Measuring People-Flow Through Doorways using Easy-to-Install IR Array Sensors

Hessam Mohammadmoradi*, Sirajum Munir†, Omprakash Gnawali*, Charles Shelton†
Computer Science Department, University of Houston, Houston, TX*
Bosch Research and Technology Center, Pittsburgh, PA †
Email: hmoradi@cs.uh.edu, sirajum.munir@us.bosch.com, gnawali@cs.uh.edu,charles.shelton@us.bosch.com

***Abstract*—People counting has many applications in smart buildings. For example, adjusting HVAC systems based on the number of occupants in each room can save a significant amount of energy. In addition, security and safety of the building can be managed by determining the number and location of occupants. Different technologies and sensing platforms have proposed for accurate and efficient people counting. However, these solutions are expensive, hard to deploy, or privacy invasive. We investigate the possibility of placing an 8×8 IR array sensor at the doorways and counting the number of people inside rooms. Our solution is real-time, inexpensive, privacy preserving with much less deployment constraints compared to its competitors. The proposed solution deals with realistic and dynamic changes in the sensing environment by leveraging a combination of Otsu's thresholding and modeling thermal noise distribution. We evaluated our solution via several controlled and uncontrolled real-world environments. The results show an average of 93% accuracy in estimating the number of occupants in rooms.**

I. Introduction

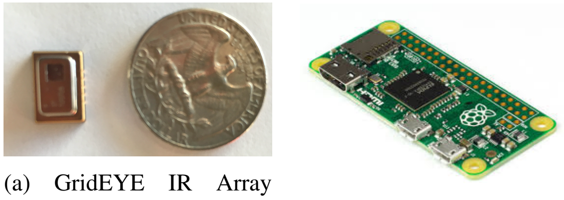
Improving energy efficiency of buildings has been an active area of research for many years and there is a global effort to reduce energy waste. Energy consumed in buildings is a large fraction of the total energy consumed by commercial and residential sectors (40% in the U.S. [5]). HVAC systems are usually the most energy consuming components in buildings (40% as in [23]). Recent advance in Internet of Things (IoT) technologies has started a new era in modern building management. Various types of sensing platforms are being deployed to understand the in-depth behavior of the occupants for efficient building energy and occupant comfort management. Technology that can accurately estimate the number of occupants in a room could become a key enabler for many applications in this space. For example, the estimated number of occupants in the building can be used to control HVAC systems and save a significant amount of energy (25% as in [3]). Occupancy estimation is also valuable in other areas such as safety and marketing.

There are several people counting solutions proposed by research community or industry sector. People counting using RGB cameras is accurate but often raises privacy concerns and may not be deployed in many residential and commercial buildings. Break-Beam sensors are the cheapest people counting solution available commercially. They use breaks in active IR signals to detect objects when they pass through a door and break the signal. However, there are tight restrictions regarding the placement of Break-Beam sensors at the doorway (specific height and pointing directly to each other) that make them hard and even impossible to deploy in some scenarios. Ultrasonic sensor based solutions require a significant amount of training to achieve reasonable occupancy estimation accuracy. Besides, ultrasonic waves usually are not pet-friendly. High-resolution thermal imagers are accurate; but price for commercial thermal imagers starts at $250 which is prohibitively expensive for large scale deployments.

In this paper, we use a low resolution (8×8 pixels) IR array sensor to count the number of people inside a room. The main idea is deploying an IR array sensor on sides or top of the doorway and counting entrance and exit events. The solution extracts and tracks humans from captured IR images using their temperature difference compared to the background. Our solution is lightweight and runs on a Raspberry Pi Zero (costs only $5) which makes it an affordable solution for large scale deployments, e.g., commercial buildings, academic buildings, hospitals, restaurants, and shopping centers. The IR array sensor costs less than $22 and consumes ∼4.5 mA in its active times. The IR array sensor can be mounted on top of the door or on either side of the door as long as people walk inside its field of view. The placement and orientation of the sensor do not have to be as constrained as a Break-Beam sensor-based solution. The solution has almost no privacy concerns as the resolution is so low and human body temperature is so similar, it is almost impossible to identify occupants using our sensor. To the best of our knowledge, this is the first work that places low-power low-resolution IR array sensors on doorways to estimate the number of people inside the room.

The contributions of this paper are as follows:

- We develop a real-time occupancy estimation solution that is easier to deploy compared to similar solutions, privacy preserving, and very inexpensive by using low resolution IR array sensors.
- We perform a range of micro-benchmarks to understand the characteristics of the sensor and analyze its performance under different deployment scenarios.
- We have deployed our solution in a commercial building and also in an academic building containing diverse types of doors and dynamic environments due to HVAC operations and movement of people. We observe that our solution achieves 93% accuracy in occupancy estimation in real-time.



(a) GridEYE IR Array Sensor

(b) Raspberry Pi Zero

Figure 1: Hardware Platform.

- We also show the potential of using an IR array sensor to determine skin temperature of individuals.

II. Related Work

Despite the simplicity of the occupancy estimation problem, finding an accurate solution has been a serious challenge for many years. There are few metrics that are important in the presence detection and people counting such as accuracy, cost, privacy, scalability and last but not least mounting flexibility. We compare the state of the art solutions in this area in Table I. In the following paragraphs, we discuss the state of the art occupancy estimation solutions and elaborate their advantages and disadvantages.

Break-Beam sensors: Break-Beam sensors use two sensors for counting people, one being the transmitter and the other one being the receiver. The transmitter unit continuously emits an IR wave in a straight line to the receiver unit. If someone passes through this line, the receiver detects a break in the line and considers it as a human passing event. Break-Beam sensors are the cheapest people counting solution available in the market. For detecting the direction of movement, 4 sensors are required in two modules. If either module fails, the complete solution does not work. There is a good chance of such a failure as the sensors are suggested to be mounted exactly facing each other. In addition, they should be placed between 125 cm to 140 cm height from the ground (according to the manual of the All-Tag bidirectional break-beam sensor [6]) and people may hit the sensors at that level inadvertently. If the sensors are mounted at a higher level, then they don't see shorter people. If the sensors are mounted at a lower level, then they just count legs. Also, they can not count accurately when multiple people move simultaneously and hence is not very useful for wider doors with traffic. Hence, they are only useful for narrow doors with single entrance/exit events and even there the narrow door increases the chance of getting an inadvertent hit.

Ultrasonic sensors: Researchers used ultrasonic sensors to count people [10], [17], [22]. [10] generates ultrasound chirps and measures features like people's height to count entrance and exit. Ultrasound based solutions require a significant amount of training even in small rooms. The room size has a direct impact on final accuracy [22]. Ultrasound signals are usually not pet-friendly and for generating higher frequencies more powerful transducers are required that are more expensive.

IR Array Sensors: Closest work to our solution is Thermosense [3] which uses a GridEYE sensor for counting people in the scene. GridEYE is mounted on top of the ceiling and records IR images. Set of features like the number of connected components and the number of hot pixels in the image is fed to neural network to accurately count the number of people in each frame (0.35 RSME). IR array sensors are very sensitive to distance and their performance dramatically degrades in longer distances which makes the height of ceiling a crucial factor in the accuracy of Thermosense. The GridEYE's maximum covered area shrinks in lower ceilings and in the higher ones, humans are not distinguishable from the background or the other hot objects. In addition, room size affects Thermosense. In the best mounting configuration, GridEYE can cover a 2.5m*2.5m area which means for larger rooms, multiple GridEYEs are required. Besides, Thermosense does not process at real-time and requires training depending on position and orientation of the sensor. Even at the same location, environmental changes caused by HVAC systems may affect the accuracy of the trained system. Our solution counts people inside the frame with much higher accuracy (0.06 RMSE) compared to Thermosense. Since we deploy the sensor at the doorway, room size and ceiling height do not have any impact on the accuracy of estimated occupancy.

RGB cameras: A number of solutions use RGB video cameras for counting people [25], [24], [21], [12], [2], [8]. RGB cameras based solutions can have relatively high accuracy, but cameras are privacy invasive. Even if a camera does not stream videos, it can be compromised while connected to the Internet and raise privacy concerns.

Other solutions: Researchers also use environmental sensors [14] and vibration sensors [19] to estimate occupancy. However, these solutions are very prone to environmental noise. [16] uses a depth sensor mounted at the ceiling near to a door to estimate room occupancy. In [13], electricity consumption data is used to detect occupancy (occupied or not) in households. Smartphones are utilized to localize and detect occupancy inside the buildings in [11], [9]. In contrast to our approach, these two are infrastructure-less solutions and require the users to carry smartphones all the time.

III. Approach

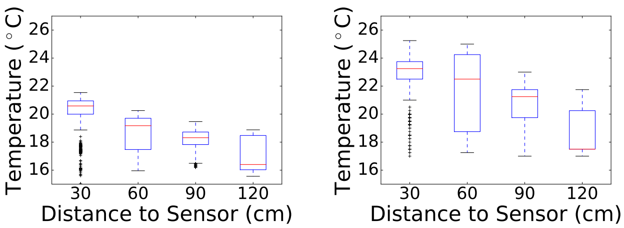
This section elaborates our design choices and technical challenges in order to develop a people counting solution.

A. Hardware Platform

IR Array Sensor: There are several IR array sensors available in the market with a diverse range of resolutions, frame rates, and prices. Scalability and low-cost are key requirements of people counting solutions. In our solution, we use a GirdEYE IR array sensor (Figure 1a) produced by the Panasonic Corporation. It has 8×8 resolution and costs around $22. As reported in GridEYE's datasheet and verified by ourselves through a simple experiment, it has $\pm 2.5°C$ temperature accuracy. GridEYE's $60°$ field of view is comparably wide considering other options with the same price range. GridEye uses 4.5 mA, 0.2 mA, and 0.8 mA on normal mode, sleep mode, and stand by mode, respectively.

Table I: State of the Art People Counting Solutions

Solution	Application	Cost ($)	Privacy Preserving Level	Scalability	Real Time	Flexibility
Break Beam Sensors	Counting	$\leq$ 10	High	Yes	Yes	No
PIR Sensors	Presence	$\leq$ 10	High	Yes	Yes	Yes
Ultrasonic Sensor	Counting	$\leq$ 100	Moderate	No	Training Required	No
RGB Cameras	Counting	$\leq$ 100	Low	Yes	Yes	No
IR Imager	Counting	$\leq$ 25	High	Yes	Training Required	No
Our Solution	Counting	$\leq$ 25	High	Yes	Yes	Yes



(a) Average Reported Body Temperature

(b) Maximum Reported Body Temperature

Figure 2: Temperature Accuracy Drop Over Distance

GridEYE's frame rate is 10 Hz which is fast enough to capture people's entrance and exit events considering the average walking speed of humans is 1.38 m/s[1].

Computational Unit: Raspberry Pi Zero (Figure 1b) provides a reasonably powerful computational environment at a significantly low price ($5). It has a 1 GHz single core CPU with 512 MB internal memory [20]. It fits our computational and cost requirements.

B. IR Array Sensors' Challenges

The low resolution GridEYE is very prone to noise. In addition, its reported temperature values are sensitive to ambient temperature and air flow in the room. Following sub-sections provide more detail about these challenges and techniques to mitigate them.

1) Accuracy Drop over Longer Distances: IR array sensors measure IR waves emitted from the surface of the objects. The amplitude of the received IR signal determines the estimated temperature. The IR wave's amplitude drops as a function of distance from the emitter. This characteristic of IR array sensors presents a serious challenge in our human counting solution. We conducted a simple experiment to evaluate the performance drop of GridEYE over distance. Figure 2 shows temperature values reported for the same person standing in front of the sensor in different distances. In this experiment, one person is standing in front of the sensor in four different distances. We take 20 temperature smaples from each distance to plot the figure.

Figure 2a shows average body temperature and Figure 2b shows maximum temperature from the body area reported by GridEYE. As expected, there is a significant accuracy drop when the sensor and human are at larger distances. The reported temperature for a person standing 120 cm distance from the sensor is close to the reported background temperature, which makes the person almost invisible. Skin temperature of a human is usually between 32°C to 34°C [4]. However, due to accuracy drop over distance, the reported temperature values are lower than that.

2) Sensor Noise: IR array sensors are prone to noise. Each pixel shows significant oscillations in its reported temperature over time. In addition to inherent noise in the sensor, change in deployment environment's characteristics like ambient temperature causes a significant impact on the sensor's reports.

We perform a simple test to characterize noise levels in GridEYE. We place the sensor at a close distance (1 cm) of a foam plate (Figure 4a) and monitor temperature values. Different pixels are supposed to report the same temperature because the foam is covering the whole sensor's field of view. However, as it is shown in Figure 4b, there are significant variations in the reported temperature values. This experiment shows that pixel value fluctuations are large (-2.5°C,2.5°C) and frequent that we can not calibrate the sensor at a pixel level.

3) Mounting Flexibility: The IR array sensor reports the temperature of objects in its field of view. The orientation or shape of a living object may vary, but the reported temperature is different from the background when the object is not too far. We leverage this property of thermal sensing to increase the mounting flexibility of our solution.

Installing an IR sensor on top of a door usually has more practical challenges than installing the sensor on the side of the door. Data collected by an IR sensor mounted at the top of a high door will have low accuracy. On the other hand, the main advantage of placing the sensor on top of a door is handling crowded scenarios. If two people walk side by side, a top view sensor sees both of them.

The GridEYE sensor can also be placed on sides of a door. In this case, the sensor can see a significant portion of the body, if the subject is relatively close to the sensor. However, if two people walk in close distance, the person closer to the sensor occludes the other person. A significant number of building doors are narrow (e.g., 90 cm) and typically a single person passes through a narrow door at a time. The side mounted solution is more appropriate for a narrow door, where two people usually do not enter simultaneously.

4) The Presence of Activities Other than Walking: One can design a system assuming straight walks through a door. However, the occupants may perform a range of other activities in or near the doorway. For instance, someone can stay at the

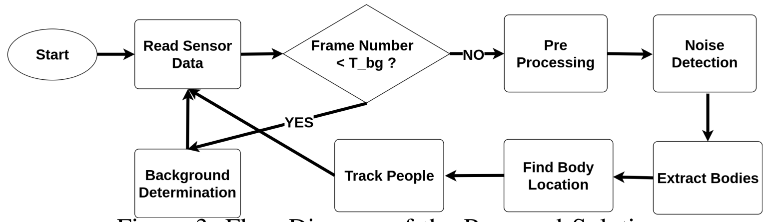
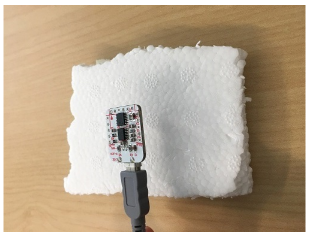
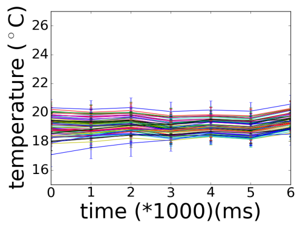


Figure 3: Flow Diagram of the Proposed Solution



(a) Experiment Setup



(b) Temperature for Each Pixel

Figure 4: GridEYE's Noise.

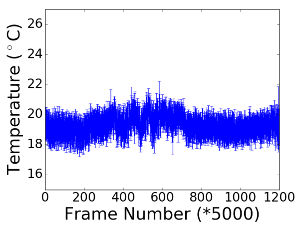


Figure 5: Background Temperature During 7 Days

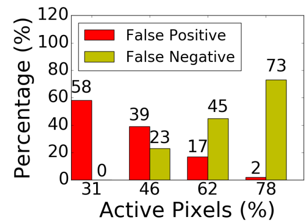
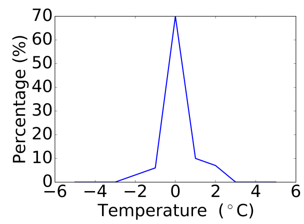
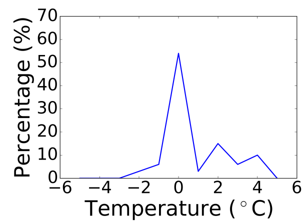


Figure 6: Performance of Using Standard Deviation in Detecting Frames with Humans Inside



(a) Frame without Human



(b) Frame with Human

Figure 7: Histogram of Differences to the Background

doorway for a long time, go back and forth, or wave hands in front of the sensor to play with it.

C. People Counting Solution Building Blocks

Our design of people counting solution is based on a detailed analysis of the GridEYE's characteristics and the challenges of using IR Array sensors. The system consists of several components, as shown in Figure 3. Each block is explained in details in the following paragraphs.

1) Read Sensor Data: A GridEYE is attached to a Raspberry Pi via a USB port. We wrote a simple driver in C++ to read frames from GridEYE. We store the frames into an internal queue at a frequency of 10 Hz.

2) Background Determination: Ability to detect small changes in a scene's temperature pattern is critical to achieve high accuracy in people counting. We first estimate the background by calculating the average of frames. The first time the system starts, it collects T_bg number of frames and calculates pixel-wise average of them. Our experiments show that 250 frames are sufficient to calculate IR pixel values that are representative of the background. The key requirement during background calculation phase is absence of temporal heat sources such as humans or hot objects in the scene. The background estimation takes place just once and the system keeps counting people in real-time after that. To evaluate our assumption that the thermal background in a building setting is stable, we collect data for 7 days from an empty scene and plot average temperature reported by the sensor (Figure 5). The variations in reported temperatures during long term do not exceed our expectations ($\pm 2.5\ ^\circ C$, GridEYE's noise).

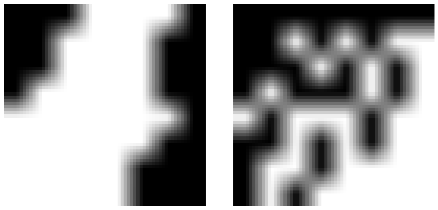
3) Preprocessing: OpenCV's linear interpolation technique is used to scale up the original frame for visualization purpose.

4) Noise Detection: We try different techniques to eliminate noise and finally use a combination of techniques to mitigate noise in the IR Array sensor data, as described below.

Standard Deviation: The standard deviation of each pixel as an indicator to find active pixels (pixels representing a human) has been shown to be effective [3]. To test the effectiveness of this approach in our setting, we collected 100 frames from the same person standing in different distances to the sensor and manually tagged pixels with human inside (active pixels). We then applied standard deviation based approach to tag active pixels. We calculated standard deviation of background pixels and considered pixels were bigger than 2 times of standard deviation as active pixels. Figure 6 shows the number of false positives and false negatives. As Figure 6 shows, standard deviation is not a reliable metric to find active pixels.

Heat Distribution: In our setting, thermal noise follows a common distribution that is significantly different from the histogram of pixel values when someone comes in front of the sensor. Figure 7 shows the distribution of pixel values in differences frames (difference from the background) with and without a human inside.

When there is a human in the scene, the distribution of pixel values is wider (compared to the distribution in a scenario without a human) and has multiple spikes. We use this



(a) Frame with Human (b) Frame without Human

Figure 8: Otsu's Binarization

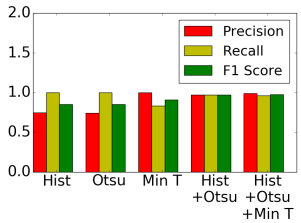


Figure 9: Noise Removal Technique Evaluation

characteristic to differentiate frames with and without humans.

The biggest spike in the histogram is due to background temperature and second biggest spike is for a human inside the frame. If the second biggest spike is larger than specific thresholds (width > 50% max width and amplitude > 40% biggest amplitude), we consider it as human. We determine these thresholds by performing extensive experiments.

Otsu's Binarization: Some frames without humans pass the above filters because the noise in the sensor data causes the pixel value distribution to be different from the modeled distribution of the frame without humans. In order to mitigate this problem, we used a thresholding technique called Otsu's Binarization [18] to divide pixels into humans and background classes. We calculate a threshold and if the difference between average temperatures in each class is not bigger than our threshold (0.75 $^\circ C$), we just consider the frame as a background. Figure 8 shows examples of frames classified as with and without humans after running Otsu's thresholding on them.

Temperature Filter: Our hypothesis is that frames with a human should have higher average temperature than frames without. In our solution, we measure the average background temperature. The average temperature of a frame that contains a human should be at least 0.25 $^\circ C$ higher than the average temperature of the background. Otherwise, we assume that it is a background frame and discard it.

Final Noise Removal Technique: We did a series of experiments to find the best noise removal technique. We collected data from different peoples' walk (10 individuals) and run each technique to identify noisy frames from the ones with humans inside. We manually tagged each frame to obtain the ground truth.

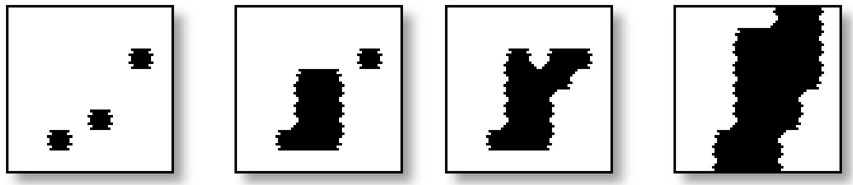
Figure 9 shows precision, recall, and F1 score of different combination of these techniques after running each technique on the same set of frames (100 frames). We find that combining heat distribution detection, Otsu's thresholding, and temperature filtering techniques achieves the best performance.

5) Extract Bodies: This building block detects multiple people in the scene and extracts their bodies for tracking. In the distance of 60 cm from the sensor, the width of field of view is 70 cm, which makes the possibility of seeing more than two people in one frame very unlikely. If two people walk very closely, due to the limited resolution and accuracy of GridEYE, the solution can not detect two people. However, if there is a small gap between them, the system detects and tracks both persons in the following frames. In order to find that gap and human bodies, we start with a temperature threshold of background temperature + 0.25$^\circ C$ and set all the pixel (temperature) values lower than that to 0 (shown as black in Figure 10) and higher than that to 1 (shown as white in Figure 10). We try to find body contours with pixels that are set to 1. If the size of the detected body is larger than a threshold (30% of frame's area), there is a possibility that there are two persons in the frame. In this case, all pixels below the minimum possible body temperature are ignored and our system tries to find two reasonably large contours (bigger than 10% of frame area) among the rest of the pixels. If our system can not find two bodies, it increases the threshold by 0.25$^\circ C$ and repeats the process. It keeps running this process until it finds two bodies or just one small body (less than 10% of frame area), which means that there is just one body in the frame. Figure 10 shows the frames after increasing the threshold value. In this example, after four iterations of increasing the threshold, two bodies emerge on the frame (white components).

6) Find Body Location: After extracting bodies, next step is finding their location. The solution calculates the sum of the values in each column and analyzes columns from left to right trying to find specific patterns corresponding to the number of found bodies. If there is just one person in the frame, starting from one side of the frame, there should be a gradual increase in the column values going up to a maximum value and after that, a gradual decrease; the location of maximum value is considered as the location of the body (Figure 11a). Figure 11b shows a scenario with two maximums corresponding to the location of two people.

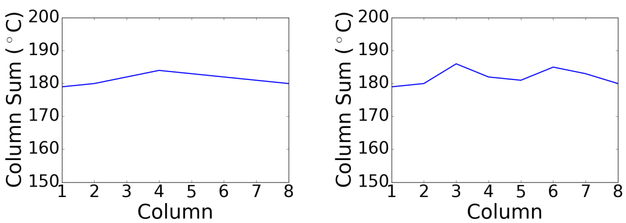
7) Track People: Due to noise and limited field of view challenges, direction of movement can not be extracted from a single frame. Our solution monitors series of frames to identify direction of movements. In order to track the same person across several frames, the solution extracts a few features from each frame and matches previously seen people to newly found bodies utilizing the following features.

- Spatial Distance: Normalized distance from the location of the previously seen body to the location of the new body. If this distance is smaller than a threshold (10% of frame width), these two bodies belong to the same person.
- Temperature Distance: Temperature difference between bodies which belong to the same person should be less than a specific threshold (less than 1 $^\circ C$).
- Temporal Distance: The time difference between two frames which contain matching bodies should be less than



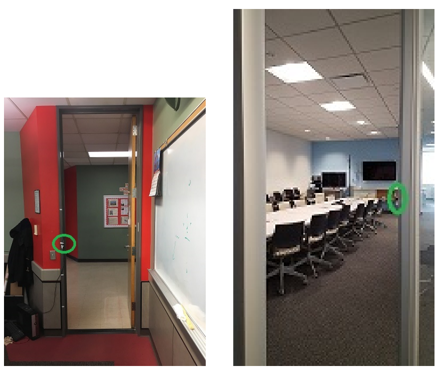
(a) Threshold = Background + 0.25°C (b) Threshold = Background + 0.5°C (c) Threshold = Background + 0.75°C (d) Threshold = Background + 1°C

Figure 10: Extracting Multiple Bodies by Applying Multilevel Thresholding



(a) One Person in the Scene (b) Two Persons in the Scene

Figure 11: Finding Locations of Extracted Bodies



(a) Controlled Experiment (b) Uncontrolled Experiment

Figure 12: Experimental Setup

a specific threshold (less than 5 frames).

IV. Performance Evaluation

We present the results from evaluating our system in controlled environment (our lab is shown in Figure 12a) and in uncontrolled environments (public areas like classrooms, computer labs and conference rooms (Figure 12b). We obtained IRB approval for this study. All the experiments have been conducted in room temperature (21°).

A. Controlled Experiments

In this section, we evaluate the accuracy of our solution in counting people and also benchmark its performance in different walking and mounting scenarios. In all the experiments, ground truth data is manually labeled.

1) People Counting Accuracy: We asked the participants to walk through two different doors (90 cm and 180 cm widths) one at a time. The doors had both our system and Break-beam sensors to provide fair comparisons of performance.

Table II and III contain results from a wide (180 cm width) and a narrow (90 cm width) door, respectively. The minimum accuracy reported in both experiments is 94%, which is very reasonable considering the advantages of the proposed solution. Performance of our solution degrades at wider doors due to an accuracy drop for longer distances and sensor noise. However, our solution has less constraint on sensor deployment compared to break beam sensors. Also, our solution can estimate skin temperature, which can be useful for improving thermal comfort of the occupants that break beam sensors can not do.

Table II: Performance Evaluation in a Door with 180 cm Width

Event	# Ground Truth	# Our Solution	# Break Beam Sensors
Entrance	315	304 (96%)	315(100%)
Exit	315	298 (94%)	315(100%)

Table III: Performance Evaluation in a Door with 90 cm Width

Event	# Ground Truth	# Our Solution	# Break Beam Sensors
Entrance	75	75 (100%)	75 (100%)
Exit	75	72 (96%)	75 (100%)

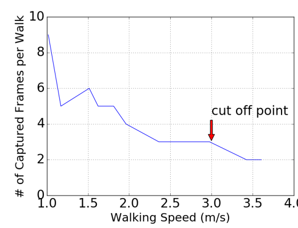
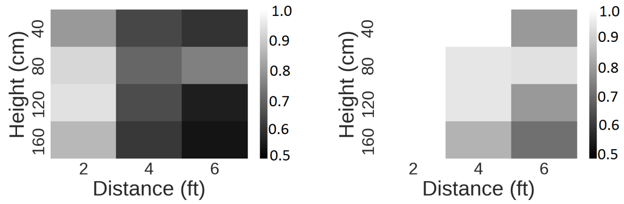


Figure 13: Walking Speed's Impact on Performance

2) Walking Speed Analysis: We mounted the sensor at the height of 120 cm from the floor and asked the participants to walk in front of the sensor at different speeds. GridEYE's maximum sampling frequency is 10 Hz, which means with higher walking speeds, the total number of frames that system can see a person during a walk is reduced. We count number of frames that system can see a person passes through a door. Figure 13 shows the results. As a person walks faster, s/he appears in fewer frames during the walk. However, as long as the system can track a person in at least 3 frames, it is able to register the walk. The average walking speed for humans is 1.38 m/s [1], which means that our solution can detect someone walking at 2.17 times of the average speed.

3) Mounting Height and Walking Distance: We mount the sensors at different heights from the floor (40 cm, 80 cm, 120 cm, and 160 cm) and asked five participants to walk at different



(a) Precision (b) Recall

Figure 14: Performance Analysis in Different Mounting Positions and Walking Distances

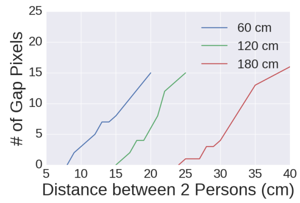
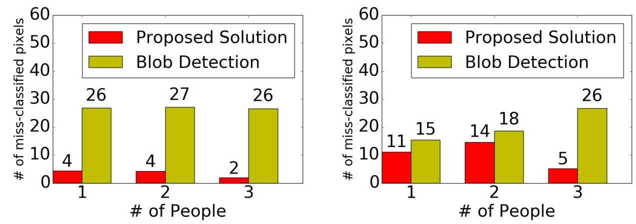


Figure 15: Minimum Identifiable Distance between 2 People Standing in 60cm, 120cm, and 180 cm away from the Sensor. Gap pixels are background pixels between 2 extracted bodies that separate them.

distances from the frame of the door (60 cm, 120 cm, and 180 cm). The critical determinant of accuracy of the proposed solution is the ability to distinguish pixels representing a human from the background pixels. To evaluate deployment variation's impact on the performance, precision, and recall for classifying pixels representing a human and background are computed. For the ground truth, all the analyzed frames (30 frames per configuration) are manually marked. Figure 14 shows the precision and recall values. The figure shows that both precision and recall values drop over longer distances but mounting height does not have a noticeable impact ($\leq$%10) on precision and recall values and our system accurately counts entrance and exit events.

4) Multiple Persons in The Scene: Distinguishing different objects in a low-resolution frame is a challenging problem, e.g., when people walk one after another through a busy doorway. To test the performance of the proposed system in crowded scenarios, GridEYE is mounted at the height of 140 cm (which is best mounting position based on previous experiments) and two persons are asked to stay in GridEYE's field of view in 60 cm, 120 cm and 180 cm distances. Our solution accurately counted multiple people in the scene, if there was at least one pixel gap between two persons in the scene. In each configuration, the minimum identifiable distance between two persons is measured (average of 20 frames in each configuration) and is shown in Figure 15. The minimum identifiable gap between two people has a direct relationship to the distance of humans from the sensor. If the distance of humans from the sensor is increased, more spacing is needed in between them to detect both of them.

5) Comparison to Blob Detection Algorithm: Blob detection algorithms are widely used to find objects in a scene. Most of these algorithms use a combination of adaptive thresholding and filtering to robustly find distinct blobs in an image. Now

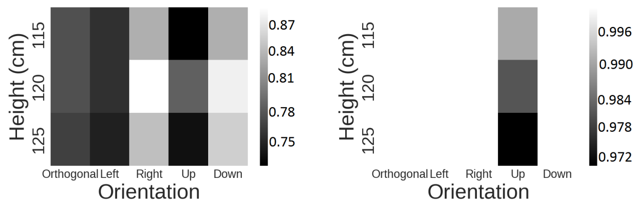


(a) 60 cm distance (b) 120 cm distance

Figure 16: Performance Comparison with Blob Detection

we compare the performance of blob detection algorithms and our proposed solution to detect walkers on low-resolution thermal images captured by a GridEYE. We use OpenCV's blob detection algorithm and set the start and end thresholds to the minimum and maximum values in the image. We asked 1,2, and 3 people to stand in front of the sensor in 60 cm and 120 cm distances, captured the frames, and used blob detector and our solution to extract pixels representing humans. Figure 16 shows number of misclassified pixels in both approaches. For ground truth, we manually tag the pixels. Our solution outperformed blob detection algorithm in all the scenarios. In 60 cm distance, blob detection algorithm misclassified 35% more pixels than our solution. During this experiment, there was no filtering or any other limitation for blob detection algorithm. In most of the cases, blob detector merged multiple bodies into a single blob. In addition, in some cases, blob detector found blobs that are hot areas caused by noise of the GridEYE sensor.

6) Mounting Flexibility and Robustness: Although our system works best when the GridEYE sensor is installed orthogonal to the door frame, by mounting error or due to mounting constraints, the sensor may be placed at an angle. To test the impact of such an angled installation on the system's performance, we mount the sensor in three different heights on side of a door and we tilt it in different directions (left, right, up, and down) by 20°. We asked five persons to walk across the door and measured precision and recall values for extracting pixels with body inside in all four directions, including orthogonal (ideal) orientation (Figure 17). The figure shows that the performance drop (precision and recall) is less than 10% compared to the orthogonal orientation, which did not affect people counting accuracy. In all the cases, our solution accurately counts the number of walks. In another set of experiments, we compared the performance of our system and a break-beam sensor based system when they are installed at different heights from the floor. Under the height of 30 cm, the break-beam-based system over-counts due to legs breaking the beams separately while our system worked correctly even with GridEYE placed as low as 15 cm from the floor. GridEYE can be installed even above the average height of the walkers as long as the passing person is within the field of view while break-beam sensors must be installed below the average height. Thus, our system is more flexible than a break-beam sensor based system considering the angle of installation as well as possible locations on the door frame where the sensors



(a) Precision (b) Recall

Figure 17: Performance Evaluation in Tilted Mountings

Table IV: Performance Evaluation in a Computer Lab

Event	# Ground Truth	# Our Solution
Entrance	13	12 (92%)
Exit	10	9 (90%)

Table V: Performance Evaluation in a Classroom

Event	# Ground Truth	# Our Solution
Entrance	54	48 (89%)
Exit	83	75 (90%)

can be mounted.

B. Experiments in Uncontrolled Environments

We deployed our solution in public areas like classrooms and computer labs to evaluate its performance in uncontrolled environments at real-time. In both cases, the door width was 90 cm and sensor is placed at a height of 140 cm from the floor. In the computer lab case, the length of experiment was four hours and in the classroom, it was two hours. During the entire evaluation, number of entrance and exit events are manually counted for ground truth collection and are compared to real-time estimates reported by our solution. During both experiments, we observe various unexpected behavior from the crowd, e.g., someone standing in front of the sensor while someone else was walking in/out. Our result includes all these abnormal behaviors. Table IV shows our evaluation results from computer lab and Table V summarizes results from evaluation in a classroom. Overall, we find that the system achieved 89-92% accuracy in estimating entrance and exit events.

V. Skin Temperature Sensing

In this section, we perform a case study to see if we can sense skin temperature of a human using an IR array sensor. If skin temperature of the occupants can be determined, it can be very useful to adjust temperature set point of HVAC systems to improve thermal comfort of the occupants.

To obtain ground-truth of skin temperature, we use the skin temperature sensor available at Microsoft Band 2 [15]. We develop our own smartphone application to retrieve temperature readings at 1 sample per second. We evaluate the performance of measuring skin temperature using GridEye and Heimann thermopile array sensor (32*31 pixels). The reason for using two different IR array sensors is to show the performance of two sensors that vary on price as a GridEye sensor costs ∼22 USD whereas a Heimann thermopile array sensor costs ∼210 USD. We mounted a GridEYE sensor and a Heimann thermopile array sensor on two sides of a table and marked distances from the sensors at every foot. We asked 6 subjects to wave their hands in front of each sensor in all the distances and recorded the reported temperature.

The experimental setup and temperature differences from the ground truth for both sensors are shown in Figure 18. We find that our system can estimate body temperature with less than 0.25 degree error if subject waves his hand in close distance (∼2cm) to the sensor. The accuracy of GridEye quickly drops as the distance from the sensor grows. However, Heimann sensor shows a significantly better performance for assessing skin temperature even when the subjects are 2-3 feet away. This experiment shows the potential of using IR array sensors for assessing skin temperature, which can be useful for improving thermal comfort of the occupants.

VI. Discussions

GridEYE sensor has a 60° field of view in both horizontal and vertical direction. If the sensor is mounted at sideways and someone stands close to the sensor, s/he blocks its line of sight and the sensor can not see the entrance and exit events of others. In addition, our real-world experiments suggest that such an occlusion is rare for a narrow (∼90 cm) door.

One of the key assumptions in our work is that people's body temperature is higher than ambient temperature. However, there are cases when this assumption is not valid. Our solution can not detect people in these cases.

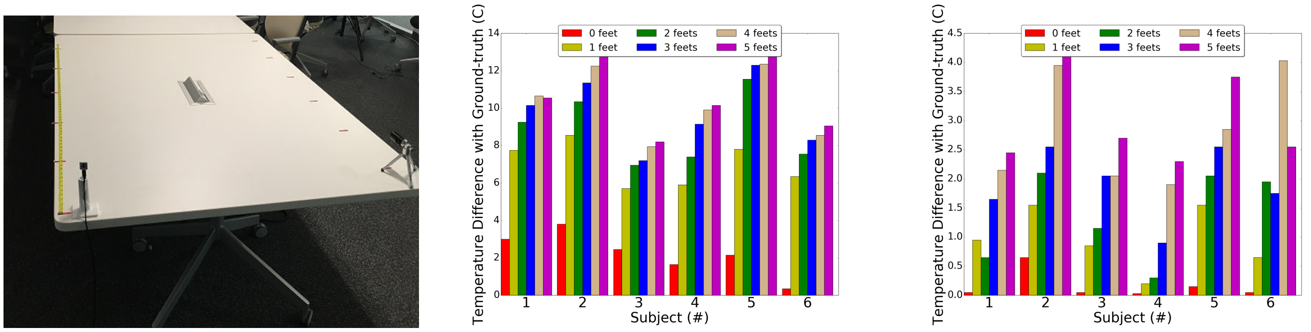
Our solution requires mounting the sensor on the side or on the top of a door. Hence, our solution can not be used in open areas where standard doors are absent.

We estimate that our system will cost less than $50 ($25 for GridEYE, $5 for Raspberry Pi, $7 for SD card, $10 for USB WiFi adapter), which is significantly cheaper than commercial break-beam people counting solutions ($250 per unit) [7].

Our solution estimates background at the time of initialization and does not update the background later. We verified that this strategy works for 7 days. However, if the background needs to be updated dynamically, a PIR sensor can be added to detect the presence of an individual and background is updated only when there is no one in the scene.

VII. Conclusions

In this paper, we propose an inexpensive and location flexible people counting solution by utilizing a low power, low resolution IR array sensor. Our solution runs on a Raspberry Pi Zero, which is an affordable powerful computing environment. We have much less restrictions on placing the sensor on doorways. There is no training required for our system and it counts people at real-time. Our solution can be placed on top of a door or at the sides. We evaluated the performance of occupancy estimation in different environments via controlled and uncontrolled experiments and showed the potential of using an IR array sensor for detecting skin temperature for enhancing thermal comfort of the occupants.



(a) Experimental Setup (b) GridEYE 8*8 IR Array (c) Heimann 32*31 Thermopile Array

Figure 18: Skin Temperature Sensing

References

[1] Average human walking speed. https://en.wikipedia.org/wiki/Walking. Accessed: 2017-01-23.

[2] J. Barandiaran, B. Murguia, and F. Boto. Real-time people counting using multiple lines. In *2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pages 159–162. IEEE, 2008.

[3] A. Beltran, V. L. Erickson, and A. E. Cerpa. Thermosense: Occupancy thermal based sensing for hvac control. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pages 1–8. ACM, 2013.

[4] Body skin temperature. http://hypertextbook.com/facts/2001/ AbantyFarzana.shtml. Accessed: 2017-01-23.

[5] B. E. D. Book. Energy efficiency and renewable energy. *US department of energy*, 2011.

[6] Bi-directional display people counter. http://www.sdinternational.nl /downloads/leaflets/People%20Counter%20-%20Display.pdf. Accessed: 2017-01-12.

[7] Break beam sensor, all tag. http://all-tag.com/product-items/display-counter-bi-directional/. Accessed: 2017-01-23.

[8] T.-H. Chen. An automatic bi-directional passing-people counting method based on color image processing. In *Security Technology, 2003. Proceedings. IEEE 37th Annual 2003 International Carnahan Conference on*, pages 200–207. IEEE, 2003.

[9] G. Fierro, O. Rehmane, A. Krioukov, and D. Culler. Zone-level occupancy counting with existing infrastructure. In *Proceedings of the Fourth ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, pages 205–206. ACM, 2012.

[10] T. W. Hnat, E. Griffiths, R. Dawson, and K. Whitehouse. Doorjamb: unobtrusive room-level tracking of people in homes using doorway sensors. In *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*, pages 309–322. ACM, 2012.

[11] M. A. A. H. Khan, H. Hossain, and N. Roy. Infrastructure-less occupancy detection and semantic localization in smart environments. In *proceedings of the 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services on 12th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, pages 51–60. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2015.

[12] J.-W. Kim, K.-S. Choi, B.-D. Choi, and S.-J. Ko. Real-time vision-based people counting system for the security door. In *International Technical Conference on Circuits/Systems Computers and Communications*, pages 1416–1419, 2002.

[13] W. Kleiminger, C. Beckel, T. Staake, and S. Santini. Occupancy detection from electricity consumption data. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*, pages 1–8. ACM, 2013.

[14] K. P. Lam, M. Höynck, B. Dong, B. Andrews, Y.-S. Chiou, R. Zhang, D. Benitez, J. Choi, et al. Occupancy detection through an extensive environmental sensor network in an open-plan office building. *IBPSA Building Simulation*, 145:1452–1459, 2009.

[15] Microsoft band 2. https://www.microsoft.com/ microsoft-band/en-us. Accessed: 2017-01-23.

[16] S. Munir, R. Arora, C. Hesling, J. Li, J. Francis, C. Shelton, C. Martin, A. Rowe, and M. Berges. Real-time fine grained occupancy estimation using depth sensors on ARM embedded platforms. In *Proceedings of the 23rd IEEE Real Time and Embedded Technology and Application Symposium*, 2017.

[17] N. Nasir, K. Palani, A. Chugh, V. C. Prakash, U. Arote, A. P. Krishnan, and K. Ramamritham. Fusing sensors for occupancy sensing in smart buildings. In *International Conference on Distributed Computing and Internet Technology*, pages 73–92. Springer, 2015.

[18] Otsu's binarization. http://docs.opencv.org/trunk/d7/d4d/tutorial_py_thresholding.html. Accessed: 2017-01-23.

[19] S. Pan, A. Bonde, J. Jing, L. Zhang, P. Zhang, and H. Y. Noh. Boes: building occupancy estimation system using sparse ambient vibration monitoring. In *SPIE Smart Structures and Materials+ Nondestructive Evaluation and Health Monitoring*, pages 90611O–90611O. International Society for Optics and Photonics, 2014.

[20] Raspberry pi zero. https://www.raspberrypi.org/products/pi-zero/. Accessed: 2017-01-12.

[21] M. Rossi and A. Bozzoli. Tracking and counting moving people. In *Image Processing, 1994. Proceedings. ICIP-94., IEEE International Conference*, volume 3, pages 212–216. IEEE, 1994.

[22] O. Shih and A. Rowe. Occupancy estimation using ultrasonic chirps. In *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*, pages 149–158. ACM, 2015.

[23] H. Teraoka, B. Balaji, R. Zhang, A. Nwokafor, B. Narayanaswamy, and Y. Agarwal. Buildingsherlock: Fault management framework for hvac systems in commercial buildings. Technical report, Technical Report, CSE, UCSD, 2014.

[24] D. B. Yang, H. H. González-Baños, and L. J. Guibas. Counting people in crowds with a real-time network of simple image sensors. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 122–129. IEEE, 2003.

[25] X. Zhao, E. Delleandrea, and L. Chen. A people counting system based on face detection and tracking in a video. In *Advanced Video and Signal Based Surveillance, 2009. AVSS'09. Sixth IEEE International Conference on*, pages 67–72. IEEE, 2009.