

# Multiple thermal sensor array fusion towards enabling privacy-preserving human monitoring applications

Abdallah Naser , Ahmad Lotfi  and Junpei Zhong 

**Abstract**—Human-centric applications of a single Thermal Sensor Array (TSA) have performed extremely well in many areas. However, most of these works have not yet reached the real applicability stage of the Internet of Things (IoT) applications. The main limitation of deploying such systems on a large scale is the challenge of fusing multiple TSAs to cover a wide inspection area, e.g. smart homes, hospitals and many other domestic environments. On the other hand, objects that appear in the low-resolution thermal images acquired from TSA have low intra-class variations and high inter-class similarities, making the identification of the overlapping regions through matching a comparable template image in multiple images very difficult. This paper proposes a motion-based approach to fuse multiple TSAs and learn the domestic environment layout to enable further human-centred IoT applications to run in the cloud. Besides, a privacy-improvement on utilising these sensors in IoT applications is proposed. The proposed approach is evaluated with comprehensive experiments on different sensor placements and domestic environment conditions. This paper shows an average performance of 92.5% accuracy using various machine learning techniques and use case scenarios.

**Index Terms**—sensor fusion, human-centred approach, Internet of Things, thermal sensor array, machine learning, optical flow, privacy-preserving approach.

## I. INTRODUCTION

THE emerging and necessity for the Internet of Things (IoT) applied for healthcare applications have gained increasing attention in recent years. Specifically, in the field of assistive technologies, various types of IoT devices enable older adults to live independently in their own homes to cope with the increasing cost of long-term care demands [1]. The sensing technologies applications on these IoT healthcare devices can be classified into three main categories: first, wearable-based sensors, which usually require the users to wear or carry a device perpetually. This is inconvenient for the elderly, and it is even more difficult to be managed by older adults with Dementia, e.g. Alzheimer’s disease, as there is a high probability of forgetting to carry these devices. On the other hand, there are still outstanding limitations in designing wearable devices such as energy-efficient, fabrication and lightweight [2]. Second, ambient sensing devices

such as Passive Infra-Red (PIR) sensors are installed in a home environment. Such devices preserve privacy but do not generally perform well in multi-occupancy home scenarios [3]. Other privacy-preserving device-free sensing methods, including Wireless Fidelity (WiFi), Radar, and Radio Frequency Identification (RFID), suffer from notable limitations in indoor human monitoring applications such as vulnerability to environmental interference [4]. Unlike conventional radars, Millimeter-Wave (mmWave) uses a short wavelength, which enables its radar to achieve high resolution and small antenna size, but makes it vulnerable to noise in indoor applications [5]. Third, vision-based sensing, for example, cameras that perform very well in real-world scenarios, although it violates users’ privacy [6], clearly in domestic environments, e.g. homes, hospitals, nursing homes etc.

Recently, there has been a growing interest in using Thermal Sensor Array (TSA) to fine-tune the trade-off between privacy and performance in domestic environment applications [7]–[13]. The motivation behind using this sensor in IoT applications is its advantage to perform well while maintaining privacy, low cost, and non-contact capabilities. Nevertheless, most of the works that use the TSA have not yet reached the deployment stage. This is because the TSA fusion has not been comprehensively considered for the implementation of IoT applications. This is a critical flaw in most of the recent work since the Fields of View (FoV) of the TSA is not wide enough to cover the entire living spaces such as a home environment (e.g. the FOV of MLX 90640 TSA sensor is  $55^\circ \times 35^\circ$ ). Hence, multiple sensors are required to monitor such environments. Identifying the overlapping regions between two or more TSA sensors’ FoV has not yet been explored. This paper explores the apparent motion pattern of moving objects from multiple sensors to integrate sensory data and identify overlapping regions. By doing so, TSA will be applicable in real-world IoT applications to cover a wide inspection area without the prerequisite for home layout or sensor placement restrictions. In summary, the main contributions of this paper include:

- A novel approach to fusing multiple TSAs to cover wide inspection area to enable further human-centred applications to run in a central cloud platform;
- The proposed approach can identify overlapping regions between two or more low-resolution TSAs;
- An improvement for the sensor privacy feature in IoT applications to avoid the human image being reconstructed by a third party during data transmission and storage in

Abdallah Naser is a PhD student within the Computational Intelligence and Applications Research Group at Nottingham Trent University, Nottingham, NG11 8NS, United Kingdom (email: abdallah.naser@ntu.ac.uk).

Ahmad Lotfi is the Head of Computational Intelligence and Applications (CIA) Research Group at Nottingham Trent University, Nottingham, NG11 8NS, United Kingdom (email: ahmad.lotfi@ntu.ac.uk).

Junpei Zhong is a research assistant professor at the Hong Kong Polytechnic University, Hong Kong (email: joni.zhong@polyu.edu.hk).

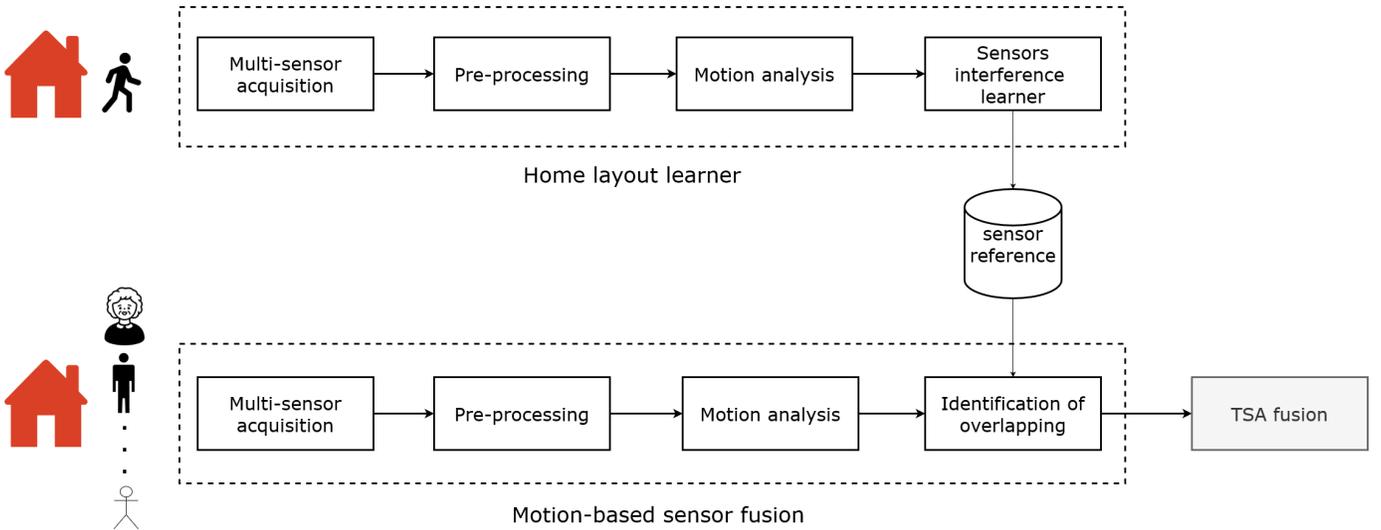


Fig. 1: A schematic diagram of the proposed approach for home layout learning and multiple TSAs fusion through the analysis of apparent motion pattern of moving subjects in the acquired sensors' signals.

the cloud;

- The proposed approach is adaptive to work in different indoor layouts and sensor placements through proposing an environmental layout learner;
- Comprehensive experiments to validate the proposed approach on various sensor placements and domestic environment conditions. In particular, to experimentally validate the feasibility of optical flow in verifying whether a subject appearing in two different sensor outputs in a domestic multi-occupancy environment is the same person.

The remaining parts of this paper are organised as follows: in Section II a summary of the related work is presented. Section III explains the proposed framework architecture. Comprehensive experiments and analysis are presented and discussed in Sections IV, and V followed by pertinent conclusions drawn in Section VI.

## II. RELATED WORK

The TSA sensors have recently come in light through emerging them in responses to global COVID-19 pandemic [14]–[18]. Unlike PIR sensors, TSA sensors can detect motionless warm objects and moving objects with the direction of their movements within their FoV through employing conventional PIR motion detectors in what is called thermopile array [19]. Therefore, TSA has also been proposed for several human-centred applications [20]–[25]. However, none of these works discovers the scenarios of using multiple TSA's. This is a critical issue, for example, in the occupancy estimation systems, which utilises TSA [26]–[31]. This is because humans located in the overlapping FoV's of two or more sensors will be counted as two subjects in the prediction stage, which leads to a wrong occupancy estimate.

Similarly, there has been recent works on using the TSA on human activity recognition, and abnormal behaviour detection [32]–[36]. The approach followed to process the TSA output is similar to image-processing approaches [37] while the

analytical techniques on individual time intervals, frames, were different for instance Support Vector Machine (SVM) [12], Adaptive Boosting [6], [10], K-Nearest Neighbour (KNN) [30], [38], decision trees [20], [39], and Kalman filtering [40], [41]. One of the notable technical challenges reported in most human-based applications, which use TSA is that external heat sources have a major negative impact on the system performance. On the other hand, TSA's fusion has not yet been investigated in such applications. Considering fall as abnormal human behaviour that could happen in an overlapped area between two sensors (One sensor would not be sufficient to cover the entire environment, e.g. older adult home). The proposed systems may trigger two separate fall alerts incorrectly. Moreover, identifying overlapped regions between multiple TSAs significantly impacts other human monitoring applications, including occupancy estimation. The impact of this can be clearly demonstrated in situations where a person could be present in the overlapped region of two sensors, and thus the system might consider them as two people in the environment rather than one.

The fusion concepts have been applied in various applications [42], including biometric authentication systems [43]–[45], air pollution monitoring [46], COVID-19 non-remedial solutions [47], surveillance networks [48], vehicle accident detection and classification [49], human emotion monitoring [50] and many other vital applications.

To summarise, TSA sensors are promising emerging sensing technologies for human-centric applications due to their low cost, privacy-preserving and non-contact capabilities. Following our previous studies reported in [6], [8] and [13] on occupancy and human distance estimation to identify the multi-occupancy environment, human-to-human distance and human-to-sensor distance using single-based TSA processing, this paper intends to address the challenge of enabling TSA to cover a wide inspection area by proposing a free-position sensor placement and fusion approach, which can identify overlapping sensor areas from low-resolution thermal images.

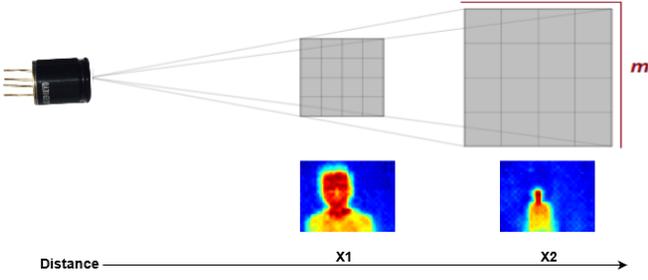


Fig. 2: A visualisation of the relationship between the sensor FoV's depth and the inspection area.

### III. THERMAL SENSOR ARRAY FUSION FOR IOT APPLICATION

The proposed approach, which is illustrated in Fig. 1 consists of two main stages. The first stage is called home layout learner, which enables the flexibility of the proposed approach to work with different and complicated environments by estimating the environment layout and the sensors' placement in the mean of finding the flow of human movement in their domestic environment. This is achieved by finding a sensor reference, which indicates the typical human movement flow in multiple sensor's inspection areas and the overlap between the installed sensors' FoV by analysing the movement patterns of a single human subject moving in the environment. Hence, the home layout learner stage will only run once at the first time after installation to identify which sensors are interfering with each other in their FoVs. By doing so, the proposed approach is not only flexible to work with different home layouts but also reduce the computational time and resources needed to cover a large domestic environment, e.g. hospital. Besides, the environment layout learner is also feasible to determine the typical human movement flow in a simple domestic environment, including rooms and large open halls that may have more than one sensor to cover the inspection area. After the layout is learnt at the first stage, the second stage is used for sensor fusion and identification of overlapping regions between two or more TSA sensors.

A detailed description of the functional phases (multi-sensor acquisition, pre-processing, motion analysis, sensors interference learner, and identifying the overlapped regions) of the two stages is described below.

#### A. Sensor acquisition

The TSA's ability to maintain users' privacy in a domestic environment, low-cost and contactless features is the impetus for proposing an enabling technology to fuse multiple TSA for an IoT application. Besides, the reference rate of the sensor is between  $0.5\text{ Hz}$  and  $64\text{ Hz}$ , and this capability makes it suitable to detect swift human movements. On the other hand, a single TSA would not be sufficient to cover a wide inspection area. Fig. 2 shows an illustrative diagram of the relationship between the sensor's FoV and the inspection area. It can be observed from this figure that the inspection area is larger from a distance. Considering the FoV of the TSA sensor used in

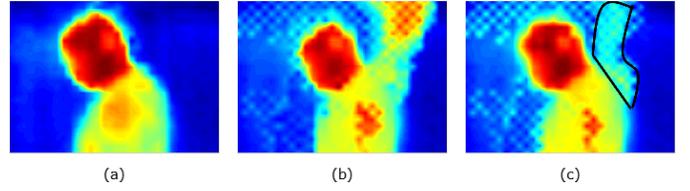


Fig. 3: The effects of human movement on the acquired thermal scene, (a) a stationary human presence, (b) human hand movement, (c) thermal noise induced by human movement.

our investigation, MLX90640<sup>1</sup>, which is  $55^\circ \times 35^\circ$ . The size of the inspection area at a different distance can be calculated as follows:

$$m = 2 \times d \times \tan\left(\frac{FoV}{2}\right) \quad (1)$$

where  $d$  is the depth distance in the FoV and  $m$  is the inspection area. Assuming  $d$  is  $2m$ , then inspection area  $m$  would be  $2.08m \times 1.26m$ . The collected data, which represents the thermal signature in the FoV, at time  $t$  with size  $C$  is transformed from  $1D$  linear vector  $\mathbf{x}_t = [x_{t_1}, \dots, x_{t_C}]^T$  to  $2D$  grid format of  $32 \times 24$  resolution. A colourmap scheme is then applied on each of the  $2D$  temperature metrics for visualisation purposes. Fig. 4 shows a visual calibration between high and low-resolution thermal images at a near and far human-sensor distance for one person and two other persons standing facing each other. It can be seen from the low-resolution thermal images shown in Fig. 4 (a) that human identifying information in the output of TSA is not as clear as the output of the used high-resolution thermal camera shown in Fig. 4 (b). Another motivation for using TSA in indoor human monitoring applications is its cost is much lower than the high-resolution camera. For instance, the price of a used TSA is about 0.00125 of the price of the used FLIR T6XX camera.

#### B. Pre-processing

There are notable limitations to proceeding with the human motion-based feature extraction on the output of this sensor. In particular, human movement generates a thermal noise in the background scene, which can bias human motion estimation by including unwanted background pixels. To illustrate this on a thermal image of a stationary human subject shown in Fig. 3 (a). Given the movement of the human hand shown in Fig. 3 (b), the acquired thermal image after movement is affected by thermal noise induced by human movement as highlighted in Fig. 3 (c).

This is a critical consideration in multiple TSA use case scenarios from different sensor positions, for example, the ceiling, where the human would be in direct contact with a background surface such as the floor would result in prolonged thermal noise lingering in multiple acquired thermal images. To overcome this, an adaptive temperature level-based thresholding technique to separate the human presence from the thermal background scene is proposed on Ostu's method

<sup>1</sup>The sensor details can be obtained from the Melexis website: <https://www.melexis.com/en/product/MLX90640/>

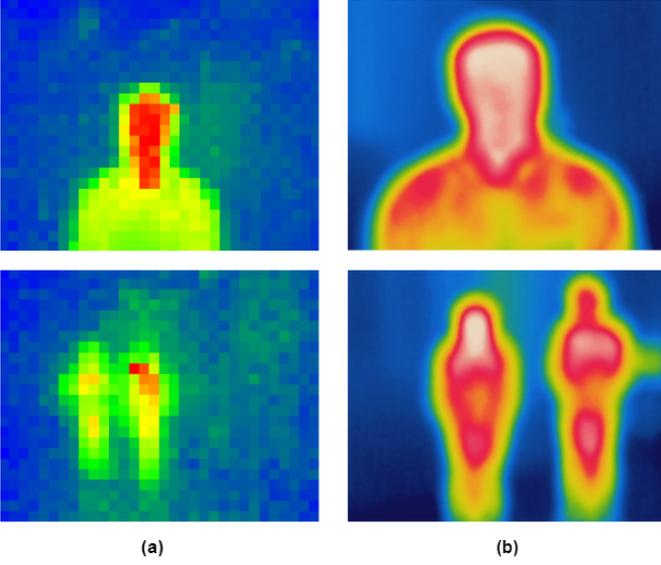


Fig. 4: A visual calibration between thermal imaging using (a) low-resolution TSA, (b) high-resolution thermal camera.

[51]. Mathematically, the threshold  $k$  of separating each  $2D$  temperature matrix  $x_t$  into foreground  $t_f$  and background  $t_b$  classes should minimise the intra-class variance  $\sigma_w^2(k)$ , defined as a weighted sum of variances of the two classes:

$$\sigma_w^2(k) = \omega_{t_f}(k)\sigma_{t_f}^2(k) + \omega_{t_b}(k)\sigma_{t_b}^2(k) \quad (2)$$

where  $\omega_{t_f}$  and  $\omega_{t_b}$  are the probabilities of the foreground and background classes, respectively separated by  $k$ , while  $\sigma_{t_f}$  and  $\sigma_{t_b}$  are the intra-class variances of these two classes. The class probability  $\omega_{t_f}, \omega_{t_b}(k)$  is found from the  $L$  bins of thermal image histogram:

$$\omega_{t_f}(k) = \sum_{i=0}^{k-1} p(i) \quad (3)$$

$$\omega_{t_b}(k) = \sum_{i=k}^{L-1} p(i) \quad (4)$$

In a binary classification problem, minimising the intra-class variance (variation between multiple samples of a class) is equivalent to maximising the inter-class variance (variation between classes):

$$\begin{aligned} \sigma_b^2(k) &= \sigma^2 - \sigma_w^2(k) = \omega_{t_f}(\mu_{t_f} - \mu_K)^2 + \omega_{t_b}(\mu_{t_b} - \mu_K)^2 \\ &= \omega_0(k)\omega_1(k)[\mu_0(t) - \mu_1(k)]^2 \end{aligned} \quad (5)$$

which is exposed in terms of class probabilities  $\omega$  and class means  $\mu$ , where  $\mu_{t_f}(k)$ ,  $\mu_{t_b}(k)$  and  $\mu_K$  represents:

$$\mu_{t_f}(k) = \frac{\sum_{i=0}^{k-1} ip(i)}{\omega_{t_f}(k)} \quad (6)$$

$$\mu_{t_b}(k) = \frac{\sum_{i=k}^{L-1} ip(i)}{\omega_{t_b}(k)} \quad (7)$$

$$\mu_K = \sum_{i=0}^{L-1} ip(i) \quad (8)$$

By computing  $\omega$  and  $\mu$  iteratively, and  $k$  for each thermal scene, the algorithm would separate the scene into background and foreground classes regardless of the human to sensor distance. Hence, the larger human to sensor distances, the lower human acquired temperatures. However, the drawback of this method is when the histogram is not bi-modal distribution and has, for example, three peaks like a temperature much higher than that of a human temperature, such as a hot cup of coffee, as shown in the representative surface diagram of Fig. 5 (a) and its corresponding thermal image. To avoid this, a modification before applying this approach is suggested using a temperature filter to convert temperatures above normal human temperatures to the minimum temperature of the thermal scene (not zero, to maintain the variance of human and background temperatures). This empirical value is set to be  $33^\circ C$  by the designated TSA. The filtered value is therefore as follows:

$$x_t = \begin{cases} x_i & \text{for } x_i \leq 33 \\ \min(x_t) & \text{otherwise} \end{cases} \quad \text{where } x_i = FIR \quad (9)$$

Fig. 5 (b) shows an illustrative result of applying this modification filter on the original acquired thermal scene shown in Fig. 5 (a), while Fig. 5 (c) shows the result of the complete pre-processing phase.

### C. Motion analysis

The employed human motion analysis is based on extracting the apparent relative motion vector between a human observer and the background scene, referred to as the optical flow. The term optical flow can also refer to the distribution of the apparent velocities of movement of brightness pattern in the scene. Technically, the optical flow aims to find the motion pattern in terms of velocity and direction between a sequence of two image frames obtained at times  $t$  and  $t + \Delta t$  at each temperature value in the acquired heat-map (at every pixel in each pre-processed thermal image). Fig. 6 shows visible illustrations of extracted motion vectors on different human movements' directions. Accordingly, each movement has a corresponding flow representation. Therefore, it is possible to use optical flow to determine the direction of human movements.

In this paper, Horn-Schunk algorithm [52] is used to estimate the flow for the output of TSA in the form of velocity and direction. The flow of this optical flow algorithm is formulated as a global energy functional  $E$ , which is then minimised. This function for the 2-D thermal image is given as:

$$E = \iint \left[ (I_x u + I_y v + I_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2) \right] dx dy \quad (10)$$

Where  $I_x, I_y, I_t$  refer to the derivatives of the thermal image temperature values along with  $x, y$ , and time dimensions, respectively,  $\vec{V} = [u(x, y), v(x, y)]^T$  is the motion vector, and

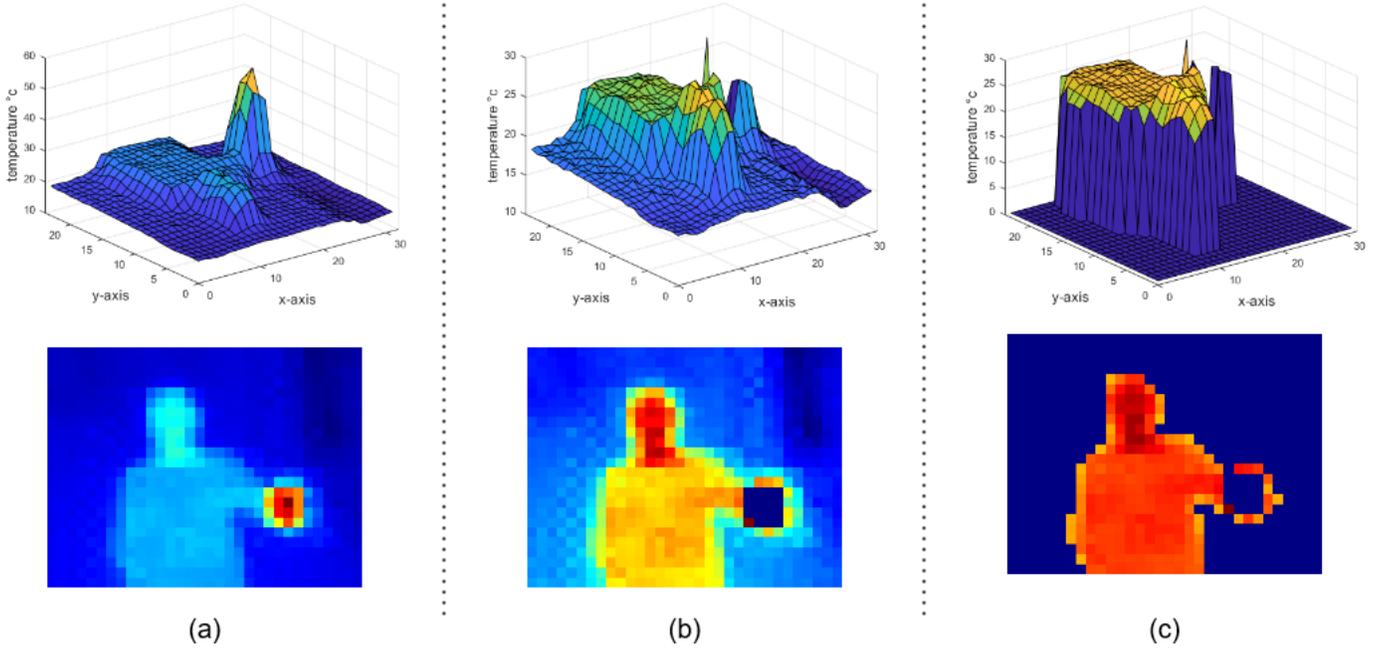


Fig. 5: Illustrative results of the proposed pre-processing phase, (a) the original temperature surface plot and its corresponding heat-map, (b) the result of applying the temperature filter, (c) the result of separating the acquired temperatures into background and foreground categories.

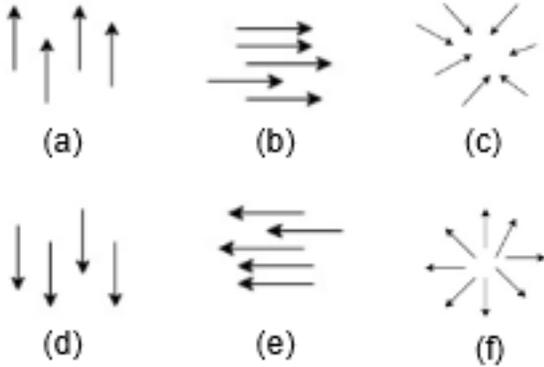


Fig. 6: An illustration of the motion vector based on optical flow, (a) represents vertical movement upwards, (b) horizontal movement to the right, (c) backward movement, (d) vertical movement downwards, (e) horizontal movement to the left, (f) forward movement.

the parameter  $\alpha$  is a regularisation constant. Hence, Larger alpha leads to a smoother flow.

The assumption behind this algorithm is that there is no significant change in the lighting between two consecutive frames, which refers to as Brightness Constancy Assumption. In other words, the colours of the moving thermal objects should remain the same, regardless of the change in temperature of these objects caused by the evolution of the sensor to human distance. This justifies our proposed adaptive thresholding in the pre-processing phase B to keep the human colours the same in the consecutive thermal image frames regardless of any potential changes in the human acquired

temperatures. Also, it is able to eliminate the background heat, which could be a moving thermal noise generated as a result of human movement.

Fig. 7 shows sample images on applying the optical flow estimation on the pre-processed thermal images. Fig. 7 (a) and Fig. 7 (b) show the estimation of the optical flow on the pre-processed TSA output when the subject is moving horizontally and when there is a hand movement. Since the proposed pre-processing technique aims to find a threshold in the temperature values of the scene regardless of human presence, the relatively high-intensity pixels (temperature) from the background scene appear in the foreground scene of the image in non-human presence scenarios. However, this does not affect the optical flow estimation because these pixels have no motion, as shown in Fig. 7 (c). Hence, the length of each arrow represents the magnitude of the velocity.

#### D. Sensors interference, overlap learning and fusion

The layout learner uses the motion vectors extracted from the previous phase to find the sensors that interfere with their FoV when enrolling on a new domestic environment (sensors installation). This paper proposes to find the sensors that interfere with each other and the time of the motion to find the order of typical human movement flow by a threshold-based motion trigger for a single occupancy movement. Considering a small home as an example, it requires four TSAs ( $S_1, S_2, S_3, S_4$ ). If  $S_1$  and  $S_3$  have motion velocities more noticeable than the non-human presence scenes at the time of acquisition, this means that these two sensors are interfering with each other. Besides, the motion sequence order extracted from the outputs of the installed sensors collected at this stage indicates sufficient information regarding the layout of the home, which

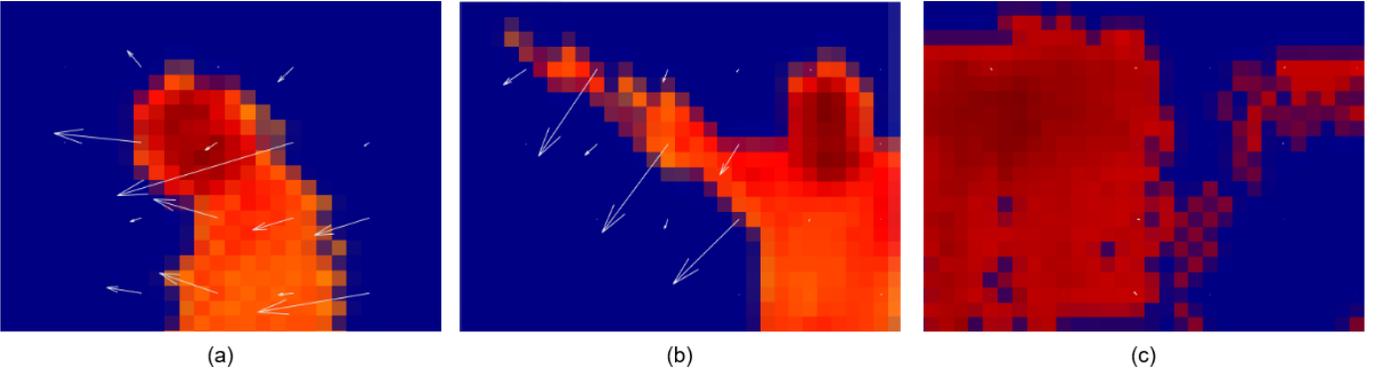


Fig. 7: The estimate of optical flow on the pre-processed TSA output, (a) human moving horizontally, (b) human hand movement, (c) empty human scene.

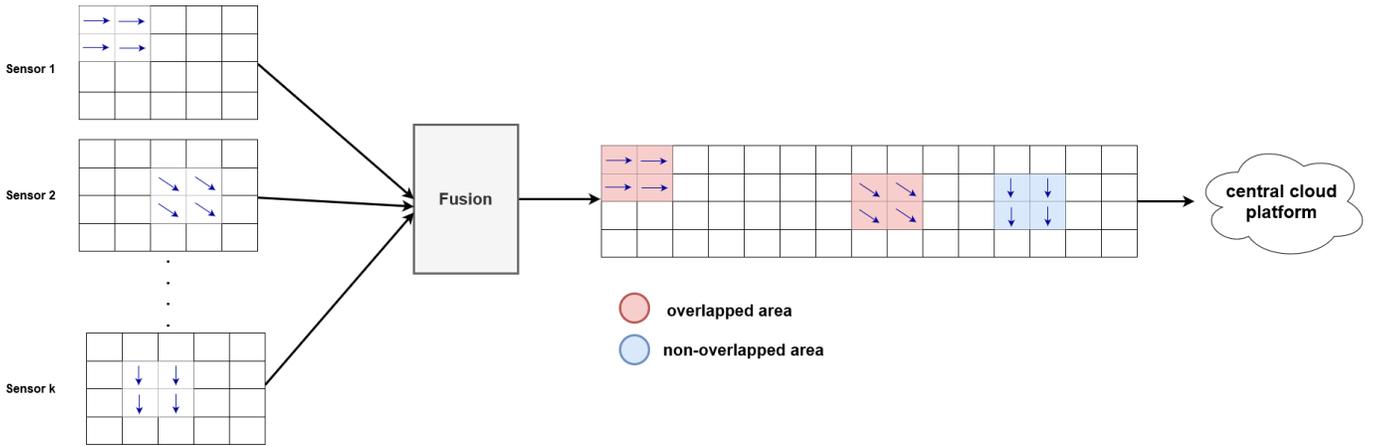


Fig. 8: The fusion approach of multiple TSAs using the extracted motion vectors. The approach proposes a privacy transmission of the monitored human subject information to a central cloud platform by replacing the temperature values with the extracted motion features.

would be used to regulate the arrangement of motion vectors in the fusion phase for potential human activity recognition applications.

Additionally, there is another situation about the fusion of the sensors in the deployment stage: identifying multiple human subjects occupying the environment is based on matching the motion vectors between interference sensors at the time of acquisition to identify the subjects that may appear in multi-FoVs. This has been achieved through finding the Euclidean distance between moving objects in the interference sensors to be used as a feature in a binary classification problem. Referring to the previous example of a small home with four sensors ( $S_1, S_2, S_3, S_4$ ), which has two interference sensors  $S_1$  and  $S_3$ . The matching between the motion vectors will only be performed for these two sensors and not, e.g.  $S_1$  and  $S_4$ , as these sensors do not interfere with each other as referenced in the enrolment stage. Doing this will reduce the required computational resources for the matching task and overcome the situation when similar human movements are performed simultaneously by different human subjects, e.g. yoga, meditation, or prayers.

Fig. 8 illustrates the proposed TSA's fusion method of the extracted motion vectors, which has the exact size of

the TSA heat-map. The output of this fusion method is a one commutative motion vector with the overlapping motion from different sensors identified. Hence, the order of these motion sequences is based on the obtained reference from the layout learner. Although the TSA sensors do not produce any identifiable information about people, it may still be a privacy risk to transfer their output to a centralised cloud platform to enable further human-centric applications. Therefore, this paper proposes replacing the TSA output with the motion vector produced from the fusion method.

#### IV. EXPERIMENTS

To evaluate the performance of the proposed methodology, comprehensive experimental work was performed with different use cases and scenarios for sensor positions to confirm the validity of the proposed approach in different sensor placements. During data collection, the experimental home environment was in multi-occupancy mode, which means that a human subject was performing different activities of daily living, including walking, sitting, standing, and laying in the bed in the overlapping region of the sensors shown in Fig. 9 and another person was in a different sensor FoV. Therefore, this section is focused on experimentally verifying

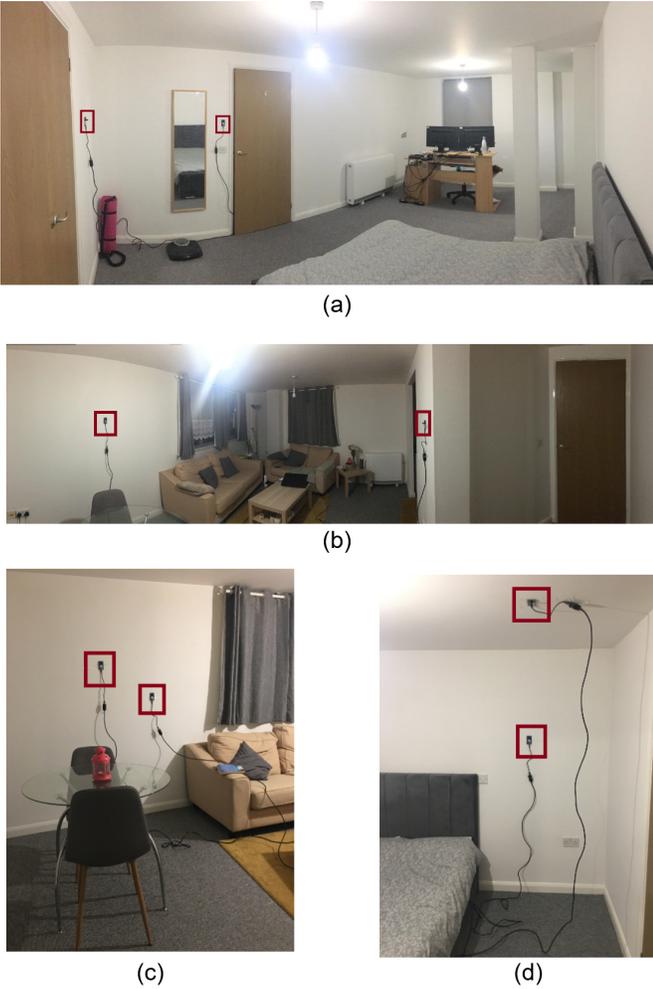


Fig. 9: Data collection stage, (a) sensors placed side by side at  $90^\circ$ , (b) sensors placed opposite each other, (c) sensors on the same wall, (d) sensor on wall and ceiling.

the capability of the proposed approach on identifying the sensor overlapped regions through matching different human presence acquired from different TSA placements and sensor to human distance on the basis of the moving object motion analysis. Building on top of this, there would not be any problem in the case of more than one human subject moving in a single sensor's FoVs as the problem arises when the same moving object (human) appear in two sensors' output due to overlapped FoVs. The size of the dataset collected to conduct these experiments was 1530 image frames. A detailed description of the experimental results is reported below.

### A. Experiment 1

This experiment contains a human subject moving in the overlapping area of two sensors placed side by side at  $90^\circ$ , as shown in Fig. 9 (a). A third sensor was installed in another room with a different human participant performing normal activities. The extracted human motion vectors from these multiple sensors were analysed by finding the distance between a set of maximum motion velocities for each human participant in the outputs of the sensors. Fig. 10 (a) shows the distance between the overlapped and non-overlapped regions based on human movements.

Finally, the distance is then used as a feature in different classification algorithms, including Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Linear discriminant analysis (LDA). The performances of these algorithms were 87.9%, 89.6%, 85.2%, 87.9%, respectively. Hence, 10-fold cross-validation is used to avoid the over-fitting problem and guarantee the trained model's generalisation ability.

### B. Experiment 2

Similar to the data collection settings of the first experiment, this experiment contains human participants in different rooms, and the aim is to assess the ability of the proposed methodology in distinguishing between overlapping and non-overlapping regions when sensing interference occurs between opposite sensor positions, as shown in Fig. 9 (b).

It can be observed from the corresponding visual analysis shown in Fig. 10 (b) is the distance between the magnitude of velocities for different human subject movements is significantly higher than those obtained for the same human subject at the opposite sensor positions. Utilising this distance measurement in the classification algorithms achieves a performance of 92.2%, 92.2%, 89.4%, 90.1% using logistic regression, SVM, KNN, and LDA, respectively.

### C. Experiment 3

The TSAs could be installed on the same wall but at different heights, as shown in Fig. 9 (c), for example, for human activity recognition for an adult occupancy with a child or to always acquire the upper body of an adult person at the different sensor to human distance.

This experiment intends to assess the proposed methodology on such a use case scenario. The performance of the proposed methodology using logistic regression was 96.1%, SVM was 96.1%, KNN was 87.9%, and LDA was 87.1% with 10-folds cross-validation. The justification behind obtaining a better performance on this sensor placement scenario is because the

TABLE I: A summary of experimental results on identifying overlapping regions between multiple sensors with different sensor placements.

	Logistic Regression	SVM	KNN	Linear Discriminant
<b>Experiment 1</b>	87.9%	89.6%	85.2%	87.9%
<b>Experiment 2</b>	92.2%	92.2%	89.4%	90.1%
<b>Experiment 3</b>	96.1%	96.1%	87.9%	87.1%
<b>Experiment 4</b>	92.1%	91.4%	89.2%	91.0%

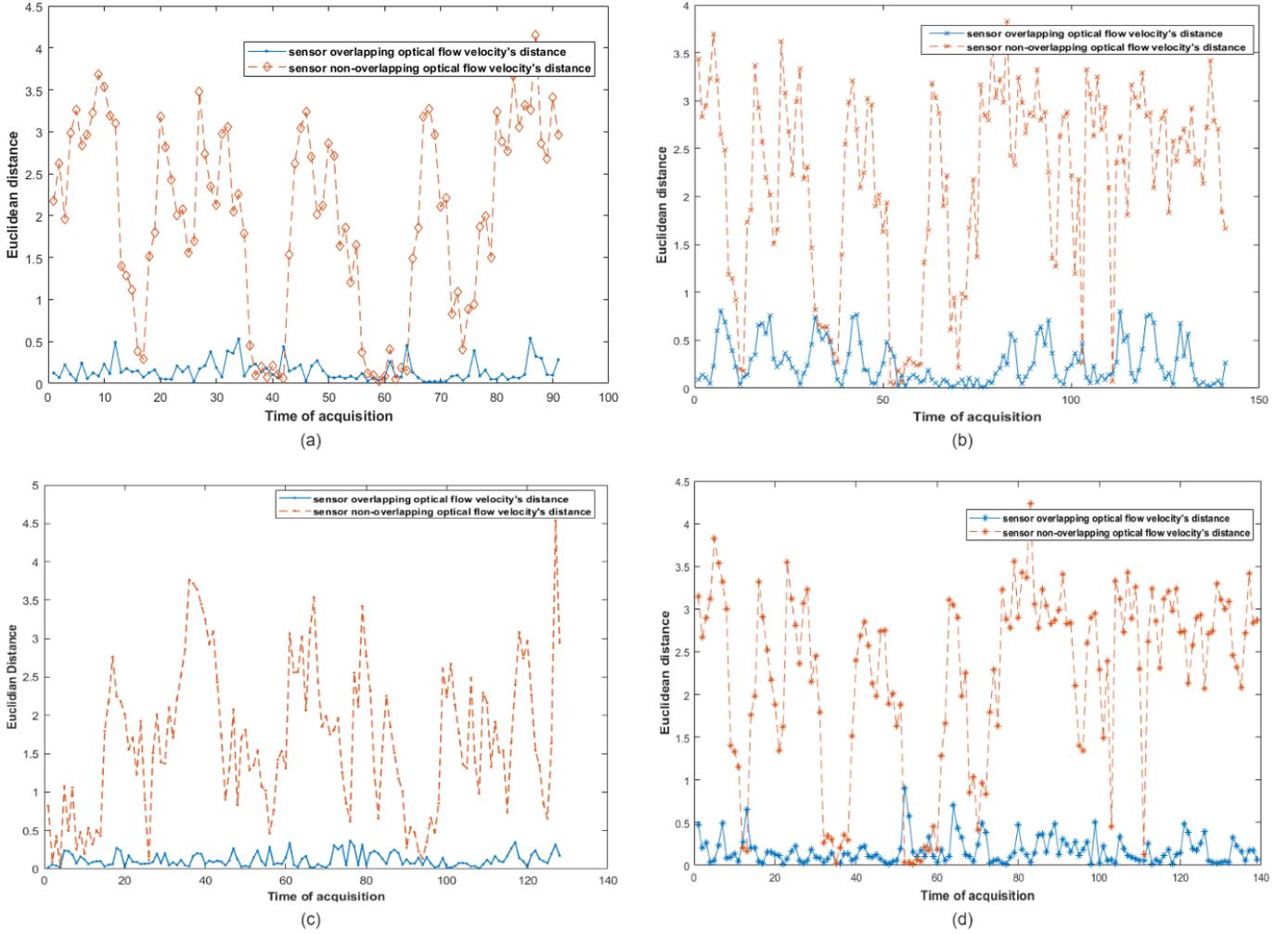


Fig. 10: Motion analysis between overlapping and non-overlapping FoV's regions with different human participants using data obtained from different sensor placements, (a) sensors placed side by side at 90°, (b) sensors placed opposite each other, (c) sensors on the same wall, (d) sensor on wall and ceiling.

intra-class similarity between the classes is high, as shown in Fig. 10 (c).

#### D. Experiment 4

The last experiment was designed to assess the proposed methodology on completely different sensor positions, specifically on wall and ceiling sensor placements, as shown in Fig. 9 (d). The performance of using the distance measurement between the magnitude of velocities of the extracted motion vectors, which is shown in 10 (d), achieves 92.1% using logistic regression, 96.1% using SVM, 89.2% using KNN, and 91.0% using LDA.

Table I provides a comparison of experimental results on identifying the overlapping regions using multiple sensors and different sensor placements. It can be concluded from these results that the proposed methodology provides excellent analytical performance with different sensor placements and human subjects. Besides, identifying the overlap regions from the same wall placement provides the best performance among others. In contrast, the worst performance result was reported

in Experiment 1. The rationale behind this lower performance in the side-by-side sensor placements is that a large portion of human presence is missing compared to the other investigated cases, which leads to a loss of motion features for the same human subject at a given time.

#### V. ROBUSTNESS ANALYSIS

This section contains three additional experiments to validate the proposed approach's robustness to enable the fusion of multiple TSAs into human monitoring applications. The first experiment aims to validate the proposed technique to enhance the TSA privacy feature in IoT applications by replacing temperature values with motion vectors for further human-centric applications running on the cloud. Since TSA is proposed in this paper for human-based applications, this experiment is intended to verify whether human presence can be determined from motion vectors without having to transfer temperature values to a central cloud platform. In this experiment, 218 non-human frame images were collected from noisy thermal scenes and scenes after the human subject left and another 209 human presence frame images. The

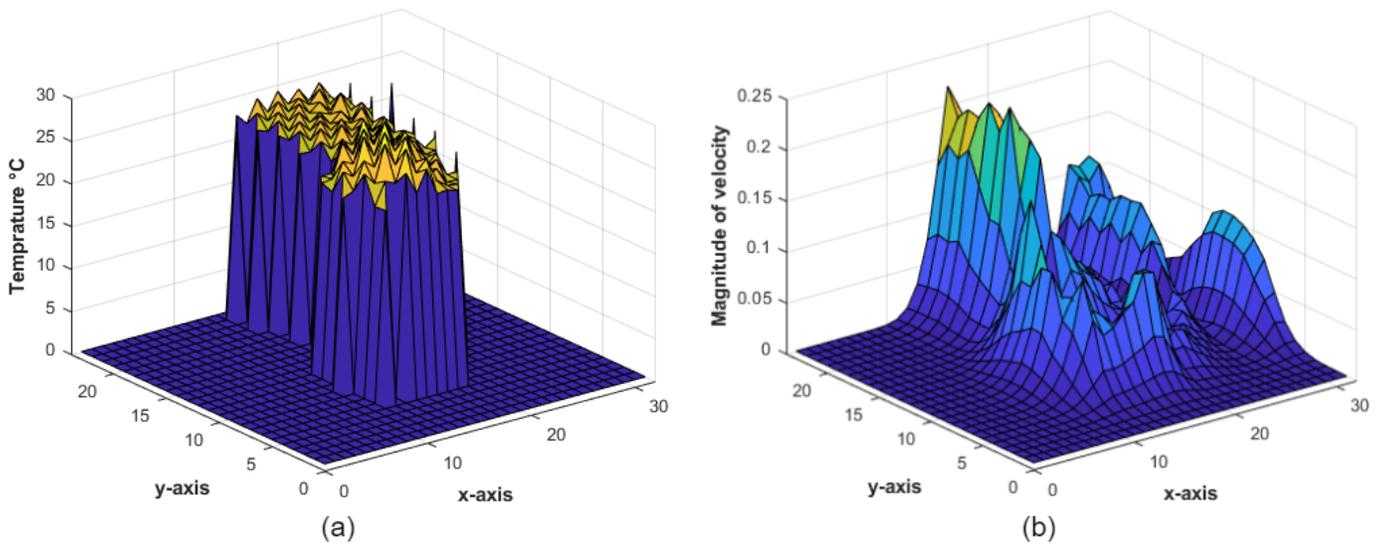


Fig. 11: A comparison of human temperature distribution versus the corresponding motion vector for human detection and localisation application, (a) the pre-processed thermal scene temperature values, (b) the velocity movement magnitudes for the same acquired thermal scene.

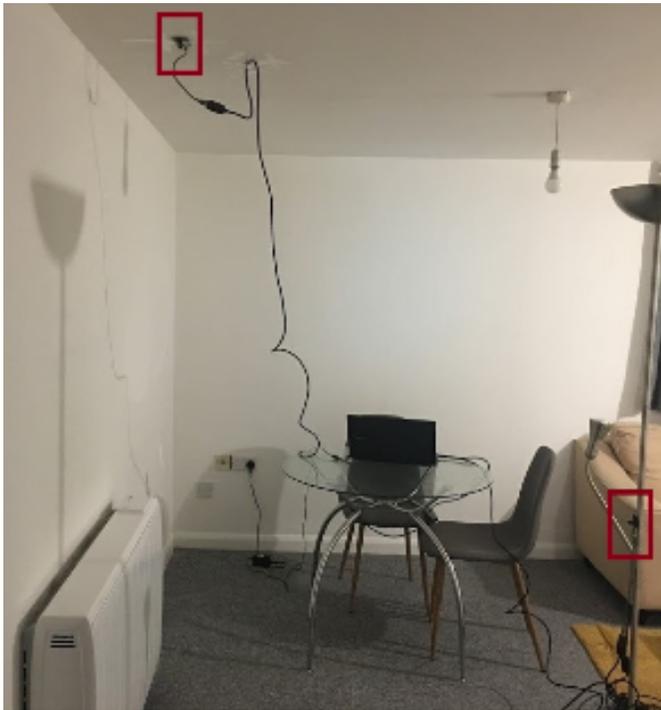


Fig. 12: A data collection stage consists of two simultaneous placements of TSAs to analyse the effects of a moving-based thermal noise generated by a domestic heater on the proposed sensor fusion approach.

performance of utilising the motion vectors in human detection using the classification approach was 94.9% achieved accuracy using SVM with 10-folds cross-validation. Hence, this result was based on the result of motion analysis of each frame image, and the performance could be boosted by using all the frame images at a specific time, e.g. every 1 second.

Fig. 11 (a) shows a sample of the distribution of the human

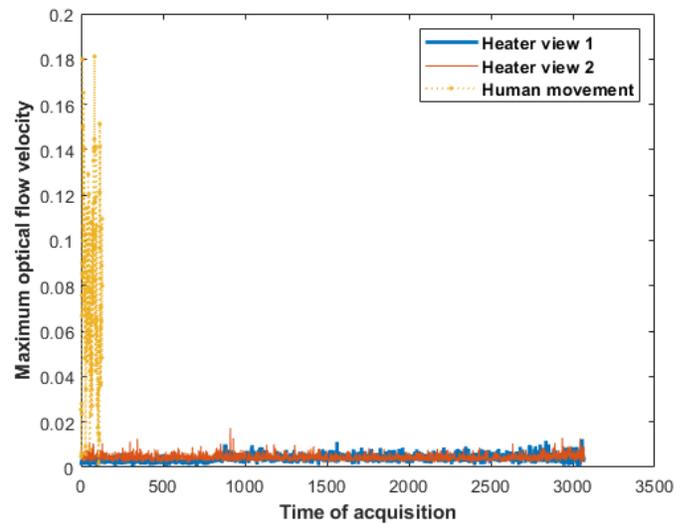


Fig. 13: The maximum velocities of the heat generated by a domestic heater from two different sensor placements versus a sample of human movement.

temperature values on  $x$  and  $y$ -axis while Fig. 11 (b) shows the corresponding magnitude of velocities for the same pre-processed image frame. It can be observed from these sample distributions and experimental results that human detection and localisation can be computed using motion vectors without transferring the complete thermal information to the cloud. On the other hand, some human-centred applications may require the provision of human temperature values. However, TSA is a low-resolution sensor, and the acquired temperature varies depending on different conditions, e.g. human to sensor distance and wearing heavy clothes. Therefore, in such applications, it is sufficient and private to provide human temperature values from the acquired scene as features (maximum, minimum, median, and average) rather than the full picture of human

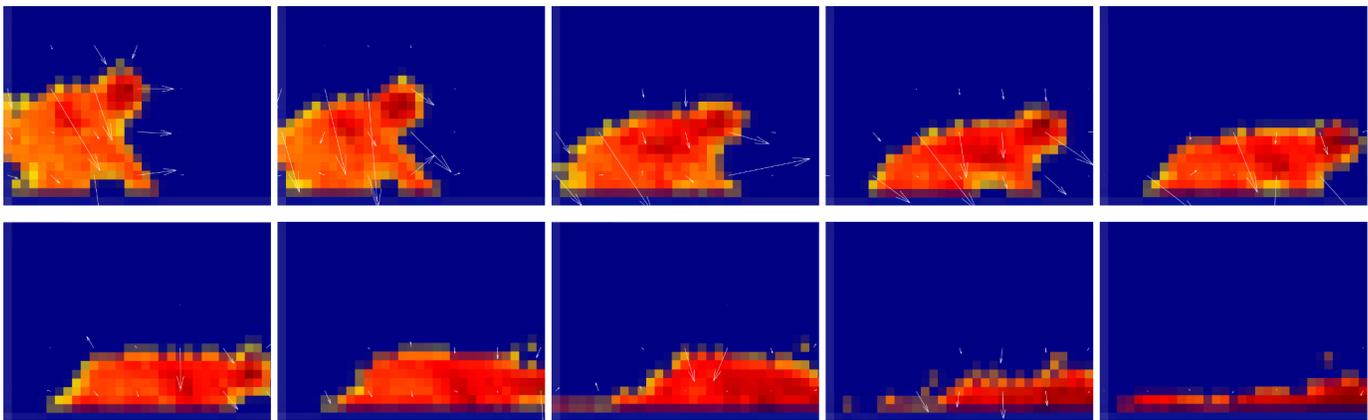


Fig. 14: A visualisation of optical flow estimation on a series of pre-processed low-resolution frames containing abnormal human fall activity.

temperature values to avoid human image construction on this low-resolution sensing approach by a third party.

The second experiment evaluates the feasibility of the proposed privacy enhancement feature to replace the obtained TSA’s temperature values with optical flow vectors. Therefore, this experiment uses only the optical flow vectors without any temperature values in Activities of Daily Living recognition. Specifically, to recognise two activities (1) normal human walking activity and (2) abnormal human fall activity. A dataset of 673 thermal frames were collected, which consisted of 226 fall frames and 447 walk frames. The dataset is divided into 70% for training and 30% for testing. Adding to this, a Bidirectional LSTM (Bi-LSTM) network [53] with Stochastic Gradient Descent with momentum (SGDM) [54] optimisation algorithm has been utilised to recognise the activity in a sequence of optical flow vectors. The performance of classifying a sequence of optical flow vectors into either fall or walking achieved 94.4% accuracy for the testing dataset.

This very good performance can be justified in representing the optical flow vectors on a series of pre-processed low-resolution frames shown in Fig. 14. From these frames, it is apparent that the optical flow can represent the human motion pattern in terms of velocity and orientation represented by the length and direction of the arrows in the frames shown in Fig. 14 respectively. In this specific example, the pattern of human movements starts from up to down until it reaches the inactivity state that is demonstrated by the decrease in human movement velocity in the last few frames. The ability of the optical flow to describe the human motion pattern post applying the proposed pre-processing techniques on the output of the TSA shows the potential to replace TSA temperature values with optical flow vectors in human monitoring applications.

The third experiment aimed to further evaluate the proposed approach’s practicality in a home environment where there could be circulating convection heat from heating systems. In this experiment, two sensors were installed on the top and opposite views of the heater, as shown in Fig. 12. This experiment considered the most complex motion-based thermal noise scenario by collecting data from the cold-heater state to the

hot state for more than six continuous hours. The total number of thermal images collected for this experiment is 145,776 frames. Fig. 13 shows the maximum motion velocities of a subset sample of the collected data from the two heater views versus a human movement.

Two important observations can be drawn from this experiment. First, motion-based thermal noises do not affect the applicability of the proposed approach to enable further human-centric IoT applications. Second, the proposed methodology for TSA overcomes considerable challenges that appear on temperature-based image processing techniques with this type of low-resolution thermal images.

## VI. CONCLUSION

Thermal Sensor Array provides low-resolution thermal imaging, making it an excellent approach for low-cost, privacy-preserving, and passive human-centric IoT applications. In contrast to high-resolution imaging, it is difficult to find a comparable image-based reference to incorporate multiple TSAs to cover a wide inspection area. This paper proposes a motion-based approach to integrate multiple TSAs and identify overlapping regions in the sensors’ field of view. The efficiency of computing time and resources is achieved through proposing an environmental layout learner. Furthermore, the paper proposes to replace the temperature values acquired by the sensors with the extracted motion vectors for further human activity recognition operating in a centralised cloud platform to avoid re-configuring the human image by a third party. Extensive experiments were performed to validate the proposed approach on different sensor placement and motion-based thermal noise and a transfer application of the proposed approach to detect human presence using motion vectors without any temperature values.

The proposed motion analysis approach is based on stationary sensor placement like wall or ceiling. Future work could be undertaken on a moving sensor placement on the example of assistive robots for the elderly. Besides, the feasibility of the proposed privacy enhancement feature to replace the TSA’s temperature values with the optical flow prior to sending them to the cloud has been assessed in recognising human

walking and fall patterns. However, fall detection is abnormal human activity that can occur from standing, sitting or even unpredicted activity. Therefore, future work can be undertaken to address a valid issue concerning the users' accountability to the system's decision in human monitoring applications.

#### ACKNOWLEDGMENT

Abdallah Naser is supported by Nottingham Trent University through a fully-funded Scholarship Scheme.

#### REFERENCES

- [1] M. Al-Khafajiy, T. Baker, C. Chalmers, M. Asim, H. Kolivand, M. Fahim, and A. Waraich, "Remote health monitoring of elderly through wearable sensors," *Multimedia Tools and Applications*, vol. 78, no. 17, pp. 24 681–24 706, 2019.
- [2] S. C. Mukhopadhyay, "Wearable sensors for human activity monitoring: A review," *IEEE sensors journal*, vol. 15, no. 3, pp. 1321–1330, 2014.
- [3] A. Lotfi, C. Langensiepen, S. M. Mahmoud, and M. J. Akhlaghinia, "Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour," *Journal of ambient intelligence and humanized computing*, vol. 3, no. 3, pp. 205–218, 2012.
- [4] Z. Hussain, Q. Z. Sheng, and W. E. Zhang, "A review and categorization of techniques on device-free human activity recognition," *Journal of Network and Computer Applications*, vol. 167, p. 102738, 2020.
- [5] H. Cui and N. Dahnoun, "High precision human detection and tracking using millimeter-wave radars," *IEEE Aerospace and Electronic Systems Magazine*, vol. 36, no. 1, pp. 22–32, 2021.
- [6] A. Naser, A. Lotfi, and J. Zhong, "Adaptive thermal sensor array placement for human segmentation and occupancy estimation," *IEEE Sensors Journal*, 2020.
- [7] S. Savazzi, V. Rampa, L. Costa, and D. Tolochenko, "Analysis of body-induced thermal signatures for social distancing monitoring," in *2020 IEEE Sensors*. IEEE, 2020, pp. 1–4.
- [8] A. Naser, A. Lotfi, and J. Zhong, "Towards human distance estimation using a thermal sensor array," *Neural Computing and Applications*, pp. 1–11, 2021.
- [9] Z. Chen and Y. Wang, "Remote recognition of in-bed postures using a thermopile array sensor with machine learning," *IEEE Sensors Journal*, vol. 21, no. 9, pp. 10 428–10 436, 2021.
- [10] A. Naser, A. Lotfi, J. Zhong, and J. He, "Heat-map based occupancy estimation using adaptive boosting," in *2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 2020, pp. 1–7.
- [11] D. Qu, B. Yang, and N. Gu, "Indoor multiple human targets localization and tracking using thermopile sensor," *Infrared Physics & Technology*, vol. 97, pp. 349–359, 2019.
- [12] Z. Chen and Y. Wang, "Infrared-ultrasonic sensor fusion for support vector machine-based fall detection," *Journal of Intelligent Material Systems and Structures*, vol. 29, no. 9, pp. 2027–2039, 2018.
- [13] A. Naser, A. Lotfi, and J. Zhong, "A novel privacy-preserving approach for physical distancing measurement using thermal sensor array," in *The 14th Pervasive Technologies Related to Assistive Environments Conference*, 2021, pp. 81–85.
- [14] V. M. Ionescu and F. M. Enescu, "Low cost thermal sensor array for wide area monitoring," in *2020 12th International conference on electronics, computers and artificial intelligence (ECAI)*. IEEE, 2020, pp. 1–4.
- [15] E. Moisello, P. Malcovati, and E. Bonizzoni, "Thermal sensors for contactless temperature measurements, occupancy detection, and automatic operation of appliances during the covid-19 pandemic: A review," *Micromachines*, vol. 12, no. 2, p. 148, 2021.
- [16] M. Arun, E. Baraneetharan, A. Kanchana, S. Prabu *et al.*, "Detection and monitoring of the asymptotic covid-19 patients using iot devices and sensors," *International Journal of Pervasive Computing and Communications*, 2020.
- [17] K. Logu, T. Devi, N. Deepa, S. R. Kumar, and N. Gayathri, "Real-time mild and moderate covid-19 human body temperature detection using artificial intelligence," *AI and IoT-Based Intelligent Automation in Robotics*, pp. 189–204, 2021.
- [18] A. Hossam, A. Magdy, A. Fawzy, and S. M. Abd El-Kader, "An integrated iot system to control the spread of covid-19 in egypt," in *International Conference on Advanced Intelligent Systems and Informatics*. Springer, 2020, pp. 336–346.
- [19] J. L. Honorato, I. Spiniak, and M. Torres-Torriti, "Human detection using thermopiles," in *2008 IEEE Latin American Robotic Symposium*. IEEE, 2008, pp. 151–157.
- [20] P. Hevesi, S. Wille, G. Pirkl, N. Wehn, and P. Lukowicz, "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*, 2014, pp. 141–145.
- [21] H. M. Ng, "Human localization and activity detection using thermopile sensors," in *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE, 2013, pp. 337–338.
- [22] M. Kuki, H. Nakajima, N. Tsuchiya, and Y. Hata, "Human movement trajectory recording for home alone by thermopile array sensor," in *2012 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2012, pp. 2042–2047.
- [23] J. Kemper and D. Hauschildt, "Passive infrared localization with a probability hypothesis density filter," in *2010 7th Workshop on Positioning, Navigation and Communication*. IEEE, 2010, pp. 68–76.
- [24] Z. Chen, Y. Wang, and H. Liu, "Unobtrusive sensor-based occupancy facing direction detection and tracking using advanced machine learning algorithms," *IEEE Sensors Journal*, vol. 18, no. 15, pp. 6360–6368, 2018.
- [25] J. Zhang, J. Chen, Q. Lin, and L. Cheng, "Moving object distance estimation method based on target extraction with a stereo camera," in *2019 IEEE 4th International Conference on Image, Vision and Computing (ICIVC)*. IEEE, 2019, pp. 572–577.
- [26] V. Chidurala and X. Li, "Occupancy estimation using thermal imaging sensors and machine learning algorithms," *IEEE Sensors Journal*, vol. 21, no. 6, pp. 8627–8638, 2021.
- [27] Y. Yuan, X. Li, Z. Liu, and X. Guan, "Occupancy estimation in buildings based on infrared array sensors detection," *IEEE Sensors Journal*, vol. 20, no. 2, pp. 1043–1053, 2019.
- [28] M. Cokbas, P. Ishwar, and J. Konrad, "Low-resolution overhead thermal tripwire for occupancy estimation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 88–89.
- [29] A. Tyndall, R. Cardell-Oliver, and A. Keating, "Occupancy estimation using a low-pixel count thermal imager," *IEEE Sensors Journal*, vol. 16, no. 10, pp. 3784–3791, 2016.
- [30] 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.
- [31] C. Perra, A. Kumar, M. Losito, P. Pirino, M. Moradpour, and G. Gatto, "Monitoring indoor people presence in buildings using low-cost infrared sensor array in doorways," *Sensors*, vol. 21, no. 12, p. 4062, 2021.
- [32] C. Yin, J. Chen, X. Miao, H. Jiang, and D. Chen, "Device-free human activity recognition with low-resolution infrared array sensor using long short-term memory neural network," *Sensors*, vol. 21, no. 10, p. 3551, 2021.
- [33] K. Naik, T. Pandit, N. Naik, and P. Shah, "Activity recognition in residential spaces with internet of things devices and thermal imaging," *Sensors*, vol. 21, no. 3, p. 988, 2021.
- [34] L. I. L. Gonzalez, M. Troost, and O. Amft, "Using a thermopile matrix sensor to recognize energy-related activities in offices," *Procedia Computer Science*, vol. 19, pp. 678–685, 2013.
- [35] S. Tateno, F. Meng, R. Qian, and Y. Hachiya, "Privacy-preserved fall detection method with three-dimensional convolutional neural network using low-resolution infrared array sensor," *Sensors*, vol. 20, no. 20, p. 5957, 2020.
- [36] Z. Chen and Y. Wang, "Sleep monitoring using an infrared thermal array sensor," in *Sensors and Smart Structures Technologies for Civil, Mechanical, and Aerospace Systems 2019*, vol. 10970. International Society for Optics and Photonics, 2019, p. 109701D.
- [37] R. Gade and T. B. Moeslund, "Thermal cameras and applications: a survey," *Machine vision and applications*, vol. 25, no. 1, pp. 245–262, 2014.
- [38] J. Tanaka, H. Imamoto, T. Seki, and M. Oba, "Low power wireless human detector utilizing thermopile infrared array sensor," in *SENSORS, 2014 IEEE*. IEEE, 2014, pp. 462–465.
- [39] L. Walmsley-Eyre and R. Cardell-Oliver, "Hierarchical classification of low resolution thermal images for occupancy estimation," in *2017 IEEE 42nd conference on local computer networks workshops (LCN Workshops)*. IEEE, 2017, pp. 9–17.
- [40] S. J. Lee, G. Shah, A. A. Bhattacharya, and Y. Motai, "Human tracking with an infrared camera using a curve matching framework," *EURASIP Journal on Advances in Signal Processing*, vol. 2012, no. 1, p. 99, 2012.

- [41] P. Zappi, E. Farella, and L. Benini, "Tracking motion direction and distance with pyroelectric ir sensors," *IEEE Sensors Journal*, vol. 10, no. 9, pp. 1486–1494, 2010.
- [42] R. C. Luo, C.-C. Yih, and K. L. Su, "Multisensor fusion and integration: approaches, applications, and future research directions," *IEEE Sensors journal*, vol. 2, no. 2, pp. 107–119, 2002.
- [43] S. Hoque, K. Sirlantzis, and M. C. Fairhurst, "A new chain-code quantization approach enabling high performance handwriting recognition based on multi-classifier schemes," in *ICDAR*, 2003, pp. 834–838.
- [44] K. Sirlantzis, M. C. Fairhurst, and M. Hoque, "Genetic algorithms for multi-classifier system configuration: a case study in character recognition," in *International Workshop on Multiple Classifier Systems*. Springer, 2001, pp. 99–108.
- [45] K. Sirlantzis, S. Hoque, and M. C. Fairhurst, "Diversity in multiple classifier ensembles based on binary feature quantisation with application to face recognition," *Applied Soft Computing*, vol. 8, no. 1, pp. 437–445, 2008.
- [46] P. Ferrer-Cid, J. M. Barcelo-Ordinas, J. Garcia-Vidal, A. Ripoll, and M. Viana, "Multisensor data fusion calibration in iot air pollution platforms," *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3124–3132, 2020.
- [47] H. Lin, S. Garg, J. Hu, X. Wang, M. J. Piran, and M. S. Hossain, "Privacy-enhanced data fusion for covid-19 applications in intelligent internet of medical things," *IEEE Internet of Things Journal*, 2020.
- [48] K. Muhammad, T. Hussain, M. Tanveer, G. Sannino, and V. H. C. de Albuquerque, "Cost-effective video summarization using deep cnn with hierarchical weighted fusion for iot surveillance networks," *IEEE Internet of Things Journal*, vol. 7, no. 5, pp. 4455–4463, 2019.
- [49] N. Kumar, D. Acharya, and D. Lohani, "An iot-based vehicle accident detection and classification system using sensor fusion," *IEEE Internet of Things Journal*, vol. 8, no. 2, pp. 869–880, 2020.
- [50] E. Kanjo, E. M. Younis, and C. S. Ang, "Deep learning analysis of mobile physiological, environmental and location sensor data for emotion detection," *Information Fusion*, vol. 49, pp. 46–56, 2019.
- [51] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.
- [52] B. K. Horn and B. G. Schunck, "Determining optical flow," *Artificial intelligence*, vol. 17, no. 1-3, pp. 185–203, 1981.
- [53] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE transactions on Signal Processing*, vol. 45, no. 11, pp. 2673–2681, 1997.
- [54] N. Qian, "On the momentum term in gradient descent learning algorithms," *Neural networks*, vol. 12, no. 1, pp. 145–151, 1999.



**Abdallah Naser** received the B.Sc. degree in computer engineering from Cyprus International University with High-Honour class and the M.Sc. degree in information security and biometrics with Distinction from the University of Kent. He is currently pursuing the PhD degree with Nottingham Trent University, U.K. He granted a patent for inventing a new way to secure online accounts using concepts of pattern recognition from the Palestinian patent office. Besides, he was one of the winning teams in the K-Hackathon competition organised by Middle East

Technical University for developing a classification IoT system to detect car accidents in addition to other academic awards. His current research interests focus on utilising privacy-preserving sensors for occupancy monitoring through image processing and machine learning approaches.



**Ahmad Lotfi** (M'96-SM'08) received his BSc. and MTech. in control systems from Isfahan University of Technology, Iran and Indian Institute of Technology Delhi, India respectively. He received his PhD degree in Learning Fuzzy Systems from the University of Queensland, Australia in 1995. He is currently a Professor of Computational Intelligence and Head of the Department of Computer Science at Nottingham Trent University, Nottingham, where he is also leading the Computational Intelligence and Applications (CIA) research group. His research

focuses on the identification of progressive changes in the behaviour of elderly people suffering from Dementia or any other cognitive impairments. Accurate identification of progressive changes through the utilisation of an unobtrusive sensor network or robotics platform will enable carers (formal and informal) to intervene when deemed necessary. Research collaboration is established with world-leading researchers. He has worked in collaboration with many healthcare commercial organisations and end-users including Tunstall Healthcare Group and Nottingham Adult Care. He has received external funding from Innovate UK, EU and industrial companies to support his research. He has authored and co-authored over 200 scientific papers in the area of computational intelligence, anomaly detection and machine learning in highly prestigious journals and international conferences. He has been invited as an Expert Evaluator and Panel Member for many European and International Research Programmes. More details are available from: <http://www.lotfi.uk>



**Junpei Zhong** (S'10-M'15-SM'20) is a research assistant professor at the Hong Kong Polytechnic University. He graduated from the South China University of Technology with a Bachelor in Automation. He obtained his MPhil degree at the Hong Kong Polytechnic University, Hong Kong, and the PhD degree at the University of Hamburg, Germany. Before returning to PolyU, he has been a researcher at the University of Hertfordshire (UK), University of Plymouth (UK), Waseda University (Japan) and the National Institute of Advanced Industrial Science

and Technology (AIST, Japan) and Nottingham Trent University (UK). He has participated in several EU and Japanese projects centring around machine learning, assistive technologies, and cognitive robots. His current research focuses on cognitive assistive robots and their usage in wellbeing and elderly healthcare. He is the organiser of the IEEE Symposium on Domestic Robotics 2019 (part of the Symposium Series of Computational Intelligence 2019) in Xiamen, China. He also organised special sessions and workshops in ICIRA 2019, IEEE-RAS Humanoids 2018, IEEE ICDL-EpiRob 2018, and guest editors of several journals such as IEEE Transaction on Cognitive and Developmental Systems, Interaction Studies and Complexity.