LOCI: Privacy-aware, Device-free, Low-power Localization of Multiple Persons using IR Sensors

S. Narayana*, V. Rao*, R. V. Prasad*, A. K. Kanthila*, K. Managundi*, L. Mottola#, T. V. Prabhakar##
{Sujay.Narayana,V.Rao}@tudelft.nl,rvprasad@acm.org,luca.mottola@ri.se,tvprabs@iisc.ac.in
*TU Delft, the Netherlands #RISE SICS, Sweden ##IISc, India

ABSTRACT

High accuracy and device-free indoor localization is still a *holy grail* to enable smart environments. With the growing privacy concerns and regulations, it is necessary to develop methods and systems that can be low-power, device-free as well as privacy-aware. While IR-based solutions fit the bill, they require many modules to be installed in the area of interest for higher accuracy, or proper planning during installation, or they may not work if the background has multiple heat-emitting objects, etc. In this paper, we propose a custom-built miniature device called LOCI that uses IR sensing. One unit of LOCI can provide three-dimensional localization at best. LOCI uses only a thermopile and a PIR sensor built within a 5x5x2 cm^3 module. Since IR-based sensing is used, LOCI consumes around 80 mW. LOCI uses analog waveform from the PIR sensor with the gain of the PIR sensor dynamically controlled through software in real-time to simulate spatial diversity. LOCI proposes low-complexity techniques with sensor fusion to eliminate the noise in the background, which has not been handled in previous works even with sophisticated signal processing techniques. Since LOCI uses raw data from the thermopile, the computations are power-efficient. We present the complete design of LOCI and the proposed methodology to estimate height and location. LOCI achieves accuracies of sub-22 cm with a confidence of 0.5 and sub-35 cm with a confidence of 0.8. The best-case location accuracy is 12.5 cm. The accuracy of height estimation is within 8 cm in majority cases. LOCI can easily be extended to recognize activities.

KEYWORDS

PIR, thermopile, infrared, device-free, passive, localization, infrared, privacy-aware

1 INTRODUCTION

Location-based context-aware services constitute an important ingredient of smart indoor environments, such as assisted living spaces (ALS) and smart buildings. Meanwhile, privacy-aware technologies are gaining more prominence than before with the introduction of the General Data Protection Regulation (GDPR) by the European Union [1]. Due to these regulations, localization techniques that do not contain any user identifiable information are preferred.

This eliminates the use of cameras, tags, smartphones or wearables in conjunction with radio-frequency (RF) based technologies such as RFID, Wi-Fi, Bluetooth and ultra-wideband (UWB) [2]. Moreover, smart semi-public buildings that are safety-critical, such as banks or airports, induce an additional constraint that localization systems must work without any dependency on infrastructure, in order to be operable in case of emergency situations such as hostage crisis. Thus there is a demand for low-power and privacy-aware systems that do not necessitate user involvement.

Figure 1: LOCI is of size 5x5x2 cm^3 housing thermopile and PIR sensors.

Device-free localization techniques that consider no tags, wearables, or smartphones with the user are well-suited for this class of applications. Several techniques using Wi-Fi, LiDARs, acoustics, ultrasound and passive infrared (IR) based techniques have been proposed [2, 3]. In general, these techniques trade-off either cost of deployment or energy-efficiency for accuracy.

Why IR? Of the device-free techniques, IR based techniques have shown the potential to be low-complexity, low power and have good accuracy of ± 65 cm [4].[1] Motion sensors, essentially IR sensors, are becoming prevalent in older buildings as the first step towards becoming smart; these sensors invariably will be part of future smart buildings as well. Instead of just using them as human presence detectors, the same sensors can be leveraged to provide user location information to aid, for instance, smart HVAC systems within the buildings to save energy. This would result in better offerings for infrastructure services and cost savings as another system/infrastructure for localization is not required.

Constraints. While IR sensors, such as thermopiles and passive infrared (PIR), seem ideal candidates to localize humans, they are, however, accompanied by several constraints. (a) IR radiation is also emitted by other hot objects including pets, lights and hot water bottles. (b) The radiation intensity is quite weak. This is aggravated if the user wears warm clothing. (c) The signal characteristics in PIRs are affected by the direction of movement, distance to the sensor and movement speed of the user. (d) Either a single thermopile or PIR cannot determine coordinates of a single user. (e) It is impossible to always detect if there are multiple people by just analyzing PIR signals, however, people counting up to a reasonable accuracy can be achieved using thermopiles.

State of the art falls short. Several systems and localization techniques have been developed that deal with the above constraints. However, they exhibit one or more of the following shortcomings.

[1] Higher accuracies can be obtained by trading off cost of deployment.

978-1-7281-5497-8/20/$31.00 ©2020 IEEE
DOI 10.1109/IPSN48710.2020.00018

(1) *Unscalable hardware.* Many proposed systems require installation of a few sensor boxes to localize with reasonable accuracy [3–5]. Moreover, as in [3] each PIR can be tuned only to a particular single range, thus it is necessary to carefully plan the deployment of multiple PIRs for each space.
(2) *Locating multiple persons.* Existing systems do not offer localization of more than one person with thermal sensors. The systems, proposed hitherto, cannot be used for localization in these cases except for people counting [6].
(3) *Inflexible.* Some works create zones that may differ from room to room; some would require multiple towers placed at prespecified angles and so on [3]. Furthermore, with thermopile arrays, many works choose to deploy their module on the ceiling in order to get a planar view of the floor for localization [6]. The accuracy of these solutions depend on the height of the ceiling, and have an increased cost of deployment to cover large areas due to limited FoV. These make the systems inflexible. Further, this type of deployment can only count people.

Our proposal. In this paper, we propose a novel localization platform, LOCI, that addresses these shortcomings by innovatively leveraging the constraints to our benefit as outlined below. LOCI is a combination of innovative and novel hardware and software modules, as shown in Figure 1.

Unlike existing systems that need multiple sensors (thermopiles or PIRs) to be carefully deployed for obtaining the user's position, we intend to deploy one LOCI per room to compute user's location with high accuracy. Our innovation is that we employ PIR as a depth sensor in conjunction with a thermopile (Melexis 32×24 thermopile array). Such a combination of sensors also enables us to localize multiple people. While this system is extensible to other applications such as user movement tracking, fall detection (for ALS), and height estimation, in this work we focus only on localization of people (tested up to three people). However, we present some results for height estimation too. Indeed fall detection is a simple extension once we estimate the height in realtime; it is done by identifying sudden height change in two successive frames.

We show that the output of a PIR sensor is a function of the user's distance, speed, direction, and the gain of the amplifier; the implicit assumption that gain cannot be changed, negatively impacts the accuracy. Another innovation of our system is that we can dynamically control the gain of the PIR sensor through software. This eliminates the need to create zones and enables us to achieve higher accuracies, even in the presence of multiple people.

A common strategy for ranging with PIRs is using set of curves of output voltage with respect to various distances, also per zone if zones are being created, during deployment [3]. However, such a strategy fails as the set of curves need to be obtained in every location. We, therefore, leverage machine learning to train once and deploy in any indoor location.

Challenges. LOCI aims to achieve significant progress in state of the art, but has to clear several technical hurdles. As we intend to deploy only one system per room, it is significantly difficult to achieve a better localization accuracy in three dimensions even with our fused sensor data from thermopiles and PIR.

- With the PIR being used as a depth sensor, it is extremely difficult to discern a fast-moving person nearby from a slow-moving person at a farther distance – even when there is only one person in the field of view. This is because the incident IR radiation falling on the PIR can be similar.
- Low-power and real-time computation – the output from thermopile array requires heavy floating-point computation to convert to usable temperature values. Higher the number of pixels, higher is the required processing. On the other hand, the temperature value from an individual element depends on its placement in the thermopile array. The challenge is to directly use the thermopile output.
- To use the low power computing platform the sampling rate needs to be low to avoid heavy computations. However, the location of the person should be computed when the person is present at that spot rather than when (s)he is gone, as in [3], which negates the whole purpose.
- Thermopile is extremely sensitive to slight disturbances in ambient heat. Thus removing the background hot objects and ambient heat noise is mandatory for effective localization and height estimation.

Contributions. We summarize our contributions as follows:

(1) This is the first work to achieve localization using two different types of infrared (IR) sensors by sensor fusion. Furthermore, this is also the first work to localize multiple people using IR sensors with ease of deployment.
(2) We follow pre-processing steps including eliminating background noise due to hot objects (e.g., candle, monitors, lights etc.) other than people, which has not been done in literature in the context of IR sensors. Unlike in literature, our algorithm is agnostic to the clothing worn (e.g., jackets, full-clothing, and semi-clad).
(3) We make all the computation on the onboard low-power microcontrollers (MCU), including handling thermopile data and the machine learning solution. To handle data from thermopile, we derive equations that enable us to utilize the raw values from thermopiles, unlike other works. We employ k-NN based machine learning technique to make our solution to work seamlessly in most indoor locations.
(4) We demonstrate the accuracy of 12.5 cm in the best case and less than 35 cm in the worst case in a large area of about 72 m^2.
(5) A miniaturized low power and low form factor and high FoV system that can be easily mounted on a wall.
(6) Use of variable gain at PIR sensor to avail the spatial diversity gain that helps in avoiding multiple PIR sensors (like in [3]) without complex deployment.

2 RELATED WORK

Localization has been a very active area of research for decades now, including privacy-aware, device-free techniques. Solutions using WiFi channel state information (CSI) [11], mmWave based [12], ultra-wideband [13] and visible light based [14] do fall in this category. As argued in Section 1, these methods require to trade-off cost of deployment and/or energy efficiency for accuracy. For instance, at least two devices are required to measure Wi-Fi CSI, one to transmit and the other to receive and process [11]. A similar approach is necessary for mmWave [12], ultra-wideband [13] and visible light

Table 1: State of the art that use thermopiles and PIR sensors for human detection, tracking, and localization.

Work	Type of sensor/s	Coverage area	Objective	No. of people	Accuracy	Notes on deployment
Narayana et al. [3]	8 PIRs	8m x 8m	Localization + Height classification	1	30 cm	Two sensor towers, having 4 sensors each, are placed at 90°
B. Mukhopadhyay et al. [4]	4 PIRs	7 m x 7.5 m	Localization	1	65 cm	4 PIRs placed on the edge of the experimental arena
W. Chen et al. [7]	2 Thermopiles	2.35 m x 3 m	Fall detection + Localization	1	13.39 cm [2]	Two sensors attached to wall at 30°
C. Basu et al. [6]	1 Thermopile	2.5 m x 2.5 m	Occupancy detection + Tracking	5 Occupancy, 1 tracking	22 cm	Ceiling mounted
Z. Chen et al. [8]	1 Thermopile	2.4 m x 1.2 m	Tracking	1	19 cm [3]	Thermopile placed on a motor across the wall
J. Kemper et al. [9]	4 Thermopiles	4.9 m x 6.2 m	Localization + Tracking	2	25 cm	4 sensors placed in 4 corners
M. N. Hock [5]	5 Thermopiles	4.6 m x 2.7 m	Human detection	1	50 cm	2 sensors on one side and 3 on adjacent side
X. Liu [10]	4 Thermopiles	7 m x 7 m	Localization	1	63 cm [4]	4 sensors placed in 4 corners
LOCI	1 Thermopile + 1 PIR	9 m x 8 m	Localization + Tracking + Height classification	3	12.5 cm	Single device mounted 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.

based systems [14]. Furthermore, none of these techniques are low-power solutions as the entire system's power consumption will be in the order of several 100s of mW.

On the other hand, IR based systems are low-power (typically in the order of μW) and simple to use systems but with lower accuracies, typically sub-meter, than other systems mentioned above (e.g., visible light based [14]). These accuracy levels are sufficient for smart building applications. Moreover, the ubiquity of IR sensors in buildings and their cost (USD 38 for thermopiles) make them a very attractive choice for developing localization system.

Table 1 summarizes the most important works on localization using IR sensors. Narayana et al. presented a novel sensor tower containing four collocated PIR sensors that can perform height classification and localization of moving objects [3]. They 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° localize the moving warm object. Mukhopadhyay et al. improved this work 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 [4]. Their algorithm exploits multi-lateration and Support Vector Regression (SVR) based techniques to fulfil the objective. Chen et al. make use of two thermopile array sensors, placed at 30°angles, and 3.3 m away for tracking elderly and to detect fall [7]. The location of the human is obtained using the angle of arrival (AOA) from each sensor. Occupancy detection and tracking of people within an FoV of 2.5 m x 2.5 m is reported in [6]. Tracking is performed using Support Vector Machine (SVM) classification on connected component based features of the thermopile data.

Z. Chen et al. propose an activity recognition and human tracking system by using a low pixel infrared thermopile array [8]. They present two feature extraction methods: manually-defined, and a pre-trained convolutional neural network (CNN) model. An implementation of Probability Hypothesis Density (PHD) filter is presented by J. Kemper et al. to perform localization and tracking using 4 thermopile array sensors. 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 [5]. In total, 5 thermopile sensors require to localize a person in 4.6 m x 2.7 m area. C. Shih et al. designed a ceiling-mounted thermopile array sensor network to track gait of a moving person [15]. The tracking algorithm is developed on a virtual run-time library called WuKong.

S. Lee et al. present a location-recognition system called PIR sensor-based indoor location-aware system (PILAS) [16]. Different sensing areas are created using 12 PIR sensors which are placed on the ceiling. A threshold was set to turn the PIR sensor On or Off depending on human movement. Using a ceiling mount 8 x 8 array thermopile sensor, D. Qu et al. present a system to perform localization and tracking [17]. The algorithm utilizes the Kalman filter to track a person in the detection region. X. Liu et al. introduce a concept of 'azimuth' change that adopts a 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 [10]. Unlike LOCI, this concept requires more than single sensor placed at pre-defined locations, hence, not flexible in deployment. Moreover, all the sensors must be at the line of sight.

3 DESIGN OF LOCI

3.1 Design requirements

We set out with the following design requirements for LOCI that would make the solution easily usable across buildings and use-cases. **D1**: The solution must be one piece of hardware that is convenient to deploy in a room. **D2**: The solution must be able to localize people in the room; it must provide a 2D localization and 3D localization at best. **D3**: The solution must work despite changing backgrounds, warm objects, clothing of the users **D4**: All the processing must be done in the box. Only the results can be communicated outside using WiFi or another radio. **D5**: The solution must be low-power and energy-efficient **D6**: The solution must be deployable easily across rooms without requiring any zone creations/data collection.

3.2 The hardware

Given the design requirement that we can have only one box of sensors, we should be able to capture where the user is and how far is (s)he from the box. 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 called **LOCI**. The proposed hardware platform

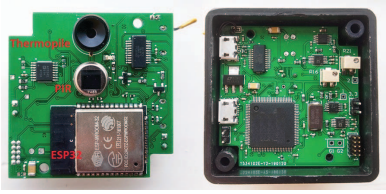


Figure 2: The proposed hardware system

Table 2: Comparison of different commercial thermopile sensors.

	HTPA80x64d	MLX90621	MLX90640	MLX90641	Grid-Eye
Manufacturer	Heimann	Melexis	Melexis	Melexis	Panasonic
Resolution	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
Approximate cost	$315	$36.74	$37.48	$40.40	$17.54

* OMRON, a leading manufacturer of thermopile array sensors such as D6T-44L-06 discontinued most of its products.

that houses two collocated infrared sensors – a thermopile array sensor and a PIR sensor placed one above the other as shown in Figure 2. The distance between the center of thermopile and PIR is fixed to be around 1 cm so that the FoV of the thermopile is not blocked by the PIR sensor.

The overall dimension of the LOCI platform is 5 x 5 x 2 cm^3 and is developed in-house. The hardware also includes an ultra-low-power ARM Cortex M0 MCU ATSAML21J18B [18], from Atmel's pico power series MCUs, for acquiring data from the infrared sensors and process them. The data comprising of detection, localization, and height of the person is transmitted to a central server using an ESP32 WiFi module [19]. We outline the sensor selection and variable gain amplifier in greater detail in the following subsections.

3.2.1 Selecting suitable sensors. We outline the reasoning for the selection of the two sensors – thermopile and PIR.

Thermopile array sensor. To cover a reasonably large room (say 8×8 m^2), we would need a large field of view (FoV) – a minimum of 90° FoV is required to avoid blind spots in a horizontal direction when the sensor platform is placed at the corner in a room – and good number of pixels for identifying location in one plane. As human movement would not be at very high speed, a refresh rate of 8 – 16 Hz would suffice. There are various commercial thermopile array sensors available from manufacturers such as Panasonic, Melexis, and Heimann Sensor GmbH, which we list in Table 2 with relevant specifications. While sensor models, such as FLIR Lepton provide high-resolution imaging and fall under thermal camera category, they need high operating and processing power and are quite expensive (≈$250). We choose MLX90640ESF-BAA FIR sensor from Melexis in our platform as it provides large FoV of 110°x 75°, good resolution of 32 x 24, a decent refresh rate of 64 Hz, and cost of $38 [20]. Furthermore, the sensor can measure object temperature between -40°C to 300°C with a frame accuracy of ±1°C.

PIR sensor. We employ Zilog's ZSBG446671, a dual-element PIR sensor in our system because of its large FoV of 132° x 222° [21]. The

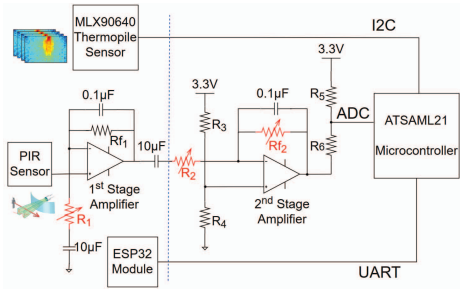


Figure 3: Circuit diagram of our system

detection area of a PIR sensor element is small and very sensitive to the infrared energy. To strengthen the incoming infrared rays, we placed a *Fresnel* lens in such a way that the center of the PIR sensor coincides with the focal point of the lens. 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, that containing multiple spot Fresnel lenses on its circumference and concentrates the incoming raysonto the PIR sensor elements. The FoV of the chosen golf ball lens is 150°x 150°, larger than that of the thermopile. Thus, 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 the thermopile sensor to meet the Nyquist criterion.

3.2.2 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 an MCU. While most of the work in the literature that use PIR involve working with binary output as provided in the vendor reference circuits [10], there are a few that utilize the PIR output in analog form [3, 22]. We make use of the analog signals from a single PIR, read by the ADC pin of the MCU, to estimate the distance of the moving object from the sensor. The traditional way, as found in the literature, is to use fixed gain amplifier stages by selecting suitable feedback resistors. On the contrary, in this work, we facilitate the MCU to vary the gain of each amplifier stage using digital potentiometers that are controlled using I2C lines[5].

Amplification gains of COTS amplifiers in one stage will not be sufficient, leading to two-stages. We therefore select Texas Instrument's LPV802 dual channel nano-power amplifier in our design as they consume only 1 μW. The circuit diagram of our hardware platform is shown in Figure 3.

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

$$V_o = -V_{in}\left(1 + \frac{Rf_1}{R_1}\right)\left(\frac{Rf_2}{R_2}\right), \quad (1)$$

where Rf_1 is fixed to 3 MΩ, Rf_2 is a 1 MΩ dual channel digital potentiometer, AD5242BRUZ1M from Analog Devices, whose resistance can be varied in 255 steps between 0 and 1 MΩ. We connect both the

[5]It 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.

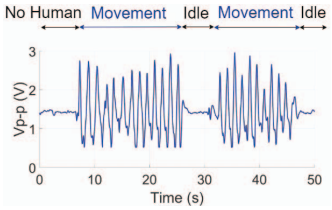


Figure 4: Amplified output from the PIR

resistor channels in series to get a broader range of up to 2 MΩ. Similarly, R_1 and R_2 are 512 kΩ digital potentiometers, AD5272BRMZ from Analog Devices, that can be varied in 1024 steps between 0 and 512 kΩ. By adjusting these resistor values dynamically, our algorithm is able to vary the overall gain of the PIR output between $\approx 2\text{x}10^{-6}$ to $6\text{x}10^{12}$ in $\approx 537\text{x}10^{6}$ 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 MCU 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, to scale and shift the PIR output voltage before fed to the ADC pin of the MCU. We fix R_3 = 510 kΩ, R_4 = 240 kΩ, R_3 = 10 kΩ, and R_3 = 2.4 kΩ to get a full PIR output swing between 0.2 V and 3.25 V. A sample amplified output from the PIR sensor is shown in Figure 4.

3.3 The software

Capturing the signals from the sensors can easily be programmed. In this section, we look at clearing the hurdles to compute location information on the onboard MCU.

3.3.1 Efficient processing of thermopile data. 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 [20],

$$T_{o(i,j)} = \sqrt[4]{\frac{V_{IR(i,j)_{COMP}}}{\alpha_{comp(i,j)} * (1 - K_{sTo2} * 273.15) + S_{x(i,j)}} + T_{a-r}} - 273.15, \quad (2)$$

where,

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

$T_{o(i,j)}$ is the temperature reading for pixel $i, j, \forall i \leq 32$, and $j \leq 24$, i,j$\in N$, $V_{IR(i,j)_{COMP}}$ is the offset compensated raw value for each pixel, $\alpha_{comp(i,j)}$ and K_{sTo2} are constants corresponding to each pixel, and T_{a-r} varies with the ambient temperature. These parameters are calculated using various constants and pixel offset values[6] that are stored in the EEPROM of the sensor.

The calculation indicated in Equation 2 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 [23]. Further, running the localization algorithm and wireless data transmissions require high power micro-controllers as the host

[6]Each pixel will be provided with a correction factor when the manufacturer calibrates the sensor.

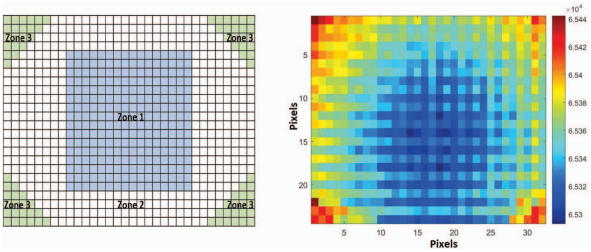


Figure 5: MLX90640ESF-BAA frame is divided into three zones depending on the measurement accuracy [20].

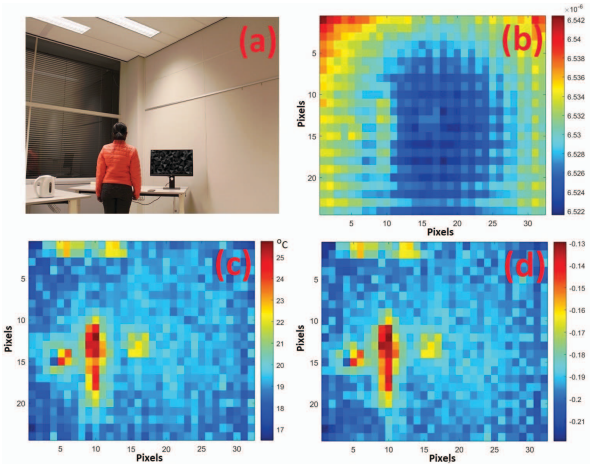


Figure 6: (a) Image with warm objects such as human, incandescent light, hot kettle, monitor (b) Raw values output from the thermopile sensor (c) Absolute temperature obtained using Equation 2 (d) Raw compensated value calculated using Equation 3.

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 computing the absolute temperature data.

The MLX90640ESF-BAA frame, containing 768 pixels, is divided into three zones, Zones 1 to 3, 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 5. For an object in front of the sensor with temperatures between 0°C and 50°C, Zone 1 has the highest accuracy of ±0.5°C, Zone 2 with ±1°C and Zone 3, the least with ±2°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 2. Due to the paucity of space, we refer the readers to [20] for details regarding the calculation of these parameters. We analyzed the range of $V_{IR(i,j)_{COMP}}$, $\alpha_{comp(i,j)}$, K_{sTo2}, and T_{a-r} for $T_{o(i,j)} \in [-20°\text{C}, 125°\text{C}]$. $V_{IR(i,j)_{COMP}}$ has the range $[-79, 427]$, $\alpha_{comp(i,j)}$ in the range $[-3\text{x}10^{-8}, 1.23\text{x}10^{-7}]$, K_{sTo2} is in the order of $-2\text{x}10^{-4}$, and T_{a-r} in the range $[4\text{x}10^{9}, 2.3\text{x}10^{10}]$ for ambient temperatures between 0° C and 100° C. Looking at the

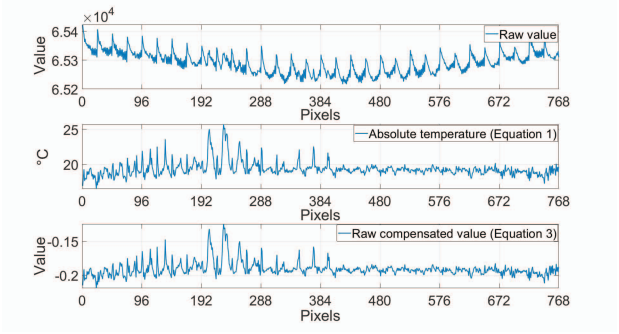


Figure 7: Values represented by all 768 pixels in a frame for the images displayed in Figure 6.

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

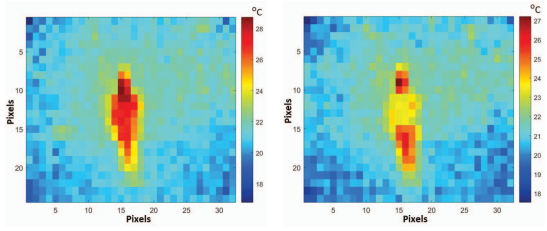
$$RAW_{(i,j)comp} = \left(\frac{V_{IR(i,j)_{COMP}}}{\alpha_{comp(i,j)}} + T_{a-r} \right) \text{x} 10^{-9}, \tag{3}$$

where $RAW_{(i,j)comp}$ is the compensated (for offset and irregularities in zones) raw value which is of much lower complexity and can be run on our onboard MCU.

We evaluate how good the gradients are preserved by our simplified raw value from Equation 3 as compared to the absolute temperature gradients as obtained by Equation 2. Figure 6 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; and (d) the same frame after raw values are processed using our technique. 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, and found that our approach indeed preserves the gradients.

Figure 7 shows the values represented by each pixel (Pixel-1 corresponds to top left of the frame and incremented column-wise) for the frame displayed in Figure 6. We observe in the figure that raw compensated value varies with the same pattern as of absolute temperature. Complete processing is done on our hardware platform – reading EEPROM data and acquiring raw data from the thermopile sensor at 8 Hz, reading PIR data at 16 Hz, execute our algorithm, and transfer the data over WiFi to a centralized server. Recall that we used an ultra-low-power MCU ATSAML21J18B, executing at 16 MHz with 32 kB RAM, and consuming ≈2 mA current. The maximum power consumption of the platform is around 80 mW without WiFi transmission, and around 460mW when WiFi chip is transmitting at 0 dBm. The power can still be reduced if the transmission power is reduced, or different protocols such as Bluetooth is used for communication.

3.3.2 Localization. The algorithm for computing the location of multiple persons is presented in detail in Section 5. Before we proceed to the algorithm, we present the characterization of the sensors, which forms the basis for the design of our algorithm.



(a) Human with thin cloth **(b) Human with thick jacket**

Figure 8: Temperature recorded by MLX90640 for a human at 2 m from the sensor with and without jacket

4 CHARACTERIZATION OF SENSORS

4.1 Thermopile sensor

From our experiments with thermopile on LOCI, we made the following observations that can be utilized in our algorithm to perform localization and height estimation.

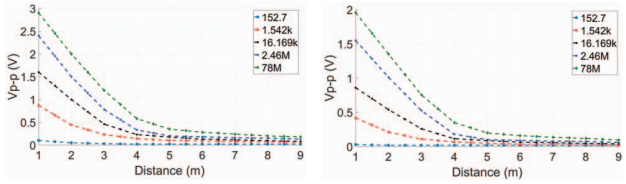
(1) As the distance between the warm body and the sensor increases, the temperature read by the sensor decreases. This holds with the theoretical model provided in [5] that says $T \propto \frac{1}{d^2}$, where T is the temperature recorded by a pixel and d is the distance of the warm body from the sensor.
(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 spatial-temporal changes in the number of pixels covered by the 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,

$$P_t(h, v) \propto \frac{s}{d}, \tag{4}$$

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 by the sensor and is considered as part of the human.
(6) A human can be detected even if (s)he is wearing thick clothing such as a jacket (see Figure 8).
(7) Head, chest and waist are the parts of the human body that exhibits high gradients compared to the rest of the body.

4.2 PIR sensor

Considering the generic models from [3] and [24], we can portray the relationship between the peak to peak voltage V_{p-p} generated



(a) Speed 0.5 m/s (b) Speed 1.2 m/s

Figure 9: Peak to peak voltage generated by the PIR sensor for different gains at different speeds.

by the PIR sensor and the amplifier gain G as,

$$V_{p-p} \propto \frac{I\,G}{s^2\,d^2}, \tag{5}$$

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. We characterized the V_{p-p} against different speed, distances and gains. In Figure 9, we show the V_{p-p} 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 speed. Similarly, as the movement speed increases at a constant distance, the gain has to be increased to observe the same levels of signal.

From our other experiments, we made the following observations that can be utilized in our algorithm.

(1) As the distance of the person from the sensor increases, the peak to peak voltage generated by the amplifier stages decrease, and vice versa, provided that the speed, body temperature and the amplifier gain remains approximately the same.
(2) Provided that a person moves at a specific 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 (see Figure 9a and Figure 9b).
(3) The peak to peak voltage output from the amplifier stages decreases as the speed of the movement at a specific distance increases (see Figure 9a and Figure 9b).
(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 $s_1 > s2$. This is because the duration and amount of IR rays falling on the sensor reduce as the speed of the moving object increases. Similarly, converse also holds.
(5) Static hot objects such as hot kettle, computer, light bulbs do not affect the output of the PIR sensor.

5 LOCALIZATION

In this section, we explain how the features from the thermopile and PIR sensor presented in Section 4 can be exploited to perform localization. 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. As we intend to ensure seamless operations in most indoor locations without (re-)collecting data in every room, we employ machine learning techniques. Since we consider k-NN, a supervised classification method, there are two steps involved in achieving this: obtaining training data that forms the training phase, and online computation in the deployment phase. Several pre-processing steps are also employed to improve the accuracy, which are explained in this section. The flowchart of our proposed steps to localize is shown in Figure 10.

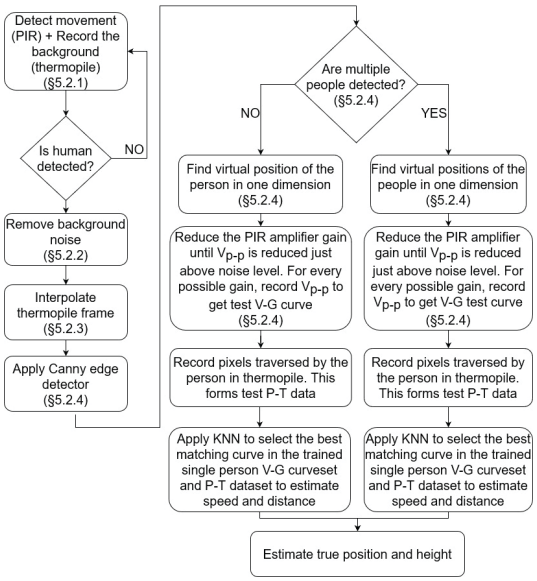


Figure 10: Flowchart of our proposed localization technique

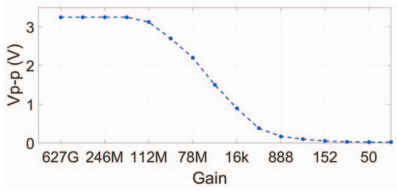


Figure 11: Voltage - Gain curve obtained for movement with speed 0.5 m/s at 2 m in front of the sensor

5.1 Training Phase

In this phase, we collect training data that comprises of curves obtained from the PIR sensor, and pixel data from thermopile array. The training data can be collected in a typical room/setting similar to deployment and need not be the target place of deployment itself. We tested LOCI in different rooms after collecting training data (see Section 6).

Curvesets from PIR sensor: We gather peak to peak voltage measurements V_{p-p} (represented by Equation 5) against all dependent parameters - all possible amplifier gains G, various speeds s, and at different distances d. We accomplish this 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) when a person is moving at a constant speed at a given distance[7]. The obtained peak to peak values V_{p-p} for various gains G are stitched together to form Voltage - Gain curve (V-G curve) (see Figure 11 for an example). Similar such curves are recorded for different people moving with speeds between 0.1 m/s and 2 m/s (we assume, 2 m/s is the maximum walking speed of human in indoor scenarios for our test cases, however, this can be extended to higher speeds if need be). Hence, a single V-G curve spans over three dimensions with speed, distance and incident infrared energy (different people with distinct clothing), that addresses all the variables in Equation 5. Since there is a clear relation between speed, distance, and the output voltage generated by the thermopile array and PIR sensor, it is possible to generate training data for some cases by interpolating between two existing training data sets [3]. However, it is tedious to perform the same for different clothing and people. We represent the training data from the PIR sensor as,

$$V_{p-p} = f_{(s,d,I)}(G), \tag{6}$$

where gain G is varied from maximum to the minimum detection level, I represents infrared energy from different people with different clothing. While direction can be a factor, it only influences the phase of the signal, not the V-G curve [3]. Hence, it has not been considered here. Moreover, the direction of the movement can be identified using the thermopile sensor.

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. 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 distances d. Hence, each P-T data is of four dimensions. The datasets from both PIR and thermopile sensor are recorded concurrently so that both the datasets represent the same event. We represent the training data from thermopile sensor as,

$$p_t(h, v) = (i, j)_{(s,d)}, \tag{7}$$

where $p_t(h, v)$ is the pixels traversed in horizontal and vertical direction with $i \leq 32$, and $j \leq 24$. While (6) represents a curve, (7) is represented by four numerical values.

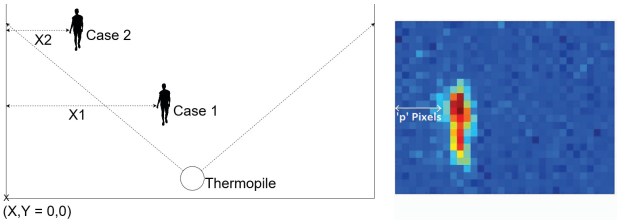
Dataset with multiple people: The training data is also collected for multiple people (tested up to 3, but can be extended) walking randomly at different speeds, direction, and distance, similar to that of a single person. This forms another set of PIR and thermopile array data. Hence, there will be two sets of PIR and thermopile training data - one for a single person, and another for multiple people.

5.2 Experiment

Once the training data is available, localization and tracking are performed with the following steps. To explain the steps, we consider a sample case wherein a human, light bulb, and a monitor is present in a room as shown in Figure 13b.

5.2.1 Detection of movement. There is no need to begin localization unless presence is detected by the PIR sensor. To capture slightest of the movements in FoV, the amplifier gain is set to the maximum.

[7]The speed is maintained ±0.1 m/s, also measured and calibrated with a smartphone app.



(a) Two cases with human in front of the sensor at two different distances **(b) Test scenario**

Figure 12: Estimation of virtual position

The human presence is indicated when $V_{p-p} > 0.02$, as 0.02 is the mean noise amplitude at the highest possible gain[8]. Simultaneously, snapshots of a frame from the thermopile sensor are taken periodically that forms the background frame. When there is a movement detection in the PIR sensor, the background estimation process is stopped. Figure 13a shows the background frame wherein light bulbs, monitor, and a hot water kettle is detected. To save power, thermopile can be duty-cycled when there is no activity detected.

5.2.2 Background and noise removal. Warm objects, such as lights, can invariably be part of the room of deployment. To be robust, we will need to eliminate these static objects as background. To this end, we utilize the background frame captured in the previous step. We represent the background frame $B_{i,j}$ as,

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

where $i \leq 32$, and $j \leq 24$, and $F_k(i, j)$ represents a single frame containing 768 pixels.

When a movement is detected (shown in Figure 13b), the background frame is subtracted from each thermopile frame that is being read. Figure 13c shows the frame from thermopile corresponding to the scenario shown in Figure 13b. The resultant background removed frame is shown in Figure 13d 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. This is usually white noise [25]. Such pixels can be neutralised using two-dimensional Gaussian filter $G(x)$ as,

$$G(i, j) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{i^2+j^2}{2\sigma^2}}, \tag{9}$$

where σ is the standard deviation, which we found out empirically to be 1.5 in our application to get better smoothing. The filtered frame is shown in Figure 13e.

5.2.3 Interpolation of thermopile data. To get better accuracy, the filtered data is interpolated once in each dimension to get a frame of size of (32 x 2) x (24 x 2) = 3072 pixels. Further interpolation leads to false detection (see next sub section). The interpolated image is shown in Figure 13f.

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

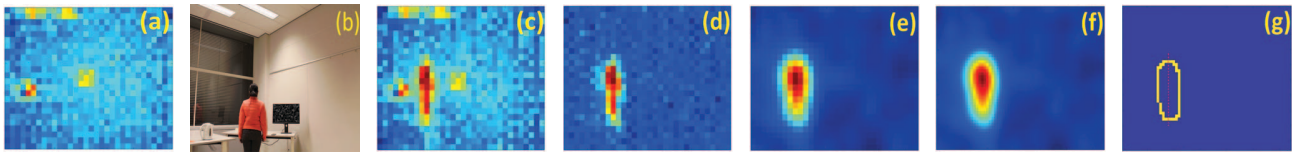


Figure 13: Steps involved in localization: (a) Background frame (b) Test scenario (c) Typical thermopile frame (d) Background subtracted image (e) Image obtained after applying Gaussian filter (f) Interpolated image (g) Image after edge detection

5.2.4 Estimation of virtual position in one dimension. The next step is to identify the position of the person in one dimension, i.e., horizontal. To perform this, the pixels forming the outline that represents the person has to be detected. This is done using Canny edge detector as it involves low complexity processing compared to other edge detection techniques, and is widely applied in various computer vision systems [26]. This is a multi-step technique that detects edges as well as suppresses noise at the same time. Figure 13g shows the position of the human in the frame after applying an edge detection algorithm. It should be noted that interpolating the frame more than once resulted in multiple closed loops for a single person. This is because of the low amplitude noise in the background multiplies with interpolation leading to false detection.

We consider the centroid of the shape as position 'X_v' (red line in the image is passing through the centroid) from Column 1 of the frame. This position is 'virtual' due to the rectangular projection that enlarges as we move farther away from the sensor. To explain this, we consider two cases: (a) the person is nearer to the sensor and (b) further away (see case 1 and case 2 in Figure 12a). Because of conical FoV of the thermopile array, the projected image on it in both the cases may be the same, i.e., the person on the frame is 'p' pixels away from the leftmost column, as shown in Figure 12b. Hence, to identify the true position of the person, it is necessary to know his distance from the sensor. We estimate this distance using the PIR sensor.

Multiple people: In the case of multiple people in the sensor FoV, more than one closed shapes are formed after passing the thermopile frame through Canny edge detector. The number of closed shapes indicates the count of the people and their virtual position is the distance between Column 1 of the frame and the centroid of the respective closed shape.

5.2.5 Distance estimation using machine learning. The next step is to find the distance of the people from the sensor platform. This is where the variable gain feature of our system is utilized. We consider two cases - one with a single person in the FoV, and the other with multiple people.

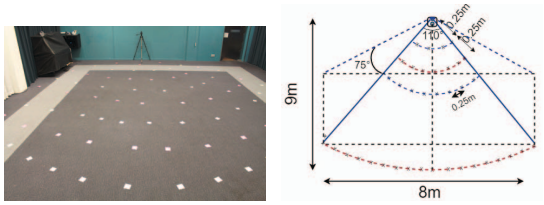
Case 1 - Single person in the sensor FoV: As soon as a human is detected (Section 5.2.1), the gain of the amplifier is reduced from the maximum to the level at which the peak to peak amplitude of the output is just above the detection level. Note that the gain can be reduced to the minimum but the outputs for gains set below the detection level contain only the noise. For each gain set in the range between the 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 pixels traversed test data (P-T data) from thermopile is recorded at the same time when the test V-G curve from PIR is computed. Now, we have a pair of V-G curve and P-T data, and a training dataset containing paired sets of V-G curves and P-T data. As explained in Section 5.1, each V-G curve in PIR training dataset has its pair in P-T data in thermopile dataset. We compare the testing V-G curve, P-T data pair with each pair in the training dataset to obtain the closest match. We perform the matching by employing machine learning technique - k-nearest neighbours (k-NN) classification [6]. First, we compute Euclidean distance between each point in testing V-G curve and each V-G curve in training data. We consider k=7 (found empirically) curves to be compared. Next, we compare the corresponding test P-T data with the data in trained P-T dataset for the number of pixels traversed. We consider the nearest match as the solution, providing distance and speed. Indeed, it is possible to estimate the location (once the range is estimated using PIR) with the subsequent thermopile images using any signal processing tools such as, least squares method, covariance of the images, or other regression techniques. However, we chose to limit the complexity of the signal processing by considering just the pixels traversed in subsequent frames. Additionally, the proposed technique provided the best accuracy and less operational overheads compared to the least squares method.

Case 2 - Multiple people in the sensor FoV: Localizing multiple people with PIRs as there is no way to know the count of people or compute their horizontal position [3]. This is the first work that realizes localizing multiple people and is done with the same complexity as of the single user case. In this case, the same procedure as in Case 1 is performed but training dataset corresponding to multiple people is considered for machine learning. The count will indicate which dataset to use. Now, with the distance of a (or each) person from the sensor is known, the virtual position 'X_v' (Section 5.2.4) can be used to find the true position 'X' using the relation,

$$X = (\text{No. of pixels covered from column 1 to } X_v) \times (2\sin(110^\circ/2) \times \text{Distance of the person from the sensor}).$$

5.3 Height classification

To obtain a 3D localization, we should also estimate the height of the person, once the person is localized in two dimensions. With the estimated range (Y) of each person from the sensor platform is known, the height of the person can be estimated using the number of pixels covered by the person in a thermopile frame in vertical direction. We know that there are 24 pixels vertically, with the thermopile sensor FoV of 75°. To obtain a right triangle, we need to bisect this angle; then we use basic trigonometric identities (tan) and divide by 12 (=24/2) to obtain P_h, the height per pixel at range Y. Thus the height 'P_h' of each pixel projected at distance 'Y' can be calculated as $\tan(75^\circ/2)/12 * Y = 0.06394 * Y$ m.



(a) Experimental setup covering the area 9 m x 8 m (b) FoV of the sensor platform

Figure 14: Experiment setup for localization and tracking

Assuming that the person is perpendicular to the FoV cone axis, the height H of the person is given by,

$$H = P_h \text{ x No. of pixels representing the person vertically,} \quad (10)$$

6 PERFORMANCE EVALUATION

We evaluated LOCI and the algorithm in real world scenarios. We also considered the factors such as background noise, static warm objects, obstacles, and more number of people, that affects the localization accuracy. In this section, we first explain our experimental setup and then present the results and observations.

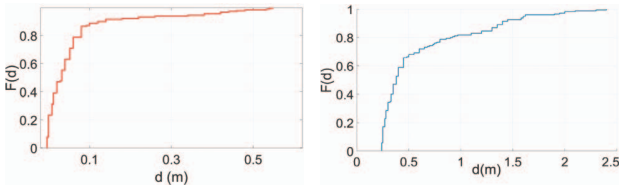
6.1 Experimental Setup

The layout of the rooms is not of a major concern for our platform. However, the number of the sensor platforms required to cover a room may differ depending on the layout and size of the room. For our experiments, we choose rectangular rooms. **Figure 14** shows one of our experimental setups with area of 9 x 8 m^2. Concentric circles from 0.25 m to 9 m with the sensor platform as the center were drawn at 0.25 m increment (see Figure 14b). Points at every 0.25 m were marked on the circumferences. The sensor platform was placed at a height of 1.2 m, however it can also be placed higher if required.

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 in increments of 0.5 m/s, V-G curves and corresponding P-T data were recorded in the FoV. Several V-G curves were recorded for single person and multiple people (up to 3) in 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 at several different locations and 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.

6.2 Localization accuracy

6.2.1 Single person in FoV. We tested LOCI in different rooms. The rooms had obstacles such as table and chairs, and static warm objects such as monitor and light bulbs. The testing was performed in the same rooms in which the training data was collected. Ten people were asked to walk randomly, not just along the circumference, inside rooms at different speeds and move from one room to another. In total, 20 recordings were done for each person. The ground truth



(a) Single person in FoV (b) Multiple people in FoV

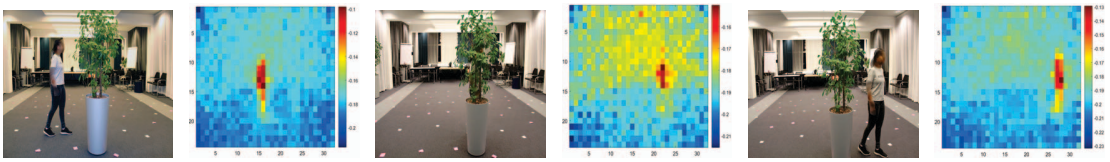
Figure 15: CDF of absolute localization error for different number of people

was recorded in a video and manually tabulated. Figure 15a shows the CDF of absolute localization error with respect to the ground truth. An undesired outlier was 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 such close 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 better accuracy (maximum error of 0.24 m at 4.5 m). However, as the person moves farther away, the person covers less number of pixels and the error starts to increase. The best localization accuracy obtained was 100 % in the grid resolution of 0.25m x0.25 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. We observe that 80% of the times, the error is within 35 cm.

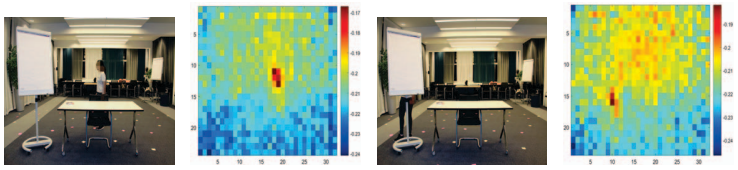
6.2.2 Multiple people in FoV. We use the proposed algorithm to also count people and use it to localize multiple people. We first present the counting the results in Figure 17. In ideal cases, as presented in the figure, we can count almost precisely. As we do not take exhaustive steps, as it is not the main focus of this work, cases when one person is behind the other, our counting results are not precise. Figure 15b shows the CDF of localization error obtained with more than one person (two and three people) in the sensor FoV. The persons were asked to randomly move around, not just along the circumference. From the plot we observe that 82% of the times the localization error is within 88 cm. Compared to single person tracking, the error in localizing multiple people is more as the signal captured by the PIR sensor depends on the speed and distance of the people, and all curvesets for such random walking cannot be captured. Furthermore, a precise counting mechanism will also help improve the accuracy. The idea we propose and demonstrate is that first the people counting step identifies 'x' number of people in the frame, then our system localizes those 'x' people. Indeed this works well when there is sparse occupation in the room. However, accuracy of people counting is less if significant portions of contour of two or more persons overlap. In literature, to avoid this problem, there are people counting algorithms using thermal images when the sensor is placed on the ceiling but not when mounted on a wall.

6.3 Impact of obstacles

To analyze the impact of obstacles, we present a case that is typical of office rooms – 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



(a) Before plant (b) Before plant - thermopile frame (c) Behind plant (d) Behind plant - thermopile frame (e) After plant (f) After plant - thermopile frame



(g) Behind chair (h) Behind chair - thermopile frame (i) Behind board (j) Behind board - thermopile frame

Figure 16: Experiment to analyze the impact of obstacles

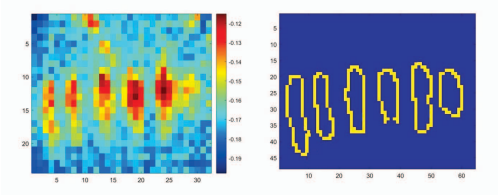


Figure 17: People counting. (a) Multiple people in FoV as captured by the thermopile; (b) Outlines of people after applying the steps outlined in Section 5.2.1.

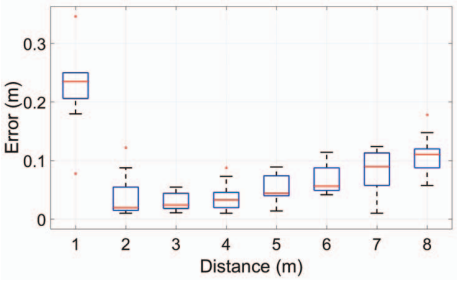


Figure 18: Box plot showing error in height classification

shown in Figure 16 along with the corresponding frame from the thermopile. 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 16d, 16h and Figure 16j, 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, the number of pixels traversed by the person changes when the person is behind an obstacle, misdetecting that the person is far 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. It is impossible to eliminate the effects of the obstacles as they cannot be detected by our sensors. Our system needs line of sight in order to function properly and provide a good accuracy. If a person is completely behind an opaque object where the body heat is blocked, the system will not recognize the person.

6.4 Height classification accuracy

To evaluate our approach on height classification, we considered 10 people with different heights in the range 1 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 18 shows the box plot of errors in height classification at different distances. 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 is 14 cm.

7 DISCUSSIONS

Our system preserves privacy as we do not use a camera. While PIR sensors are fully privacy preserving, a thermopile array with a low resolution (such as 768 pixels) may solely provide the signature/gait of the person, but not his/her actual image. As seen in Figure 6, just by sensing the temperature it is impossible to identify or even get a *decent* identifiable contour of the person. Hence it does not reveal any private information.

Tracking: An additional feature of tracking users can be implemented. Once localization is completed, tracking can be performed by localizing the person in subsequent thermopile frames. From Section 5.2.5, we get the absolute positions of people in two dimensions. The trajectory of the movement can be estimated using the subsequent localization snapshots. This scheme works well for a single person but may result in erroneous results for more persons, for e.g., (i) when two persons cross each other, or when they walk towards each other, meet at a point, and turn back – in these scenarios, we lose the track with respect to the particular person as the system is privacy aware. (ii) when one person is walking behind another – here, it is difficult to distinguish between the two.

Device lifetime: The entire hardware is powered by a 1600 mAh, 3.7 V battery embedded inside the enclosure. The idea is to charge the battery using Photo Voltaic cells mounted on the LOCI enclosure. A simple calculation shows that it is possible maintain the working of LOCI. When all the subsystems are ON continuously (including transmission), the battery would last for 11 hours (covering most of the day time) with single charge. Considering 33 cm^2 of solar panel area at 20 % efficiency, the battery would be fully recharged in approximately 5 hours [27] even in indoors. However, if we duty-cycle WiFi (100 ms ON every second), then the battery would last for 43 hours and thus even continuous operation is possible.

Deployment: Our system is robust to deployment variations as long as the the box is placed perpendicular to the ground plane and that there is line of sight (placing at corner of walls is also possible). Varying the height may lower the height classification accuracy. However, the height of deployment does not affect the accuracy of localization much. Our system is flexible enough to mount on the ceiling or on the wall. Our system is the first to provide such flexibility. As seen in the literature, localization, tracking, and people counting by placing multiple thermopiles or PIRs on the ceiling have been attempted. Our system, being compact, low power and multi-tasked, can also be used to perform these tasks. The evaluations that we presented are with the wall deployment. For ceiling deployment, we can implement the algorithms presented in [15, 17, 28] on our sensor platform.

8 CONCLUSIONS

Due to the high demand for smart buildings and smart personalized applications the context-awareness is highly sought after. Localization, one of the major component of context-aware services, which is being addressed in the last two decades. In this paper, we addressed the need for a device-free, privacy-aware, low-power localization in semi-public places. We custom-built a highly miniaturized platform (5x5x2 cm^3) called LOCI that could be easily deployed on a wall. We addressed many challenges such as a single system to localize as well as estimate height in 9 m x 8 m room. Using sensor data from PIR and thermopile jointly we showed that we can accurately localize and track persons in real-time. LOCI uses 80 mW for estimating the location and height. Using thermopile we removed the background hot objects and ambient heat noise. Comparing with the literature, LOCI is agnostic to the clothes worn. In this work, we focused on joint height estimation and localization with a single platform. We proposed many novel techniques such as variable gain for creating spatial diversity gain with a single PIR sensor. LOCI achieves 50% of the times <22 cm accuracy and 80% of the times <35 cm. By showing the best-case location accuracy of 12.5 cm, we outperform the state of the art location accuracy by 1 cm. The height estimation accuracy is within 8 cm in majority cases. The next step is to make the system robust and provide easy installation and usage methods including other services.

REFERENCES

[1] E. Commission. (2019) 2018 reform of eu data protection rules. [Online]. Available: https://ec.europa.eu/commission/priorities/justice-and-fundamental-rights/data-protection/2018-reform-eu-data-protection-rules_en

[2] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, 2019.

[3] S. Narayana, R. V. Prasad, V. S. Rao, T. V. Prabhakar, S. S. Kowshik, and M. S. Iyer, "Pir sensors: Characterization and novel localization technique," in *ACM/IEEE IPSN*. ACM, 2015, pp. 142–153.

[4] B. Mukhopadhyay, S. Sarangi, S. Srirangarajan, and S. Kar, "Indoor localization using analog output of pyroelectric infrared sensors," in *2018 IEEE WCNC*, April 2018, pp. 1–6.

[5] 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)*, April 2013, pp. 337–338.

[6] C. Basu and A. Rowe, "Tracking motion and proxemics using thermal-sensor array," *CoRR*, vol. abs/1511.08166, 2015. [Online]. Available: http://arxiv.org/abs/1511.08166

[7] W.-H. Chen and H.-P. 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)*, Oct 2015, pp. 428–434.

[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*, vol. 18, no. 15, pp. 6360–6368, Aug 2018.

[9] J. Kemper and D. Hauschildt, "Passive infrared localization with a probability hypothesis density filter," in *2010 7th Workshop on Positioning, Navigation and Communication*, March 2010, pp. 68–76.

[10] X. Liu, T. Yang, S. Tang, P. Guo, and J. Niu, "From relative azimuth to absolute location: Pushing the limit of pir sensor based localization," in *ACM MobiCom*, 2020.

[11] Y. Ma, G. Zhou, and S. Wang, "Wifi sensing with channel state information: A survey," *ACM Computing Surveys (CSUR)*, vol. 52, no. 3, p. 46, 2019.

[12] T. Wei and X. Zhang, "mtrack: High-precision passive tracking using millimeter wave radios," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*. ACM, 2015, pp. 117–129.

[13] F. Adib, Z. Kabelac, D. Katabi, and R. C. Miller, "3d tracking via body radio reflections," in *11th {USENIX} Symposium on Networked Systems Design and Implementation ({NSDI} 14)*, 2014, pp. 317–329.

[14] C. Zhang and X. Zhang, "Pulsar: Towards ubiquitous visible light localization," in *Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking*. ACM, 2017, pp. 208–221.

[15] C.-S. Shih, T.-Y. Wang, J.-J. Chou, Z.-Y. Chuang, C.-C. Chuang, K.-J. Lin, W.-D. Wang, and K.-C. Huang, "Collaborative sensing for privacy preserving gait tracking using iot middleware," in *RACS*, ser. RACS '17, 2017, pp. 152–159.

[16] 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*, vol. 52, no. 4, pp. 1311–1317, Nov 2006.

[17] D. Qu, B. Yang, and N. Gu, "Indoor multiple human targets localization and tracking using thermopile sensor," *Infrared Physics and Technology*, vol. 97, pp. 349 – 359, 2019.

[18] Microchip. (2019) Microchip atsaml21 based flash microcontroller (mcu). [Online]. Available: https://www.microchip.com/wwwproducts/en/ATSAML21J18B

[19] Espressif. (2019) Espressif esp32. [Online]. Available: https://www.espressif.com/en/products/hardware/esp32/overview

[20] Melexis. (2019) Mlx90640 32x24 ir array. [Online; accessed 01-April-2019]. [Online]. Available: https://www.melexis.com/-/media/files/documents/datasheets/mlx90640-datasheet-melexis.pdf

[21] Zilog. (2019) Zmotion pyroelectric sensors. product specification. [accessed 01-April-2019]. [Online]. Available: http://www.zilog.com/docs/PS0336.pdf

[22] H. Gami, "Movement direction and distance classification using a single pir sensor," *IEEE Sensors Letters*, vol. 2, no. 1, pp. 1–4, March 2018.

[23] SparkFun. (2019) Sparkfun ir array breakout - 55 degree fov, mlx90640 (qwiic). [Online; accessed 01-April-2019]. [Online]. Available: https://www.sparkfun.com/products/14844

[24] B. Mukhopadhyay, S. Srirangarajan, and S. Kar, "Modeling the analog response of passive infrared sensor," *Sensors and Actuators A: Physical*, vol. 279, 2018.

[25] R. Haddad and A. Akansu, "A class of fast gaussian binomial filters for speech and image processing," *Trans. Sig. Proc.*, vol. 39, no. 3, pp. 723–727, Mar. 1991.

[26] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Jun. 1986.

[27] M. Gorlatova, M. Zapas, E. Xu, M. Bahlke, I. J. Kymissis, and G. Zussman, "CRAWDAD dataset columbia/enhants (v. 2011-04-07)," Downloaded from http://crawdad.org/columbia/enhants/20110407/energy, Apr. 2011, traceset: energy.

[28] 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)*, Oct 2012, pp. 2042–2047.