<https://helda.helsinki.fi>

How Low Can You Go? Performance Trade-offs in Low-Resolution Thermal Sensors for Occupancy Detection: A Systematic Evaluation

Rinta-Homi, Mikko

2021-09

Rinta-Homi , M , Hossein Motlagh , N , Zuniga Corrales , A , Flores , H & Nurmi , P 2021 , ' How Low Can You Go? Performance Trade-offs in Low-Resolution Thermal Sensors for Occupancy Detection: A Systematic Evaluation ' , Proceedings of ACM on interactive, pömobile, wearable and ubiquitous technologies , vol. 5 , no. 3 , 126 , p

<http://hdl.handle.net/10138/338822>

<https://doi.org/10.1145/3478104>

unspecified

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

How Low Can You Go? Performance Trade-offs in Low-Resolution Thermal Sensors for Occupancy Detection: A Systematic Evaluation

MIKKO RINTA-HOMI, University of Helsinki, Finland
NASER HOSSEIN MOTLAGH, University of Helsinki, Finland
AGUSTIN ZUNIGA, University of Helsinki, Finland
HUBER FLORES, University of Tartu, Estonia
PETTERI NURMI*, University of Helsinki, Finland

We contribute by systematically analysing the performance trade-offs, costs (privacy loss and deployment cost) and limits of low-resolution thermal array sensors for occupancy detection. First, to assess performance limits, we manipulate the frame rate and resolution of images to establish the lowest possible values where reliable occupancy information can be captured. We also assess the effect of different viewing angles on the performance. We analyse performance using two datasets, an open-source dataset of thermal array sensor measurements (TIDOS) and a proprietary dataset that is used to validate the generality of the findings and to study the effect of different viewing angles. Our results show that even cameras with a 4×2 resolution – significantly lower than what has been used in previous research – can support reliable detection, as long as the frame rate is at least 4 frames per second. The lowest tested resolution, 2×2 , can also offer reliable detection rates but requires higher frame rates (at least 16 frames per second) and careful adjustment of the camera viewing angle. We also show that the performance is sensitive to the viewing angle of the sensor, suggesting that the camera’s field-of-view needs to be carefully adjusted to maximize the performance of low-resolution cameras. Second, in terms of costs, using a camera with only 4×2 resolution reveals very few insights about the occupants’ identity or behaviour, and thus helps to preserve their privacy. Besides privacy, lowering the resolution and frame rate decreases manufacturing and operating costs and helps to make the solution easier to adopt. Based on our results, we derive guidelines on how to choose sensor resolution in real-world deployments by carrying out a small-scale trade-off analysis that considers two representative buildings as potential deployment areas and compares the cost, privacy and accuracy trade-offs of different resolutions.

CCS Concepts: • **Human-centered computing** → *Ubiquitous and mobile computing design and evaluation methods*.

Additional Key Words and Phrases: Thermal Sensing, Occupancy Detection, Internet of Things, Smart homes, HVAC

ACM Reference Format:

Mikko Rinta-Homi, Naser Hossein Motlagh, Agustin Zuniga, Huber Flores, and Petteri Nurmi. 2021. How Low Can You Go? Performance Trade-offs in Low-Resolution Thermal Sensors for Occupancy Detection: A Systematic Evaluation. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 3, Article 126 (September 2021), 22 pages. <https://doi.org/10.1145/3478104>

*Corresponding author

Authors’ addresses: Mikko Rinta-Homi, University of Helsinki, Helsinki, Finland, mikko.rinta-homi@helsinki.fi; Naser Hossein Motlagh, University of Helsinki, Helsinki, Finland, naser.motlagh@helsinki.fi; Agustin Zuniga, University of Helsinki, Helsinki, Finland, agustin.zuniga@helsinki.fi; Huber Flores, University of Tartu, Tartu, Estonia, huber.flores@ut.ee; Petteri Nurmi, University of Helsinki, Helsinki, Finland, petteri.nurmi@helsinki.fi.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2021 Association for Computing Machinery.
2474-9567/2021/9-ART126 \$15.00
<https://doi.org/10.1145/3478104>

Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., Vol. 5, No. 3, Article 126. Publication date: September 2021.

1 INTRODUCTION

Human occupancy information is essential for building management systems, providing a basis for optimizing energy use and ensuring comfortable and healthy indoor environment. Estimates suggest that buildings account for over 40% of overall energy consumption in the developed countries, with heating, ventilation and air conditioning (HVAC) being responsible for over 60% of this [14]. Having accurate information about human occupancy can be used to adapt HVAC functionality on-demand [2] and potentially achieve significant savings [16, 18, 34]. Beyond energy savings, adaptive HVAC is also essential for maintaining a healthy and productive indoor environment. Indeed, workplace productivity has long been linked with the state of comfort in environmental conditions, with factors such as air quality and temperature being directly linked to work performance, susceptibility to errors, creativity, and even work satisfaction [10, 17].

Human occupancy information can be captured at different levels of granularity with occupancy detection, a binary indicator of whether the space is occupied or not, and counting, estimating the number of people inside a space, being the coarsest and most common ones. In the literature, a wide range of technologies for occupancy detection have been proposed, ranging from the use of reed switches and passive infrared to camera-based solutions [9, 44]. Generally, a single solution rarely fits all needs as the requirements depend on the nature of the space and the needs of the building management solution. For example, residential homes benefit from solutions that are reliable, low-cost, privacy-preserving, and non-intrusive, whereas larger office spaces or commercial buildings can benefit from more granular tracking as it enables additional applications, such as security auditing.

Low-resolution thermal array sensors are emerging as a promising solution for occupancy detection and counting in residential homes as the sensors are inexpensive, have low deployment cost, and are both energy-efficient and non-intrusive to occupants [6, 7, 13, 32, 46]. A significant challenge with low-resolution sensors, however, is that their performance is highly sensitive to the overall resolution, frame rate, and the field-of-view of the camera [46]. Existing research on low-resolution imaging has focused on developing prototype systems for occupancy detection and studying the performance of these systems without attempting to quantify the performance limits of such systems or to systematically assess the impact of different camera parameters. For example, Beltran et al. [6] proposed a system that integrates a 8×8 low-resolution imager with passive infrared sensors. Tyndall et al. [46] consider 4×16 sensor arrays whereas Cokbas et al. [13] rely on 32×16 sensor arrays. Understanding the different performance trade-offs of such systems is essential for generalizing these systems and offering solutions that can be adopted to different building environments.

In this article, we contribute by systematically evaluating the performance trade-offs and costs (loss of privacy and cost of deployments) of low-resolution thermal array sensors with differing sensor resolutions. We identify lower limits for resolution and frame rate to establish a minimum point where reliable detection is possible. We also assess the effect of different viewing angles on the performance. We conduct our analysis through systematic benchmarks using image manipulations. We consider two datasets, an open-source dataset with tripwire-triggered thermal images (TIDOS [13]) and a proprietary dataset that is used to validate the generality of the findings and to study the effect of differences in the camera's fields-of-view. The results of our evaluation show that even cameras with a 4×2 resolution – significantly lower than what has been used in previous research – can support reliable detection, as long as the frame rate is at least 4 frames per second. The lowest tested resolution, 2×2 , can also offer reliable detection rates but requires significantly higher frame rates. In each test case, the performance is sensitive to the viewing angle, suggesting that the camera's field-of-view needs to be carefully adjusted to maximize the performance of low-resolution cameras. Using a camera with only 4×2 resolution reveals very few insights about the occupants' identity or behaviour, and thus helps to preserve their privacy. Besides privacy, lowering the resolution and frame rate decreases manufacturing and operating costs and helps to make the solution easier to adopt. Finally, based on our results, we derive guidelines on how to choose sensor resolution in real-world deployments by carrying out a small-scale trade-off analysis that considers two representative

Fig. 1. (a) Person walking comfortable speed below the camera recording 8 frames per second. (b) Maximum temperature difference ($^{\circ}\text{C}$) per pixel caused by noise in 16 seconds video.

buildings as potential deployment areas and compares the cost, privacy and accuracy trade-offs of different resolutions. Taken together, our results offer novel insights on the performance and limits of low-resolution occupancy detection, offer guidelines on how to select the best sensor technology for a given deployment and with respect to different performance and cost criteria, and pave way toward establishing low-resolution thermal sensor arrays as a potential solution for supporting building management systems.

2 EXPERIMENTAL SETUP

The focus of our work is on understanding performance trade-offs in the use of low-resolution thermal array sensors for occupancy detection at different frame rates and resolutions. We also establish a lower limit for these parameters where reliable tracking remains feasible. We conduct our experiments using two datasets, the TIDOS open source dataset, and a proprietary dataset from a controlled testbed. The two datasets were collected using thermal sensor arrays of the same make, but differing fields-of-view. For both datasets, ground truth was based on manual annotation of the measurements. We next describe the thermal sensor array and the two datasets used in our experiments.

2.1 Thermal Sensor Array

Both of our datasets have been collected using Melexis MLX90640¹ thermal array sensors, which have a 32×24 pixel resolution. The device is available with two different fields-of-view (FOV): standard ($55^{\circ} \times 35^{\circ}$) and wide angle ($110^{\circ} \times 75^{\circ}$). The experiments consider both variants to be able to evaluate the effect of viewing angle. Specifically, the primary dataset (TIDOS) has been collected using the standard FOV, whereas the secondary (controlled) dataset has been collected with the wide-angle sensor. In both datasets the framerate was 16 frames per second (FPS). The MLX90640 camera uses subpixel sampling with a chess reading pattern [30], meaning that between each frame the camera alternates the pixels which are being read from the camera in a chess pattern. This can be clearly seen in Figure 1(a), as the moving object has a clear checkerboard pattern to the direction it is heading. Thus, in practice, half of the pixels get updated only on even frames, and the other half only on odd frames. In the frames the pixels from the previous frame are used to complete the image for those pixels that are not captured.

While higher frame rate does provide more data, it also introduces more noise. According to the datasheet of the sensor, the root mean square noise value in $110^{\circ} \times 75^{\circ}$ sensor is approximately 0.4°C for 8 FPS, and little bit over 0.5°C for 16 FPS [30]. To demonstrate this, we compared two 16-seconds motionless videos captured at 8 and 16 FPS, and calculated the maximum temperature difference for each pixel. Figure 1(b) shows the results. The edges and especially the corners present higher level of noise which coincides with the MLX90640's datasheet specifications. The average difference was lower when recording at 8 FPS (1.13°C) than at 16 FPS (1.8°C). The noise is higher in the sensor used by the TIDOS dataset, according to the datasheets nearly twice as much [30].

¹<https://www.melexis.com/en/product/MLX90640/Far-Infrared-Thermal-Sensor-Array>

Fig. 2. Testing environment for the proprietary dataset (a) and the paths used to record measurements (b): 1) walking to or from the neighboring door, 2) walking in a straight line.

However, this does not have any impact on occupancy detection as the temperature difference between human and background is typically sufficiently large.

2.2 Primary Dataset: TIDOS

As our primary dataset we use the Thermal Images for Door-based Occupancy Sensing (TIDOS) dataset [13], which was collected to provide a privacy-preserving method to estimate the amount of people indoors through thermal images. TIDOS contains 118 208 frames (123 minutes and 8 seconds of video). In total, 261 transitions (in or out from the room) happened in the samples. The measurements in TIDOS have been collected by installing the MLX90640 camera above the doorway of a classroom in different occupation scenarios (lecture, lunch meetings, high activity and edge cases). The thermal camera was located at a height of 240 cm, the side of the larger viewing angle was parallel to the door frame and the vision area was $249.9 \text{ cm} \times 151.34 \text{ cm}$. The videos were recorded at 32×24 pixel resolution, 16 FPS refresh rate and viewing angles of $55^\circ \times 35^\circ$.

For the experiments, the TIDOS' data is divided into scenes where the expected result is the change in the number of people in the rooms. The scenes are formed by finding one-second frames where (a) no transitions happen, and (b) only background can be detected. The approach used for background detection is based on thresholding and is described in Section 3.2. Besides scene separation we deleted sequences where no long activity was detected. Four seconds of background data is preserved in the beginning on each scene to establish a baseline for background estimation. The resulting scenes were manually verified against counting errors. For example, one of the recordings ended on a picture where a person was halfway moving out of the camera vision, so this ending was removed. After the pre-processing, the data contained 254 scenes, consisting of 42 minutes and 45 seconds of video material. In total, 65% of generic background data was removed from the measurements.

2.3 Secondary Dataset: Controlled Testbed Measurements

To validate the generality of the findings and to investigate the effect of having a different field-of-view, we supplement the TIDOS dataset with measurements collected from a controlled testbed. These measurements were collected using a wide angle variant of the Melexis MLX90640 thermal array sensor, which has a field of view of $110^\circ \times 75^\circ$. The measurements were collected using the same 16 FPS frame rate as in the TIDOS dataset.

Table 1. Self-recorded dataset. 2+2 indicates that the test is done in both directions twice.

Path	Door State	Walking Type	Single person	Two persons
1	Closed	Calm	2+2	2+2
1	Closed	Brisk	2+2	2+2
1	Open	Calm	2+2	2+2
1	Open	Brisk	2+2	2+2
1	Open	Calm, opposite directions		4
2	Closed	Calm	2+2	2+2
2	Closed	Brisk	2+2	2+2
2	Open	Calm	2+2	2+2
2	Open	Brisk	2+2	2+2
Total		32	36	68

Figure 2(a) shows the setup, the measurement angles, and the resulting overall vision area. The camera was positioned at the top of the door frame and the larger viewing axis was aligned parallel to the walking path through the doorway as this provided better coverage of the activity around the door. The dimensions of the door frame are 205 cm $\times$ 81 cm and the vision area on both sides of the door was 313.1 cm $\times$ 291.1 cm. The area was calculated considering the height of the door frame, the sensor height and that the larger viewing angle is parallel to the fully opened door.

The data recording targeted everyday occasions and consisted of one or two persons walking through the sensor's coverage area. The scenarios covered going through open or closed doors, walking from different sides of the doorway, as a single person or in a pair, walking to or from the neighboring door or walking in a straight line, and walking calmly or briskly. All variations were recorded twice. In addition, four extra recordings where two people walk opposite directions at the same time were done. In total, 68 different samples were recorded, which have a total transition of 104. The paths used for recording are illustrated in Figure 2(b) and the measurements are summarized in Table 1.

3 METHODOLOGY

We investigate the effect of different camera configurations by systematically manipulating the camera resolution and frame rate and assessing their impact on occupancy detection. We use a generic processing pipeline that has been inspired by state-of-the-art systems for occupancy counting using thermal array sensor data [12, 13]. Indeed, the main contribution of our work is the systematic evaluation of camera configurations and the resulting insights about trade-offs in the use of thermal array sensors with differing resolutions, not the processing of thermal images itself. We focus exclusively on occupancy detection and counting as they form the basis for most energy saving operations. Thermal array sensors can be used to detect occupancy information at finer resolution, e.g., Hevesi et al. [19] demonstrated how thermal array sensors can be used for coarse grained activity recognition. However, this requires dense deployments and complex processing, which are contrary to the goal of achieving energy savings at a low cost. In the following we describe the overall processing pipeline, and the operations that are used to manipulate the images.

3.1 Pre-processing

Several factors can affect the performance of thermal sensing devices. Firstly, internal heating can cause inaccurate measurements in commercial off-the-shelf thermal devices (e.g., cameras based on uncooled FLIR - Forward

Fig. 3. Pre-processing techniques.

Looking InfraRed technology) [27]. Secondly, camera resolution and alignment between thermal and RGB images can affect the quality of measurements. We reduce these effects by applying common pre-processing techniques for thermal images: background separation [7, 13, 33], image rescaling [7] and Gaussian blur [7, 13]. Figure 3 illustrates the use of these techniques for pre-processing the thermal images. Background removal is useful to ensure the processing focuses on areas of interest only. Rescaling helps to ensure a consistent processing resolution, whereas Gaussian blur reduces the influence of outlier pixels that could be caused by sensor errors or thermal reflection. Background separation uses the first four seconds of the recordings to count the average temperature for each pixel and uses that information to remove background from the image. If the pixel value is below its average plus a threshold value given as a parameter, the value of the pixel is set to black (i.e., 0). Image rescaling uses bi-cubic interpolation and multiplies the image dimensions (horizontally and vertically) to get a new image size and resolution. Gaussian blur is a low-pass filter that helps to reduce the checkered noise from frames based on a pixel's neighbourhood. Besides reducing the effects of noise, the pre-processing techniques also improve the robustness of occupancy tracking and counting.

3.2 Occupancy Detection and Counting

We estimate the number of people from the pre-processed images using blob detection. The detected blobs are given as input to a blob tracker which connects blobs from sequential frames, and a flow counter which calculates the number of transitions using the start and destination coordinates of blobs. Intuitively, the blob detection identifies clusters that correspond to different people. The tracking monitors the movement of these people over time, whereas flow counting estimates movements in and out from a particular area. Figure 4 summarizes and illustrates the workings of the pipeline we use to count occupancy.

Blob Detection: We use the open source simple blob detector algorithm included as part of OpenCV to extract blobs from images. The algorithm creates multiple binary images with different thresholds and extract blobs location and size based on pixel's proximity, inertia ratio, circularity, convexity, and size. In our case, the maximum threshold corresponds to a white pixel. The algorithm essentially performs clustering on the pre-processed image using different threshold values to determine regions of interest. Since our pipeline operates on thermal images, the differences in intensity values between the foreground and background are significant and simple clustering is sufficient for detecting human occupants. The use of multiple thresholds, in turn, helps to ensure the detection can operate robustly at different distances (i.e., different blob size) and image resolutions. More complex blob detection algorithms (such as those based on Gaussian operations) are only necessary if the intensity differences between areas are more subtle, or if the blobs are relatively small compared to the resolution of the input image.

Fig. 4. Processing pipeline for occupancy estimation.

As we operate on low resolution images, the blobs are necessarily large compared to the overall image resolution. We have also separately evaluated the use of Difference of Gaussian (DoG) blob detection, which achieved largely similar (slightly lower) results.

Blob Tracking: We use a state-of-the-art approach [33] for tracking blobs. The output of the blob detection is used to build a log of the changes in the positions and centres of blobs. Based on this log, blobs in subsequent frames are connected and linked with the same occupant. As the tracking operates based on blob location information, temperature changes do not affect tracking results [32].

Flow Counter: Counting is determined from the starting and ending coordinates of the logical paths resulting from blob tracking. For this, the view is divided into three areas: left zone, transition area and right zone. If the starting coordinates are in the left zone and ending coordinates are in the right zone, then a transition from left to right has happened. For other directions the transitions are defined analogously. The transition zone can also be non-existent, leading to only left and right zones. The number of persons is estimated from the difference in transition areas.

3.3 Feature Selection and Extraction

The evaluation considers three characteristics which have a significant impact on the processing cost and the cost of the cameras themselves: image resolution, frame rate and field-of-view [46]. The images are obtained from video data with specified frame rate, resolution, and horizontal and vertical viewing angle. We vary original features of the images and determine the optimal parameters that maximize counting accuracy. Resolution is changed by calculating the average temperature of multiple pixels and then combining them according a drop factor. Image distortion due to pixel division is avoided by using only factors that are multiples of the original image resolution. Frame rate is transformed by considering only every n number of frames from the original video. Viewing angle is modified by cutting the pixels from the sides of the image so that wideness of field of view will change. Figure 5 shows an example of features modification. We use tree of the Parzen estimator (TPE) algorithm to find the right combination of parameter values that maximize the accuracy of occupancy counting from thermal images using our pipeline. TPE suggests new parameter adjustments using probabilistic models obtained from historical iterations. We run the algorithm using the hyperopt library².

4 RESULTS

We next analyse the performance of thermal array sensors considering different image resolutions, frame rates and viewing angles. Highlights of our results are as follows:

²<https://hyperopt.github.io/hyperopt>

Fig. 5. Examples of image modifications: resolution (a), frame rate (b), viewing angle (c).

- Resolutions higher than 4×2 do not have significant advantage, and they all perform very similarly. Resolution of 2×2 achieves respectable accuracy with frame rate of 16 but performs significantly worse with lower frame rates than the higher resolutions.
- For any resolution starting from 4×2 , raising frame rate above approximately 4 frames per second has only a minor positive effect on counting accuracy. Lowering frame rate below this point can have notable decrease in performance. Resolution 2×2 benefits from any frame rate increase, with 16 frames per second being the highest tested frame rate.
- Viewing angle is an extremely relevant factor especially in lower resolutions. A favourable viewing angle can be determined for both axes, and this viewing angle works well even with different resolutions and frame rates. The favourable viewing angle depends on installation height of the camera and potentially other environment dimensions such as doorway width and requires further research to be determined precisely. The only exception is viewing angle parallel to the walking path through the door in resolution 2×2 . Here low viewing angle works best with high frame rates, and lower frame rates require higher viewing angles.

4.1 Impact of Camera Parameters

Frame Rate: We consider 6 frame rates: 16, 8, $5\frac{1}{3}$, 4, $3\frac{1}{5}$, $2\frac{2}{3}$, $2\frac{2}{7}$ and 2. The original frame rate in the data is 16 frames per second and these frame rates were derived by considering every n :th frame. The frame rate generally depends on the viewing angle, with viewing angles covering the entrance area to a space being best suited for lower frame rates. In our experiments, frame rates below 2 resulted in significant decreases in accuracy. Figure 6 shows the results for the controlled testbed. Frame rates between 4 and 16 generally perform very similarly and do not have a significant impact on the overall performance of the occupancy counting algorithm.

Resolution: We next analyse the effect of image resolution on counting accuracy. We fix the frame rate at 16 frames per second as the results in the previous section showed this frame rate to produce good results across all resolutions. Indeed, of all the tested configuration parameters, frame rate had the least overall impact, with only the lowest possible resolution (2×2) being impacted by it. Figure 7 shows the results for different resolutions. The counting accuracy is consistently around 90% with the lowest resolution being the sole exception. The average counting accuracies of the different resolutions are: 91.2%, 92.8%, 93.1% and 92.0% (for 32×24 , 8×8 , 4×4 and 2×2 , respectively). Reducing the resolution to 2×2 decreases accuracy by around 10%, resulting in an average accuracy of 84.1%.

Privacy: Besides lower processing cost, the main benefit of lower resolution is improved privacy protection. Indeed, the higher the resolution, the more identifiable information it captures with the highest resolution being capable of identifying the occupant reliably [12]. The lowest resolution (i.e., 2×2) suffers from being unable to capture more than one person at a time. Indeed, the minimum size of a blob is 2 pixels and thermal measurements

Fig. 6. Self-recorded dataset: (a) counting accuracy for different frame rates in different resolutions, (b) average counting accuracy of different frame rates.

Fig. 7. Counting accuracy at different resolutions and for different frame rates in the self-recorded dataset.

result in blobs that are located next to each other being merged. Another weakness of the lowest resolution is that it requires occupants to pass sufficiently centrally under the sensor as otherwise the measurements can be discarded as noise.

4.2 Common Viewing Angle

We next analyse the impact of viewing angle using the collected measurements. The camera used in the testbed can be configured with a high spectrum of different angles ($110^\circ \times 75^\circ$). Thus, it is possible to analyse how different angles in measurements influence overall counting performance. As the analysis has shown that 8×8 and below

Fig. 8. Results of counting errors for resolution 2×2 in frame rate of 8 in self-recorded dataset: (a) mean counting error for different viewing angles, (b) counting error for the combination of viewing angles parallel to the path (vertical angle) and parallel to the door frame (horizontal angle).

are sufficient for reliable counting while not violating the privacy of the occupants, we only consider the lower resolutions: 2×2 , 4×2 , 4×4 and 8×8 . To analyse a specific angle, we rely on the average of their measurements as it reduces the noise in measurements. Figure 8 shows the results for different angles using a specific resolution and frame rate. We separately consider the errors for angles parallel to the (door) frame (i.e., horizontal angle) and for those parallel to the walking path (i.e., vertical angle); see Figure 2(a) for an illustration of the angles.

Viewing Angle Parallel to the Door Frame (Horizontal Angle): Figure 9 shows the results for angles parallel to the door frame. Different angles tend to perform better in a specific resolution. A viewing angle of 50° works optimally with resolutions of 2×2 and 4×4 whereas viewing angles between 50° and 75° provide the best results when the resolution is 8×8 . In contrast, when using a resolution of 4×2 , the viewing angles that provide the best counting performance are between 37.5° and 56.25° .

Viewing Angle Parallel to the Walking Path (Vertical Angle): Figure 10 shows the results for different resolutions and viewing angles parallel to the walking path. Generally all resolutions perform well when the viewing angle is between 42° and 55° as long as the frame rate is at least 4 FPS. The optimal viewing angle range depends on the resolution and frame rate so that higher resolutions support a broader range of viewing angles as the frame rate drops. Indeed, the 8×8 resolutions supports the full range of viewing angles at 4 FPS whereas the 4×2 resolution requires viewing angle to be between 41.25° and 68.75° when the frame rate is 4 FPS. The 2×2 resolution performs best with much lower viewing angles, reaching optimal performance when the angle is between 20.65° and 48.125° when frame rate is 16 FPS. On lower frame rates, higher viewing angles tend to provide the best performance. Overall, the results thus suggest that frame rate of 4 FPS is sufficient for all except the lowest resolution, and generally a wide range of viewing angles can be supported. The lowest resolution can also support accurate counting, but requires frame rate to be higher than with the other resolutions.

Practical Impact: In practice, the viewing angle depends on the configuration of the door area and the height of the door frame. The best results are obtained when the camera captures the immediate area after the door frame, as this monitors the physical entry and exit point. For the higher resolutions, a wider range of viewing angles is

Fig. 9. Counting errors for different degrees parallel to the door frame in the self-recorded dataset using different resolutions: (a) mean counting errors, (b) counting errors for frames per second - Viewing angle combinations

Fig. 10. Counting errors for different degrees parallel to the walking path (vertical angle) in the self-recorded dataset: (a) mean counting error for different viewing angles and resolutions, (b) counting error for different combinations of viewing angles parallel to the walking path and frames per second.

possible as the resolution ensures the sensor covers a wider area. This means that low resolution is more than sufficient, as long as the sensor can be aligned with the entry/exit area, and higher resolutions can provide more flexibility with respect to viewing angle.

4.3 Evaluation of Counting Errors

We next analyse the scenarios to identify where counting errors are more likely to occur, and show that a resolution of 8×8 and frame rates equal or higher than 4 frames per second provide a lower counting error. We do this by evaluating the effect of different camera resolution and frame rate on the amount of failed scenes under different testing scenarios from the TIDOS dataset (edge cases, high activity, lecture and lunch meeting).

Fig. 11. Counting errors for different viewing angles parallel to the walking path (vertical angle) and resolution in the TIDOS dataset: (a) mean counting error, (b) counting error for different combinations of viewing angles and frames per second.

The results are shown in Table 2. The number of failed scenes behaves similarly for different resolutions, but the number of missing scenes is higher when the frame rate is lower. Resolution 8×8 offers a lower average of percentage of failures for different frame rates ($\mu = 5.6$, $\sigma = 4.5$) compared to 32×24 ($\mu = 6.7$, $\sigma = 4.3$) and 4×4 ($\mu = 6.5$, $\sigma = 4.7$). A smaller error with respect to the total number of scenes occurs for frame rates above $5\frac{1}{3}$ frames per second (less than 2%), while the opposite is true for frame rates below $2\frac{2}{3}$ frames per second (more than 8%). As expected, when comparing individual tests, the cases with a higher number of failed scenes are high activity ($\mu = 9.5$, $\sigma = 5.0$) and lunch meeting ($\mu = 5.8$, $\sigma = 5.8$). We confirm that a lower failure percentage occurs for an image resolution of 8×8 (edge cases: $\mu = 4.8$, $\sigma = 4.5$; high activity: $\mu = 6.1$, $\sigma = 4.1$; lunch meeting: $\mu = 5.2$, $\sigma = 5.3$) compared to resolutions of 32×4 (edge cases: $\mu = 5.3$, $\sigma = 3.5$; high activity: $\mu = 7.9$, $\sigma = 3.6$; lunch meeting: $\mu = 5.7$, $\sigma = 6.3$) and 4×4 (edge cases: $\mu = 5.3$, $\sigma = 6.1$; high activity: $\mu = 7.2$, $\sigma = 3.8$; lunch meeting: $\mu = 6.1$, $\sigma = 6$). The evaluation suggests that most of the errors during counting occur in scenes that involve a higher number of persons. A low frame rate (below 4 frames per second) produces a higher number of failed scenes, which can be explained by the missing of fast blob transitions or the merging of multiple blobs during the detection. In the latter case, these errors are more visible for resolutions smaller than 8×8 . For frame rates equal or higher than 4 frames per second, errors can be explained for example by the occurrence of false negatives due to a higher resolution footage.

4.4 Trade-off Analysis

The results thus far have highlighted different performance and cost trade-offs with different sensor resolutions. We next carry out a small-scale trade-off analysis of thermal array sensors considering examples of different real-life smart deployments.

Deployment Scenarios: The optimal density of sensor deployments depends on the application scenario. For example, adaptive thermostat adjustment for single occupant apartments may be able to operate with one or two sensors that are placed at the entry point, whereas multiple occupant smart spaces may require dense deployments to monitor continuous occupancy in a room [6]. For example, Murao et al. [37] showed that continuous occupancy monitoring for the entire space requires at least one sensor per each 5 m^2 area. Objects that are distributed within the area, such as furniture, cubicle separator panels, ornaments and so on, further increase the needed density as otherwise there can be blind spots in the space.

Table 2. Amount of failed scenes for different resolution and frame rates in the TIDOS dataset.

Resolution	Frame rate	Edge cases	High activity	Lecture	Lunch Meeting	Total
32×24	16	0	4	0	1	5
	8	1	8	0	0	9
	$5\frac{1}{3}$	1	8	0	1	10
	4	1	8	0	1	10
	$3\frac{1}{5}$	3	9	0	5	17
	$2\frac{2}{3}$	2	14	0	10	26
	$2\frac{2}{7}$	2	17	0	11	30
	2	1	17	0	18	36
4×4	16	0	4	0	0	4
	8	1	8	0	2	11
	$5\frac{1}{3}$	1	4	0	1	6
	4	0	8	0	4	12
	$3\frac{1}{5}$	1	8	0	7	16
	$2\frac{2}{3}$	1	14	0	7	22
	$2\frac{2}{7}$	2	12	0	10	24
	2	5	19	0	19	43
8×8	16	0	3	0	1	4
	8	1	4	0	0	5
	$5\frac{1}{3}$	0	4	0	0	4
	4	1	4	0	3	8
	$3\frac{1}{5}$	1	7	0	7	15
	$2\frac{2}{3}$	1	13	0	8	22
	$2\frac{2}{7}$	3	12	0	8	23
	2	3	18	0	16	37

To carry out a trade-off analysis of different sensors, we next consider the characteristics of two different (real) smart spaces. The first space is the Tellus smart space (TELLUS), which is in the University of Oulu, Finland in the Linnanmaa campus. The size of the smart office space area is $66.8 \text{ m} \times 36.4 \text{ m}$. It includes different smaller spaces such as for meetings rooms, open stage, study, and rest areas. The space includes 352 infrastructure based multisensor devices capable of measuring temperature, humidity, CO_2 , motion, and light [35]. The second space we consider is the Ubikampus smart space (UBIKAMPUS) at the University of Helsinki, Finland in the Kumpula campus. This space has an area of 400 m^2 and it also is separated into smaller spaces for meeting rooms, rest areas and kitchen. The space is equipped with 10 motion detectors which are installed on the ceiling of the space with equal distances [36]. We analyze the deployment cost of sensors required to monitor continuously occupancy in each smart space. We estimate recommended optimal deployment for each area (without obstacles/separations) and compare with the existing deployment. We also estimate the deployment cost when using low and high cost sensors, e.g., from \$30 sensors (low cost) to up to \$1000 thermal cameras (high cost) [19].

Results: To monitor occupancy continuously, the recommended deployment for TELLUS (2431.52 m^2) would comprise of 486 sensors. Similarly, for UBIKAMPUS, the recommended optimal deployment comprises of 80 sensors to ensure continuous coverage. For both smart spaces, the optimal number of sensors required to monitor occupancy continuously is higher than what is supported by the existing deployment, suggesting that the current deployments are likely to contain blind spots. For TELLUS, the optimal 486 sensor deployment would incur an additional cost between \$4000 and \$134 000 depending on the sensor technology. For thermal sensor arrays, the highest additional cost would be around \$17 420 (\$130 per sensor for higher resolution thermal array sensors). Similarly, UBIKAMPUS requires approximately 80 sensors, which incurs a cost of (\$14 400) and (\$480 000) when doing the same comparison.

The comparison thus suggests that low resolution sensors can bring significant savings, while helping to protect the privacy of the occupants. Naturally, this result only considers occupancy detection and detecting activities or other types of information could require adopting higher resolution sensors, or developing improved processing pipelines that can extract the needed information from low resolution sensors. In practice we would expect hybrid deployments where a smaller set of higher resolution monitoring technologies are deployed at areas that are actively used, whereas low resolution sensors are used to exhaustively cover the remaining areas of the space. While the comparison only considered smart offices, similar savings could be obtained for residential buildings. For example, an apartment building containing 30 apartments would only incur around \$1000 cost if it deployed a low-resolution sensor inside each apartment (e.g., above the doorway).

4.5 Comparison and Discussion

Occupancy Detection Performance: Our experiments focused on analysing different performance and cost trade-offs without performing a detailed analysis of the processing pipeline itself. The pipeline we use is similar to those used in prior works [12, 13] and the results in the original TIDOS paper showed this type of pipeline to improve performance contrast to simple baselines [13]. As a result, we have focused to evaluating performance trade-offs and comparing against the results in the TIDOS article. Direct comparison against the work of Chidurala et al. [12] is not possible as the authors use a different dataset and different sensors for collecting the data.

Training the counting pipeline with TIDOS data resulted in 10 miscalculations across 261 transitions with full resolution. Two false negatives were identified, which were caused by the tracking algorithm confusing a person who just left with a new person who had just arrived. No false positives were found when reviewing a set of random samples. After correcting the false negative, the total accuracy was 96.8% (8 miscalculations in 247 transitions). This is slightly better than what is reported in the original article, which is likely due to small variations in the parameters of the preprocessing phase. The overall performance is also in line with more advanced solutions, e.g., Chidurala et al. [12] demonstrated similar performance (95% for 8×8 resolution, and 99% for higher resolutions) using a more complex processing pipeline whereas Zou et al. [50] report 92.8% accuracy for occupancy counting using a Wi-Fi transmitter-receiver pair.

Failure Analysis: To obtain insights into the performance and limitations, we have manually reviewed the reasons for failures. We found that people who walked close to each other so that the floor could not be seen between them at all caused the system to malfunction and lose track of the occupants. Second, waving a hand under the camera can cause the counting algorithm to get confused. Third, if a person is not significantly warmer than the background, this can lead to the person not being detected – for example, heavy clothing can block thermal radiation being observed. Finally, high overall level of activity can confuse the blob tracker, particularly at higher frame rates.

Computational Performance: In terms of computational performance, occupancy counting results with different number of iterations can be seen in Figure 12. The performance rapidly increases during the first 1000 iterations, but slows down rapidly once this point is reached. The probabilities for finding a new and better result in every 250 iterations can be seen in Figure 13. When gathering data, iterations of 4096 were used because lower values caused more visible noise in results whereas higher amounts would require unnecessary amounts of computational resources without providing significant improvements.

Comparison between Datasets: The camera parameter experiments were largely conducted using the controlled testbed due to having better control and wider range of parameters available. We next briefly summarize the results for TIDOS and compare them to those obtained for our testbed.

In terms of resolution, we found similar results across both datasets. When using the same resolution as in our primary experiments, the resolutions 24×32 , 8×8 , 4×4 and 4×2 have similar average results of 92.3%, 93.0%, 92.1% and 91.9%, respectively. In contrast, the average result for resolution 2×2 is only 84.4%. The results are illustrated

Fig. 12. Average best results over iterations when calculating results for TIDOS dataset (sample size 2328).

Fig. 13. Average probability for finding better results on the TIDOS dataset (sample size 2328).

Fig. 14. Counting accuracy for different resolutions are frame rates in the TIDOS dataset.

in Figure 14. Regarding frame rate, Figure 15 shows average performance per frame rate of measurements available in TIDOS. From the results, Overall, frame rates from 16 to 5.33 does not have a significant difference in performance, but reducing frame rate below 5.3 starts degrading performance significantly. For frame rates of 4 and 3.2 the performance drop is marginal at higher resolutions, but significant at the lowest resolution (2×2).

For the TIDOS dataset, the results for viewing angle are largely similar to those for the controlled testbed despite the camera having a smaller range of viewing angles. For the angle parallel to the door frame (i.e., vertical angle), the results are shown in Figure 16. Angles ranging from 37.8° to 41.25° provide the highest accuracy. For viewing angles parallel to the walking path (i.e., horizontal angle), the best results were obtained with a viewing angle of 23.33° for frame rates ranging from 4 to 16 FPS. On slower frame rates, higher viewing angles tend to result in better performance. As before, the results for the 2×2 resolution are generally worse for all viewing angles. Average results for both resolutions 2×2 and 4×2 can be seen in Figure 11.

Fig. 15. TIDOS: (a) counting accuracy for different frame rates are resolutions, (b) average counting accuracy for different frame rates.

Fig. 16. Counting errors for different degrees parallel to the door frame and resolution in TIDOS dataset: (a) mean counting error, (b) counting error for different combinations of viewing angles and frames per second.

5 DISCUSSION

Implications for Occupancy Detection: We rely on low-resolution thermal array sensor footage to detect and count people indoors. The results demonstrated that the resolution and frame rate are intrinsically linked and must be optimized together to achieve robust detection performance – especially when resource usage needs to be minimized. The performance is also impacted by the field of view of the camera, suggesting that best energy savings are only achievable when the placement of the sensor can be optimized for the target space. The optimal angle is close to 45°, with the small variability around the viewing angle depending on the exact characteristics of the deployment. Lower door frames require slightly lower angles than higher frames, and generally the best

performance results when the immediate area under the door is captured. Low resolution thermal sensor arrays are best suited for deployments where each sensor monitors a single entry point – or one side of an entry point. This makes the technology best suited for residential homes and office spaces that have clear entry points rather than open spaces with many entry and exit points. As outlined in the trade-off analysis, if the sensors are deployed densely into the environment, they can also be used to support continuous occupancy monitoring for complex spaces, such as collaborative work environments. The power consumption of the solution can be further reduced using a secondary modality, such as a laser tripwire or a reed switch, to trigger the sensor duty cycle.

Room for Improvement: The main limitation of thermal sensor arrays is the need to observe thermal radiation. Heavy clothing, e.g., for winter purposes, can block most of the thermal radiation that skin would normally emit. However, in such cases it may be possible to use alternative means to recognize humans. For example, people often touch furniture or other objects when they enter indoor areas and this touch results in a temporary thermal footprint [15] that can be captured using a thermal sensor array. Other limitations of thermal arrays include the need for separate power source, even if the power draw is small, and the need for sufficiently unobstructed view of the entrance area. Indeed, furniture, clothing racks, and other objects may block parts of the sensor’s view and decrease recognition performance.

Processing Pipeline: We relied on a relatively simple processing pipeline that integrated standard off-the-shelf image processing techniques. As we are operating in the thermal spectrum, differences in temperature are typically significant and easy to observe, and hence a simple processing framework is sufficient for most practical use cases. Indeed, the main errors relate to (i) only partially observing the person entering the scene and (ii) multiple people being close to each other in which case the blobs of the individual persons can accidentally be merged. These errors have negligible impact on occupancy detection, but they can have larger impact on occupancy counting and more fine-grained processing, such as activity recognition. As the resolution of the input images is very small in all experiments, the main improvements in performance would result from considering additional information, such as absolute temperature measurements, as part of the blob detection, or using features that characterize the blob features to filter out incorrect cases.

Other Technologies: Thermal array sensors have several benefits, including low-cost, non-intrusiveness, and energy-efficiency. A potential alternative would be to use passive infrared sensors instead of thermal arrays, but this likely requires deploying PIR arrays rather than individual devices. Indeed, previous research has shown that individual PIR sensors are not sufficiently accurate for occupancy detection – let alone counting [1]. Another option would be to use low resolution cameras operating within the visible light spectrum (i.e., regular cameras) as they are capable of preserving privacy. The main disadvantage of low resolution cameras is that it is difficult to get them to work accurately. In the case of thermal array sensors foreground - background segmentation is close to trivial due to the intensity values reflecting temperature, and similarly blob detection is reliable as long as thick clothing is not used. Low resolution cameras, in turn, are sensitive to luminosity and require high contrast between the occupants and the background to be able to work accurately.

Energy-Efficiency in Building Management Systems: We demonstrated that the counting of people can be performed with very low resolution (higher than 2×2) and frame rate (5Hz or higher) thermal images. This result has significant potential for building management systems and especially intelligent HVAC. Indeed, our results show that intelligent HVAC systems can be configured to work with minimal computing resources [21]. Likewise, effective counting of individuals can also be used to regulate the energy consumption of devices and appliances in indoor environments. This can potentially be used to reduce electricity bills as well as to reduce carbon emission from smart buildings [20].

Indoor Productivity: Considering that humans spend about 90% of their time indoors [26, 29], making sure that indoor air quality is fresh and healthy is critical. Efficient counting of people can be used to estimate accurate

levels of CO₂ indoors and regulate ventilation. This can be used to ensure that the environment indoor foster productivity and comfort rather than produce drop in work performance and discomfort.

Privacy: Relying on images with low resolution and captured at low frame rate reduces risks of recording private information about individuals. These benefits do not result merely from the use of a different sensing modality – indeed, thermal footprints have been used to identify individuals. However, low resolution footage that provides just enough information to detect the presence of an individual mitigates risks of identifying individuals and helps to protect privacy.

6 RELATED WORK

Closest to our work are previous efforts in using thermal array sensors. Existing works in these methods have focused on developing occupancy detection pipelines and evaluating them in isolated environments. Our work extends on these works by providing a systematic analysis of the performance and cost trade-offs associated with thermal array sensors, and providing insights into how to choose the optimal sensors for deployments.

Air quality sensors: Can be used as proxies to detect occupancy and people movements inside spaces [23, 35, 36, 49]. The main drawback with these sensors is delay in detection. For example, while occupancy increases CO₂ concentrations, it can take 10 to 20 minutes for these changes to be observable [31]. The delay in detecting the first occupant is usually shortest and detecting subsequent people usually takes much longer. Similarly, there is often a significant delay in detecting when a person leaves the space. For example, Javad et al. [22] reported first occupant detection to take about 9 minutes whereas detecting people leaving took even up to 25 minutes.

Passive infrared (PIR) sensors: Monitor infrared radiation changes in the environment caused by movement of a person [48]. The main benefits of PIR sensors are low cost and low power draw. PIR sensors generally have lower accuracy, and they also cannot detect people that do not move [1]. For example, a sleeping person can not be detected reliably [43]. PIR sensors can also be used to support detecting direction of movements, e.g., detecting people entering or leaving a space. For example, placing PIR sensors on both sides of the door can be used to detect which side the door was accessed [1] or in which direction a person is moving to [47]. This approach, however, fails when there are multiple people walking through the door at the same time. PIR sensors can also be used for occupancy counting. For example, regression analysis can be used to count occupants based on motion patterns. A single PIR sensor can count the people in the space with an error of one person, for up to 14 people. Crowds larger than 14 people limit the view of PIR sensors and might require installing multiple PIR sensors [41].

Wireless sensing: Refers to technologies like IEEE 802.11 (Wi-Fi) and Bluetooth. Occupancy detection based on wireless sensing method can be in the forms of *device-based* and *device-free* detection. The former uses personal devices such as mobile phones and laptops to detect occupancy. The main drawback of these technologies is that they require occupants to carry their device with, and they also require integrating the personal devices with the building management system. Even if people carry the device, system optimization techniques, such as the device's power saving mode, can influence the connections. For instance, based on a research iPhones tend to force sleep mode that results disconnecting the Wi-Fi; while Mac OS X devices cause remain connected to Wi-Fi despite being turned off [3]. Wi-Fi-devices regularly send probe requests which are proven to be reliable in tracking in the past [38], but specifically to avoid tracking, modern systems use MAC address randomization which randomizes the MAC address used in identification [28]. Therefore, it is unknown if the probe requests are from the same device or from multiple devices. Although, this prevents counting, occupancy can still be detected. Note that this can also become a drawback if an adversary fakes occupancy to increase energy costs in an area. Bluetooth devices also use similar MAC address randomization [4], which their frequency of randomization allows using them for tracking devices [24].

Device-free detection refers to analyzing signal fluctuations in an environment while considering channel state information without taking into account the device which the signal is propagated from. Indeed, the presence

and movement of people inside spaces can block and impact signal propagation that results to detecting presence of people [50]. The studies train algorithm, consider channel state information and utilize space features such as size, shape, and layout to perform occupancy sensing and counting people [11]. For example, a study uses a Wi-Fi-transmitter and receiver in conference rooms to detect up to 11 persons. As a result, the study depicts an accuracy of 99.1% in occupancy detection, and 92.8% in occupancy counting [50]. Another study trains a deep neural network using ten samples of approximately 3 minutes each and as a result detects the presence of up to 9 persons, and manages to achieve an accuracy of 88.66% [11].

Cameras: Can offer very precise occupancy counting, but require very invasive monitoring that rarely is acceptable to occupants – or even in line with privacy legislation. This requires that cameras are needed to be placed and directed properly to the vision zone. The privacy loss associated with cameras is also related to the camera resolution. High-resolution cameras involve privacy concerns and they need a constant light source to capture images. The accuracy of cameras generally is very good, e.g., Zou et al. [51] showed 95.3% accuracy with only 0.7 delay in detection. Another alternative is to use *infrared cameras* which are able to see in the dark. *Infrared cameras* not only have proven to be suitable solutions for occupancy sensing [8], they also support activity monitoring [19] and can be used to estimate thermal comfort [40]. The main concern with high-resolution infrared cameras relates to their high costs which makes them an undesirable solution for occupancy sensing. Similarly to regular cameras, infrared cameras also are privacy intrusive [7, 32].

Thermal Array Sensors: Are low-cost and low resolution thermal cameras which are both energy-efficient and non-intrusive to occupants [6, 7, 7, 13, 32, 46]. The main challenge with low-resolution sensors relates to their performance, since they are highly sensitive to the overall resolution, frame rate, and the field-of-view of the camera [46]. Although, the low-resolution is these infrared cameras limitation, this limitation enables them involving less privacy concerns, which this is seen as an advantage for using them for occupancy sensing [7, 32]. In the images taken by these cameras, people are often seen as "blobs", and algorithms detects blobs instead of people. For example, using a 16×16 pixel infrared camera in a room, despite low resolution the counting accuracy was high as up to mistakes occurred when counting up to 17 people [7]. Using a camera with 8×8 can also achieve over 90% accuracy, even when tested in uncontrolled environments [12, 33].

Other occupancy sensing methods: Many other methods for occupancy monitoring have been proposed. For example, occupancy can be detected using electricity smart meters [45]. The main limitation with this method relates to the need for precise and case-by-case training models for buildings. Another limitation is the lack of interaction with the power consuming devices, especially sensing the activities at night-times [5]. Another approach is the use of ultrasonic sound that can result an average counting error below 10% in different-sized spaces. This method is vulnerable to changes in the space such as opening windows or doors impacts the results, and requires calibration in larger spaces to remain accurate [42]. Geophone sensors can be used to detect vibrations on the floor, which are then filtered to get only the vibrations caused by stepping. With overlapping sensing areas, the system requires multiple geophone sensors to detect the walking direction, hence, the method can achieve 85% of counting accuracy [39]. Micro-switches are also seen to be a solution to detect occupancy counting, since in workplaces people tend to sit long times at their desks, the micro-switches can be attached to chairs to determine which chairs were in use in the place. However, this approach is vulnerable to moving the chairs away from their intended place which can impact the accuracy of occupancy sensing [25].

7 SUMMARY

We contributed by systematically evaluating the performance trade-offs of low-resolution thermal imagers for occupancy detection. We also established a lower limit for the frame rate and resolution that can support reliable occupancy detection. Through our experiments, conducted on two datasets, we demonstrated that even cameras with a 4×2 resolution – significantly lower than what has been used in previous research [12, 13] – can support

reliable detection, as long as the frame rate is at least 4 frames per second. The lowest tested resolution, 2×2 , can also offer reliable detection rates but requires significantly higher frame rates (at least 16 FPS) and careful adjustment of the viewing angle to ensure all transitions are captured by the sensor. The results are sensitive to the viewing angle, suggesting that the camera's field-of-view needs to be carefully adjusted to maximize the performance of low resolution cameras – or the resolution needs to be selected to be sufficiently high to account for possible variations in viewing angle. As shown in our trade-off analysis, using a camera with a resolution of 4×2 reveals very few insights about the occupants' identity or behaviour, and thus helps to preserve their privacy. Besides privacy, lowering the resolution and frame rate decreases manufacturing and operating costs and helps to make the solution easier to adopt. Our results pave way for widespread adoption of very low resolution thermal sensor arrays as a occupancy detection technology. The main limitation is the need to configure the viewing angle of the cameras carefully, with the optimal placement being directly above doorways. As a result, low resolution thermal sensor arrays are best suited for residential homes and other spaces with clearly identifiable entry points.

ACKNOWLEDGMENTS

This research is supported by the European Regional Funds through the IT Academy Programme.

REFERENCES

- [1] Yuvraj Agarwal, Bharathan Balaji, Rajesh Gupta, Jacob Lyles, Michael Wei, and Thomas Weng. 2010. Occupancy-Driven Energy Management for Smart Building Automation. In *Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Building (BuildSys '10)*. Association for Computing Machinery, New York, NY, USA, 1–6. <https://doi.org/10.1145/1878431.1878433>
- [2] Kaiser Ahmed, Jarek Kurnitski, and Piia Sormunen. 2015. Demand controlled ventilation indoor climate and energy performance in a high performance building with air flow rate controlled chilled beams. *Energy and Buildings* 109 (2015), 115–126. <https://doi.org/10.1016/j.enbuild.2015.09.052>
- [3] Bharathan Balaji, Jian Xu, Anthony Nwokafor, Rajesh Gupta, and Yuvraj Agarwal. 2013. Sentinel: Occupancy Based HVAC Actuation Using Existing WiFi Infrastructure within Commercial Buildings. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems (SenSys '13)*. Association for Computing Machinery, New York, NY, USA, Article Article 17, 14 pages. <https://doi.org/10.1145/2517351.2517370>
- [4] Johannes K Becker, David Li, and David Starobinski. 2019. Tracking Anonymized Bluetooth Devices. *Proceedings on Privacy Enhancing Technologies* 2019, 3 (2019), 50 – 65. <https://doi.org/10.2478/popets-2019-0036>
- [5] Vincent Becker and Wilhelm Kleiminger. 2018. Exploring zero-training algorithms for occupancy detection based on smart meter measurements. *Computer Science-Research and Development* 33, 1-2 (2018), 25–36. <https://doi.org/10.1007/s00450-017-0344-9>
- [6] Alex Beltran, Varick L Erickson, and Alberto E Cerpa. 2013. Thermosense: Occupancy thermal based sensing for hvac control. In *Proceedings of the 5th ACM Workshop on Embedded Systems For Energy-Efficient Buildings*. Association for Computing Machinery, New York, NY, USA, 1–8. <https://doi.org/10.1145/2528282.2528301>
- [7] M. Berger and A. Armitage. 2010. Room occupancy measurement using low-resolution infrared cameras. In *IET Irish Signals and Systems Conference (ISSC 2010)*. IET Digital Library, Cork, Ireland, 249–254. <https://doi.org/10.1049/cp.2010.0521>
- [8] Ningyuan Cao, Justin Ting, Shreyas Sen, and Arijit Raychowdhury. 2018. Smart sensing for HVAC control: Collaborative intelligence in optical and IR cameras. *IEEE Transactions on Industrial Electronics* 65, 12 (2018), 9785–9794. <https://doi.org/10.1109/TIE.2018.2818665>
- [9] Salvatore Carlucci, Marilena De Simone, Steven K Firth, Mikkel B Kjærgaard, Romana Markovic, Mohammad Saiedur Rahaman, Masab Khalid Annaqeeb, Silvia Biandrate, Anooshmita Das, Jakub Wladyslaw Dziedzic, et al. 2020. Modeling occupant behavior in buildings. *Building and Environment* 174 (2020), 106768. <https://doi.org/10.1016/j.buildenv.2020.106768>
- [10] Heather Chappells. 2010. Comfort, well-being and the socio-technical dynamics of everyday life. *Intelligent Buildings International* 2, 4 (2010), 286–298. <https://doi.org/10.3763/inbi.2010.0003>
- [11] Y. Cheng and R. Y. Chang. 2017. Device-free indoor people counting using Wi-Fi channel state information for Internet of Things. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, Singapore, 1–6. <https://doi.org/10.1109/GLOCOM.2017.8254522>
- [12] Veena Chidurala and Xinrong Li. 2021. Occupancy Estimation Using Thermal Imaging Sensors and Machine Learning Algorithms. *IEEE Sensors Journal* 21, 6 (2021), 8627–8638. <https://doi.org/10.1109/JSEN.2021.3049311>
- [13] Mertcan Cokbas, Prakash Ishwar, and Janusz Konrad. 2020. Low-resolution overhead thermal tripwire for occupancy estimation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*. CVF/IEEE, USA, 88–89.
- [14] European Commission. 2016. An EU Strategy on Heating and Cooling. *SWD(2016) 24 final* 1, 1 (2016), 1–13. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016DC0051>

- [15] Hilary Emenike, Farooq Dar, Mohan Liyanage, Rajesh Sharma, Agustin Zuniga, Mohammad A Hoque, Marko Radeta, Petteri Nurmi, and Huber Flores. 2021. Characterizing Everyday Objects using Human Touch: Thermal Dissipation as a Sensing Modality. In *2021 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, IEEE, Kassel, Germany, 1–8. <https://doi.org/10.1109/PERCOM50583.2021.9439120>
- [16] Yunqing Fan, Keiji Kameishi, Shigeki Onishi, and Kazuhide Ito. 2014. Field-based study on the energy-saving effects of CO2 demand controlled ventilation in an office with application of energy recovery ventilators. *Energy and buildings* 68 (2014), 412–422. <https://doi.org/10.1016/j.enbuild.2013.09.043>
- [17] Alan Hedge and Daniel E Gaygen. 2010. Indoor environment conditions and computer work in an office. *Hvac&R Research* 16, 2 (2010), 123–138. <https://doi.org/10.1080/10789669.2010.10390897>
- [18] Arefeh Hesarak and Sture Holmberg. 2015. Demand-controlled ventilation in new residential buildings: consequences on indoor air quality and energy savings. *Indoor and Built Environment* 24, 2 (2015), 162–173. <https://doi.org/10.1177/1420326X13508565>
- [19] Peter Hevesi, Sebastian Wille, Gerald Pirk, Norbert Wehn, and Paul Lukowicz. 2014. Monitoring household activities and user location with a cheap, unobtrusive thermal sensor array. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Association for Computing Machinery, New York, NY, USA, 141–145. <https://doi.org/10.1145/2632048.2636084>
- [20] Naser Hossein Motlagh, Ali Khatibi, and Alireza Aslani. 2020. Toward Sustainable Energy-Independent Buildings Using Internet of Things. *Energies* 13, 22 (2020), 5954. <https://doi.org/10.3390/en13225954>
- [21] Naser Hossein Motlagh, Mahsa Mohammadrezaei, Julian Hunt, and Behnam Zakeri. 2020. Internet of Things (IoT) and the energy sector. *Energies* 13, 2 (2020), 494. <https://doi.org/10.3390/en13020494>
- [22] A. Javed, H. Larijani, A. Ahmadiania, R. Emmanuel, M. Mannion, and D. Gibson. 2017. Design and Implementation of a Cloud Enabled Random Neural Network-Based Decentralized Smart Controller With Intelligent Sensor Nodes for HVAC. *IEEE Internet of Things Journal* 4, 2 (2017), 393–403. <https://doi.org/10.1109/JIOT.2016.2627403>
- [23] Ming Jin, Nikolaos Bekiaris-Liberis, Kevin Weekly, Costas J Spanos, and Alexandre M Bayen. 2016. Occupancy detection via environmental sensing. *IEEE Transactions on Automation Science and Engineering* 15, 2 (2016), 443–455. [10.1109/TASE.2016.2619720](https://doi.org/10.1109/TASE.2016.2619720)
- [24] G. Kalantar, A. Mohammadi, and S. N. Sadrieh. 2018. Analyzing the Effect of Bluetooth Low Energy (BLE) with Randomized MAC Addresses in IoT Applications. In *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*. IEEE, Halifax, NS, Canada, 27–34. <https://doi.org/10.1109/TASE.2016.2619720>
- [25] Timilehin Labeodan, Wim Zeiler, Gert Boxem, and Yang Zhao. 2015. Occupancy measurement in commercial office buildings for demand-driven control applications—A survey and detection system evaluation. *Energy and Buildings* 93 (2015), 303 – 314. <https://doi.org/10.1016/j.enbuild.2015.02.028>
- [26] Judith A Leech, William C Nelson, RICHARD T BURNETT, Shawn Aaron, and Mark E Raizenne. 2002. It's about time: a comparison of Canadian and American time-activity patterns. *Journal of Exposure Science & Environmental Epidemiology* 12, 6 (2002), 427–432. <https://doi.org/10.1038/sj.jea.7500244>
- [27] Titti Malmivirta, Jonatan Hamberg, Eemil Lagerspetz, Xin Li, Ella Peltonen, Huber Flores, and Petteri Nurmi. 2019. Hot or not? robust and accurate continuous thermal imaging on flir cameras. In *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, IEEE, Kyoto, Japan, 1–9. <https://doi.org/10.1109/PERCOM.2019.8767423>
- [28] Célestin Matte, Mathieu Cunche, Franck Rousseau, and Mathy Vanhoef. 2016. Defeating MAC Address Randomization Through Timing Attacks. In *Proceedings of the 9th ACM Conference on Security & Privacy in Wireless and Mobile Networks (WiSec '16)*. Association for Computing Machinery, New York, NY, USA, 15–20. <https://doi.org/10.1145/2939918.2939930>
- [29] Carlyn J Matz, David M Stieb, Karelyn Davis, Marika Egyed, Andreas Rose, Benedito Chou, and Orly Brion. 2014. Effects of age, season, gender and urban-rural status on time-activity: Canadian Human Activity Pattern Survey 2 (CHAPS 2). *International journal of environmental research and public health* 11, 2 (2014), 2108–2124. <https://doi.org/10.3390/ijerph110202108>
- [30] Melexis NV 2019. *Melexis MLX90640 Thermal IR array datasheet*. Melexis NV. <https://www.melexis.com/-/media/files/documents/datasheets/mlx90640-datasheet-melexis.pdf> Rev. 12.
- [31] Sean Meyn, Amit Surana, Yiqing Lin, Stella Oggianu, Satish Narayanan, and Thomas Frewen. 2009. A sensor-utility-network method for estimation of occupancy in buildings. In *Proceedings of the IEEE Conference on Decision and Control*. IEEE, Shanghai, China, 1494–1500. <https://doi.org/10.1109/CDC.2009.5400442>
- [32] Aravind K Mikkilineni, Jin Dong, Teja Kuruganti, and David Fugate. 2019. A novel occupancy detection solution using low-power IR-FPA based wireless occupancy sensor. *Energy and Buildings* 192 (2019), 63–74. <https://doi.org/10.1016/j.enbuild.2019.03.022>
- [33] H. Mohammadmoradi, S. Munir, O. Gnawali, and C. Shelton. 2017. Measuring People-Flow through Doorways Using Easy-to-Install IR Array Sensors. In *2017 13th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, Ottawa, ON, Canada, 35–43. <https://doi.org/10.1109/DCOSS.2017.26>
- [34] Dorte Kragssig Mortensen and Toke Rammer Nielsen. 2011. System design for demand controlled ventilation in multi-family dwellings. *International Journal of Ventilation* 10, 3 (2011), 205–216. <https://doi.org/10.1080/14733315.2011.11683949>

- [35] Naser Hossein Motlagh, Pupu Toivonen, Martha Arbayani Zaidan, Emil Lagerspetz, Ella Peltonen, Ekaterina Gilman, Petteri Nurmi, and Sasu Tarkoma. 2021. Monitoring Social Distancing in Smart Spaces using Infrastructure-Based Sensors. In *IEEE 7th World Forum on Internet of Things (WF-IoT 2021)*. IEEE, New York, NY, USA, 1–6.
- [36] Naser Hossein Motlagh, Martha Arbayani Zaidan, Emil Lagerspetz, Samu Varjonen, Juhani Toivonen, Julien Mineraud, Andrew Rebeiro-Hargrave, Matti Siekkinen, Tareq Hussein, Petteri Nurmi, et al. 2019. Indoor air quality monitoring using infrastructure-based motion detectors. In *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, Vol. 1. IEEE, Helsinki, Finland, 902–907. <https://doi.org/10.1109/INDIN41052.2019.8972332>
- [37] Kazuya Murao, Tsutomu Terada, Ai Yano, and Ryuichi Matsukura. 2012. Detecting room-to-room movement by passive infrared sensors in home environments. In *First Workshop on recent advances in behavior prediction and pro-active pervasive computing*. Association for Computing Machinery, New York, NY, USA.
- [38] T. A. Myrvoll, J. E. Håkegård, T. Matsui, and F. Septier. 2017. Counting public transport passenger using WiFi signatures of mobile devices. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, Yokohama, Japan, 1–6. <https://doi.org/10.1109/ITSC.2017.8317687>
- [39] Shijia Pan, Amelie Bonde, Jie Jing, Lin Zhang, Pei Zhang, and Hae Young Noh. 2014. Boes: building occupancy estimation system using sparse ambient vibration monitoring. In *Sensors and smart structures technologies for civil, mechanical, and aerospace systems 2014*, Vol. 9061. International Society for Optics and Photonics, SPIE, San Diego, California, USA, 90611O. <https://doi.org/10.1117/12.2046510>
- [40] Juhi Ranjan and James Scott. 2016. ThermalSense: determining dynamic thermal comfort preferences using thermographic imaging. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Association for Computing Machinery, New York, NY, USA, 1212–1222. <https://doi.org/10.1145/2971648.2971659>
- [41] Yordan P. Raykov, Emre Ozer, Ganesh Dasika, Alexis Boukouvalas, and Max A. Little. 2016. Predicting Room Occupancy with a Single Passive Infrared (PIR) Sensor through Behavior Extraction. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp '16)*. Association for Computing Machinery, New York, NY, USA, 1016–1027. <https://doi.org/10.1145/2971648.2971746>
- [42] Oliver Shih and Anthony Rowe. 2015. Occupancy Estimation Using Ultrasonic Chirps. In *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems (ICCPS '15)*. Association for Computing Machinery, New York, NY, USA, 149–158. <https://doi.org/10.1145/2735960.2735969>
- [43] Elahe Soltanaghaei and Kamin Whitehouse. 2016. WalkSense: Classifying Home Occupancy States Using Walkway Sensing. In *Proceedings of the 3rd ACM International Conference on Systems for Energy-Efficient Built Environments (BuildSys '16)*. Association for Computing Machinery, New York, NY, USA, 167–176. <https://doi.org/10.1145/2993422.2993576>
- [44] Kailai Sun, Qianchuan Zhao, and Jianhong Zou. 2020. A review of building occupancy measurement systems. *Energy and Buildings* 216 (2020), 109965. <https://doi.org/10.1016/j.enbuild.2020.109965>
- [45] G. Tang, K. Wu, J. Lei, and W. Xiao. 2015. The meter tells you are at home! Non-intrusive occupancy detection via load curve data. In *2015 IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, Miami, FL, USA, 897–902. <https://doi.org/10.1109/SmartGridComm.2015.7436415>
- [46] Ash Tyndall, Rachel Cardell-Oliver, and Adrian Keating. 2016. Occupancy estimation using a low-pixel count thermal imager. *IEEE Sensors Journal* 16, 10 (2016), 3784–3791. <https://doi.org/10.1109/JSEN.2016.2530824>
- [47] Florian Wahl, Marija Milenkovic, and Oliver Amft. 2012. A distributed PIR-based approach for estimating people count in office environments. In *2012 IEEE 15th International Conference on Computational Science and Engineering*. IEEE, Paphos, Cyprus, 640–647. <https://doi.org/10.1109/ICCSE.2012.92>
- [48] Wei Wang, Jiayu Chen, and Tianzhen Hong. 2018. Modeling occupancy distribution in large spaces with multi-feature classification algorithm. *Building and Environment* 137 (2018), 108–117. <https://doi.org/10.1016/j.buildenv.2018.04.002>
- [49] Kevin Weekly, Donghyun Rim, Lin Zhang, Alexandre M Bayen, William W Nazaroff, and Costas J Spanos. 2013. Low-cost coarse airborne particulate matter sensing for indoor occupancy detection. In *2013 IEEE international conference on automation science and engineering (CASE)*. IEEE, Madison, WI, USA, 32–37. <https://doi.org/10.1109/CoASE.2013.6653970>
- [50] Han Zou, Yuxun Zhou, Jianfei Yang, and Costas J. Spanos. 2018. Device-free occupancy detection and crowd counting in smart buildings with WiFi-enabled IoT. *Energy and Buildings* 174 (2018), 309 – 322. <https://doi.org/10.1016/j.enbuild.2018.06.040>
- [51] Jianhong Zou, Qianchuan Zhao, Wen Yang, and Fulin Wang. 2017. Occupancy detection in the office by analyzing surveillance videos and its application to building energy conservation. *Energy and Buildings* 152 (2017), 385–398. <https://doi.org/10.1016/j.enbuild.2017.07.064>