Occupancy Estimation using a Low-Pixel Count Thermal Imager

Ash Tyndall, Rachel Cardell-Oliver, and Adrian Keating, *SM, IEEE*

***Abstract*—An occupancy estimation sensor system based on low-pixel count sensor arrays is proposed. We evaluate a system comprising a 4 by 16 thermal detector array. Machine learning classifiers are used to interpret the raw data from the detector array for deducing the number of occupants in the sensor's field of view. We observe that nominal classifiers provide more robust classification than numerical classifiers. Further, entropy-based classification is applied for the first time for occupancy estimation and found to produce the lowest root mean squared errors (RMSE) and highest correlation coefficients compared to previously trialled classifiers.**

I. Introduction

OCCUPANCY ESTIMATION is the problem of determining the number of people in a given space. Continually knowing the number of occupants in a space can help reduce energy consumption and greenhouse gas emissions in homes or office buildings through more efficient climate control. Occupancy estimation in homes for the elderly or disabled can be used to improve the security and safety of the residents. Robust occupancy estimation is also important for security monitoring, emergency evacuation, and rescue operations.

The ideal occupancy estimation sensor system should be reliable, privacy-preserving, low-cost, non-invasive and energy efficient. Low pixel count sensor arrays satisfy these requirements since they are either the first of a new technology to be developed, or are the lowest cost implementation of a given technology. However, sensor arrays cannot directly sense occupancy, but must be coupled with intelligent software algorithms to infer occupancy counts.

Existing occupancy estimation algorithms for thermal array sensors are either image-based or flow-based [1]. The former analyse a two dimensional image of a scene whereas the latter analyse movement through the monitored area. The image-based approach is investigated in this paper because it is most suitable for occupancy estimation in homes or offices, where people may be stationary and accurate counting is of high priority. Flow-based approaches are best suited to public spaces such as shops and railway stations where large numbers of people are in motion and the proneness of these systems to cumulative counting errors is not so important [2]. Since occupancy estimation with a single sensor modality tends to have low accuracy, ensemble methods have been proposed that combine the results from multiple sensors and classifiers [3].

Ash Tyndall is with the School of Computer Science and Software Engineering, University of Western Australia, Perth, Australia (ash@ash.id.au)

Rachel Cardell-Oliver is with the School of Computer Science and Software Engineering, University of Western Australia, Perth, Australia (rachel.cardell-oliver@uwa.edu.au)

A. Keating is with the School of Mechanical and Chemical Engineering, University of Western Australia, Perth, Australia (adrian.keating@uwa.edu.au)

Manuscript received XX, XX; revised XX, XX.

In this paper we present the development of a complete occupancy estimation system from start to finish, based on low pixel count thermal imaging technology. The design is based on the results of a recently reported thermal array sensor system, Thermosense [4]. Our system differs from Thermosense in the choice of thermal sensor; positioning of the sensor; and selection of classification algorithms. By examining the effects of each of these choices on system performance we contribute new insights for developing robust and general purpose occupancy sensors. The main contributions of this paper are as follows. First, we validate the feasibility and accuracy of occupancy estimation using thermal array sensors. Second, we determine notable properties of these sensing technologies that affect their utility as occupant detectors including detection threshold levels, angled imaging and the sampling rate-noise trade off. Finally, we evaluate the robustness and generality of existing methods when translated to different settings. We were not able to replicate the classification accuracies of the Thermosense study [4], but we did discover that entropy-measure classification algorithms K* and C4.5 gave more accurate predictions than the algorithms used in previous studies. Our system is shown to be fit for purpose: privacy-preserving, low-cost, non-invasive, energy efficient and sufficiently accurate for applications such as controlling climate and lighting in buildings. The results also demonstrate that there are many open challenges for achieving the goal of next generation occupancy estimation sensor systems with very high reliability and the ability to configure themselves in new settings.

II. Background and Related Work

This paper addresses the occupancy estimation problem which is to determine the number of people in an area. Occupancy estimation is a more complex problem than occupancy detection which is to determine if there are any people in an area or not. Occupancy estimation is related to the optical flow or optical turnstile problem which is to detect and count people moving through a space.

Many different human traits and sensor modalities have been investigated for solving the occupancy estimation and detection problems [1]. Either intrinsic or extrinsic traits of people can be used for both tasks. Intrinsic traits are direct measurements of human occupants. They can be either static traits such as thermal emissions [4]–[8], CO2 emissions [9] or still images captured by cameras [5], [10], [11]. Alternatively, dynamic traits of human occupancy can be used, such as

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2016.2530824, IEEE Sensors Journal

movement detected by ultrasound or PIR sensors [12], pressure or movement sensors. Extrinsic traits are environmental factors that are correlated with human occupancy. For example, occupancy can be inferred from human-wearable sensors such as WiFi enabled smart phones [13] or RFID devices [14], if these devices are carried by all occupants. Ambient sensor systems [3], [15] use collections of sensors such as temperature, light, door switches, or electricity consumption or computer use [16] to infer the presence of people in a room.

Our goal is to estimate occupancy for small numbers of people with high accuracy using a low cost thermal array sensor. The recent commercial availability of inexpensive thermal sensor arrays offers new opportunities for achieving this goal. For example, the Melexis MLX90620 [17] is a 16 by 4 thermal array, around US$55, that can detect the thermal signature of people in a room up at distances of 8 m to 15 m [6]. Panasonic's Grid-EYE [4] is an 8 by 8 sensor array, around US$35, that can monitor a 2.5 m square area when mounted on a 3.0 m ceiling, giving pixels of approximately 60 cm square. Research systems utilising low-pixel sensor arrays with motors to enable larger images to be pieced together show good accuracy for controlled test settings [7], [8]. Commercial thermal camera systems such as the Irisys Gazelle people counter [2] and the FLIR ONE module for mobile phones, US$250 [18], are currently too expensive for our low cost requirement, but prices are falling.

Thermal array sensors cannot directly sense occupancy, but must be coupled with intelligent software algorithms to infer occupancy counts. A typical software pipeline for occupancy estimation comprises:

1) pre-processing (noise removal and thresholding);
2) feature extraction (from the thermal image);
3) training one or more classification algorithms to infer occupancy counts from the image features; and
4) combining the results of different classification algorithms, when available.

A few occupancy estimation systems have been proposed using thermal arrays. The ThermoSense system used the Panasonic Grid-EYE and a Passive Infrared Sensor (PIR) [4], [19]. The PIR is used for basic motion detection to identify when the sensing area is empty and the thermal array is then used for occupancy estimation when there is at least one person in the room. Thermosense was tested with a ceiling mounted setting to detect up to 3 people sitting at office desks.

Amin *et al.* [5] used an optical camera and a thermal array for sensing with a back-propagation neural network classifier trained to estimate the number of occupants. A ceiling mounted system was tested for occupancy estimation in the waiting area outside an elevator. It was found that sensor modalities complemented each other with the thermal array being more accurate for more than 6 people and the camera for fewer than 6. An ensemble classifier that combined the results from both systems gave the best performance. Kahn *et al.* presented a hierarchical approach that first determines occupancy or not, next it estimates broad classes of occupant numbers, and finally exact numbers of occupants [15]. Although the system used ambient sensing modalities such as temperature and light, its hierarchical machine learning approach is applicable for other sensor modalities.

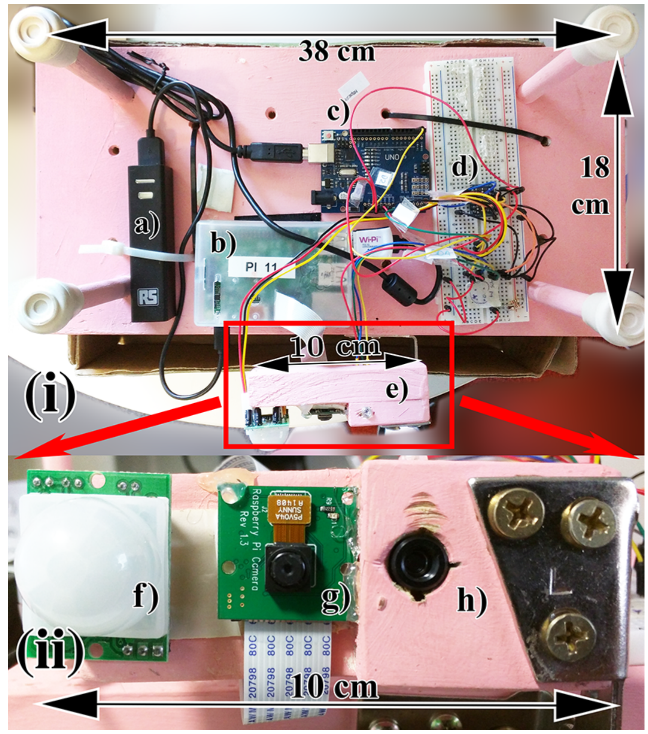


Fig. 1. (i) top view layout of the circuit containing a) Battery pack, b) Raspberry Pi, c) Arduino, d) Level-shifting circuitry, e) Movable sensor mount, f) PIR, g) visible camera, h) MLX90620; (ii) front expanded view of the 3 sensors

Existing systems are typically designed, trained and evaluated for a single space and sensor setting. It is not known how well these systems perform in different settings. Our aim is to support the development of more accurate, robust, and general purpose occupancy counting systems by investigating the contributions of individual steps of the software pipeline to the performance of the whole system.

III. Implementation

Our estimation system consists of a 4×16 thermal detector array (MLX90620) and Passive Infrared Sensor (PIR). A visible camera is used in experiments to evaluate the accuracy of the sensor. All sensors interrogate the same area. The camera and thermal detector array were fixed and subsequently registered so that their images could be overlaid. The arrangement of the sensors is shown in Fig. 1(i), with closer detail on the sensors in Fig. 1(ii). Both thermal and visual data were collected to determine the accuracy of the machine learning classification algorithms used. The camera and passive infrared sensor are directly interfaced to a microcontroller (Arduino), which is subsequently connected into a local computer (Raspberry Pi) for data storage. Registered visual and thermal images are captured simultaneously frame by frame.

Our sensor system uses supervised learning algorithms in which the system is "trained" to identify the number of occupants using labelled training data. The trained algorithm

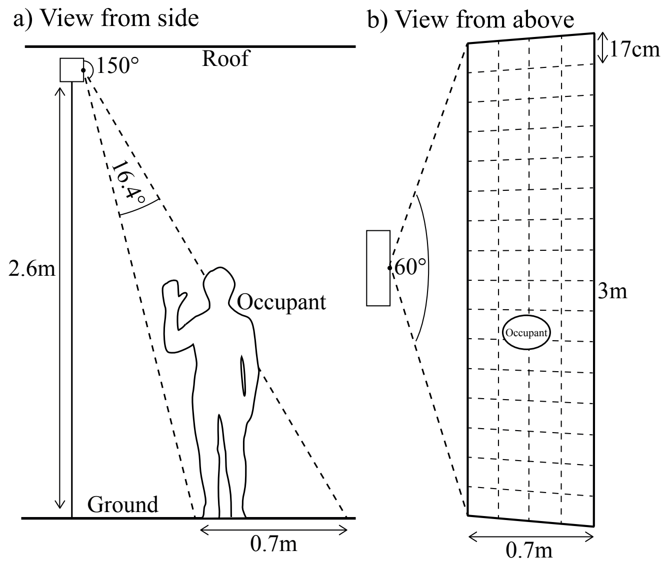


Fig. 2. Physical arrangement of the classifier experiment including (a) side view showing test rig in Fig. 1 mounted at ceiling height, (b) physical extent of the mapped area [20]

can then be used on any unknown (unlabelled) data. During experiments, the video sequence was labelled manually with the number of people in the sensor's field of view. This information is used as ground truth for training purposes. Two types of classification algorithms were considered: numeric and nominal. Numeric algorithms are based on linear regression. A prediction of 2.7 people can be interpreted as 3 people, while a prediction of 2.9 is also interpreted as 3 people but with stronger confidence. Nominal algorithms assign a captured frame to one of a fixed number of classes, for example, 1-person-present, 2-people-present, 3-people-present. Nominal algorithms have additional advantages in situations where the classes may not be identical to the number of people. For example, 1-to-3-people-present, 4-to-7-people-present, 8-or-more-people-present could be more useful categories in some circumstances, while also being potentially easier to train.

In order to make their predictions, machine learning algorithms use feature vectors that describe relevant properties of each captured frame. For example, the number of sufficiently "hot" pixels in a captured frame is one feature that is correlated with the number of occupants, and so can be used for occupancy estimation.

Since the data input to the classification algorithms is based on experimentally obtained images, the training data may be unbalanced in the number of frames belonging to each category. Significantly unbalanced data can degrade the performance of machine learning algorithms. We found that in most cases, the zero-people case was the most frequent, and this case unbalanced the data sets. The thermal array is used only to distinguish between the cases with a positive numbers of occupants, and these cases give a reasonably balanced dataset suitable for machine learning.

For evaluation, our system was mounted on a 2.6 m ceiling, pointing down at a slight angle in order to prevent the interference of the mounting pole and to increase thermal mass detected as shown in Fig. 2(a). The sensor detects occupants in a rectangular area of approximately 3 m by 0.7 m in which up to three people may be present. This is similar to the spaces monitored in previous studies [3], [4]. Multiple thermal arrays or incorporating a motor to move the sensor [7], [8] could be used to monitor larger areas. Our mounting gave pixels of approximately $17\,\mathrm{cm}^2$. People may be only partially visible in the field of view and this edge-case is an important consideration for the detection system. Temporal sequences of captured images alternate between periods where the number of occupants is changing (people entering or leaving) and stable periods.

The scenario of an MLX sensor mounted in a corridor or large laboratory was considered in [6]. Using the 60 by 15 degree option, the field of view at 8 meters from the sensor would be 9.2 m by 2.1 m giving pixels of approximately $56\,\mathrm{cm}^2$. For this scenario a standing person would occupy an area of around 4 by 4 pixels [6].

All experiments involved between one and three people, entering from the left and exiting from the right. Experiments included walk through, persons adding to a scene by standing or sitting, or persons exiting a scene, sequentially or simultaneously. The people involved were of average height, wearing various clothing. The room was monitored to be 18 °C during these experiments. Each experiment was recorded using the visible camera, the thermal sensor array and the passive infrared sensor (PIR sensor), synchronized at 1 Hz over approximately 10×60-120-second intervals. Each experiment had 10-15 frames at the beginning where nothing was within the view of the sensor to allow the thermal background to be calculated. Each frame generated from these experiments was manually tagged with the ground truth value of its number of occupants.

IV. Features

The raw thermal data is transformed into a set of features to be used by the classification algorithms. The first task is to distinguish between background infra-red radiation in the room from static objects such as computers or televisions, and radiation from humans. We used the background separation algorithm and features suggested by Beltran *et al.* [4] to enable comparison. Occupants are separated from background infra-red radiation through the use of an image subtraction algorithm maintaining per-pixel mean and standard deviation values to update a thermal background map. If no motion is detected for 15 minutes, this map is updated using a slow-moving Exponential Weighted Moving Average (EMWA). Additionally, if the room remains occupied for a long period, a more complex scaling algorithm is used which considers the coldest points in the room empty, and averages them against the new background, then performs an EMWA with a lower weighting. The result of this background differencing stage is an array of pixels which are either "occupied" or "empty". To control the experimental scope more tightly, we opted to exclude the use of this more complex scaling algorithm and PIR-based non-occupancy detection from our data, and instead used a manual indication of non-occupancy.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2016.2530824, IEEE Sensors Journal

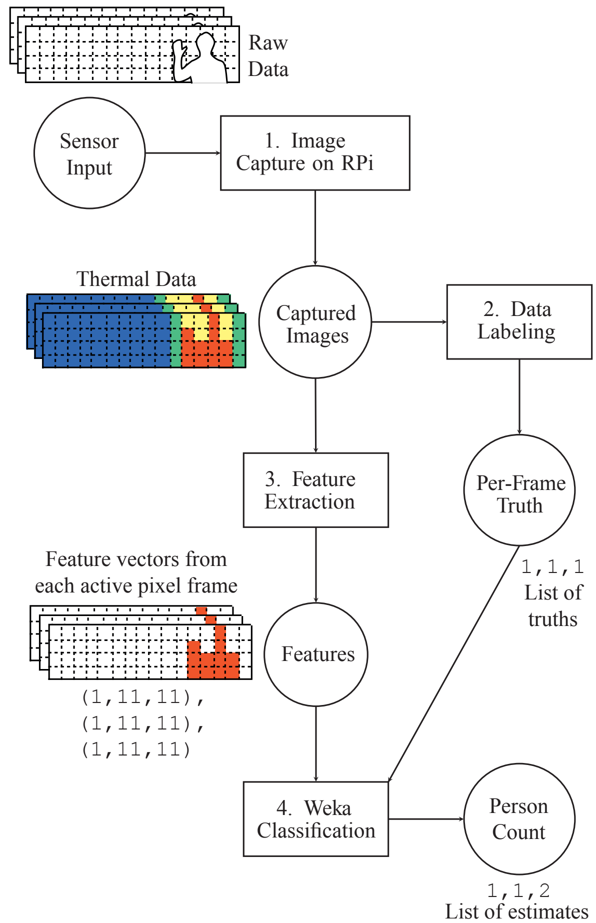


Fig. 3. Process flow diagram for turning raw sensor input into occupancy estimates

The second task is to extract features that can be used to infer occupancy from the pixel array. The goal is to do this in such a way that the algorithms work well when applied to different spaces. Beltran et al. [4] identified three features that were correlated with the number of human occupants and not too susceptible to individual room conditions, the size of people, or the mounting position of the sensor. Each feature is correlated with the number of occupants in the field of view. Using multiple features makes detection more robust than with a single feature. Three different features were used:

1) Number of active pixels: The total number of "occupied" pixels in a given frame
2) Number of connected components: If each active pixel is joined with its immediate active neighbors, how many "islands" of active pixels (termed connected components in graph theory) exist.
3) Size of the largest connected component: The number of active pixels contained within the largest connected component.

V. CLASSIFIERS

A selection of machine learning algorithms were tested to classify occupancy based on the features. To perform classification tests we use version 3.7.12 of the open-source Weka toolkit [21], which provides easy access to a variety of machine learning algorithms and the tools necessary to analyze their effectiveness. A process flow which describes the process from the collection of the 4×16 thermal sensor data through the Weka classification algorithm to determine occupancy is show in Fig 3.

For comparison, we replicated three machine learning classifiers used by Beltran *et al.* [4], namely K-Nearest Neighbors (KNN), a Multilayer Perceptron Artificial Neural Network (MLP) and a Linear Regression model. When evaluating their test cases, for best comparison we used their approach of limiting the maximum number of people to 3 and for the MLP, using 70% of the data for training the neural net, 15% for testing the net and the remaining 15% for validating results. For the Artificial Neural Network we used Weka's "MultilayerPerceptron" neural network which creates a hidden layer of five neurons given as:

$$0.5(\text{attributes} + \text{classes})$$

It uses a sigmoid activation function for all neurons, except in the case that a numerical answer is to be predicted, in which case, it uses a linear activation function for the output layer.

For the K-Nearest Neighbors (KNN) implementation we used $k = 5$ with Euclidean distance as the distance metric between feature vectors. The Linear Regression model used is $y = \beta_A A + \beta_S S + \beta$, where A is the number of active pixels, S is the size of the largest connected component, and the β values represent the corresponding coefficients. Following Beltran *et al.* [4], we excluded the number of connected components feature in the Linear Regression model since they found it decreased classification accuracy.

As well as replicating the classification algorithms used by Beltran *et al.* [4] we evaluated several other classification algorithms, as summarised in Table I. 0-R is used as a lower bound case. For nominal prediction it classifies all new data as belonging to the category that is most common in the training data. For numeric prediction it classifies all new data as the mean of the test data.

Most of these techniques are well known machine learning approaches [22], but the KStar (K*) algorithm is less common and so we summarise the approach. The K* algorithm, developed by Cleary and Trigg [23] is a type of K-Nearest Neighbors algorithm. Instead of Euclidean distance to compare neighbours, K* uses entropic distance: a measure of how much effort is required to convert one example into another. This has several positive effects: it makes the algorithm more robust to missing values and enables the classifier to output either a numeric or nominal result. We chose K* because it allows the investigation of KNN-like techniques for the numeric classification and provides a second entropy based method in addition to C4.5. For all tests where not specifically stated otherwise, we use 10-fold cross-validation to validate our results.

Table II shows the accuracy results for each classification algorithm on our experimental data set. For comparison, the results reported by Beltran *et al.* [4] are also shown. We replicated the parameters and classification algorithms reported by Beltran. Our results were within 5% of RMSE

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2016.2530824, IEEE Sensors Journal

TABLE I
WEKA PARAMETERS USED FOR SELECTED CLASSIFICATIONS ALGORITHMS

Type	Attribute	Weka Class & Parameters
Neural Network (ANN)	Nominal, Numeric	weka.classifiers.functions.Multilayer Perceptron -L 0.3 -M 0.2 -N 500 -V 15 -S 0 -E 20 -H 5
K-Nearest Neighbours (KNN)	Nominal, Numeric	weka.classifiers.lazy.IBk -K 5 -W 0 -F -A "weka.core.neighboursearch.LinearNNSearch -A \"weka.core.EuclideanDistance -R first-last\""
Naive Bayes	Nominal	weka.classifiers.bayes.NaiveBayes
Support Vector Machine (SVM)	Nominal	weka.classifiers.functions.SMO -C 1.0 -L 0.001 -P 1.0E-12 -N 0 -V -1 -W 1 -K "weka.classifiers.functions.supportVector.Poly Kernel -C 250007 -E 1.0"
C4.5 Decision Tree	Nominal	weka.classifiers.trees.J48 -C 0.25 -M 2
K*	Nominal, Numeric	weka.classifiers.lazy.KStar -B 20 -M a
Linear Regression	Numeric	weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8
0-R	Nominal, Numeric	weka.classifiers.rules.ZeroR

TABLE II
RESULTS OF CLASSIFICATION EXPERIMENTS

Classifier	RMSE[1]	Precision (%)	Correlation (r)
Thermosense Reported Results [4]			
KNN[2]	0.346		
Lin Reg[3]	0.385		0.926
MLP	0.409		0.945
Thermosense Replication on our Data			
KNN (Nom)[2]	0.364	65.65	
MLP	0.592		0.687
Lin Reg[3]	0.525		0.589
KNN (Num)[2]	1.123		0.377
Numeric Classification Algorithms			
K*	0.423		0.760
0-R	0.651		-0.118
Nominal Classification Algorithms			
K*	0.304	82.56	
C4.5	0.314	82.39	
MLP	0.362	77.14	
SVM	0.398	67.18	
N. Bayes	0.405	63.59	
0-R	0.442	49.74	

[1]: Model deviation from occupant ground truth
[2]: Includes zero occupant cases in training data
[3]: Excludes number of connected components feature
%: Precision, measuring a nominal test result
r: Pearson's r, measuring a numeric test result

for the KNN model using nominal data, but our best attempts yielded RMSE 36% and 44% higher for the Linear Regression and MLP models, respectively. This agrees with the much lower correlation coefficients (r) value in our replicated data. We believe these differences are due to the angle inherent in our imaging system shown in Fig. 2 compared with ceiling mounted sensors, as well as the rectangular field of our sensor rather than a square field of view. Our setting allowed for more situations in which people were partially included in the scene than the Thermosense experiments, which may also have affected accuracy. Nevertheless, such physical arrangements and occupancies are typical in installations and the robustness of classifiers should take this situation into account.

For our classification algorithms, accuracy was significantly improved for nominal classification, where in some cases we exceeded the RMSEs reported by Beltran *et al.* [4]. Interestingly, within our dataset, the K* and C4.5 algorithms were most accurate, with accuracies of 82.56% and 82.39% respectively. The RMSEs of 0.304 and 0.314 (for K* and C4.5 respectively) were a significant improvement on the best case Thermosense KNN RMSE of 0.346. Further numeric classification with K*, showed better correlation than any replicated Thermosense technique, with $r = 0.760$. Additionally, the K* RMSE of 0.423 was also superior.

For all the results reported in Table II, occupancy count is predicted based on a single thermal snapshot of the field of view. It does not take into account the past history of occupancy. Snapshots are recorded once a second. In practice, predicting the number of people in a room during an interval of seconds or even minutes is sufficient for most occupancy estimation applications such as climate control and security. So we investigated whether the fine-grained experimental data could be used to obtain more accurate results at a coarser time granularity. Erikson *et al.* approached this problem by using a weighted moving average filter over the time series of frame-by-frame predictions to define the final prediction count. They did not, however, give separate results for the effect of this step over the unfiltered classification results. A weighted average approach can correct for one-off point anomalies in the predictions (that is, a single incorrect prediction within a series of correct ones). But weighted averaging also delays response to true changes of occupancy in the ground truth. We tested window sizes from 2 seconds to 1 minute and found that weighted average filters were not effective on our experimental data, but that it actually decreased prediction accuracy. Another approach is to use a coarser temporal grain for both ground truth and prediction. We defined ground truth accuracy per W-second window as the most frequent occupancy count occurring in that window. Predicted occupancy over a W-second window was defined in the same way, based on the 1-second prediction data. With this filtering approach the best case prediction accuracy was over 12% higher using a 40-second window, compared with a 1-second window. In future work we plan a more detailed investigation of which types of temporal features are best for improving the accuracy and robustness of occupancy sensing.

VI. SENSOR CHARACTERISTICS

This section considers how sensor characteristics of the thermal detector array affect occupancy estimation. Fig. 4

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2016.2530824, IEEE Sensors Journal

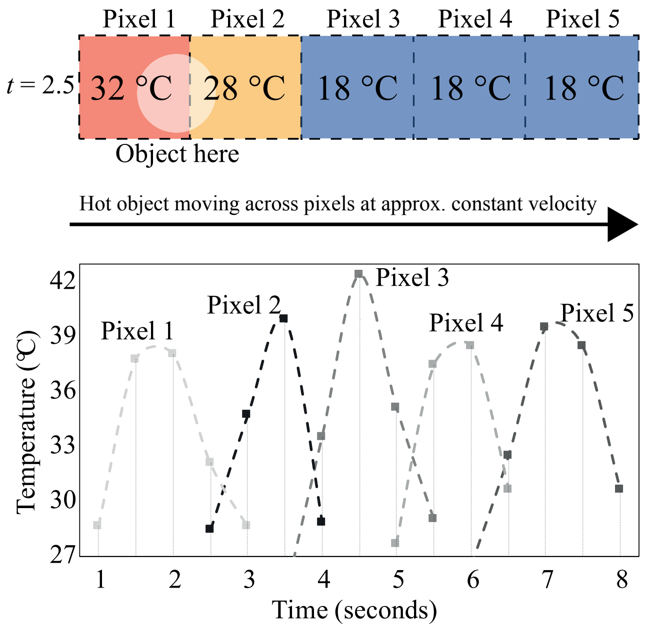


Fig. 4. Temperature plot a singe line of pixels in the MLX90620's thermal array as a hot object moves across them at a constant velocity

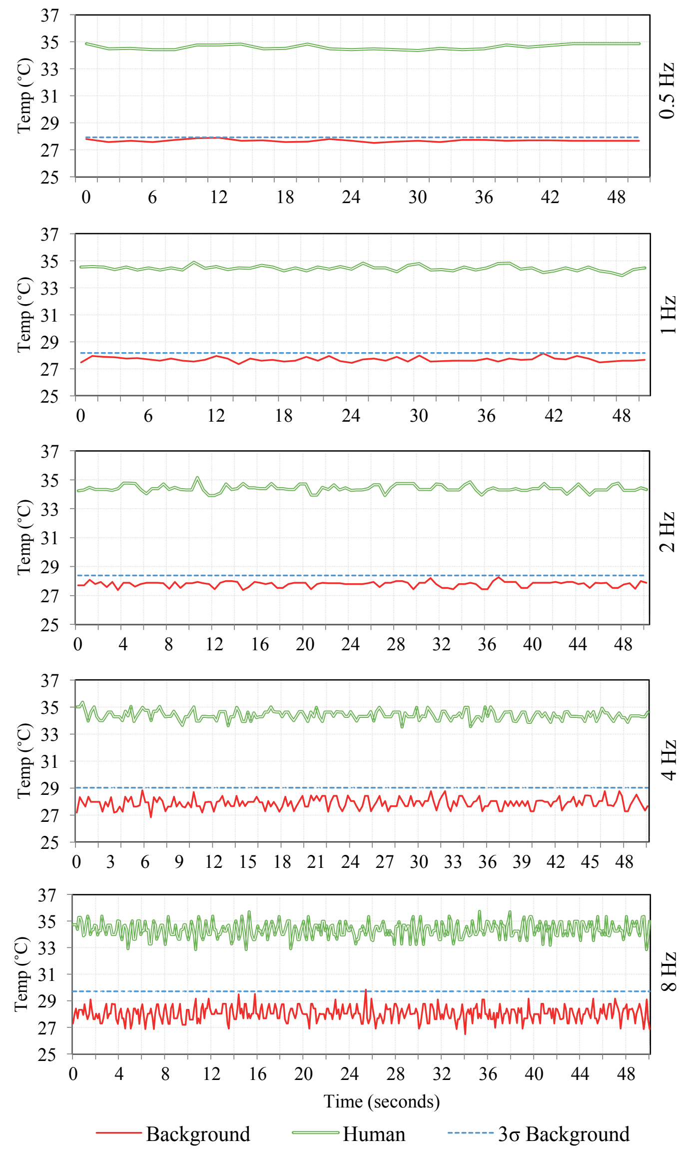


Fig. 5. Plot of occupant and background sensor noise at sampling speeds 0.5 Hz – 8 Hz

shows the output from six of the sensor's central pixels as a hot object is moved from left to right at approximately constant speed. One of the most interesting phenomena in this graph is the variability of the object's detected temperature as it moves "between" two different pixels. Each pixel has a peak in the detected temperature, which rolls off as the object passes the field of view. The data fits a Lorentzian with a full-width half maximum of 1.6 s. The time axis can also be interpreted as a spatial extent, given the constant velocity at which the hot object moves pass the detector. For the 60° field of view for the sensor we calculated that a hot object traveling past at 0.2 m/s at a distance of 4 m would give rise to a thermal signal on the detector with a 1.44 s FWHM, consistent with the measurement. This signal variation across a single detector is not anticipated to impact our intended use as an occupancy detector, as it only arises for hot objects which are very close or very small (small angular extent).

One of the features of the MLX is the ability to sample the thermal data at a variety of sample rates between 0.5 Hz and 512 Hz, where a higher sample rate results in larger thermal fluctuations. Since our algorithm separates occupants from a thermal background, it is important to determine if the noise could affect our ability to do this accurately. Fig. 5 shows the temperature of a central pixel of the sensor at different frame rates. Each measurement contains two datasets, one from the pixel viewing a person and the other viewing only the thermal background. As expected, the noise in the thermal data increases significantly with the sample rate. Further, rather than simply downsampling to a lower rate, the sensor integrates the data over the period of the sampling rate, averaging the data over longer periods at lower sample rates which decreases the variation in data. In each dataset we include the 3σ level of the background thermal variation. Above the set level, the probability that the thermal noise will be detected as a person is less than 0.13%. The observation here is that the sampling rate and threshold level to determine occupancy are fundamentally linked in this type of sensor array.

In the 0.5 Hz case, the third standard deviation above background (3σ) is $\Delta T = 6.4\,°C$ below the minimum occupancy value (o) detected, which we define as a guard-band (ΔT). As sampling rate increases, this guard-band between the occupancy value and the 3σ-background threshold slowly decreases, becoming 5.75 °C for 1 Hz, 5.53 °C for 2 Hz, 4.48 °C for 4 Hz, and only 3.15 °C for 8 Hz. This data, shown in Fig. 6, indicates the variation of the thermal guard-band and suggesting a theoretical upper limit of 25 Hz for the sampling rate. However, in none of the cases studied was $3\sigma \geq \min(o)$, which would cause a false positive. The variation of the output signal with the square root of the frequency is consistent with a thermal sensor which is thermal or Johnson noise limited [24].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2016.2530824, IEEE Sensors Journal

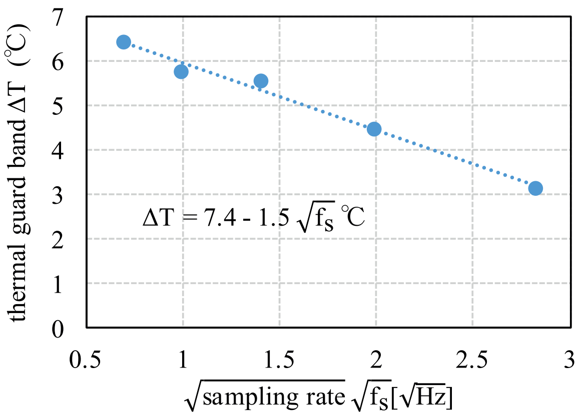


Fig. 6. Thermal guard band variation as a function of the sampling rate

VII. Discussion

Creating a system that is wholly automated and can detect occupants with a high level of accuracy is important to ensure that climate control, patient security and other occupancy-estimation tasks are reliably executed. Our results differ from the RMSE values of Thermosense with the K-Nearest Neighbors, Linear Regression, and Multi-Layer Perceptron classifiers. We believe this was because the classifiers Thermosense used were highly sensitive to their sensor's specific properties such as pixel arrangement (Grid-EYE 8×8 pixel array), the inclusion of partial people within the sensed region, and a flat sense region (no projection angle). Our sensor contained 4×16 pixels which better suits rectangular spaces and corridors, and we considered only integer values of people in our models as this appeared the most appropriate metric for decision making in applications such as climate control. However among our own selected machine learning algorithms, K* and C4.5 achieved accuracies in the 80%+ range. These algorithms also improved upon Thermosense's best RMSE by 12% and 9% respectively. Both of these algorithms leverage entropy measures as a way of partitioning data, suggesting for the first time that that entropy-based approaches may be more suited to this application. Using the K* or C4.5 machine learning algorithm, we are confident that this prototype could achieve similar levels in practical applications.

VIII. Conclusions

This paper investigates the performance of an occupancy estimation system based on a thermal detector array. Different machine learning classifiers were compared, using training data obtained from a 4 by 16 thermal detector array and a passive infrared sensor. The results showed poor agreement with previously reported systems that used a square-grid for estimation. Our experiments considered implementation challenges such as detection threshold levels (noise), angled occupancy imaging and non-square (rectangular) fields of view. Understanding the sampling rate-noise trade off of thermal detector array sensors was important to ensure robust estimation. We observed that nominal classification algorithms performed better than numerical ones and that employing additional features, such as recent occupancy history, could improve classification performance. For the first time, entropy based classifiers were considered for occupancy estimation, and found to give the best performance in terms of the lowest root mean squared errors (RMSE) and highest correlation coefficients. In particular, the K* algorithm was found to be the best classifier, with an accuracy of 82.56% and a RMSE of 0.304.

References

[1] T. Teixeira, G. Dublon, and A. Savvides, "A survey of human-sensing: Methods for detecting presence, count, location, track, and identity," Embedded Networks and Applications Lab (ENALAB), Yale University, Tech. Rep., 2010, retrieved September 1, 2015 from http://www.eng.yale.edu/enalab/publications/human_sensing_enalabWIP.pdf.

[2] Irisys, "Gazelle dualview people counter from irisys," http://www.irisys.net/products-gazelle-dualview-people-counter, retrieved January 13, 2016.

[3] Z. Yang and B. Becerik-Gerber, "Cross-space building occupancy modeling by contextual information based learning," in *Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, ser. BuildSys '15. New York, NY, USA: ACM, 2015, pp. 177–186. [Online]. Available: http://doi.acm.org/10.1145/2821650.2821668

[4] 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*. ACM, 2013, pp. 1–8.

[5] *Automated people-counting by using low-resolution infrared and visual cameras*, vol. 41, no. 6, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0263224107001078

[6] "White paper mlx90620 for smart & green HVAC," http://www.melexis.com/Assets/MLX90620-for-smart-and-green-HVAC-6256.aspx, April 2013, retrieved January 13, 2016.

[7] K. Hashimoto, T. Tsuruta, K. Morinaka, and N. Yoshiike, "High performance human information sensor," *Sensors and Actuators A: Physical*, vol. 79, no. 1, pp. 46 – 52, 2000. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924424799002472

[8] N. Yoshiike, K. Morinaka, K. Hashimoto, M. Kawaguri, and S. Tanaka, "360-degree direction type human information sensor," *Sensors and Actuators A: Physical*, vol. 77, no. 3, pp. 199 – 208, 1999. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0924424798003677

[9] E. Hailemariam, R. Goldstein, R. Attar, and A. Khan, "Real-time occupancy detection using decision trees with multiple sensor types," in *Proceedings of the 2011 Symposium on Simulation for Architecture and Urban Design*. Society for Computer Simulation International, 2011, pp. 141–148.

[10] V. L. Erickson, S. Achleitner, and A. E. Cerpa, "POEM: Power-efficient occupancy-based energy management system," in *Proceedings of the 12th international conference on Information processing in sensor networks*. ACM, 2013, pp. 203–216.

[11] J. Serrano-Cuerda, J. C. Castillo, M. V. Sokolova, and A. Fernández-Caballero, "Efficient people counting from indoor overhead video camera," in *Trends in Practical Applications of Agents and Multiagent Systems*. Springer, 2013, pp. 129–137.

[12] 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*. ACM, 2012, pp. 309–322.

[13] B. Balaji, J. Xu, A. Nwokafor, R. Gupta, and Y. Agarwal, "Sentinel: occupancy based HVAC actuation using existing WiFi infrastructure within commercial buildings," in *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2013, p. 17.

[14] N. Li, G. Calis, and B. Becerik-Gerber, "Measuring and monitoring occupancy with an RFID based system for demand-driven HVAC operations," *Automation in construction*, vol. 24, pp. 89–99, 2012.

[15] A. Khan, J. Nicholson, S. Mellor, D. Jackson, K. Ladha, C. Ladha, J. Hand, J. Clarke, P. Olivier, and T. Plötz, "Occupancy monitoring using environmental & context sensors and a hierarchical analysis framework," in *Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, ser. BuildSys '14. New York, NY, USA: ACM, 2014, pp. 90–99. [Online]. Available: http://doi.acm.org/10.1145/2674061.2674080

[16] 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*. ACM, 2013, pp. 1–8.

[17] Melexis, "Datasheet IR thermometer 16X4 sensor array MLX90620," http://www.melexis.com/Infrared-Thermometer-Sensors/Infrared-Thermometer-Sensors/MLX90620-776.aspx, 2012, retrieved January 7, 2015.

[18] SparkFun, "FLiR Dev Kit," https://www.sparkfun.com/products/13233, retrieved April 8, 2015.

[19] V. L. Erickson, "Usage based building management through wireless sensor networks," Ph.D. dissertation, Electrical Engineering and Computer Science, University of California Merced, 2014. [Online]. Available: https://escholarship.org/uc/item/5md4c017

[20] L. S. Sagan, Tompw, and Holek, "Human outline," Wikimedia Commons, Creative Commons BY-SA licensed, adapted for figure. [Online]. Available: https://commons.wikimedia.org/wiki/File:Human_outline.svg

[21] University of Waikato, "Weka," http://www.cs.waikato.ac.nz/ml/weka/, retrieved September 1, 2015.

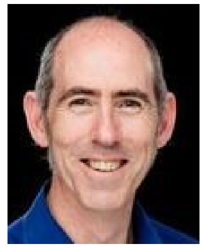
[22] J. Han, M. Kamber, and J. Pei, *Data Mining: Concepts and Techniques*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2011.

[23] J. G. Cleary, L. E. Trigg *et al.*, "K*: An instance-based learner using an entropic distance measure," in *Proceedings of the 12th International Conference on Machine learning*, vol. 5, 1995, pp. 108–114.

[24] F. Niklaus, C. Jansson, A. Decharat, J.-E. Källhammer, H. Pettersson, and G. Stemme, "Uncooled infrared bolometer arrays operating in a low to medium vacuum atmosphere: performance model and tradeoffs," in *Proceedings of SPIE, Infrared Technology and Applications XXXIII*, vol. 6542, 2007, p. 65421M.

Ash Tyndall received the BCompSc degree in Systems, Computation and Web Design, and the BCompSc(Hons) degree in thermal occupancy sensor systems from the University of Western Australia. He is a recent graduate currently working in the field of anomaly detection in video surveillance footage. His research interests include wireless sensor networks, the Internet of Things, thermal sensing and occupancy estimation.

Rachel Cardell-Oliver received the MSc degree in distributed systems software from the University of Western Australia and the PhD degree in protocol verification from the Computer Laboratory, University of Cambridge. She is an Associate Professor in the School of Computer Science and Software Engineering, University of Western Australia, Perth, Australia. Her research interests include wireless sensor networks for environmental applications, smart metering and data mining.



Adrian Keating (M'90-SM'07) was born Melbourne, Australia in 1967. He received his B.E. (Hon) and Ph.D. (Photonics) degrees in electrical and electronic engineering from the University of Melbourne, Australia in 1990 and 1995, respectively. Since 1996 he has worked at NTT Research Labs (Musashino-shi, Japan), the University of California, Santa Barbara and at Calient Networks as the Fiber Optics Technology Manager. He joined the School of Electrical, Electronic and Computer Engineering at the University of Western Australia (UWA) in 2004 and later the School of Mechanical Engineering where he is currently an Associate Professor. His current research activities are in infrared optics sensors, sensor systems, optical microelectro-mechanical systems (MEMS), and porous silicon based sensor technologies.