximum of 4 people with 80% accuracy can be counted. 4D features instead of 3D for clustering gives a remarkable accuracy however this work fails to recognize that the temperature intensity depends on the height, size and distance of the person

from the sensor. Moreover this work fails to adapt the background in real time.

Z. Chen et al. propose an activity recognition and a face direction tracking system by using a low pixel infrared thermopile array and a Time-of-flight sensor [8]. As a pre-processing step the 8×8 data matrix is interpolated to a 32×32 matrix and then binary image is extracted by using an adaptive threshold to extract the human detected information. Adaptive threshold is calculated by sorting the maximum temperature difference of every column and using the minimum value of these maximum temperature differences. Features extraction is done either by convolution neural network (CNN) or manually by defining the features. These features are fed to support vector machine (SVM) for the direction detection. A servo motor is used to rotate the sensor module depending on the distance of the heat center and the angle received from the feedback of the servo motor. The work has produced a face tracking accuracy of above 80% with the root mean square error of 0.19m. However this work uses the active TOF sensor to detect the distance between the sensor and a person.

An implementation of Probability Hypothesis Density (PHD) filter is presented by J. Kemper et al. to perform localization and tracking of multiple people using 4 thermopile array sensors placed in corners of a room[22]. A PHD filter estimate an unknown, time varying number of targets and their states from noisy observations available at discrete intervals of time [36]. PHD uses a set of weighted particles to approximate the number of estimated targets. As an initial step location of target in x and y co-ordinates is estimated by using angle of arrival(AOA) of the highest temperature pixels of all the 4 sensors or by using triangulation. Further, multiple people localization is carried out by particle implementation of PHD which is based on motion and sensor model. Motion model is based on target's former movement and Gaussian white noise, this motion model gives an estimate about the target velocity in sensor model error correction of measured location using AOA is performed.

A method of using a network of thermopile sensors distributed along the walls of a room to locate a person within the room is studied by M. Hock [35]. In total, 5 thermopile sensors are required to localize a person in $4.6 \text{ m} \times 2.7 \text{ m}$ area. The data set is fit to an equation where sensed temperature is inversely proportional to square of distance of a human from sensor. When a person enters the FOV of any of these sensors, based on which sensor senses maximum temperature the vertical and horizontal distances are estimated with an accuracy of 50cm. The body temperature and area of sensed object are assumed to be constant in this work which is not a real time scenario. Additionally other warm objects are not taken into account which is a major drawback of this system.

C. Shih et al. designed a ceiling mounted thermopile array sensor network to track gait of a moving person [40]. The tracking algorithm is developed on a virtual run-time library called WuKong[23]. A thermopile array deployed on the ceiling is used to monitor activity index in home environment. An actual human target is detected by computing z-scores for every pixel followed by z transform to reduce the effect of temperature changes. z-score is calculated using mov-

ing average and standard deviation of a data point in a particular time interval. Later human is identified if this z-score is greater than the threshold. Then gait velocity is calculated by tracking the human target.

A brief summary on rest of the significant relevant works on detection, localization and tracking are given in the following sections.

2.2.1 Detection

A. Beltran et al. designed a battery operated node called 'ThermoSense' using a combination of both PIR and Grid eye sensors to determine the occupancy count of the room [2]. ThermoSense was mounted on a ceiling and the digital output of PIR sensor was mainly used to determine whether the room was occupied or not then a Grideye sensor is used to count how many people are present in the FOV. Whenever PIR indicated the room as empty, the Grideye output at that time was saved as a reference thermal background map in order to differentiate humans from other warm objects. A 15 minute window is used to update the background based on an assumption that it is not possible for occupants to stay stationary for more than 15 minutes. Further to update the background when there are occupants in a space for a long time, a new background is formed by multiplying a scaling parameter to all the current background pixels. The scaling parameter is formed by using the minimum temperature values of all pixels in the current frame and dividing them by previous background pixel values. Further features like the number of warm pixels, size of all connected warm pixels, size of the largest connect component are used in the supervised machine learning based classification algorithms like K nearest neighbor, Artificial neural networks and linear regression to count the number of people in a room.

A low resolution 4 x 16 thermopile array sensor is mounted on the ceiling of a room at an angle to estimate occupancy in the area of 3m x 0.7m, where upto 3 people can be present [43]. The authors make use of number of pixels, connected components and size of the largest connected component as features in different classification algorithms to count people in a rectangular space. Human detection is performed by image subtraction as performed in [2]. Further features of this human detected pixels are extracted and used in KStar (K) algorithm which is similar to k-nearest neighbours (KNN) but uses entropic distance instead of euclidean distance to classify the number of people. The classification accuracy of 82% was achieved using this system.

Human detection using thermal sensors outdoor is very complicated because of the varying environment which influences the difference in human and background temperatures. To solve this issue different computer vision techniques like gradient difference, cross correlation and frame difference etc., were implemented in [6] using Grideye sensor. These methods were checked for performance, time taken and complexity. It is noted that the human detection by a simple frame to frame subtraction has real time performance but it cannot identify static targets. On the other hand, static humans could be identified us-

ing a background model to update the background over time but this was very complicated to realize in real time. Furthermore, the authors show that a human can be detected in a changing environment with 97% accuracy if an optimal threshold is used instead of a fixed threshold.

S. Mashiyama et al. mount Grid-eye thermopile sensors on ceiling to classify human activities such as walking, sitting, and falling [28]. The algorithm is implemented in three stages: human detection, feature extraction and classification. In human detection stage, a time window is fixed and a threshold temperature value is used to identify human targets. Later, features such as active frame, maximum number of reacting pixels, maximum temperature variance, moving distance are used in SVM classification algorithm for activity detection. The accuracy for all the activities other than sitting is above 94% in the experimental setting. However all the thresholds are decided in preliminary experiments and are a constant which doesn't adapt to the changing environmental conditions. Most of the features focus on fall detection; however to achieve more accuracy for sitting and walking other features like temperature variation and pixel cell variation can be considered.

Single 8 x 8 pixel thermopile array sensor is mounted on the ceiling of office pantry for recognizing objects and human-object interactions to monitor the energy related activities in the office in [13]. Pixels with temperature higher than the ambient temperature and the adjacent pixels with same temperature values are grouped as one object. A prior knowledge on static object's location is used to identify the usage of objects in the FOV. Each object is classified into different states for example refrigerator is either open or close, faucet is off, hot or cold etc. And the interactions are either static-dynamic or dynamic-dynamic. In static-dynamic static object is taken as reference and in dynamic-dynamic either of the one is considered as reference. An activity interest list is employed which contains all possible interactions with the reference object. The current activity is determined by selecting the highest ranked activity from all the detected interactions for that particular reference object. A total of 8 features were used to determine the state and interaction class. Hidden Markov model (HMM) was used to implement the state filtering. While static-dynamic interactions were classified with the accuracy higher than that of the dynamic-dynamic interactions like meeting of two persons.

2.2.2 Tracking

A thermo-spatial and conventional histogram based technique to track humans in complex environments is discussed by H. Takashi et al [20]. The temperature of target is calculated by using the median of the pixels in a bounding box. Thermo-Spatial histogram is the combination of conventional histogram which preserves the spatial information and thermal histogram which reduces the influence of background pixels. The initial location of human is given as input to the system, then the tracking in case of different occlusion scenarios like black

board, chair, human, doll and lockers are evaluated. A target human region in the current frames is detected by checking the similarity between input template and the human detected region using the histogram intersection function. Although the tracking is achieved with improved results over spatial histogram and conventional histogram, this work fails to provide the exact location estimate or the tracking results like x,y co-ordinates and the distance at which the target is from the sensor.

M. Kuki et al. use a ceiling mount 4 x 4 grid thermopile array sensor to record human movement trajectory [24]. The authors implement fuzzy rules and Connected Component Labeling (CCL) to identify and track humans. Multi-human trajectory extraction was done by attaching 16 x 16 pixel thermopile array sensor to the ceiling [25]. The trajectories are estimated by using the Mahalanobis distance algorithm. Q. Hao et al. build a wireless distributed PIR sensor system that houses 8 PIR sensors to track a single human target [16]. Kalman filter is mainly used to estimate the trajectory.

2.2.3 Localization

S. Lee et al. present a location-recognition system called PIR sensor-based indoor location-aware system (PILAS) [26]. Different sensing areas of 2m size are created using 12 PIR sensors which are placed on the ceiling of a room of size 4m x 4m and at a height of 2.5m. A threshold was set to turn the PIR sensor on or off depending on human movement. The system recognizes the resident's location by combining outputs from all the sensors belonging to one area with the accuracy of 0.5m. Since the output of the PIR depends on the incident thermal energy setting up a hard threshold to turn the sensors on and off becomes a major drawback of the system as different humans have different body temperatures and setting up lower threshold may consider noise can humans and higher threshold may not detect low temperature human signals as noise.

Using a ceiling mount 8 x 8 array thermopile sensor, D. Qu et al. present a system to perform Multiple human localization and tracking [38]. As a first step interpolation is performed to increase the resolution of the thermal image, a Gaussian filter is applied to this interpolated image to remove noise. Further, a threshold which adapts to the background changes is calculated using the mean and max temperatures of every frame. This adaptive threshold is used for human detection. Relative location of human is converted to real locations by finding the cell numbers and calculating the angle of view of these cells. Then the tracking is done using Kalman filter. Multiple human trajectories are considered and the results of tracking are shown for the same. Even though this work has addressed the main issues like pixel deviation error introduced while manufacturing silicon chips and the false detection of human due to fixed threshold the main drawback of the system is that the height of the roof differs and hence the pixel size, due to which the error is introduced in the initial step of the algorithm as mentioned earlier blind stops are produced by mounting sensor on the ceil-

ing, requiring more number of sensors to cover small area.

T. Yang et al. introduce a concept of 'azimuth' change that adopt particle filter to solve the issue of abundant training data collection in a PIR based system for localization. The system involves 4 PIR sensors located at corners of a 7 m x 7 m area [44]. The concept of fresnel lens dividing the PIR output into different zones is exploited here. The azimuth change of a moving person is defined as an angle between the targets previous location and the current location. This angle is calculated by counting the number of rising and falling edges of the output of a PIR sensor. However only the higher peak to peak voltage values are considered for processing and lower peak to peak voltages are regarded as noise but it becomes difficult to distinguish between the actual signal and noise when a person is moving at a farther distance from the sensor. To avoid this problem of false interpretation, difference between heat fluxes (DHF) of the two slots of a PIR is calculated and is used to get better estimate of number of zones covered(N). Further a particle filter is used for localization with 0.63m accuracy.

The analog output of a PIR sensor in conjunction with signal processing and machine learning algorithms is used by H. Gami to estimate the presence, direction, and distance of the human movements [12] in 2.7 m x 4.6 m area. Features used in machine learning are Peak to peak voltage value to estimate the distance and speed of movement, phase of PIR output to estimate the direction. Since the data is labelled basic supervised machine learning algorithms like Tree, KNN, Support Vector Machine (SVM) and Artificial Neural Network (ANN) are used to check the accuracy of the system. 93% of accuracy is observed using single hidden layer based ANN algorithm for distance estimate and 99% accuracy for direction estimate. However since only peak to peak voltage is used to find the distance of a person from the sensor it is prone to change with different thermal insulation or different body temperatures.

Limitations of previous work and our improvements

Of the works listed above, Narayana et al. [34], B. Mukhopadhyay et al. [32], and W. Chen et al. [7] are the state of the work (as listed in Table 2.1) done on localization, tracking, and height classification that is the closest to our work.

Deployment: While these works requires to deploy multiple sensor platforms on a region of interest (as high as 4 platforms on corners of the square region) to localize people, our approach requires only one platform to perform multiple tasks. Further, our system is portable, battery operated, and wireless, connecting all the platforms to a centralized server so that tracking can be performed at room level, as well as the building level. The system provides the flexibility of mounting on ceiling or on walls.

Noise cancellation: To the best of our knowledge, none of the above works have considered factors such as types of clothing, and background noise from household hot objects such as heaters, kettle, computer, and light bulbs. Our approach eliminate such noises and works irrespective of the ambient temperature and

the type of clothing worn to some extent.

Coverage area: While Narayana et al. [34] provide the coverage area of 8 m x 8 m, the largest in the literature, we achieve 9 m x 8 m without compromising the accuracy much, as compared.

Scalability: Further, their approach require one PIR sensor for every 1 m distance in the FoV, while our system has single PIR sensor that can adjust the detection range dynamically – as far as 9 m, hence, scalable over distance.

Applications: Unlike all the works found in the literature, our sensor platform is multi-modal – can be used for multiple applications such as localization, tracking, activity and fall monitoring, and height classification.

We also evaluate the performance of our work with the state of the art work and present the results in Chapter 5.

Chapter 3

The Hardware Platform

In this chapter, we will give a detailed explanation of our hardware platform, explain the choice of our sensors and other components and why they are needed, nature of data from these sensors.

3.1 Hardware Platform

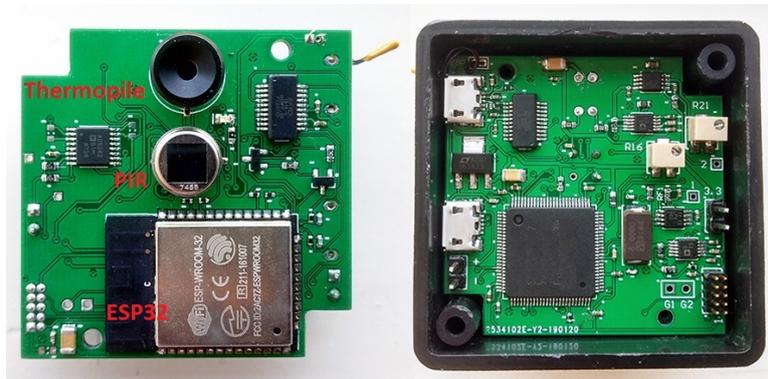


Figure 3.1: The proposed hardware system showing collocated thermopile array sensor and PIR sensor.

While a single thermopile array sensor can provide information on the direction of movement of a person, it is challenging to localize in two dimensions with reasonable accuracy because the number of pixels occupied by the target depends on its height, width and distance from the sensor. A second thermopile array sensor, placed at a certain distance in another place and at a particular angle, can assist in achieving localization but it demands additional cost, power, and synchronization [7]. To achieve 2d localization with low power and low cost system we innovate by using an inexpensive low-power PIR sensor placed adjacent to the thermopile, to provide depth information. Hence, we have designed a custom-made system for our experiments. The proposed hardware platform

Table 3.1: Comparison of different commercial thermopile sensors. We choose MLX90640 because of its large FoV (110°x 75°) and refresh rate.

Sensors	HTPA80x64d	MLX90621	MLX90640	MLX90641	Grid-Eye
Manufacturer	Heimann	Melexis	Melexis	Melexis	Panasonic
Array size	80x64	16x4	32x24	16x12	8x8
Field of View (FoV)	88°x 70°	120°x 25°	110°x 75°	110°x 75°	60°x 60°
Maximum refresh rate	200 Hz	512 Hz	64 Hz	4 Hz	10 Hz
Power consumption	82.5 mW	23.4 mW	66 mW	39.6 mW	14.85 mW

houses two collocated infrared sensors – a thermopile array sensor and a PIR sensor placed one above the other as shown in Figure 3.1. The distance between the center of thermopile and PIR is fixed to be around 1 cm so that the FoV of thermopile does not blocked by the PIR sensor. The overall dimension of the platform is 5 x 5 x 2 cm³. The maximum power consumption of the platform is 80 mW without WiFi transmission, and the entire hardware is powered by a 3.7 V battery embedded inside the enclosure. The hardware also includes an ultra low-power ARM Cortex M0 microcontroller ATSAML21J18B [31], from Atmel’s pico power series microcontrollers, for acquiring data from the infrared sensors and process them. The information comprising of detection, localization, and height of the person is transmitted to a central server using an ESP32 WiFi module [10].

3.1.1 Selecting suitable sensors

Thermopile array sensor. There are various commercial thermopile array sensors available from manufacturers such as Panasonic [37], Melexis [30], and Heimann Sensor GmbH [18]. Sensor models, such as FLIR Lepton [11] provide high-resolution imaging and fall under thermal camera category but, they need high operating and processing power and are quite expensive (\approx \$250). We list different potential thermopile array sensor modules with their relevant specifications in Table 3.1. We choose MLX90640ESF-BAA FIR sensor from Melexis in our platform as it provides large FoV of 110°x 75°, a refresh rate of 64 Hz, and cost of \$45 [29]. A minimum of 90° FoV is required to avoid blind spots in a horizontal direction when the sensor platform is placed in a corner of the room. Melexis sensor meets this requirement. The sensor can measure object temperature between -40°C to 300°C with a temperature accuracy of $\pm 1^\circ\text{C}$. Furthermore, resolution of 32 x 24 = 768 pixels (IR sensors), covering 110°x 75° FoV offered by this sensor is appropriate for human localization (which we discuss and evaluate in Section 5.1) compared to all other sensor modules. The refresh rate of the sensor was set to 8 Hz, as the walking speed of humans indoor is less than 7.2 kmph. So, this refresh rate is enough to capture all the information needed even at the highest walking speed of 7.2 kmph.

PIR sensor. We employ Zilog's ZSBG446671, a dual-element PIR sensor in our system because of its large FoV of 132°(from the center of the element on x -axis) x 222°(from the center of the element on 45°) [45]. The detection area of a PIR sensor element is small and is not very sensitive to the infrared energy. To increase the range of detection, strengthen the incoming infrared rays and change the sensing pattern, we implanted a *Fresnel* lens in such a way that the center of the PIR sensor coincides with the focal point of the lens.

Fresnel lens plays a significant role in condensing the infrared rays onto the sensor elements. The output pattern of the PIR sensor can be altered using different configurations/mouldings of Fresnel lenses [34]. In our application, we require identical detection from all the directions in the FoV of the PIR sensor to estimate the distance of the moving object from the hardware platform. Hence, we selected a generic golf ball lens, shown in Figure 1.2, that concentrates the incoming heat rays from the moving object onto the PIR sensor elements. The golf ball lens is a semi-sphere containing multiple spot Fresnel lenses on its circumference. The FoV of the chosen golf ball lens is 150°x 150°, larger than that of the PIR. However, the overall FoV of the platform is limited by the FoV of the thermopile sensor, which is the maximum FoV that our system offers. The ADC sampling rate for PIR sensor was set to 16 Hz, twice as that of thermopile sensor to meet the Nyquist criterion.

3.1.2 Efficient processing of thermopile data

In this section, we provide an efficient approach to process raw data from the thermopile without compromising the accuracy of the output much. This approach enables the algorithm to be executed on a microcontroller with low computation power.

The MLX90640 does not output the absolute temperature values but 16-bit raw values read by each pixel. These values correspond to the amount of infrared energy falling on each pixel. The raw values can be converted to the temperature values using,

$$T_{o(i,j)} = \sqrt[4]{\frac{V_{IR(i,j)_{COMP}}}{\alpha_{comp(i,j)} * (1 - K_{sTo2} * 273.15) + S_{x(i,j)}}} + T_{a-r} - 273.15, \quad (3.1)$$

Where,

$$S_{x(i,j)} = K_{sTo2} \sqrt[4]{\alpha_{comp(i,j)}^3 * V_{IR(i,j)_{COMP}} + \alpha_{comp(i,j)}^4 * T_{a-r}}. \quad (3.2)$$

Where $T_{o(i,j)}$ is the temperature reading for pixel $i, j, \forall i \leq 32, \text{ and } j \leq 24, i, j \in N, V_{IR(i,j)_{COMP}}$ is the offset compensated raw value for each pixel, K_{sTo2} is a constant and $\alpha_{comp(i,j)}$ and is a constant corresponding to each pixel, and T_{a-r} varies with the ambient temperature. These parameters are calculated using

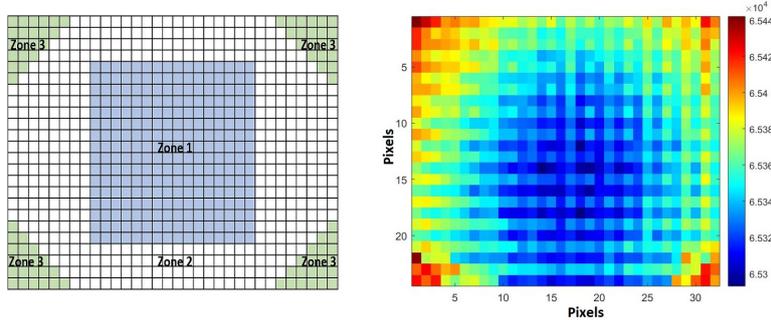


Figure 3.2: MLX90640ESF-BAA frame is divided into three zones depending on the measurement accuracy [29]. Zone 3 produces more noise than Zone 1. The image on the right shows a sensor frame with raw values from the sensor when there is no warm body in front of the sensor.

various constants and pixel offset values¹ that are stored in the EEPROM of the sensor [29].

Further, the MLX90640ESF-BAA frame, containing 768 pixels, is divided into three zones, Zones 1 to 3 by the manufacturer, based on the measurement accuracy of the pixels. The different zones of a frame associated with raw values from the sensor when there is no warm body in front of the sensor are shown in Figure 3.2. For an object in front of the sensor with temperatures between 0°C and 50°C, Zone 1 has the highest accuracy of $\pm 0.5^\circ\text{C}$, Zone 2 with $\pm 1^\circ\text{C}$ and Zone 3, the least with $\pm 2^\circ\text{C}$. Additionally, IR sensors in Zone 3 produce more noise compared to that in Zone 1. We observe in the image with raw values that there is a temperature gradient from Zone 1 to Zone 2 even if there is no hot object in front of the sensor. However, these pixel offsets or errors between different zones can be corrected using offset calculations that result in Equation 3.1.

The calculation indicated in Equation 3.1 involves complex computation of multiple floating point numbers to turn the raw pixel data into temperature data. This demands a minimum SRAM of 150 kb and ≈ 100 MHz processing power to process the raw data at 8 Hz [42]. Further, running the localization algorithm and wireless data transmissions require high power microcontrollers as the host platform. As we desire to design a low-power solution, we simplify the calculation by working with the relative difference between the pixel data rather than the absolute temperature data. Hence, we apply only that part of the calculation that involves pixel offset compensation and zone error adjustments. We analyzed the range of $V_{IR(i,j)COMP}$, $\alpha_{comp(i,j)}$, K_{sTo2} , and T_{a-r} for ambient temperatures $T_{o(i,j)} \in [-20^\circ\text{C}, 125^\circ\text{C}]$. $\alpha_{comp(i,j)}$ in the range $[-3 \times 10^{-8}, 1.23 \times 10^{-7}]$, $V_{IR(i,j)COMP}$ has the range $[-79, 427]$, K_{sTo2} is in the order of -2×10^{-4} , and T_{a-r} in the range $[4 \times 10^9, 2.3 \times 10^{10}]$ for ambient temperatures between 0° C and 100° C.

¹Each pixel is provided with a correction factor from the manufacturer, when the sensor is calibrated.

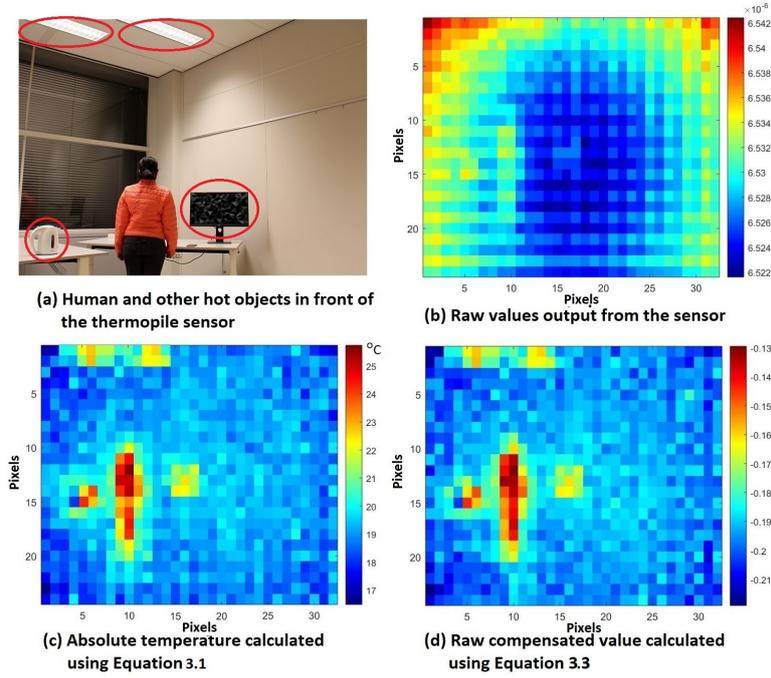


Figure 3.3: (a) Image with warm objects such as human, incandescent light, hot kettle, monitor (b) Raw values output from the thermopile sensor (c) Absolute temperature obtained using Equation 3.1 (d) Raw compensated value calculated using Equation 3.3.

Looking at the above ranges, K_{sTo2} and $S_{x(i,j)}$ can be neglected. Finally, we arrive at the relation,

$$RAW_{comp} = \left(\frac{V_{IR(i,j)COMP}}{\alpha_{comp(i,j)}} + T_{a-r} \right) \times 10^{-9}, \quad (3.3)$$

where $RAW_{(i,j)comp}$ is the compensated (for offset and irregularities in zones) raw value which is directly used in our algorithm.

Figure 3.3 shows an image with (a) warm objects such as human, incandescent light, hot kettle, and a monitor, (b) a frame displaying raw values in the form of gradient map, (c) the same frame with absolute temperature calculated using Equation 3.1, and (d) the same frame after raw values are processed using our technique using Equation 3.3. We observe from the images that our approach preserves the temperature gradient as if the raw data is converted to temperature. We also tested our approach in the presence of the light bulb at the top left corner, hot kettle and monitor, and person in front of the sensor. Figure 3.4 shows the raw compensated values represented by each pixel (Pixel-1 corresponds to top left of the frame and incremented column-wise) for the frame displayed in Figure 3.3. We observe in the figure that raw compensated value is a scaled down version of the absolute temperature. Complete processing is done

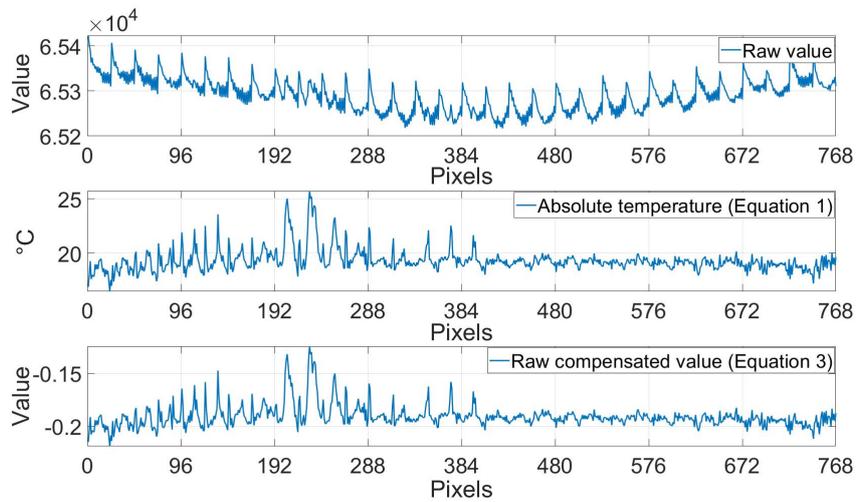


Figure 3.4: Values represented by all 768 pixels in a frame for the images displayed in Figure 3.3.

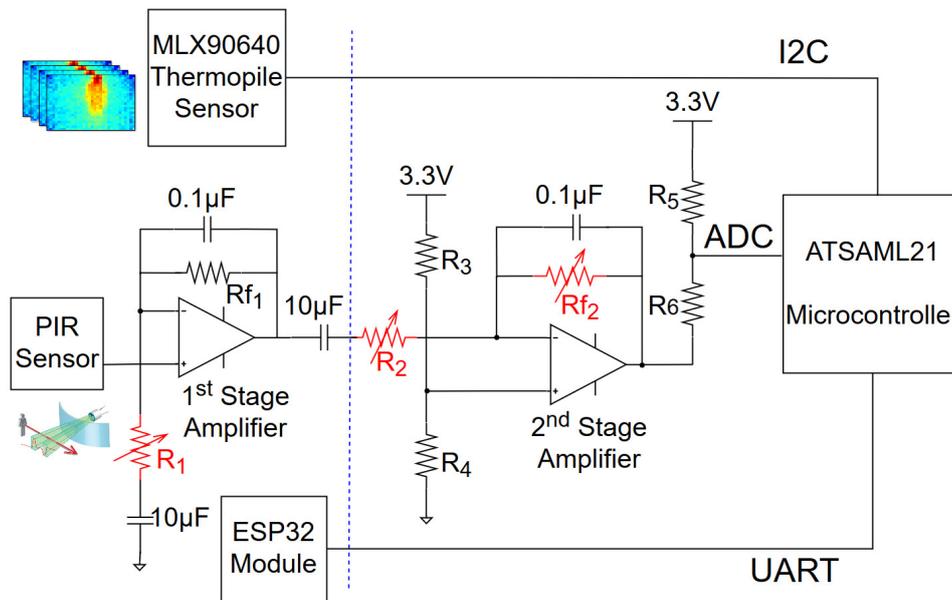


Figure 3.5: Circuit diagram of our system, displaying the hardware elements, and the two-stage amplifier design and variable resistors for the PIR sensor.

on our hardware platform – reading EEPROM data and acquiring raw data from the thermopile sensor at 64 Hz, reading PIR data at 16 Hz, execute our algorithm, and transfer the data over WiFi to a centralized server.

3.1.3 Variable gain setup for PIR sensor

The output voltage from a PIR sensor is very weak and is in the order of μV . Hence, they need to be amplified several thousand times in order to get a reasonable signal that can be measured by a microcontroller. Generally, this is done using operational amplifiers in two stages. We adopt Texas Instrument's LPV802 dual channel nano-power amplifier [21] in our design as they consume only $1\ \mu\text{W}$. The circuit diagram of our hardware platform is shown in Figure 3.5. While most of the work in literature that use PIR involve working with binary output as provided in the vendor reference circuits [44], there are a few that utilize the PIR output in analog form [34, 12]. Since analog output provides more information than the binary, we make use of the analog signals from a single PIR, read by the ADC pin of the microcontroller, to estimate the distance of the moving object from the sensor. The traditional way, as found in literature, is to use a fixed gain amplifier stages by selecting suitable feedback resistors. On the contrary, in this work, we facilitate the microcontroller to vary the gain of each amplifier stage using digital potentiometers that are controlled using I2C lines².

The amplified output V_o of the PIR sensor read by the microcontroller is proportional to the overall gain given as,

$$V_o = -V_{in} \left(1 + \frac{R_{f1}}{R_1} \right) \left(\frac{R_{f2}}{R_2} \right), \quad (3.4)$$

where R_{f1} is fixed to $3\ \text{M}\Omega$, R_{f2} is a $1\ \text{M}\Omega$ dual channel digital potentiometer, AD5242BRUZ1M from Analog Devices, whose resistance can be varied in 255 steps between 0 and $1\ \text{M}\Omega$. We connect both the resistor channels in series to get a broader range of upto $2\ \text{M}\Omega$. Similarly, R_1 and R_2 are $512\ \text{k}\Omega$ digital potentiometers, AD5272BRMZ from Analog Devices, that can be varied in 1024 steps between 0 and $512\ \text{k}\Omega$. By adjusting these resistor values dynamically, it is possible to vary the overall gain of the PIR output between $\approx 2 \times 10^{-6}$ to 6×10^{12} in $\approx 537 \times 10^6$ steps, hence customizing the detection range of the PIR sensor. The analog output from the PIR is similar to a sine wave and produces negative voltage. As the microcontroller is not capable of measuring negative voltages on its ADC pins, we introduce fixed resistors R_3 , R_4 , and R_5 , R_6 , as shown in Figure 3.5, to scale and shift the PIR output before feeding to the ADC pin of the microcontroller. We fix $R_3 = 510\ \text{k}\Omega$, $R_4 = 240\ \text{k}\Omega$, $R_5 = 10\ \text{k}\Omega$, and $R_6 = 2.4\ \text{k}\Omega$ to get a full PIR output swing between $0.2\ \text{V}$ and $3.25\ \text{V}$. A sample amplified output from the PIR sensor is shown in Figure 3.6.

3.1.4 Software

Serial communication takes place at $1/f_s$ time where f_s is the frame rate, sending 452, 16 bit unsigned integer values from micro-controller to computer using

²It should be noted that there are programmable gain amplifiers available commercially but not in the gain range that we require. Hence, we used the variable resistors.

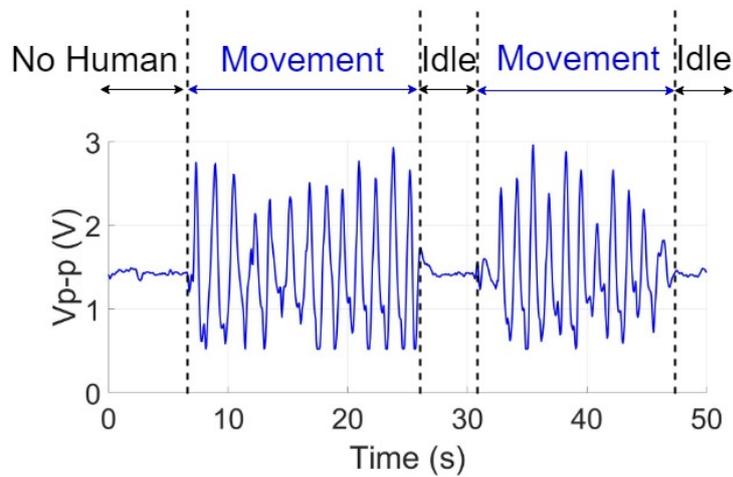


Figure 3.6: Amplified output from the PIR

Matlab. Where, the first 450 data belong to the thermopile array sensor (thermopile array sensors updates 768 pixel data in two sub-pages each sub-pages has 384 pixel values and 66 calibration values), 451 and 452 data packets are the ADC value from PIR sensor and thermistor respectively. These raw data are calibrated and stored in .CSV file format for further analysis. Simultaneously, the pixel output from thermopile sensor and peak to peak analog output from PIR sensor are graphically interfaced using Matlab as shown in Figure 3.7. The GUI also shows the values ambient temperature and different resistors values we have set to.

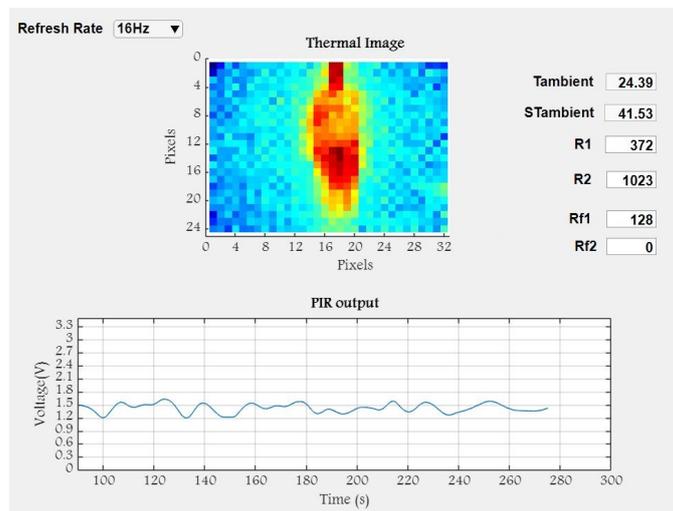


Figure 3.7: GUI to display both Sensors' output and to set different resistor values.

Chapter 4

Localization Algorithm

In this chapter, we present the observations from thermopile and PIR sensor in Section 4.1 and derive features based on these observations then we explain how the classification features can be exploited to perform localization and tracking.

The principal idea behind the fusion of two sensors - thermopile array and PIR in our system is that thermopile sensor can be used to estimate the location in two dimensions (across the FoV cone axis, and height of the object, when deployed on the walls), and PIR can be used to estimate the range between the sensor platform and the object, thus providing the location information in the third dimension.

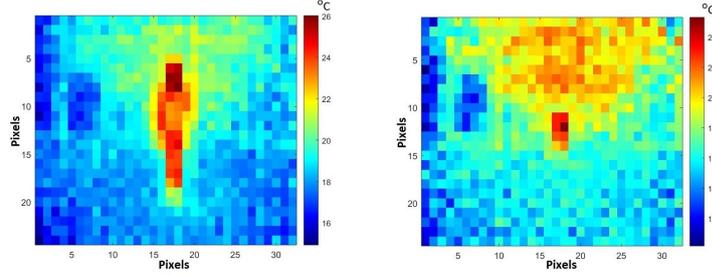
The output of sensors depend only on the thermal signatures of target. These thermal signatures are not unique to a particular target and it is highly dependent on the target size, thermal insulation and how far away from the sensor is a target moving. For instance, the pixel temperature of a thermopile array sensor due to the presence of a person wearing a jacket and is nearer to the sensor will be same as presence of a person not wearing a jacket and who is far away from the sensor. Another ambiguity would be the coverage of number of pixels by a shorter person who is near to the sensor and a taller person who is far from the sensor. Because of such limitations of the sensors, it is not possible to generalize the system to find the location of a target which directs us to the world of machine learning. In machine learning system, we can collect data for different target size, temperature and at different distance and train the model to estimate the location. The most important step to get better accuracy using machine learning algorithms is identifying the unique features of your sensor output and extracting them.

4.1 Features for machine learning

Feature extraction is the most important step to any machine learning algorithm to perform accurately. Hence, to identify the features to make use for our system we profiled the sensors for various conditions like targets of different body tem-

perature, size, height and distance of the target from the sensor, their walking speed etc.,

4.1.1 Observations from thermopile array sensor



(a) Human in front of the sensor at 2 m (b) Human in front of the sensor at 7 m

Figure 4.1: Temperature recorded by MLX90640 for a human at two different distances. The actual temperature of the human (on forehead) is around 34 °C and the sensor reported 25.8 °C at 2 m, and 21.8 °C at 7 m.

Though thermopiles provide information on the absolute temperature that each pixels see, it is not equal to the actual temperature of the object. For a warm object having constant temperature, the temperature recorded by a thermopile is inversely proportional to the distance between the sensor and the warm object. For instance, Figure 4.1 shows temperatures recorded by the thermopile sensor for a human at two different distances.

It is worth analyzing the impact of different types of clothing on the temperature measured by the thermopile sensor as it is a trivial consideration in offices and households. Hence, to investigate this, we conduct an experiment with thin clothing, and a thick jacket. Figure 4.2 shows the temperature recorded by the thermopile sensor for a human at 2 m in front of the sensor with thin clothing (Figure 4.2a) and with jacket (Figure 4.2b). It can be observed from the figures that clothing is an important factor that affects the temperature reading output by the sensor. It should be noted that the body heat from inside the jacket is recognized by the sensor.

From our experiments, we conclude that the following features can be utilized in our algorithm to perform localization, tracking, and height estimation.

Summary of thermopile features

1. As the distance between the warm body and the sensor increases, the temperature read by the sensor decreases. This holds with the theoretical model provided in [35] that says

$$T \propto \frac{1}{d^2},$$

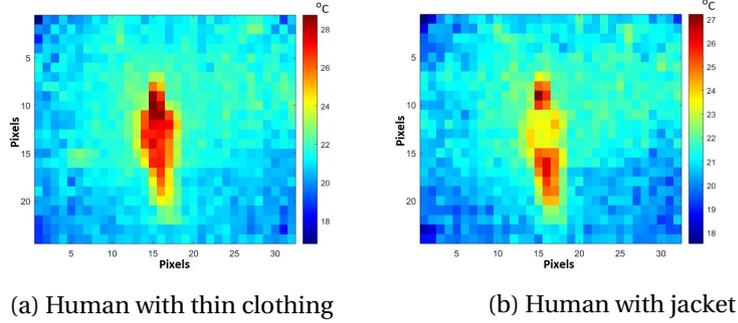


Figure 4.2: Temperature recorded by MLX90640 for a human at 2 m with and without jacket

where T is the temperature recorded by a pixel and d is the distance of the warm body from the sensor.

2. As the person moves away from the sensor, the number of pixels used to indicate the object decreases. Hence, if the height of the person is known, the distance between the object and the sensor can be estimated as the vertical FoV and the total pixel count in the vertical dimension is known. Similarly, if the distance is known, the height of the person can be estimated.
3. One of the important observations is the spatio-temporal changes in the number of pixels covered by a moving object. When a person walks in front of the thermopile array, in any direction, the number of pixels traversed horizontally and vertically in a specific duration is proportional to the speed of the movement and distance of movement from the sensor. This provides a new relation,

$$P_t(h, v) \propto \frac{s}{d}, \quad (4.1)$$

where $P_t(h, v)$ is the number of pixels traversed horizontally (h) and vertically (v) in time t , s is the speed of the movement, and d is the distance of person from the sensor.

4. The static hot objects such as bulbs, computers, and heaters that contributes to the background noise do appear in the thermopile output frames.
5. Any hot object such as kettle carried by a human is detected as a warm object by the sensor and we consider that hot object also as part of the human body.
6. A human can be detected even if s/he is wearing thick clothing such as a jacket.
7. Head, chest and waist are the parts of the human body that exhibits high gradients compared to the rest of the body.

4.1.2 Observation from PIR sensor

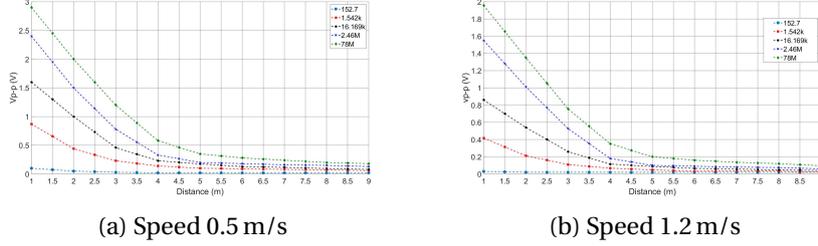


Figure 4.3: Peak to peak voltage generated by the PIR sensor for different gains at different speeds.

Considering the generic models from [34] and [33], we can portray the relation between the peak to peak voltage V_{p-p} generated by the PIR sensor and the amplifier gain G as,

$$V_{p-p} \propto \frac{I G}{s^2 d^2}, \quad (4.2)$$

where I is the infrared energy from the moving person incident on the PIR sensor, s is the speed, and d is the distance of the person from the sensor.

In Figure 4.3, we show the peak to peak voltage generated by the PIR sensor for different gains at different speeds. It must be observed that large gain is required to see the same person at a farther distance at a constant walking speed. Similarly, as the movement speed increases at a constant distance, the gain has to be increased to observe the same levels of signal.

Summary of PIR features

1. As the distance of the person from the sensor increases, the peak to peak voltage output from the PIR sensor decreases, provided that the speed of the movement, body temperature and the amplifier gain remains approximately the same.
2. When a person is moving at a particular distance with a constant speed, the peak to peak voltage generated by the amplifier output can be varied by changing the overall gain G of the amplifier stages. This can be observed in Figure 4.3a or Figure 4.3b, where speed is constant in both the cases.
3. The peak to peak voltage output from the amplifier stages decrease as the speed of the movement at a specific distance increases. This is evident from Figure 4.3a and Figure 4.3b.
4. The peak to peak voltage output from the amplifier stages for a person moving at distance d_1 with speed s_1 may be same as the output for the same person moving with speed s_2 at distance d_2 , where $d_1 < d_2$ and $s_1 > s_2$. This is

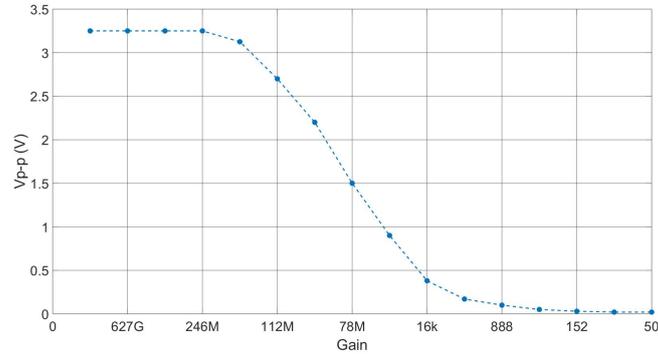


Figure 4.4: Voltage - Gain curve obtained for movement with speed 0.5 m/s at 2 m in front of the sensor

because the duration and amount of IR rays falling on the sensor reduces as the speed of the moving object increases. Similarly, when the object moves slowly at farther distance, the exposure time of the object is longer, providing higher chance of absorbing all the IR rays emitted by the object, and thus, generating relatively higher amplitude.

5. Static hot objects such as hot kettle, computer, light bulbs do not affect the output of the PIR sensor.

4.1.3 Feature extraction

The training data comprises of curves obtained from PIR sensor, and pixel data from thermopile array:

Curvesets from PIR sensor: Peak to peak voltage measurements V_{p-p} (represented by Equation 4.2) against all dependent parameters - amplifier gains G , various speeds s , and at different distances d is one of the features that will be used in machine learning algorithm. When a person is moving at a constant speed at a given distance, the peak to peak voltage output from the PIR sensor can be recorded by varying the amplifier gain from the maximum to the value at which there will not be any detection (peak to peak voltage equivalent to the noise level). This gain can be varied by decreasing the resistor values by 1 step (i.e., R_1 , R_2 by 98Ω and Rf_2 by $3.9k\Omega$) but for practical purposes we decrease gain using binary search method.

The obtained peak to peak values V_{p-p} for various gains G are stitched together to form Voltage - Gain curve (V-G curve). A sample curve for movement with speed 0.5 m/s at 2 m in front of the sensor is shown in Figure 4.4. Similar such curves can be recorded for different people moving with speeds and at different distances. So, a single V-G curve spans over three dimensions with speed, distance and different people and clothing, that addresses all the variables in

Equation 4.2. We represent the training data from PIR sensor as

$$V_{p-p} = f_{(s,d,I)}(G), \quad (4.3)$$

where gain G is varied from maximum to the minimum detection level, I represents infrared energy from different people with different clothing.

Datasets from thermopile array: In the case of thermopile sensor, there are two factors - movement speed and distance - that affect the number of pixels traversed by the moving person as indicated by Equation 4.1. This pixels traversed in horizontal and vertical direction is used as a feature for machine learning. Figure 4.5 illustrates how the pixels traversed vary as the speed and distance of person varies. When a person is at point 'A' which is at the distance near the sensor, the number of pixels traversed by the target are 20 in vertical direction and 7 in horizontal direction as the person moves far from the sensor to location 'B' the number of pixels traversed are reduced to 9 in vertical direction and 3 in horizontal, however when the person moves from 'B' to 'C' at the same distance from the sensor but parallel to the sensor the number of pixels traversed remains the same in vertical direction. Hence, the dataset from thermopile array comprises of Pixel Traversed data (P-T data) in horizontal and vertical direction recorded for different speeds s at different distances d . Hence, each P-T data is of four dimensions. The datasets from both PIR and thermopile sensor should be recorded concurrently so that both the datasets represent the same event. We represent the training data from thermopile sensor as

$$p_t(h, v) = (i, j)_{(s,d)} \quad (4.4)$$

where $p_t(h, v)$ is the pixels traversed in horizontal and vertical direction with $i \leq 32$, and $j \leq 24$.

4.2 Localization and tracking steps

Once the training data is available, localization and tracking can be performed with the following steps. We consider a sample case wherein a human, light bulb, and a monitor are present as shown in Figure 4.6, and explain the steps for better understanding.

4.2.1 Human detection and horizontal location estimate

Detection of movement

Initially, amplifier gain of the PIR sensor is set to the highest to detect the human movement in the FoV. Simultaneously, snapshots of a frame from the thermopile sensor is taken every second (as the refresh rate is 8 Hz, a frame snapshot is the pixel to pixel average of 8 frames) that forms the background frame. This frame will be subtracted from the detection frame (frame in which a human is detected) in later stages. This helps to eliminate static warm objects such as monitors

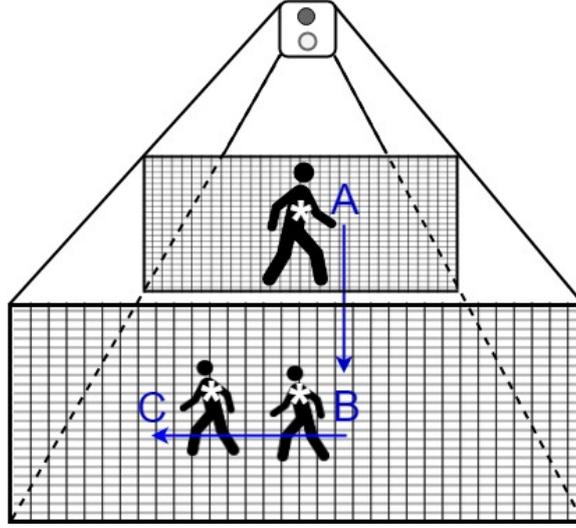


Figure 4.5: Number of pixel occupied when the target is at different locations

and light bulbs from the detection frame. We represent the background frame $B_{i,j}$ as

$$B_{i,j} = \frac{1}{8} \sum_{k=1}^8 F_k(i, j) \quad (4.5)$$

where $i \leq 32$, and $j \leq 24$, and $F_k(i, j)$ represents a single frame containing 768 pixels. The human presence is indicated when $V_{p-p} > 0.02$, as 0.02 is the mean noise amplitude at the highest possible gain¹. When there is a movement detection in the PIR sensor, the background estimation process is stopped. Figure 4.6a shows the background frame.

Background and noise removal

When there is a human presence (shown in Figure 4.6b), the background frame is subtracted from each thermopile frame that is being read. Figure 4.6c shows the frame from thermopile corresponding to the scenario shown in Figure 4.6b. The resultant background removed frame is shown in Figure 4.6d wherein the human presence is persistent. Even though the static warm bodies and background noise are removed from the data, there may be a few pixels present that do not represent the human. Such pixels can be neutralized using two-dimensional Gaussian filter $G(x)$, given in [14] as

$$G(i, j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{i^2+j^2}{2\sigma^2}}, \quad (4.6)$$

¹We observed during our experiments that the amplified PIR output does not include high frequency noise that can trigger the false detection. This may not hold always with PIR sensors from other manufacturers. In this case, a low pass filter may be required.

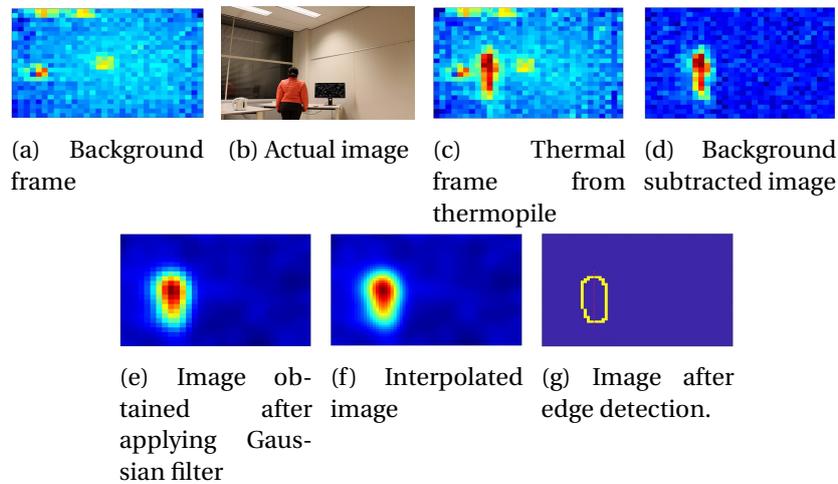


Figure 4.6: Steps involved in localization and tracking

where σ is the standard deviation, which we found out empirically to be 1.5 in our application to get better smoothing. The filtered data is shown in Figure 4.6e.

Interpolation of thermopile data

The area covered by each pixel increases as the distance from the sensor increases. To increase the resolution or decrease the area covered by each pixel, the number of pixels can be increased by interpolating the image. In our system, The filtered data is interpolated by 2 times to get a frame of size $32 \times 24 \times 2 = 3072$ pixels to get better accuracy. The interpolated image is shown in Figure 4.6f.

Estimation of position in one dimension

The next step is to identify the position of the person in one dimension, i.e, direction perpendicular to the FoV axis of the sensor platform. To perform this, the outline of the pixels representing the person has to be detected.

This can be achieved using any of the edge detection methods. They are image processing techniques to outline the boundary of objects in an image by finding the difference in the image brightness. In our system since the background noise is already subtracted, the target is indicated by different pixel values when compared to rest of the image. Hence, there will be a clear distinction in brightness between human and background due to which edge detection can accurately identify the edges of human target.

We use Canny edge detector in our algorithm as this involves low complexity processing compared to other edge detection techniques and widely applied in various computer vision systems [5]. This is a multi-step technique that detects

edges as well as suppresses noise at the same time. Once the edges are detected by using pixel magnitude gradient a double thresholding (using two different thresholds one high and another low threshold) is applied to find the potential edges of an object.

Applying edge detection filter performs better in terms of accuracy than using one particular threshold or adaptive threshold to detect humans like the works mentioned in the literature. This is because the edges are detected by checking the changes in brightness of the image and the difference between the brightness values of human is huge when compared to background image and hence there is no misclassification of noise as human and human as noise due to the usage of wrong threshold value in the algorithm.

Figure 4.6g shows the position of the human in the frame. We consider the centroid of the shape as position 'X' from the column 1 of the frame (Red line in the image is passing through the centroid). However, this position cannot be mapped onto the physical location unless the depth of the person from the sensor platform is known as the thermopile frame is the rectangular projection that enlarges away from the sensor. Hence, we call this position as the virtual position.

4.2.2 Distance estimate

Machine learning - Estimating location using PIR

The main objective of this step is to find the distance of the person from the sensor platform. As mentioned earlier in Section 4.1.3 we have to get the V-G curve and P-T data i.e., the pixels traversed in horizontal and vertical direction.

V-G curves can be obtained by using the variable gain feature in our system. As soon as a human is detected (step 4.2.1, the gain of the amplifier is reduced from maximum to the level at which the peak to peak amplitude of the output is just above the detection level. Note that the gain can be reduced to the minimum but the outputs for gains set below the detection level contain only noise. For each gain set in the range between detection level and the maximum, the V_{p-p} is recorded to form a V-G curve. This has to be performed as soon as possible (within a second or two), before the movement speed and/or distance is changed by the person. The V-G curve thus obtained forms the test data. Similarly, the test data (P-T data) from thermopile is recorded at the same time when the test V-G curve from PIR is computed.

In the training phase, V-G curves and P-T data for different distances are used as input-output pairs to make the system learn different distance depending on the V-G curves and P-T data inputs. To train the system using example input-output pairs, we should use different supervised learning algorithms. We tested different supervised machine learning algorithms for our dataset and from the Table 4.1 it is evident that KNN has highest accuracy and can be used for classification of input to different distances. On the other hand, with Artificial neural

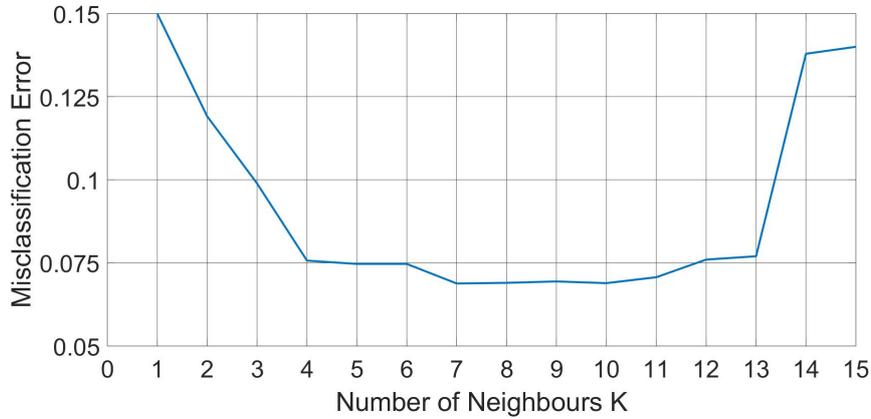


Figure 4.7: Misclassification error in KNN for different values of k

networks (ANN) there is always a trade between high accuracy and low speed and high resource utilization. Hence, we employ KNN with Euclidean distance

Algorithm	Accuracy
Tree	92.36%
SVC	72.3%
KNN	93.1%

Table 4.1: Performance of supervised machine learning algorithms

to find the distance of each point in testing V-G curve and P-T test data. k (number of neighbours) is an important parameter in KNN. It decides the accuracy of the algorithm and must be picked in such a way that we get the best possible fit for the dataset. A 10 fold cross validation was done on different values of k, the results of which are shown in Figure 4.7, from which it can be seen that above k=7 the curve is overfitting and below that it is underfitting. Hence, we set k = 7, where 7 V-G curve instances in the training dataset are most similar to the test V-G curve. Similarly for P-T dataset.

As explained in Section 4.1.3, each V-G curve in PIR training dataset has its pair in P-T data in thermopile dataset. Hence, out of the resultant 7 nearest neighbors in both PIR and P-T dataset, there is a unique V-G curve in PIR dataset that pairs with corresponding P-T data in thermopile dataset. We prove this uniqueness using contradiction.

Let us assume that there are two V-G curve – P-T data pairs. We know that these data correspond to specific speed and distance. We observe from Equation 4.2 that, as speed and/or distance increases, V_{p-p} decreases. However, from Equation 4.1, we see that $P_t(h, v)$ increases as speed increases. Further, at constant speeds, V_{p-p} varies quadratically and $P_t(h, v)$ varies linearly with distance. Hence, there cannot be multiple pairs in K-nearest neighbours (where k=7 in our

case) classified for the test data in PIR - Thermopile training dataset.

It should be noted that, the nearest training datapoint to the respective test data in a dataset (either PIR or thermopile) may not be the solution as it may not have its pair in the other K-Nearest Neighbours set. For instance, in Figure 4.8, the nearest datapoint to the test data in thermopile training data indicated by brown line is not the solution as it does not have its pair in the 7-Nearest Neighbours in the test data of PIR training data.

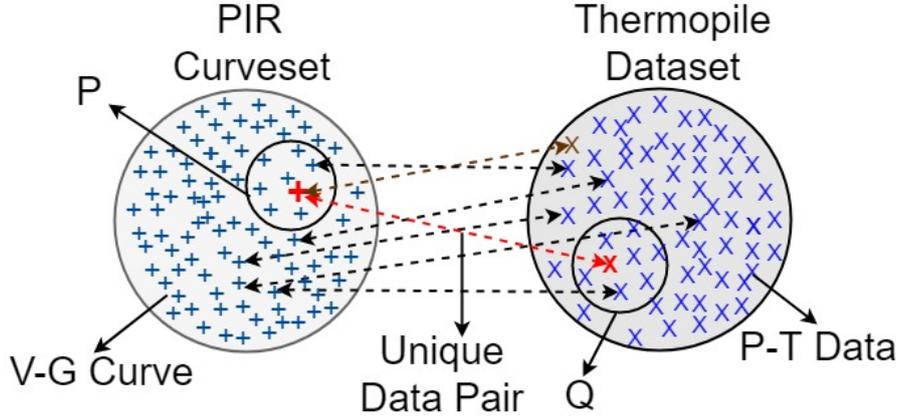


Figure 4.8: PIR and thermopile training dataset comprising of V-G curves and P-T data

The solution having distance d forms the second dimension 'Y', the distance of the person from the sensor platform. Using 'Y', we update the virtual position 'X' to the absolute position as the frame projection height and width at distance 'Y' from the sensor is known using Equation 4.7.

$$absolute\ position = P_w \cdot virtual\ position, \quad (4.7)$$

where

$$P_w = Y \left(\frac{\tan(55)}{\frac{\text{Number of pixels in horizontal direction}}{2}} \right)$$

From Figure 4.9 we know that the horizontal FOV of thermopile array sensor is 110° which spans over 32 pixels. And at a particular distance 'Y', using the $\tan(\theta)$ formula in a right angle triangle (where θ is half of the horizontal FOV) we can find the area covered by each pixel and then multiply area covered by each pixel with the virtual position to get the actual position.

4.2.3 Tracking

Tracking is important in this system to reduce the error introduced when the target is occluded by other objects. Hence, once localization is completed, tracking

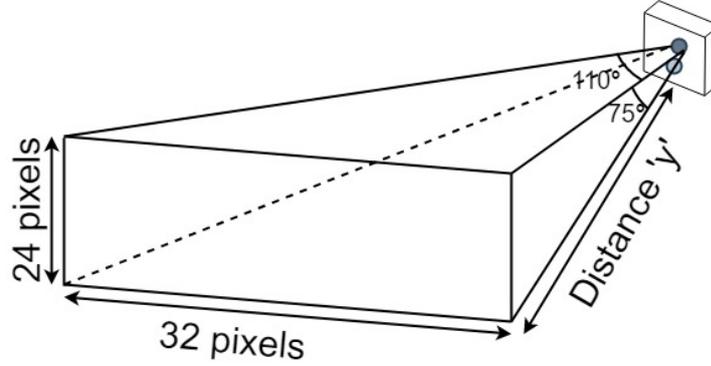


Figure 4.9: FOV of sensor system

is performed from subsequent thermopile frames. From step 4.2.2, we get the absolute position of the person in two dimensions. For every thermopile frame recorded, the difference in locations is calculated using distance formula continuously. These steps are continued until there is human detection after that the gain of the PIR sensor is reset to the highest and snapshots of a frame from the thermopile sensor is taken every second until a human is detected to form the background frame.

4.2.4 Height classification

Now that the absolute position of the person is known, the height can be estimated using the number of pixels in a frame in vertical direction, representing the person in the vertical FoV of 75°. Assuming that the person is perpendicular to the FoV cone axis, the height H of the person is given by

$$H = P_h \cdot \text{Number of pixels representing the person vertically}, \quad (4.8)$$

where

$$P_h = Y \left(\frac{\tan(37.5)}{\frac{\text{Number of pixels in vertical direction}}{2}} \right)$$

From Figure 4.9 we know that the vertical FOV of thermopile array sensor is 75° which spans over 24 pixels. And at a particular distance 'Y', using the $\tan(\theta)$ formula in a right angle triangle (where θ is half of the horizontal FOV) we can find the area covered by each pixel and then multiply area covered by each pixel with the number of pixels that are occupied by the person to calculate the height of the person. However since the resolution is very less in case of the original image i.e., 32 x 24 the area covered by each pixel is high which leads to the higher error therefore we increase the resolution by interpolating the pixels.

Algorithm 1 Location and Height estimation algorithm

$G \leftarrow$ Gain of PIR
 $V_{p-p} \leftarrow$ Peak to peak voltage output from PIR
 $h \leftarrow$ Pixel occupied by target in horizontal direction
 $v \leftarrow$ Pixel occupied by target in vertical direction
while $V_{p-p} \leq 0.02$ **do**
 $B_{i,j} \leftarrow$ Save background frame every second (using Equation 4.5)
end while
 $X \leftarrow$ Estimate virtual horizontal location (as in Section 4.2.1)
 $P_t(h, v) \leftarrow (h_{max} - h_{min}, v_{max} - v_{min})$
 $i \leftarrow 0$
while $V_{p-p} > 0.02$ **do**
 $i \leftarrow i+1$
 $G_i \leftarrow$ Decrease G_{i-1} (as mentioned in Section 4.1.3)
 $V_{p-p_i} \leftarrow V_{p-p}$ at G_i
end while
V-G curve $\leftarrow \{(V_{p-p_i}, G_i)\}_{i=1}^i$
 $d_{vg} \leftarrow$ Classify V-G curve to a distance using KNN classifier
 $d_{pt} \leftarrow$ Classify $P_t(h, v)$ to a distance using KNN classifier
 $Y \leftarrow$ Estimate distance using d_{vg} and d_{pt}
 $x \leftarrow$ Estimate absolute horizontal location using Y (in Equation 4.7)
 $H \leftarrow$ Estimate height of target using Y (in Equation 4.8)

Chapter 5

Results

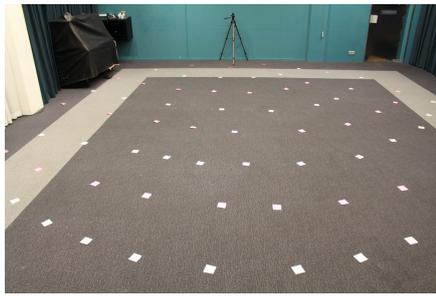
5.1 Performance Evaluation

We evaluated our sensor platform and the algorithm in real time scenario. We also considered the factors such as background noise, static warm objects, obstacles, and more number of people, that affects the localization accuracy. Further we also compared our algorithm with the state of the art approaches. In this section, we first explain our experimental setup and then present the results and observations.

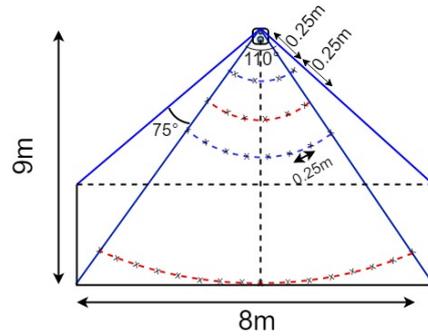
5.1.1 Experimental Setup

Figure 5.1 shows our experimental setup covering the area of 9 m x 8 m. Several circles from 0.25 m to 9 m with the sensor platform as the center were drawn at 0.25 m increment in radius to form equal distances from the sensor platform. This is shown in Figure 5.1b. Several locations were marked at every 0.25 m distance along the circumference of all the circles. The distance of 0.25 m was considered as an average width of a person. The sensor platform was placed at a height of 1.2 m, an acceptable height to place it in indoor locations and measure the height of a person.

The training data was collected from 20 different people with and without wearing jackets, and different types of clothing on different days. At different speeds from 0.2 m/s to 2 m/s, V-G curves and corresponding P-T data were recorded in the FoV. The movement direction included walking perpendicular to the FoV cone axis, parallel to the cone axis, along circumference of the circles, diagonal to the FoV cone axis, and random. Using this training dataset, we tested the performance of our algorithm in different locations and with different people. The deployment of the sensor platform was preferred to be at the corners of the wall so that the FoV of the sensor can cover a larger area.



(a) Experimental setup covering the area 9 m x 8 m



(b) Diagram showing the experimental setup

Figure 5.1: Experiment setup for localization and tracking

5.1.2 Localization and tracking accuracy

To test the localization error in different walking paths different experiments were done in a completely different setting from that of the training environment and the results for them are discussed below.

In the most basic test, the person is moving perpendicular to the FOV cone axis. In this case both the sensing elements of the PIR sensor receives different amount of incident energy, hence the output signal of PIR sensor has clear sinusoidal waves which results in the higher accuracy of the system irrespective of the speed of movement. From the Figure 5.2 it can be seen that when moving in radii from 2m till 7m the error is less than 25cm but is increasing that is mainly because of the increase in area covered by each pixel with increase in distance. At 1m the error is high because the whole body of the person is not in the FOV of the sensors and hence the peak to peak output of PIR sensor is less which results in classifying them at the higher distance. Furthermore, when a target is moving at slow speed that is at 1.8kmph and is at farther distances that is from 7m onwards at the error is high because of the diminishing analog wave.

Different walking paths as shown in Figure 5.3 have been evaluated at different distances. In walking path 1 the target is walking across the FOV cone, starting at a distance of 2m from the sensor system till 8m distance. The actual path traversed and the path estimated are shown in Figure 5.4a. As expected the er-

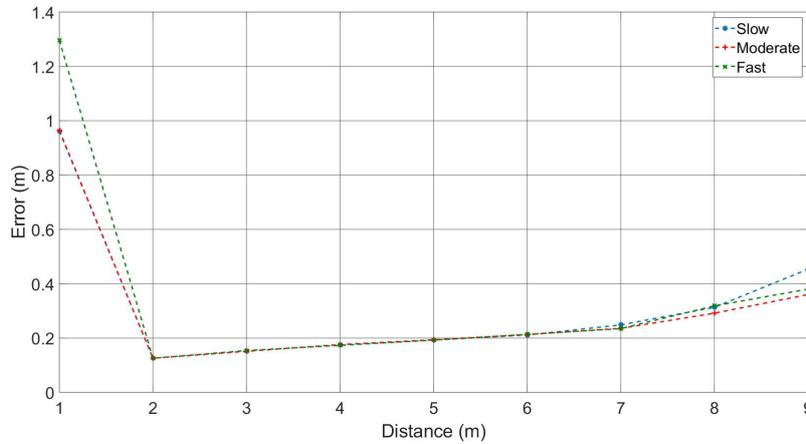


Figure 5.2: Localization accuracy while walking perpendicular to the FOV cone at different distances

ror from 2m till 4m is less than 15cm but the error at 5m and 6m is 28cm and 40cm which is higher than expected and it is because of the misclassification of distance which influenced the error in x-location calculation as well. At 8m however the pixels covered are less which results in the higher error in x-axis. However the average tracking error for this path was 18.6cm which is still acceptable.

When walking parallel to the FOV cone axis as shown in path 2 the distance estimate is more error prone because, it is possible that both the PIR sensing elements might be receiving the same energy and hence the output peak to peak voltage variation is very low, which is evident looking at the actual path and estimated path difference in Figure 5.4b. The total error in this path was 42cm.

In path 3, the target is walking from one side to another in a straight path which is same as the basic case of walking perpendicular to the FOV cone axis but walking in straight path instead of radii. And the results are similar as well which can be seen in Figure 5.4c that the error is quite low, however at distance 4.12, it can be either classified as 4.25 or 4, hence there is an error introduction in x-location calculation as well.

In path 4, target is moving across the FOV axis like in path 1 from 5m till 8m and it can be seen in Figure 5.4d that the error in tracking is mainly due to the error is x-axis estimation which is due to the increase in area covered by each pixel of thermopile array sensor with the increase in distance from the sensor.

To test the overall localization and tracking accuracy, different sensors were placed in rooms with different dimensions in a building, shown in Figure 5.1 and Figure 5.5. The rooms had obstacles such as table and chairs, and static warm objects such as monitor and light bulbs. Different people were asked to walk randomly inside rooms at different speeds. In total, 50 recordings were done.

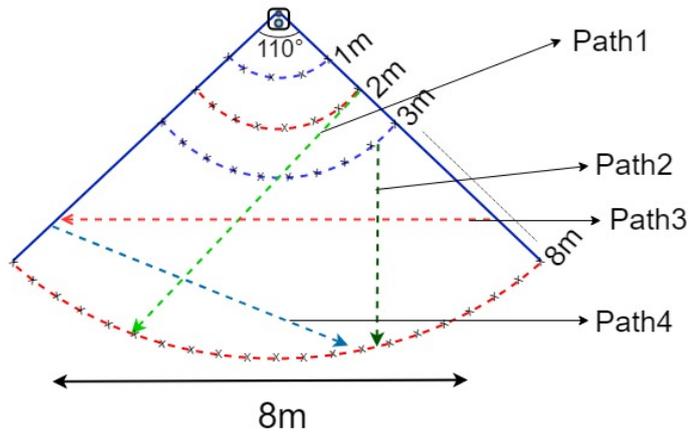
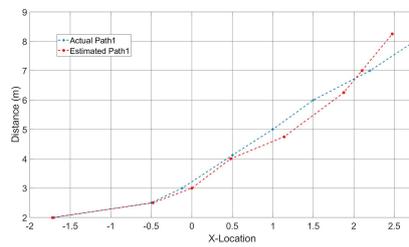
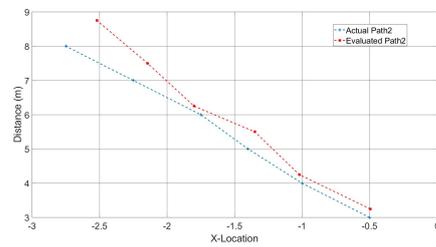


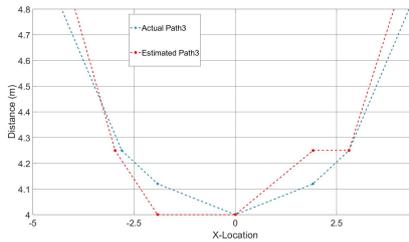
Figure 5.3: Different walking paths for evaluation



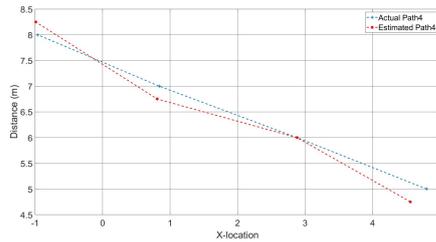
(a) Path 1 tracking accuracy.



(b) Path 2 tracking accuracy.



(c) Path 3 tracking accuracy.



(d) Path 4 tracking accuracy.

Figure 5.4: Different walking paths and their tracking accuracy.

To know the ground truth, markings were also made on all the rooms where testing was done. The error in localization was calculated by finding the dis-



(a) Room with dimensions 6 m x 5 m (b) Room with dimensions 4 m x 5 m

Figure 5.5: Additional locations considered for evaluations

tance between the estimated X, Y location and the ground truth using distance formula. The CDF graph of localization accuracy of our system is shown in Figure 5.6. We observed from the results that the maximum error of 1.62 m was seen at 0.5 m distance from the sensor platform. The reason is that, at proximity, the entire human body is not visible to the sensors and are out of the conical FoV. Hence, the pixels traversed by the person and his speed cannot be measured by the sensor accurately. As the distance increases, the human movement is completely visible to the sensors, resulting in good accuracy (maximum error of 0.24 m at 4.5 m). As the movement distance increases, much information about the person is not available to the thermopile sensor as the person covers less number of pixels. The best localization accuracy obtain was 100 % in the grid resolution of 0.25m x 0.25 m. The best case tracking accuracy obtained was 100 % in the grid of 0.25 m x 0.25 m. The maximum deviation obtained was 0.6 m.

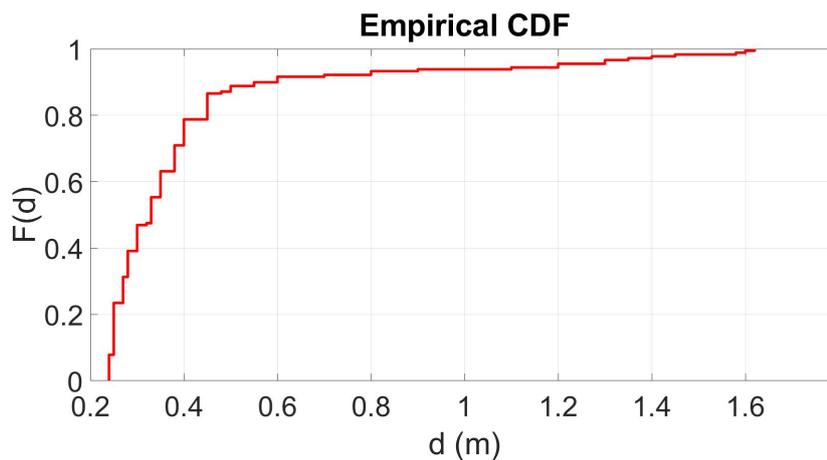


Figure 5.6: CDF of localization error

From these results it can be seen that using the V-G curves in combination with P-T data in machine learning algorithms, instead of using just on peak to peak voltage value to find the distance gave better accuracy than the previous works and also it is independent of human body temperature. Since the system is deployed in the corner of the room there are no blind spots which otherwise would have caused error in localization. This is one of the major problems the works who deploy their system on ceiling face. However since the system is deployed on the wall there can be error in localization due to occlusion of target by other objects which is elaborated in the next section.

5.1.3 Impact of obstacles

To analyze the impact of obstacles, we considered a plant and a chair that covers the lower part of the person, and a board that covers upper part of the person. The obstacles were placed 2.5 m away from the sensor platform as shown in Figure 5.7. The corresponding data from the thermopile is also shown in the figure. A person walking behind the obstacles was tracked and localized continuously. In all the cases when the person was completely visible to the sensor, the person was tracked and localized with almost no error (best case, 100 % accuracy). In the presence of obstacles, such as in Figure 5.7d, Figure 5.7h and Figure 5.7j, there was an abrupt increase in the error in 'Y' direction, i.e, the distance between the sensor platform and the person. This is because, as soon as a part of the person becomes invisible because of obstacles, the number of pixels traversed by the person changes suddenly giving the impression that the person has moved away from the sensor. This also affects the accuracy of height estimation. As we observed, the highest localization error of 42 % in accuracy was observed in the case of board, followed by the plant with 31 %, and chair with 22 %, the least. In the case of chair, the infrared energy from the person was not completely blocked.

5.1.4 Height classification accuracy

To evaluate our approach on height classification, we considered 10 people with different heights in the range 1.5 m to 1.94 m. The height classification was performed for each person at different distances varying between 1 m and 8 m in steps of 1 m. Figure 5.8 shows the 3d bar plot of errors in height classification at different distances and for different heights. We observe from the figure that the average error is 8 cm and the maximum error was seen at the extremes - i.e, at 1 m and 8 m. It is obvious that at proximity, the complete height of the person is not visible to the thermopile. Similarly, at farther distances, the number of pixels covering the person decreases. For distances between 2 m to 7 m, the maximum error was 0.14 m. It was also observed that interpolation of the thermopile data to certain extent (4 times for our system) helps to get better accuracy. The errors shown in Figure 5.8 was obtained by interpolating the thermopile data 4 times.

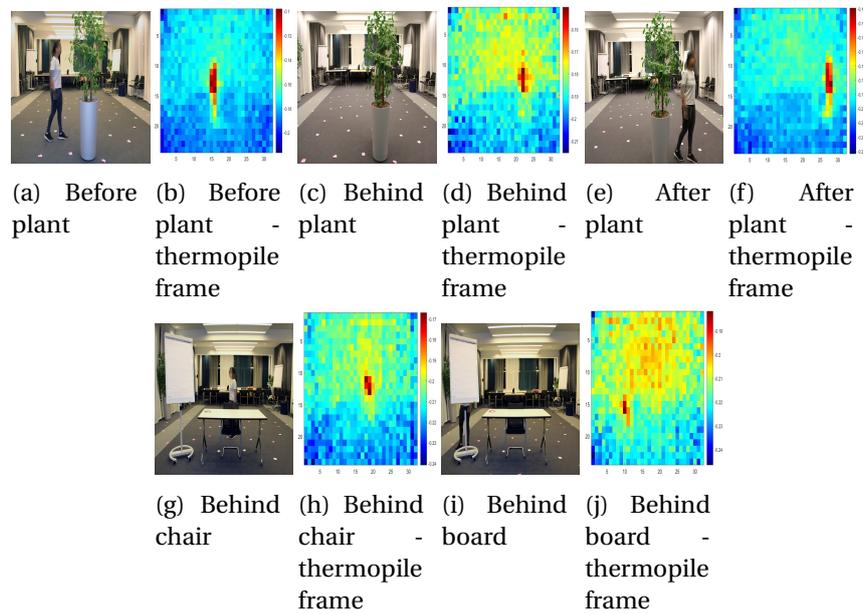


Figure 5.7: Experiment to analyze the impact of obstacles

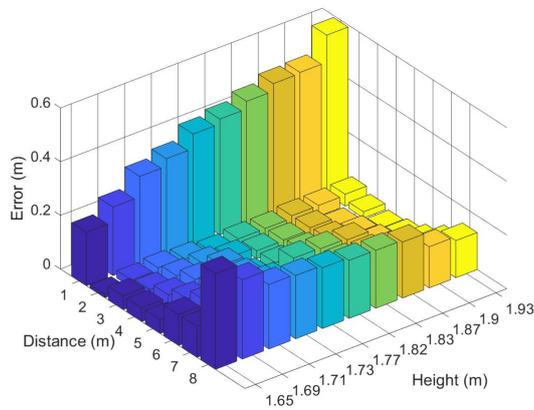


Figure 5.8: 3d bar plot showing error in height classification

5.1.5 Raw data compensation accuracy

As mentioned in Section 3.1.2 we are using compensated raw values instead of actual calculated temperature in order to enable the algorithm to execute on a microcontroller with low computation power without compromising the much accuracy. However it is important to know how much of the accuracy we are missing on and is it acceptable.

Figure 5.9 shows the variation of raw compensated values from that of actual

temperature calculated, and it can be seen that in between 20° C and 40° C the values are overlapping and deviate for temperature values below 20° C and above 40° C.

In order to check the loss of accuracy due to the use of raw compensated values, the pixels occupied by target at different body temperatures were counted for both raw compensated and actual temperature calculations. It was observed that when the target is in between the temperature range of 17° C and 19° C there was 2 pixel variation after 4 times interpolation which means the error ranges from 3cm at 2m distance to 10.8cm at 9m distance in horizontal location estimation. In case of height estimate the error ranges from 1.6cm to 11.2 cm again depending on at what distance the target is. Further, there was no variation in number of pixels excited due to the presence of target whose temperature was from 20° C to 30° C.

It should be noted that, it was practically not possible to check the error due to body temperatures below 17° C and above 35° C.

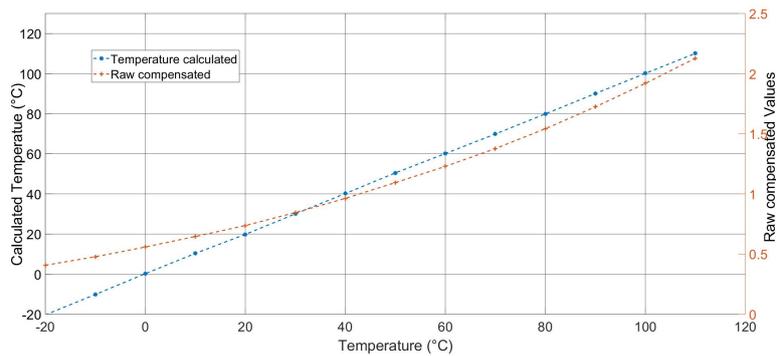
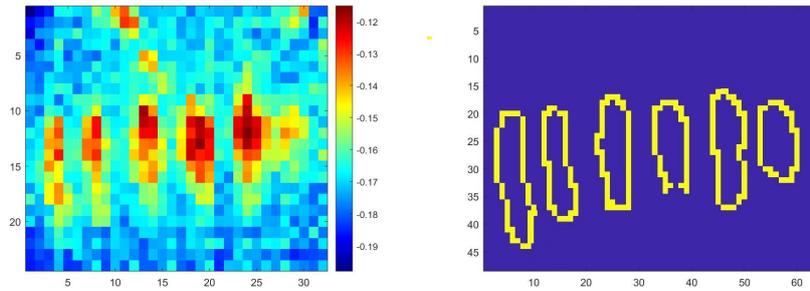


Figure 5.9: Variation of raw compensated value from calculated temperature

5.1.6 Occupancy detection

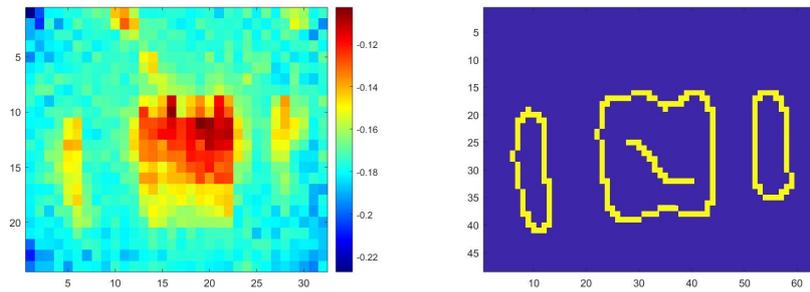
Multiple people can be detected and tracked using thermopile array sensor as shown in Figure 5.10a however, if more than one target is walking in the FOV the PIR produces arbitrary mixed waveform pattern, because of which it becomes hard to localize more than one person with this system. On the other hand, following the same steps as localization, people counting can be performed by using the number of connected components after edge detection as shown in Figure 5.10b that is when people in FOV are far from each other. But, there are multiple limitations with this approach for instance when one person is walking in front of the other, then the pixels with high temperature overlap and will be counted as a single person, similarly when group of people are together as shown in Figure 5.11a it becomes difficult to count people, since most of the connected pixels are of same temperature value, the edge detection doesn't dif-

ferentiate them as multiple people as shown in Figure 5.11b. However the size of largest edge detected component, along with blob detection can be used in clustering algorithms as mentioned in [1] to find the number of people in complex scenarios.



(a) Thermopile output showing 6 people in a frame (b) Edge detection output of 6 people

Figure 5.10: Occupancy detection for 6 people



(a) Thermopile output showing 6 people in a frame where 4 are in a group (b) Edge detection output of 6 people where 4 in a group as detected as 1 group

Figure 5.11: Occupancy detection for 6 people where 4 are in a group

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Localization is an important problem that is being continuously addressed in the last two decades. In this work, we addressed passive device-free, privacy-aware localization. We custom-built a miniature platform that could be deployed on the ceiling or the wall. Research into localization using passive infrared sensors so far, has used multiple sensors, deployed in a particular way to achieve higher accuracy in smaller areas. We addressed many such challenges for example, a single system with 2 sensors is used to localize in 9 m x 8 m room. We proposed many novel techniques such as variable gain for PIR sensor to simulate a spatial diversity gain. Using sensor data from PIR and thermopile we showed that we can accurately localize and track persons in real-time. Our sensor system uses 80 mW for estimating the location and height. Using thermopile we removed the background hot objects and ambient heat noise to detect human. Features like gain, peak to peak voltage, pixels traversed were used in knn based classification to estimate the distance of target from sensor. Which makes our system agnostic to the clothes worn, compared to the literature. In this work, we focused on joint height and localization with a single platform. The distance estimated in location estimation was used along with number of pixels covered to estimate the height of the target. This system achieves 50% of the times < 22 cm accuracy and 80% of the times < 35 cm compared to 13.5 cm best case accuracy. The height estimation is within 8 cm in majority cases.

6.2 Future Work

Although this work achieves localization and height estimation with higher accuracy there is still much research that can be done to improve the accuracy of the system. The following are some of the areas where we feel improvements can be made:

1. **Other Applications** - Since the system has thermopile array sensor, it can easily be used in other applications like activity recognition, improve the occupancy detection for complex scenarios, for elderly health care and monitoring like fall detection etc.
2. **Multiple people localization** - As mentioned early, our system can localize only one person. However, it would be better if multiple people can be located and tracked.
3. **Efficient implementations of neural networks** - There are some modified neural networks which claim to be efficient on low power micro controllers. These techniques can be implemented and tested to see if the accuracy increases.
4. **Flexible system** - Now the system has fixed PIR and thermopile array sensors. However, it can be made more flexible by giving system users to choose the thermopile array sensor of their requirement. Also to give the user an option to vary the range of detection.

Bibliography

- [1] Chandrayee Basu and Anthony Rowe. Tracking motion and proxemics using thermal-sensor array. *CoRR*, abs/1511.08166, 2015.
- [2] Alex Beltran, Varick L. Erickson, and Alberto E. Cerpa. Thermosense: Occupancy thermal based sensing for hvac control. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, BuildSys'13, pages 11:1–11:8, 2013.
- [3] Hana Godrich ; Alexander M. Haimovich ; Rick S. Blum. Target localization accuracy gain in mimo radar-based systems. In *IEEE Transactions on Information Theory*, volume 56, pages 2783 – 2803. IEEE, 2010.
- [4] J. HvizdoÅa ; J. VaÅaÅŃÅak ; A. Brezina. Object identification and localization by smart floors. In *Conference on Intelligent Engineering Systems (INES)*. IEEE, Sept. 2015.
- [5] J Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698, June 1986.
- [6] G. Cerutti, B. Milosevic, and E. Farella. Outdoor people detection in low resolution thermal images. In *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)*, pages 1–6, June 2018.
- [7] Wei-Han Chen and Hsi-Pin Ma. A fall detection system based on infrared array sensors with tracking capability for the elderly at home. In *2015 17th International Conference on E-health Networking, Application Services (HealthCom)*, pages 428–434, Oct 2015.
- [8] Z. Chen, Y. Wang, and H. Liu. Unobtrusive sensor-based occupancy facing direction detection and tracking using advanced machine learning algorithms. *IEEE Sensors Journal*, 18(15):6360–6368, Aug 2018.
- [9] Designspark. engineers-guide-to-using-pir-sensors-for-highly-reliable-security-applications, 2019. [Online; accessed 10-May-2019].
- [10] Espressif. Espressif esp32, 2019.
- [11] FLIR. Lepton lwir micro thermal camera module, 2019. [Online; accessed 01-April-2019].
- [12] H. Gami. Movement direction and distance classification using a single pir sensor. *IEEE Sensors Letters*, 2(1):1–4, March 2018.
- [13] Luis Ignacio Lopera Gonzalez, Marc Troost, and Oliver Amft. Using a thermopile matrix sensor to recognize energy-related activities in offices. *Procedia Computer Science*, 19(Seit):678–685, 2013.
- [14] R.A. Haddad and A.N. Akansu. A class of fast gaussian binomial filters for speech and image processing. *Trans. Sig. Proc.*, 39(3):723–727, Mar. 1991.
- [15] Rongbing Li ; Jianye Liu ; Ling Zhang ; Yijun Hang. Lidar/mems imu integrated navigation (slam) method for a small uav in indoor environments. In *2014 DGON Inertial Sensors and Systems (ISS)*. IEEE, 2014.

- [16] Q. Hao, F. Hu, and Y. Xiao. Multiple human tracking and identification with wireless distributed pyroelectric sensor systems. *IEEE Systems Journal*, 3(4):428–439, Dec 2009.
- [17] Moustafa Youssef Heba Aly. New insights into wifi-based device-free localization. In *UbiComp '13 Adjunct Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 541–548. ACM, 2013.
- [18] Heimann. Heimann sensors, 2019. [Online; accessed 01-April-2019].
- [19] Jie Yang Simon Sidhom Yan Wang Yingying Chen Fan Ye Hongbo Liu, Yu Gan. Push the limit of wifi based localization for smartphones. In *Mobicom '12 Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 305–316. ACM, 2012.
- [20] Takashi Hosono, Tomokazu Takahashi, Daisuke Deguchi, Ichiro Ide, Hiroshi Murase, Tomoyoshi Aizawa, and Masato Kawade. Human tracking using a far-infrared sensor array and a thermo-spatial sensitive histogram. In C.V. Jawahar and Shiguang Shan, editors, *Computer Vision - ACCV 2014 Workshops*, pages 262–274. Springer International Publishing, 2015.
- [21] Texas Instruments. Lpv802 op-amps, 2019.
- [22] J. Kemper and D. Hauschildt. Passive infrared localization with a probability hypothesis density filter. In *2010 7th Workshop on Positioning, Navigation and Communication*, pages 68–76, March 2010.
- [23] Wu Kong. Wukong github, 2019. [Online; accessed 01-April-2019].
- [24] M. Kuki, H. Nakajima, N. Tsuchiya, and Y. Hata. Human movement trajectory recording for home alone by thermopile array sensor. In *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2042–2047, Oct 2012.
- [25] Masato Kuki, Hiroshi Nakajima, Naoki Tsuchiya, Junichi Tanaka, and Yutaka Hata. Multi human movement trajectory extraction by thermal sensor. In Yong Soo Kim, Young J. Ryoo, Moon-soo Jang, and Young-Chul Bae, editors, *Advanced Intelligent Systems*, pages 35–49. Springer International Publishing, 2014.
- [26] S. Lee, K. N. Ha, and K. C. Lee. A pyroelectric infrared sensor-based indoor location-aware system for the smart home. *IEEE Transactions on Consumer Electronics*, 52(4):1311–1317, Nov 2006.
- [27] Kejie Qiu ; Fangyi Zhang ; Ming Liu. Let the light guide us: Vlc-based localization. In *IEEE Robotics & Automation Magazine*, volume 23, pages 174 – 183. IEEE, 2016.
- [28] S. Mashiyama, J. Hong, and T. Ohtsuki. Activity recognition using low resolution infrared array sensor. In *2015 IEEE International Conference on Communications (ICC)*, pages 495–500, June 2015.
- [29] Melexis. Mlx90640 32x24 ir array sensors, 2019. [Online; accessed 01-April-2019].
- [30] Melexis. Mlx90640 ir array sensors, 2019. [Online; accessed 01-April-2019].
- [31] Microchip. Microchip atsaml21 based flash microcontroller (mcu), 2019.
- [32] B. Mukhopadhyay, S. Sarangi, S. Srirangarajan, and S. Kar. Indoor localization using analog output of pyroelectric infrared sensors. In *2018 IEEE Wireless Communications and Networking Conference (WCNC)*, pages 1–6, April 2018.
- [33] Bodhibrata Mukhopadhyay, Seshan Srirangarajan, and Subrat Kar. Modeling the analog response of passive infrared sensor. *Sensors and Actuators A: Physical*, 279:65 – 74, 2018.
- [34] Sujay Narayana, R Venkatesha Prasad, Vijay S Rao, Tamma V Prabhakar, Sripad S Kowshik, and Madhuri Sheethala Iyer. Pir sensors: Characterization and novel localization technique. In *Proceedings of the 14th international conference on information processing in sensor networks*, pages 142–153. ACM, 2015.

- [35] H. M. Ng. Poster abstract: Human localization and activity detection using thermopile sensors. In *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, pages 337–338, April 2013.
- [36] Michele Pace. Multi-target tracking with phd filters, 2019. [Online; accessed 01-April-2019].
- [37] Panasonic. Panasonic infrared array sensor, 2019. [Online; accessed 01-April-2019].
- [38] Dongning Qu, Bo Yang, and Nanhao Gu. Indoor multiple human targets localization and tracking using thermopile sensor. *Infrared Physics and Technology*, 97:349 – 359, 2019.
- [39] RedSearch. Thermopile array sensor, 2019. [Online; accessed 10-May-2019].
- [40] Chi-Sheng Shih, Tim-Yuao Wang, Jyun-Jhe Chou, Ze-Yu Chuang, Ching-Chi Chuang, Kwei-Jay Lin, Wei-Dean Wang, and Kuo-Chin Huang. Collaborative sensing for privacy preserving gait tracking using iot middleware. In *Proceedings of the International Conference on Research in Adaptive and Convergent Systems, RACS '17*, pages 152–159, 2017.
- [41] Lionel M. Ni; Dian Zhang; Michael R. Souryal. Rfid-based localization and tracking technologies. *IEEE Wireless Communications*, 18(2):45 – 51, 2011.
- [42] SparkFun. Sparkfun ir array breakout - 55 degree fov, mlx90640 (qwiic), 2019. [Online; accessed 01-April-2019].
- [43] A. Tyndall, R. Cardell-Oliver, and A. Keating. Occupancy estimation using a low-pixel count thermal imager. *IEEE Sensors Journal*, 16(10):3784–3791, May 2016.
- [44] Tianye Yang, Xuefeng Liu, Shaojie Tang, Jianwei Niu, and Peng Guo. Push the limit of pir sensor based localization. 2019.
- [45] Zilog. Zmotion pyroelectric sensors. product specification, 2019. [Online; accessed 01-April-2019].