DEVICE-FREE INDOOR LOCALISATION WITH NON-WIRELESS SENSING TECHNIQUES

A THESIS BY PUBLICATIONS PRESENTED IN PARTIAL FULFILMENT OF THE
REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

ELECTRONICS AND COMPUTER ENGINEERING

MASSEY UNIVERSITY, ALBANY, NEW ZEALAND

NATHANIEL FAULKNER

2021

ABSTRACT

Global Navigation Satellite Systems provide accurate and reliable outdoor positioning to support a large number of applications across many sectors. Unfortunately, such systems do not operate reliably inside buildings due to the signal degradation caused by the absence of a clear line of sight with the satellites. The past two decades have therefore seen intensive research into the development of *Indoor Positioning System* (IPS). While considerable progress has been made in the indoor localisation discipline, there is still no widely adopted solution. The proliferation of *Internet of Things* (IoT) devices within the modern built environment provides an opportunity to localise human subjects by utilising such ubiquitous networked devices. This thesis presents the development, implementation and evaluation of several passive indoor positioning systems using ambient *Visible Light Positioning* (VLP), capacitive-flooring, and thermopile sensors (low-resolution thermal cameras). These systems position the human subject in a device-free manner (i.e., the subject is not required to be instrumented). The developed systems improve upon the state-of-the-art solutions by offering superior position accuracy whilst also using more robust and generalised test setups. The developed passive VLP system is one of the first reported solutions making use of ambient light to position a moving human subject. The capacitive-floor based system improves upon the accuracy of existing flooring solutions as well as demonstrates the potential for automated fall detection. The system also requires very little calibration, i.e., variations of the environment or subject have very little impact upon it. The thermopile positioning system is also shown to be robust to changes in the environment and subjects. Improvements are made over the current literature by testing across multiple environments and subjects whilst using a robust ground truth system. Finally, advanced machine learning methods were implemented and benchmarked against a thermopile dataset which has been made available for other researchers to use.

AUTHOR'S DECLARATION

This thesis has been produced in accordance with Massey University's "Doctoral Thesis with Publications" guidelines. Chapters 2-5 of this thesis consist of works already published, or under review in an IEEE Journal. In accordance with IEEE copyright policy, the accepted versions of the published works are presented which may have stylistic differences from the final published versions. Furthermore, due to chapters being published in different journals, there are stylistic differences between the chapters.

ACKNOWLEDGEMENTS

There are many friends that have supported me throughout the past few years, and I would like to extend a huge thankyou to those that have been on this journey with me. Finally, a huge amount of credit goes to my parents – Adrian and Karina Faulkner. Due to their upbringing and support, I have been able to come this far.

FUNDING

TABLE OF CONTENTS

CHAPTER 1 – INTRODUCTION

Indoor localisation has been an active research topic for more than two decades. While outdoor positioning (e.g., GPS) is a mature technology, it does not work reliably inside buildings and there is no standardised solution for *Indoor Positioning System* (IPS) yet. *Location-Based Services* (LBS) within the built environment require robust and affordable positioning. The ongoing adoption of the *Internet of Things* (IoT) [1] provides access to a ubiquitous network of devices and ambient sensors. This is opening up new opportunities for deploying IPS while leveraging pre-existing infrastructure.

Indoor localisation[1] can be broadly categorised into active and passive. Active localisation is similar to GPS-based positioning where the target carries a special device (or a tag). The positioning of the tag results in localising the target. Active localisation is useful for asset tracking, navigation of mobile robots, and many other applications. where the end user is carrying a device (e.g., a mobile phone). A multitude of options, e.g., computer vision [2], light detection and ranging [3], ultrasound [4], acoustic [5], geomagnetic fingerprinting [6], wireless or *Radio Frequency* (RF) [7], visible light [8], aroma fingerprinting [9], etc., have been investigated for active indoor localisation. However, for many applications relying on the target to carry a tag may not be feasible. For example, if the goal is to unobtrusively track an elderly, forgetful person, one cannot expect them to wear a tag (e.g., a bracelet) every day. Wearable devices could also be forgotten, misplaced, or damaged. Besides, they also require regular battery charging (or changing). Having to carry a tracking device can be perceived as stigmatising thus leading to reluctance to wear one.

The passive approach, also known as *Device Free Localisation* (DFL), is the method of locating an untagged target. It does not require the target to actively participate in the localisation process by carrying a tag. Passive localisation is the key to providing *Ambient Assisted Living* (AAL) in smart buildings. The application of DFL ranges from intrusion detection, fall detection, and remote monitoring of the elderly, to occupancy detection for energy-efficient *Heating, Ventilation and Air Conditioning* (HVAC) and lighting as well as occupancy counting for emergencies (e.g., office and

---

[1] In the context of this document, the terms "location, localisation, positioning, tracking" are interchangeable and present the identification of the coordinates of a subject or object.

building evacuation), business analytics for retail applications, accessibility aids for visually impaired individuals, etc.

Camera or vision-based techniques are widely used for surveillance [10], crowd counting [11] etc. Several studies have indicated that low-cost 3D cameras (e.g. Kinect) may be suitable for DFL purposes [12, 13]. However, vision-based techniques require favourable lighting conditions. and have coverage blind spots due to occlusion. Most importantly, privacy concerns [14] make camera-based IPS in residential environments impractical. This is somewhat ironic given the ubiquitous presence of cameras embedded in smartphones, computing devices, gaming consoles etc. in modern households.

Localisation based on RF sensing can potentially utilise the pre-existing wireless network of a building. It can also operate in "through walls" scenarios. DFL solutions employing customised hardware and the *Channel State Information* (CSI) metric can be quite accurate with median error reaching below half a meter [15, 16]. However, CSI is only available with Wi-Fi and as such, precludes the majority of wireless technologies (e.g., Bluetooth, Zigbee, etc.). The utilisation of the universally available *Received Signal Strength Indicator* (RSSI), lowers the localisation accuracy and requires a significant number of wireless nodes to function [17]. This takes away the appeal of localisation with pre-existing infrastructure. Surveys of the wireless DFL literature show a vast discipline with a crowded research landscape [15, 18-21].

Due to the shortcomings of the camera-based and RF-based passive positioning techniques, several other passive sensing modalities have been investigated (Fig. 1) such as the application of visible light, infrared and capacitive sensing [22].

1.1 Motivation

The existing body of literature has shown that there is still significant work required before ubiquitous indoor positioning or localisation can be achieved. The goal of this thesis is to provide reliable and practical indoor positioning methods that can be used across diverse and dynamic environments. Applications of visible light, capacitive flooring, and low-resolution thermopiles have been investigated as potential solutions along with analysis of their strengths and weaknesses.
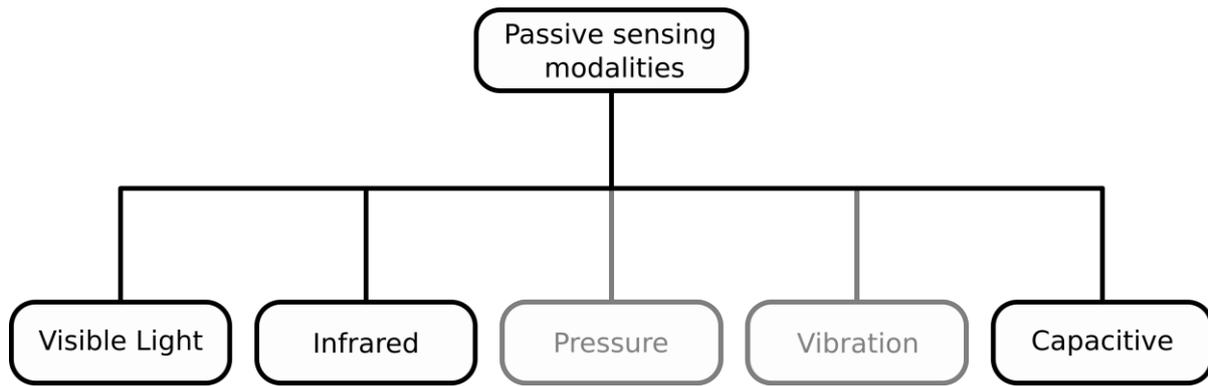
Figure 1: Passive sensing modalities which do not use RF or camera vision. Visible light, infrared and capacitive sensing are further investigated in this thesis.

While active *Visible Light Positioning* (VLP) has been a popular area of research over recent years [23], passive VLP has been largely overlooked. First of all, it is important to formulate the distinction between active and passive VLP. Active VLP requires a tracked object to carry a physical receiving device, which is actually tracked. In contrast, in passive VLP, the impact the tracked object has on its environment is observed to carry out the localisation. As the object moves, it casts shadows. This can be detected as changes in the light level at sensors placed in known locations. Whilst limited research has been done with modulated light for passive VLP, there is no prior research reported in the literature where the ambient light was used. This presents a significant opportunity for novel investigation.

There have been studies dealing with the use of floor sensing (pressure [24-27], vibration [28-30], and capacitive [31-34] types) for passive localisation. The floor-pressure-type sensing systems are largely electro-mechanical, and thus they are quite complex to construct. In addition to that, their mechanical components are subject to degradation over time. Vibration-based sensing systems show promise. However, they are highly dependent upon flooring materials that are not necessarily heterogenous. The flooring medium can vary between buildings or even within rooms. Thus, the systems would normally require a significant volume of calibration for every operating environment. The capacitive systems are not impacted by the aforementioned issues associated with vibration- and pressure-based sensing methods.

Human bodies are a source of *Infrared* (IR) radiation. It can be measured by sensors and used to localise a person. There are two main sensor types for measuring IR for localisation – *Passive IR* (PIR) [35] and

thermopile [36] sensors. PIR sensors use pyroelectric sensing elements, which output a voltage when there is a change in incident IR. Consequently, they can only detect a moving target and not a stationary one. Generally, off-the-shelf PIR sensors for motion detection output a binary signal corresponding to the presence of a moving human subject within their field of view. Multiple sensors placed at known locations can then be used for localisation [37-40]. This can also be further improved upon by using customised hardware outputting an analog signal proportional to the change in the incident IR radiation [41-43].

Instead of using pyroelectric sensing elements, an array of thermocouples can be used to make a thermopile – effectively a very low-resolution thermal camera. Thermopiles are able to measure the absolute IR value and can therefore detect both mobile and stationary subjects. Both wall- and ceiling-mounted thermopiles have been used for subject localisation [44-47].

Throughout the PhD research, improvements were made to the data collection methods and ground truth measurements. The literature shows significant variation in the methods of collecting ground truth data. The simplest approach involves a subject standing at a known point. However, this is only useful for evaluating a system's ability at locating a static target. A subject can follow a pre-determined path while walking at a fixed pace. The time taken to walk such a path is measured. Alternatively, a metronome can be employed to synchronise the subject walking. However, this can cause the subject to walk in an unnatural manner. When computational intelligence models are trained upon the data, the regular path can also become a feature. Consequently, the model may struggle to position subjects who are not walking along a similar path. Ideally, subjects should be able to move about the test area unhindered and in a natural manner while their positions are accurately measured by other systems. Several such systems have been used in literature. They are based on various tools such as a hat connected to pulleys and rotational encoders with cables [32], Xbox Kinect [48], OptiTrack [49], and optical camera [50]. During the course of the work reported here, an HTC Vive *Virtual Reality* (VR) system was repurposed for the ground truth measurements. The HTC Vive provides high-quality ground truth recordings with a precision of 0.65 mm and an accuracy of 5 mm [51].

1.2 Contribution

Several localisation systems have been developed and evaluated in the reported study. The main results of it are as follows.

- **A novel passive VLP system that uses only ambient light and employs cheap photodiode-based light-sensors embedded in a wall**. The system does not require modification to the existing lighting infrastructure. A mobile target was tracked accurately for several routes. The impact of the distance metric on the performance of the *Weighted K-Nearest Neighbour* (WKNN) classifier for localisation was investigated. It is the first reported work that uses only ambient light to localise a moving target. An article based on this work was published in IEEE Transactions on Instrumentation and Measurement (Q1).

- **A capacitive flooring system (CapLoc) that determines a position of a mobile target in real-time.** The CapLoc is not data driven. Therefore, it requires minimal calibration for the localisation thus making it more invariant to changes in the setting. It is robust and not vulnerable to factors (like wireless multipath propagation, changes of illumination conditions, different clothes worn by the target, etc.) that adversely affect other DFL systems. The experimental results showed the accurate localisation of a mobile target for multiple trajectories. CapLoc, achieved the highest reported localisation accuracy while tracking a moving subject. It is also illustrated that poses of a person lying on the floor can be captured and identified easily, adding further novelty. In future work, this can potentially be further developed for non-obtrusive automated fall detection. An article based on this work was published in IEEE Access (Q1).

- **A localisation system using wall-mounted thermopile sensors with the application of** *Machine Learning* **(ML)–based sensor models**. The sensor models can be trained once for each sensor and then transferred to other sensors. This leads to a robust and reconfigurable indoor positioning system that does not need to be retrained when deployed outside the training environment. The system can operate with only a single thermopile sensor. Alternatively, extra sensors can be added to increase the localisation accuracy. The results show that the proposed

approach is largely invariant to the subjects, system configuration, and deployment environments. These practical considerations had not been reported in the literature. An article based on this work was published in IEEE Access (Q1).

- **A robust dataset that allows benchmarking of multiple algorithms**. It contains data taken across multiple environments while employing four different test subjects. The accurate ground truth was automatically recorded by using an HTC Vive VR system. Several supervised ML techniques were trained and tested using the dataset with a particular emphasis on the generalisability of the algorithms between subjects and environments – something not widely reported in the literature. It also provides the first available dataset for thermopile-based positioning, thereby addressing a critical need. In addition, the impact of wearing insulating apparel on localisation accuracy was investigated. An article based on this work was submitted to the IEEE Internet of Things Journal (Q1) and is currently under review.

- **An accurate ground truth recording system that employs virtual reality technology**. It ensures that the localisation error is accurately measured. It also allows data collection for a large number of trajectories and path while the subject is walking naturally. This cost-effective solution addresses a critical weakness of the localisation discipline.

- **A novel VLP hardware.** The hardware developed during this PhD, was utilised in two papers [52, 53] that were published in Q1 journals. The literature review that was undertaken to perform the study and development work also contributed towards the paper [22] published in a high-impact Q1 journal. While these papers are not included in the thesis, the candidate is a named second co-author in all three publications. Several other datasets for the thermopile and capacitive floor sensors have been collected. They will be used in planned future work and will be presented in the associated publications.

1.3 Methodology

Over the course of this thesis an experimental methodology was followed. Several prototype systems were developed and implemented, each building upon the strengths and weaknesses of the previous system, in addition to those found in the literature. The performance of these systems was evaluated in

terms of their measured accuracy. The accuracy was determined by comparing the positions of each subject as given by the prototype systems against the actual position of the subject. This is then used to compute the positional error. As per suggestion of the literature [54, 55], median, mean, percentile errors and Empirical Cumulative Distribution Function (ECDF) were used as the accuracy metrics.

1.4 Thesis Structure

Chapter 2 details the development, implementation and evaluation of a passive VLP system. The system is characterised by a comparatively high localisation accuracy, especially when assessed against equivalent RF-based systems. The work was published in a Q1 ranked journal – the IEEE Transactions on Instrumentation and Measurement. A smaller-scale prototype of this system was initially reported in the peer-reviewed conference article that can be found in Appendix 1. Chapter 2 improves upon this initial implementation by investigating the effect of various parameters on the localisation error. The effects of the height of the sensors' placement, sensors' density, the fingerprint database sizes, and the distance metrics were investigated.

Whilst novel, the approach presented in Chapter 2 faced significant challenges. In order to overcome them, a flooring system was developed for passive tracking of subjects. A major advantage of this system is that it does not require a training corpus of data. Thus, it can be installed in a new environment with minimal calibration. Furthermore, the testing was performed on multiple subjects. It was found that the system provided considerably lower localisation errors compared to the passive VLP solution. A much-improved data collection method was developed. The paths of the moving subjects were tracked using the HTC Vive thus not requiring the subjects to follow some predetermined paths. That allowed for more natural movement of the subjects. The system description and the obtained results were published in the Q1 ranked journal – the IEEE Access.

It was found that the system presented in Chapter 2 required a significant training corpus for any new environment, and it was very sensitive to changes in the environment. The new system presented in Chapter 3 improves upon the environment dependence. It was achieved at the expense of a significantly higher infrastructure cost. In a similar fashion as in Chapter 2, the preliminary results associated with

the system were initially disseminated through a peer-reviewed conference article that is given in Appendix 2.

Chapter 4 provides further improvement upon developed systems. The proposed solution is environment-agnostic and of a low infrastructure cost. This was achieved by employing thermopile sensors detecting IR radiation emitted by a subject. The position of the subject within the frame of the image can be mapped to XY coordinates within a room. The system employed thermopiles affixed to the room walls. A trained model on each thermopile is used. The model is sufficiently generalised so that it is able to detect previously unseen subjects across multiple environments. This means that the model can be trained once and then be used in different deployments with very small configurational changes needed. The system can operate with a single sensor. Alternatively, more sensors can be added to decrease the localization error. This work was published in the Q1 ranked journal - the IEEE Access.

Results presented in Chapter 4 were associated with the use of several ML methods. However, there was still scope for a further and more thorough investigation into machine learning methods for thermopile positioning. In addition to this, no thermopile dataset was available in the public domain for use in positioning tasks. Therefore, in Chapter 5 a variety of machine learning methods were compared, with a particular emphasis on advanced neural networks. In addition, a dataset was collected (using the same system as in Chapter 4, albeit with the thermopile affixed to the ceiling as opposed to the walls). The large dataset will also be made publicly available. The paper that is based on this work is currently under review with the Q1 ranked periodical - the IEEE Internet of Things Journal.

Finally, Chapter 6 concludes this thesis by summarising the main contributions. In addition, some of the shortcomings of the systems detailed here within are addressed. These are accompanied by thoughts on potential future works to further develop the systems.

References:

[1]   A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things journal,* vol. 1, no. 1, pp. 22-32, 2014.
[2]   J. Z. Liang, N. Corso, E. Turner, and A. Zakhor, "Image based localization in indoor environments," in *2013 Fourth International Conference on Computing for Geospatial Research and Application*, 2013: IEEE, pp. 70-75.

[3]   R. Li, J. Liu, L. Zhang, and Y. Hang, "LIDAR/MEMS IMU integrated navigation (SLAM) method for a small UAV in indoor environments," in *2014 DGON inertial sensors and systems (ISS)*, 2014: IEEE, pp. 1-15.

[4]   F. Ijaz, H. K. Yang, A. W. Ahmad, and C. Lee, "Indoor positioning: A review of indoor ultrasonic positioning systems," in *2013 15th International Conference on Advanced Communications Technology (ICACT)*, 2013: IEEE, pp. 1146-1150.

[5]   A. Maddumabandara, H. Leung, and M. Liu, "Experimental evaluation of indoor localization using wireless sensor networks," *IEEE Sensors Journal,* vol. 15, no. 9, pp. 5228-5237, 2015.

[6]   H. Xie, T. Gu, X. Tao, H. Ye, and J. Lv, "MaLoc: A practical magnetic fingerprinting approach to indoor localization using smartphones," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 2014, pp. 243-253.

[7]   A. Yassin *et al.*, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Communications Surveys & Tutorials,* vol. 19, no. 2, pp. 1327-1346, 2016.

[8]   M. Maheepala, A. Z. Kouzani, and M. A. Joordens, "Light-based indoor positioning systems: A review," *IEEE Sensors Journal,* vol. 20, no. 8, pp. 3971-3995, 2020.

[9]   P. Müller, S. Ali-Löytty, J. Lekkala, and R. Piché, "Indoor localisation using aroma fingerprints: A first sniff," in *2017 14th Workshop on Positioning, Navigation and Communications (WPNC)*, 2017: IEEE, pp. 1-5.

[10]   P. Remagnino, G. A. Jones, N. Paragios, and C. S. Regazzoni, *Video-based surveillance systems: computer vision and distributed processing*. Boston, MA: Springer, 2002.

[11]   Y.-L. Hou and G. K. Pang, "People counting and human detection in a challenging situation," *IEEE transactions on systems, man, and cybernetics-part a: systems and humans,* vol. 41, no. 1, pp. 24-33, 2010.

[12]   L. Xia, C.-C. Chen, and J. K. Aggarwal, "Human detection using depth information by kinect," in *CVPR 2011 workshops*, 2011: IEEE, pp. 15-22.

[13]   M. R. U. Saputra, G. D. Putra, and P. I. Santosa, "Indoor human tracking application using multiple depth-cameras," in *2012 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 2012: IEEE, pp. 307-312.

[14]   M. M. Rapoport, "The home under surveillance: A tripartite assemblage," *Surveillance & Society,* vol. 10, no. 3/4, pp. 320-333, 2012.

[15]   Y. Ma, G. Zhou, and S. Wang, "WiFi sensing with channel state information: A survey," *ACM Computing Surveys (CSUR),* vol. 52, no. 3, pp. 1-36, 2019.

[16]   K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2. 0: Passive human tracking with a single Wi-Fi link," in *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*, 2018, pp. 350-361.

[17]   D. Konings, F. Alam, F. Noble, and E. M. Lai, "Device-free localization systems utilizing wireless RSSI: A comparative practical investigation," *IEEE Sensors Journal,* vol. 19, no. 7, pp. 2747-2757, 2018.

[18]   R. C. Shit *et al.*, "Ubiquitous localization (UbiLoc): A survey and taxonomy on device free localization for smart world," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 4, pp. 3532-3564, 2019.

[19]   F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 3, pp. 2568-2599, 2019.

[20]   S. Denis, R. Berkvens, and M. Weyn, "A survey on detection, tracking and identification in radio frequency-based device-free localization," *Sensors,* vol. 19, no. 23, p. 5329, 2019.

[21]   M. A. Al-Qaness *et al.*, "Channel state information from pure communication to sense and track human motion: A survey," *Sensors,* vol. 19, no. 15, p. 3329, 2019.

[22]   F. Alam, N. Faulkner, and B. Parr, "Device free localization: A review of non-RF techniques for unobtrusive indoor positioning," *IEEE Internet of Things Journal,* vol. 8, no. 6, pp. 4228-4249, 2020.

[23]   M. Afzalan and F. Jazizadeh, "Indoor positioning based on visible light communication: A performance-based survey of real-world prototypes," *ACM Computing Surveys (CSUR),* vol. 52, no. 2, pp. 1-36, 2019.

[24] M. D. Addlesee, A. Jones, F. Livesey, and F. Samaria, "The ORL active floor [sensor system]," *IEEE Personal Communications,* vol. 4, no. 5, pp. 35-41, 1997.

[25] R. J. Orr and G. D. Abowd, "The smart floor: A mechanism for natural user identification and tracking," in *CHI'00 extended abstracts on Human factors in computing systems*, 2000, pp. 275-276.

[26] M. Andries, O. Simonin, and F. Charpillet, "Localization of humans, objects, and robots interacting on load-sensing floors," *IEEE Sensors Journal,* vol. 16, no. 4, pp. 1026-1037, 2015.

[27] I. Al-Naimi and C. B. Wong, "Indoor human detection and tracking using advanced smart floor," in *2017 8th International Conference on Information and Communication Systems (ICICS)*, 2017: IEEE, pp. 34-39.

[28] M. Mirshekari, S. Pan, J. Fagert, E. M. Schooler, P. Zhang, and H. Y. Noh, "Occupant localization using footstep-induced structural vibration," *Mechanical Systems and Signal Processing,* vol. 112, pp. 77-97, 2018.

[29] S. e. Alajlouni and P. Tarazaga, "A new fast and calibration-free method for footstep impact localization in an instrumented floor," *Journal of Vibration and Control,* vol. 25, no. 10, pp. 1629-1638, 2019.

[30] J. D. Poston, R. M. Buehrer, and P. A. Tarazaga, "Indoor footstep localization from structural dynamics instrumentation," *Mechanical Systems and Signal Processing,* vol. 88, pp. 224-239, 2017.

[31] D. Savio and T. Ludwig, "Smart carpet: A footstep tracking interface," in *21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07)*, 2007, vol. 2: IEEE, pp. 754-760.

[32] H. Rimminen, J. Lindström, and R. Sepponen, "Positioning accuracy and multi-target separation with a human tracking system using near field imaging," *International Journal on Smart Sensing and Intelligent Systems,* vol. 2, no. 1, 2017.

[33] A. Braun, H. Heggen, and R. Wichert, "CapFloor–a flexible capacitive indoor localization system," in *International Competition on Evaluating AAL Systems through Competitive Benchmarking*, 2011: Springer, pp. 26-35.

[34] M. Valtonen, T. Vuorela, L. Kaila, and J. Vanhala, "Capacitive indoor positioning and contact sensing for activity recognition in smart homes," *Journal of Ambient Intelligence and Smart Environments,* vol. 4, no. 4, pp. 305-334, 2012.

[35] H. H. Kim, K. N. Ha, S. Lee, and K. C. Lee, "Resident location-recognition algorithm using a Bayesian classifier in the PIR sensor-based indoor location-aware system," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews),* vol. 39, no. 2, pp. 240-245, 2009.

[36] H. M. Ng, "Human localization and activity detection using thermopile sensors," in *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2013: IEEE, pp. 337-338.

[37] B. Song, H. Choi, and H. S. Lee, "Surveillance tracking system using passive infrared motion sensors in wireless sensor network," in *2008 International Conference on Information Networking*, 2008: IEEE, pp. 1-5.

[38] Q. Hao, F. Hu, and Y. Xiao, "Multiple human tracking and identification with wireless distributed pyroelectric sensor systems," *IEEE Systems Journal,* vol. 3, no. 4, pp. 428-439, 2009.

[39] C. Jing, B. Zhou, N. Kim, and Y. Kim, "Performance evaluation of an indoor positioning scheme using infrared motion sensors," *Information,* vol. 5, no. 4, pp. 548-557, 2014.

[40] B. Yang, Q. Wei, and L. Yuan, "Location ambiguity resolution and tracking method of human targets in wireless infrared sensor network," *Infrared Physics & Technology,* vol. 96, pp. 174-183, 2019.

[41] G. Monaci and A. Pandharipande, "Indoor user zoning and tracking in passive infrared sensing systems," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, 2012: IEEE, pp. 1089-1093.

[42] X. Liu, T. Yang, S. Tang, P. Guo, and J. Niu, "From relative azimuth to absolute location: Pushing the limit of pir sensor based localization," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking*, 2020, pp. 1-14.

[43] B. Mukhopadhyay, S. Srirangarajan, and S. Kar, "Modeling the analog response of passive infrared sensor," *Sensors and Actuators A: Physical,* vol. 279, pp. 65-74, 2018.

[44] 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.

[45] M. Kuki, H. Nakajima, N. Tsuchiya, K. Kuramoto, S. Kobashi, and Y. Hata, "Mining multi human locations using thermopile array sensors," in *2013 IEEE 43rd International Symposium on Multiple-Valued Logic*, 2013: IEEE, pp. 59-64.

[46] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural Networks for Indoor Person Tracking With Infrared Sensors," *IEEE Sensors Letters,* vol. 5, no. 1, pp. 1-4, 2021.

[47] S. Narayana *et al.*, "LOCI: privacy-aware, device-free, low-power localization of multiple persons using IR sensors," in *2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2020: IEEE, pp. 121-132.

[48] T. Grosse-Puppendahl *et al.*, "Platypus: Indoor localization and identification through sensing of electric potential changes in human bodies," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, 2016, pp. 17-30.

[49] D. Yang, B. Xu, K. Rao, and W. Sheng, "Passive infrared (PIR)-based indoor position tracking for smart homes using accessibility maps and a-star algorithm," *Sensors,* vol. 18, no. 2, p. 332, 2018.

[50] 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)*, 2012: IEEE, pp. 2042-2047.

[51] T. Glass, F. Alam, M. Legg, and F. Noble, "Autonomous Fingerprinting and Large Experimental Data Set for Visible Light Positioning," *Sensors,* vol. 21, no. 9, p. 3256, 2021.

[52] F. Alam, N. Faulkner, M. Legg, and S. Demidenko, "Indoor visible light positioning using spring-relaxation technique in real-world setting," *IEEE Access,* vol. 7, pp. 91347-91359, 2019.

[53] D. Konings, N. Faulkner, F. Alam, E. M.-K. Lai, and S. Demidenko, "FieldLight: Device-Free Indoor Human Localization Using Passive Visible Light Positioning and Artificial Potential Fields," *IEEE Sensors Journal,* vol. 20, no. 2, pp. 1054-1066, 2019.

[54] ISO/IEC JTC 1/SC 31, *Information technology–real time locating systems–test. and evaluation of localization and tracking systems (ISO/IEC 18305: 2016)*. International Organization for Standardization Geneva, Switzerland, 2016.

[55] F. Potortì *et al.*, "Comparing the performance of indoor localization systems through the EvAAL framework," *Sensors,* vol. 17, no. 10, p. 2327, 2017.

CHAPTER 2 – WATCHERS ON THE WALL

This chapter is republished in accordance with IEEE's copyright policy. This chapter is the accepted version of the published work, and as such may have stylistic differences from the final published article.

Watchers on the Wall: Passive Visible Light-Based Positioning and Tracking with Embedded Light-Sensors on the Wall

Nathaniel Faulkner, Fakhrul Alam, *Senior Member, IEEE*,
Mathew Legg, *Member, IEEE* and Serge Demidenko, *Fellow, IEEE*

Abstract— This paper reports a novel visible light positioning (VLP) system and associated experimental results. The developed VLP system is completely passive as it does not require a tracked object to carry any active device or tag. At the same time, it does not require any modification to the existing lighting infrastructure. The positioning system, termed Watchers on the Wall (WoW), localizes a target based on measuring the change it creates in the received signal strength (RSS) of the ambient light recorded at an array of light-sensors embedded in the wall. A prototype system has been implemented and tested to investigate the performance of the proposed approach with regard to localization and tracking. The experimental results show that median and 90-percentile localization errors of 7 cm and 21 cm respectively can be achieved for a 2 m x 3.6 m testbed. The effect of various parameters like the height of placement and number of the light-sensors, as well as the size of the fingerprint database, have also been studied. The impact of various distance metrics on the localization performance of the Weighted *K*-Nearest Neighbor (WKNN) classifier has been investigated. It has been found that two distance metrics outperform the commonly employed Euclidean metric. The experimental results also demonstrated that the developed system could track a mobile target along multiple routes with a median error of 12 cm.

Index Terms— Indoor localization, Device-Free Localization, Indoor Positioning System, Visible Light Positioning (VLP), Passive VLP, Weighted *K*-Nearest Neighbor Classifier

I. INTRODUCTION

Indoor positioning has been a burgeoning area of research over the past decades. In terms of outdoor positioning, GPS [1] is the de facto solution, due to it being both ubiquitous and free to use. However, it has limitations, especially in built-up areas or indoors [2]. The GPS signal is adversely impacted by multipath reflections and struggles to penetrate walls. Furthermore, the offered accuracy of several meters [3] is not good enough for indoor applications. For these reasons, other methods have been proposed. They have been based on the use of Radio Frequency Identification [4], Bluetooth [5], Wi-Fi [6], ZigBee [7], Ultra-Wideband [8], Magnetic Fingerprinting [9], Ultrasonic [10] to mention the most popular. Whilst the majority of these represent an improvement over GPS for indoor localization, they often do not provide the desired levels of accuracy, reliability or simplicity. With *Light Emitting Diodes* (LEDs) steadily replacing traditional lighting sources, a new method of positioning has come to the fore – *Visible Light Positioning* (VLP) [11]. Visible light has the benefit of being far less susceptible to multi-path interference and flat fading due to its vastly higher frequency than radio frequency signals [12]. LED lighting can also perform multiple roles – illumination, communication, and positioning. Active VLP has been well researched. It relies on a mobile object having a receiver containing either a photodiode or image sensor [13]. There are several active VLP methods that have been implemented on indoor testbeds, with the main approaches being *Received Signal Strength* (RSS) *Lateration* [14, 15], *Angle of Arrival Angulation* [16], and *Fingerprint Matching* [17].

Passive positioning or *Device-Free Localization* (DFL) [18] allows for the object detection without the need to have any receiving device attached to the tracked object. Potential applications of such localization systems could include location-based services in smart buildings, business analytics for retail applications, emergency evacuations, accessibility aids for visually impaired persons, as well as fall detection in rest homes. DFL systems based on wireless technologies have been investigated extensively in the last decade. The current wireless-based DFL solutions using *Commercial-Off-The-Shelf* (COTS) equipment require a significant number of wireless nodes while offering a median accuracy of approximately 1 m [19]. Recent works employing the customized hardware and *channel state information metric* have shown promising results with the median localization error as low as 0.35 m in *line of sight* human tracking scenarios [20, 21].

While DFL systems based on wireless technologies have

TABLE I: Comparison of WoW with other passive VLP systems

Research	Results Obtained	Receiver Sensor	Modified Lighting infrastructure	Tracking moving target	Limitations
Ibrahim *et. al* [22]	Primarily detected whether a door was open or closed.	PD collocated with LED luminaires	Yes	No	Does not track or localize target.
EyeLight [23]	93.7% occupancy count accuracy, 94 cm median localization error.	PD collocated with LED luminaires	Yes	No	Only works in controlled environment as sunlight saturates the receivers.
LiSense [24]	Mean angular accuracy of 10^0 for the 5 main body joints	Floor inlaid with PD	Yes	No	Does not track or localize target. Needs a large number of PDs
StarLight [25]	Mean angular accuracy of 13.6^0 for the 5 main body joints	Floor inlaid with PD	Yes	No	Does not track or localize target
Zhang *et. al.*[26]	Median error of 8 cm	Floor inlaid with PD	Yes	No	Localization results obtained via simulation only.
Okuli[27]	Position a finger in a 9 cm x 7 cm grid with 0.7 cm median error	Two PD around a tablet	No	No	Does not track or localize a human target. Only positions a user's finger while using a tablet.
CeilingSee [28]	Detected Room occupancy	Reverse biased LED luminaires as PD.	Yes	No	Does not track or localize target.
Hu *et. al.*[29]	Simulation results show the developed algorithm can sense change in environment.	PD embedded in the ceiling	Not specified	No	Does not track or localize. Suggested that environment change can help infer occupancy and position.
Majeed *et.al.* [30]	RMS error is 5 cm	PD collocated with LED luminaires	Yes	No	Requires fingerprinting. Localization results are obtained by simulation only.
WoW (proposed)	Median error of 7 cm for stationary and 13 cm for moving target.	PD embedded within walls.	No	Yes	Requires fingerprinting. Needs to be further developed to work in changing ambient light.

been widely reported in the literature, there are only a handful of existing works dedicated to passive VLP [22-30]. However, just like its active counterpart, passive VLP can potentially be significantly more accurate than the wireless passive positioning techniques. Consequently, there is a need to develop advanced passive VLP solutions.

Collocated LED luminaires and photodiodes have been applied to passively detect humans in [22]. The light from the LED luminaires was multiplexed using *Time-Division Multiple-Access* to identify the source of incoming light at each photodiode. The work primarily focused on investigating the ability to detect whether a door was open or closed. The work was further extended in [23] to track human movement and detect room occupancy. In the research [24], the floor was inlaid with 324 photodiodes, with 5 LED luminaires placed on the ceiling above. That setup was then used to detect the position of a human body and limbs from the shadows cast onto the floor. The work was further extended in [25] using only 20 photodiodes, albeit with a much larger number of LED panels on the ceiling. That simplified the infrastructure at the cost of a slight decrease of the accuracy. Similarly, the authors in [26] also used a grid of photodiodes embedded into the floor. LED luminaires on the ceiling cast shadows from test subjects onto the said photodiodes. However, the paper reported results that were mostly based on simulation. The only experimental result reported in the paper was a single point to point LED to photodiode link to gather parameters for a larger-scale simulation. In simulations, the authors were able to achieve a median error of 8 cm in an 8 m × 8 m × 4 m room with 4 LED luminaires, and photodiodes uniformly spaced at the distance of 0.5 m on the floor. In [27], the authors used a passive VLP approach for mobile device input using one LED and two photodiodes to detect a user's finger. The application of LED

improved the reliability in the presence of changing ambient light. The *CeilingSee* approach [28] employed reverse-biased LED luminaires as photodiodes for occupancy sensing. However, the authors did not use the system for positioning of test subjects or objects. Therefore, they did not report any results on the localization accuracy. Research [29] proposed to use ceiling mounted photodetectors for accurately sensing the indoor environment change. While this technique can potentially be adopted for occupancy inference and position estimation, no localization and tracking algorithms were reported in the paper. In addition, theoretical development was validated with simulation study only, and no practical implementation was done. Another group of researchers reported a passive VLP system that utilized a network of VLC luminaires and PD-based receivers on the ceiling [30]. The system measured the *impulse responses* (IRs) between each transmitter-receiver pair similar to the channel sounding approach [31]. The target was localized based on the measured changes of the IRs. The reported RMS localization error was based on simulation only, and no porotype development or physical system implementation was done.

Table I summarizes the reported research in the area of passive VLP.

This paper focuses on achieving accurate positioning of an object in ambient light conditions without the need for any modification to the existing lighting infrastructure (unlike the majority of VLP solutions). Table I frames the work presented in this paper with respect to the state-of-the-art in the field.

The work presented here extends the preliminary results reported in [32] and makes the following original contributions:
1. *Novel passive VLP system based on ambient light only*. This is the first implemented passive VLP system that the authors are aware of that does not require any
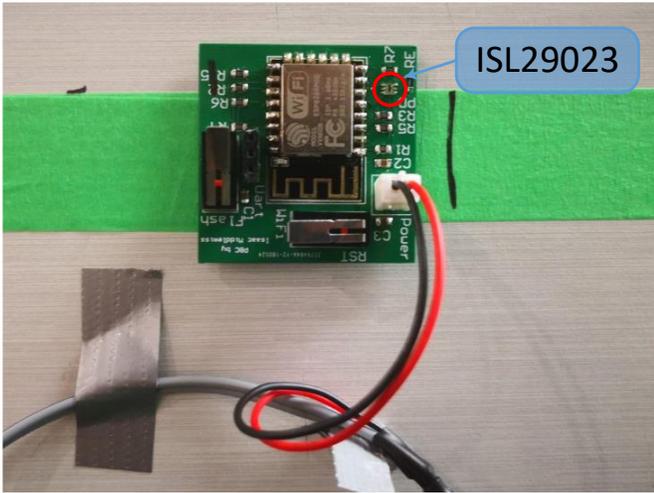
Fig 1. Custom designed light-sensor.

(a) [32]

(b)

Fig. 2. Proof of concept system: actual setup (a), and layout diagram (b).

Fig. 3. Received power at each light-sensor for three scenarios: empty test bed (first bar), test subject at the left-hand side close to the wall with the light-sensors affixed (middle bar), and the right-hand side away from wall with light-sensors affixed (end bar) [32].

modification to the lighting infrastructure.

2. *Functional passive VLP system with the associated experimental results*. The developed system, termed the WoW (shorthand for *Watchers on the Wall*) requires only cheap *photodiode* (PD)-based light-sensors embedded in a wall to operate. The developed system was extensively tested to study the impact of various factors on the localization accuracy.

3. *Moving target tracking*. The ability of the developed system to track a moving object was investigated for several routes. As far as the authors are aware, this is the first work that reports the localization accuracy of a passive VLP system while tracking a moving target traversing multiple routes.

4. *Impact of distance metric on the performance of the Weighted K-Nearest Neighbor (WKNN) classifier*. The impact of various distance metrics on localization performance was investigated. It was found that two distance metrics outperformed the commonly used *Euclidean* metric. To the best of the authors' knowledge, this is the first publication that explores the impact of the distance metric on the performance of the WKNN classifier for passive VLP.

The rest of the paper is organized as follows. Section II describes the hardware and data acquisition system of the developed VLP system, introduces the key concept of using RSS as a fingerprint with the aid of a simple proof of concept system, and proposes utilizing the WKNN algorithm for positioning. Section IV presents the localization performance of the developed system. The section also reports the impact of various parameters on the localization accuracy. Section V concludes the manuscript with suggestions for future work.

II. SYSTEM DEVELOPMENT

A. Key Concept

In a room, there are generally multiple light sources: windows, doors, and interior lights. Walls of the room are often lightly tinted therefore causing a portion of the light to be reflected. A person moving around the room produces several shadows of different intensities projected onto the floor and walls. The major shadows are results of blocking the direct paths from the ambient light sources. However, many other shadows are generated due to multiple reflections and artificial light sources. The shadows can be detected by light-sensors placed around a room as a change in the observed ambient light level, i.e., a change in RSS.
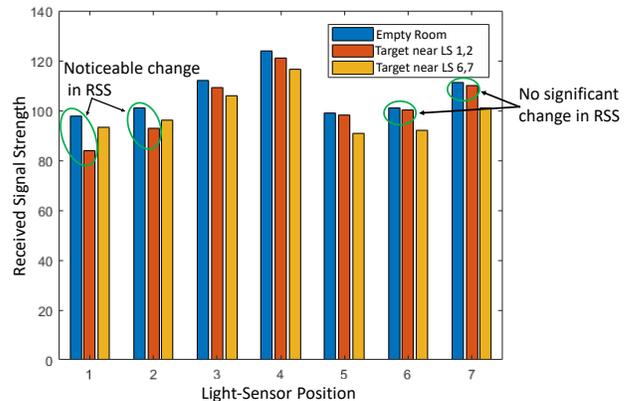
B. Hardware for RSS Data Acquisition

In order to explore the possibility of using the change in RSS
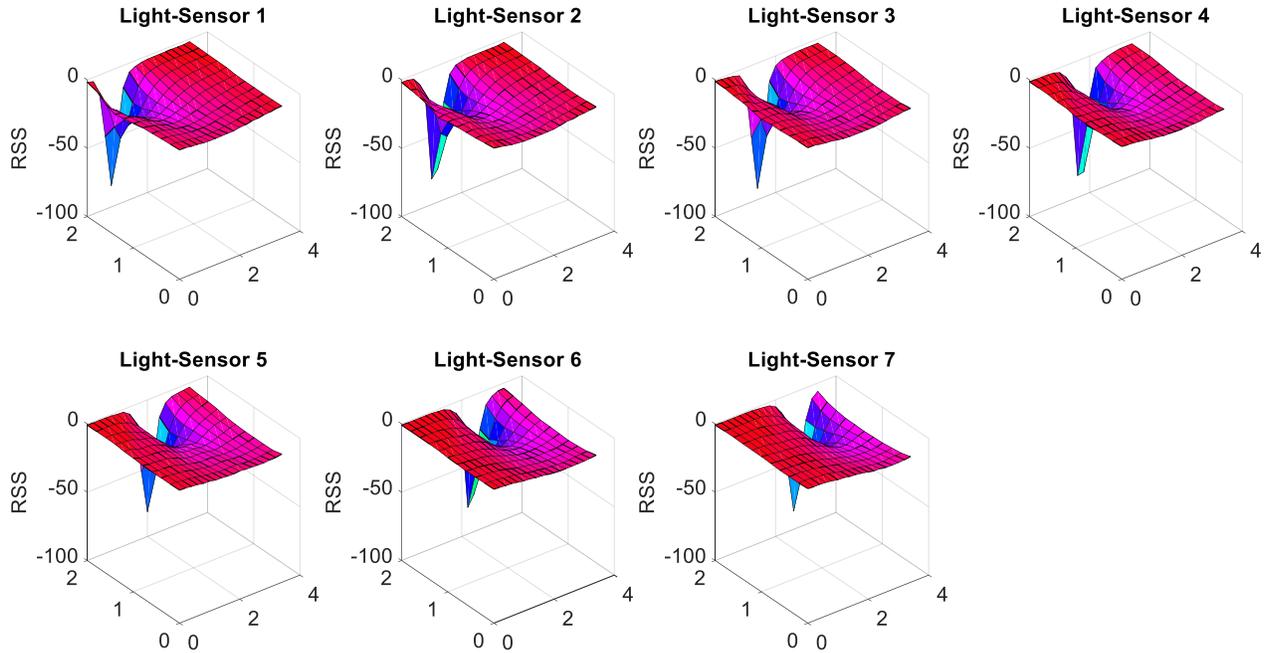
Fig. 4. RSS fingerprint at each light sensor for the proof of concept setup. The XY plane represents the floor (units in m) [32].

for positioning purposes, a simple proof of concept system was set up using ISL29023 [33] integrated digital light sensors (Fig. 1). The light-sensors were comprised of a photodiode, trans-impedance amplifier, and *Analog-to-Digital Converter* (ADC) located in the same package. Each light sensor was connected to a low-cost Wi-Fi microchip (ESP8266 [34]). The ambient light produces a DC signal at the output of the trans-impedance amplifier. The DC level is a measure of the RSS of the ambient light. It is sampled by the embedded ADC and retrieved by the microcontroller of the ESP8266. The latest 100 samples are stored in the internal memory until they are retrieved over Wi-Fi. The data can then be requested in 100-value packets from a computer and saved to a non-volatile memory.

C. Proof of Concept Setup

The sensors were placed on a board at a height of 1.05 m from the floor level. A 3.4 m x 2.2 m grid with 0.2 m squares was marked out using masking tape and a laser straight edge. The sensors were positioned along the side of the grid furthest from the wall, with the photodiodes pointing back towards the wall. The RSS data were collected at each grid intersection for a total of 198 locations. Each measurement consisted of 100 RSS readings over 10 seconds at each sensor. The layout of the proof of concept system is shown in Fig. 2.

D. RSS as Fingerprint

The RSS can be used as a fingerprint to locate mobile objects. This can be observed in Fig. 3 for the proof of concept setup. The blue bars are the RSS at the seven light-sensors when the test area is free from obstructions (i.e., moving or stationary target objects). The red and orange bars present two cases when a person is standing in the front left area (that is close to the first two sensors on the wall – 1 and 2), and then – in the back right position (i.e., opposite the last two sensors – 6 and 7), respectively. Naturally, greater drops in the RSS can be

observed at the sensors that are closer to the test subject. For example, when the test subject is in the front left position, the RSS drop is more significant for the sensors 1 and 2, while there are very little drops for the sensors 6 and 7. When the person is at the back right position, the opposite is true – the sensors 6 and 7 are affected more than the sensors 1 and 2.

The measured RSS values at each light sensor are shown in Fig. 4. These plots show the change in the RSS with a test subject (a person of 1.8 m height) standing at each individual point on the grid. A very large dip can be seen on the top-left edge of each plot where the test subject stood immediately in front of the light-sensor causing a strong shadow. This shows the possibility of taking the RSS value from the same location on each plot to construct a fingerprint ID for that position. The proof of concept along with some preliminary results using that simple setup were reported in [32].

E. WKNN Classifier for Localization using RSS

There are many classifiers to choose from when it comes to positioning that utilizes a fingerprint database. While classifiers like *Support Vector Machines* [35] and *Neural Networks* [36] have been extensively employed for indoor localization using wireless technology, they have not been commonly applied for VLP. The use of the *Weighted K-Nearest Neighbors* (WKNN) [37] is another option to classify the online readings while employing an offline fingerprint database. The recently published research [17] has shown that WKNN is well suited for active VLP system utilizing RSS. Therefore, for this work, WKNN is applied to classify live RSS readings using the fingerprint database.

Let there be N light sensors in the localization system. During the offline measurement, when the target is at a location (x_i, y_i), the corresponding ID for that location based on the RSS at the sensors can be defined as an Nx1 vector:

$$\underline{\mathbf{R}}_i = [R_{1,i}, R_{2,i} \quad \cdots, \quad R_{N,i}]^T \tag{1}$$

Here $R_{n,i}$ refers to the RSS at the nth light-sensor (DC level measured at the output of its trans-impedance amplifier) with the target being at the location i with the coordinates (x_i, y_i). During the offline stage, RSS measurements are taken at all the sensors for M predefined locations, and M x N RSS fingerprint database is constructed. Now the target can be localized in the live phase using the WKNN classifier.

During the live stage, the RSS vector at the N light sensors for a target at the location (x_j, y_j) is given by

$$\underline{\mathbf{R}}_j^{live} = [R_{1,j}^{live}, R_{2,j}^{live} \quad \cdots, \quad R_{N,j}^{live}]^T \tag{2}$$

Here, $R_{n,j}^{live}$ is RSS at the nth light sensor during the live stage. The proximity of the live location to an offline location on the fingerprint database can be determined by computing the distance $d_{j,i}$ between the vectors $\underline{\mathbf{R}}_j^{live}$ and $\underline{\mathbf{R}}_i$. By sorting the distances in ascending order, the "*nearest neighbours*" to the current live location in the offline fingerprint database are identified. The WKNN algorithm estimates the location of the target as the weighted average of the location of the first K nearest neighbours as

$$\tilde{x}_J = \frac{\sum_{k=1}^{K} w_{j,k} \times x_k}{\sum_{k=1}^{K} w_{j,k}}$$
$$\tilde{y}_J = \frac{\sum_{k=1}^{K} w_{j,k} \times y_k}{\sum_{k=1}^{K} w_{j,k}} \tag{3}$$

Here $(\tilde{x}_J, \tilde{y}_J)$ is the estimated position of the target and (x_k, y_k) is the position of the kth neighbour. The weight $w_{j,k}$ is the reciprocal of the distance $d_{j,k}$, thus giving larger weights to nearer neighbours. The value of $K = 3$ was empirically chosen in the WKNN algorithm for the system as it provided a good balance for optimizing both the median and maximum localization errors.

III. LOCALIZATION PERFORMANCE

This section investigates the localization performance of the system and reports the impact of various parameters on the localization accuracy.

A. Experimental Setup

The initial proof of concept setup described in Section II C was found to have large positioning errors at the extremities of

Fig. 5. Experimental setup showing the wall mounted light-sensors and the test space.

the test space. A slightly modified experimental setup was therefore used. The room was set up with 14 light-sensors (please see Section II B for the description of the hardware). It was decided to place the sensors on two opposing walls (as opposed to only one wall described in the proof of concept setup). Also, the spacing of the light-sensors was increased from 0.3 m (of the proof of concept setup) to an interval of 0.6 m. This was done to ensure that there were enough light-sensors near the ends of the test space. The physical setup can be seen in Fig. 5. Seven sensors were placed along each wall giving in total 14 sensing nodes ($N = 14$ in (1) and (2)). A grid of 3.6 m $\times$ 2 m dimension with 0.2 m squares was marked out with the sensors being 0.4 m back from the grid on each side and in-line with the end of the grid at both the ends of the experimental space. The width of the space was therefore 2.8 m and the length 3.6 m (however, with no walls across the ends). Data were collected at each grid intersection for a total of 209 locations using the acquisition method described in Section II B. The data were split into two parts: 1) offline fingerprint database, and 2) online RSS measurements.

All the experiments were conducted at night when the ambient light could be controlled. Multiple datasets were collected for the entire test space as shown in Fig. 6. In each case, the data was collected starting at the first position located at the corner of the grid closest to the light sensor 7 (marked as LS7 in Fig. 6). A test subject then stood at each point on the

(a) Measurement locations - "Database 1"

(b) Measurement locations - "Database 2"
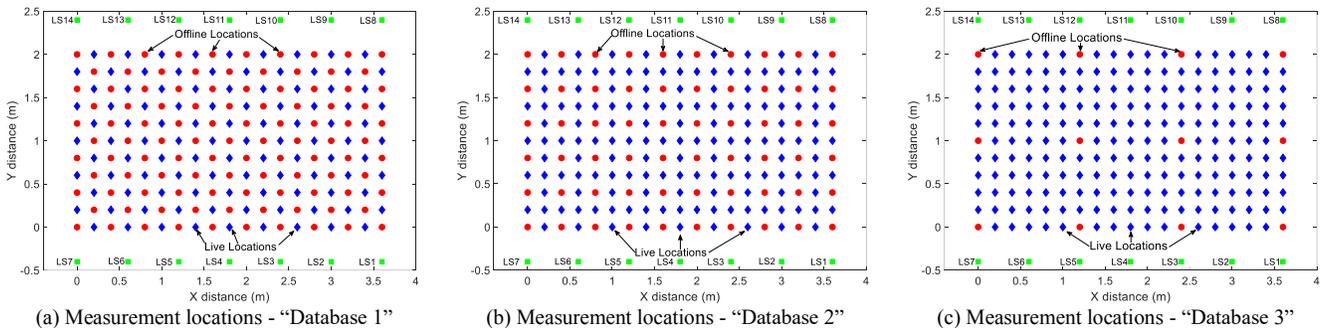

(c) Measurement locations - "Database 3"

Fig. 6. Experimental Layout showing the location of the light sensors and measurement locations for three different fingerprint databases.
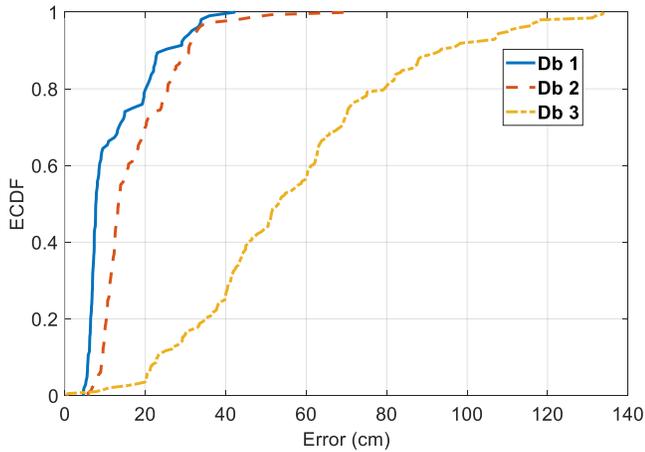
Fig. 7. ECDF of localization error for three fingerprint database sizes.

Fig. 8. ECDF of localization error for three sensor heights.

TABLE II: MEDIAN AND 90-PERCENTILE ERRORS FOR DIFFERENT
DATABASE SIZES

Error in cm					
Database 1		Database 2		Database 3	
Median	90-perc.	Median	90-perc.	Median	90-perc.
8	26	13	31	53	93

TABLE III: DEFINITION OF DISTANCE METRICS

Distance Metric	Definition
Euclidean	$d_{j,i} = \sqrt{\sum_{n=1}^{N} \left(R_{n,j}^{live} - R_{n,i} \right)^2}$
Manhattan	$d_{j,i} = \sum_{n=1}^{N} \left| R_{n,j}^{live} - R_{n,i} \right|$
Canberra	$d_{j,i} = \sum_{n} \dfrac{\left| R_{n,j}^{live} - R_{n,i} \right|}{\left| R_{n,j}^{live} + R_{n,i} \right|}$

grid in sequence whilst the reading was taken. Each reading consisted of taking measurements from each sensor simultaneously over a period of 5 seconds with the data being sampled at 10 times per second, giving an array of 50 samples per sensor per reading. The readings were taken starting at (0, 0) and proceeding in the *x* direction, i.e., (0, 0) to (1, 0) … to (19, 0) before starting the readings at the next row. Each reading was taken for a 1.8 m tall human subject facing the wall, i.e., the line of the subject's shoulders was parallel to the wall. Readings were also taken with no test subjects being present (i.e., background reading representing effectively an empty room). Background readings were taken before and after each dataset to verify that the ambient light level stayed constant.

B. Impact of the Fingerprint Database Size

Table II shows the median and the 90-percentile localization errors for the three fingerprint databases of different sizes. Offline measurement locations can be seen in Fig. 6. The *Empirical Cumulative Distribution Function* (ECDF) of the localization error is shown in Fig. 7. It can be observed that there is a clear trade-off between the localization accuracy and size or resolution of the grid. The localization accuracy degrades as the fingerprint database becomes smaller with a sparser grid. The degradation of accuracy in going from the Database 1 (111 offline measurements, *M* = 111) to the Database 2 (60 offline measurements, *M* = 60) is relatively small. Also for the Database 3, with only 12 offline measurements, the 90-percentile localization error (93 cm) is still less than 1 m thus making the WoW more accurate than many state-of-the-art wireless DFL systems [19].

It should be noted that the selection of the database is not optimized. The optimum fingerprint locations are dependent on many parameters with some being dynamic and also varying with the site. The offline measurement locations optimized for one test environment may not be ideal for another situation.

Fig. 9. ECDF of localization error for various distance metrics.

TABLE IV: LOCALIZATION ERRORS FOR VARIOUS DISTANCE
METRICS

Error in cm (Database 1)					
Euclidean		Manhattan		Canberra	
Median	90-perc.	Median	90-perc.	Median	90-perc.
8	26	7	21	7	21

Therefore simple, regular patterns for offline locations were utilized. The presented localization results could potentially be improved by using more favorable fingerprint databases found through trial and error of various offline location sets. However, this can lead to an over-trained system. Besides this may not be an objective representation of a real-world scenario where it is

(a) 14 Sensors   (b) 8 Sensors

(c) 6 Sensors   (d) 4 Sensors
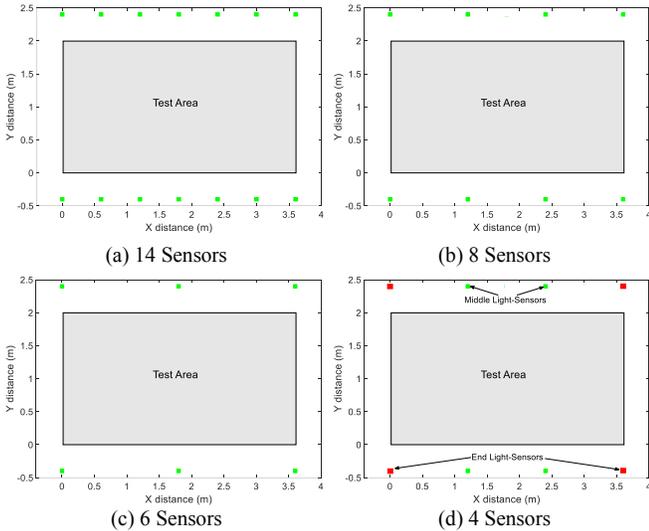
Fig. 10. Experimental layout showing the location of the light-sensors for varying sensor numbers. Note that two different layouts for the 4-sensors arrangements are shown in (d).
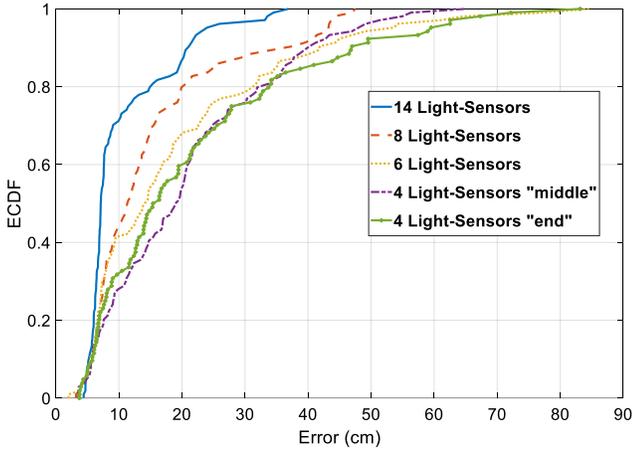
Fig. 11. ECDF of localization error for various number of light-sensors.

not always feasible to optimize the location of the offline measurement or fingerprint locations.

C. Impact of the Sensor Placement Height

Figure 8 shows the impact on the localization accuracy of the sensors placement height on the wall. Three different heights were investigated with the light-sensors set at heights of 0.75 m, 1.2 m, and 1.45 m from the floor level. It can be observed that the height of the sensor placement does not have any noticeable impact on the localization accuracy of the WoW. Thus the experimental results shown for the rest of the paper are for sensors placement height of 1.45 m.

D. Impact of the Distance Metric

Euclidean distance is commonly utilized for identifying the nearest neighbours and computing the weight of the WKNN algorithm [37]. However, several alternative distance metrics are known from the literature [38]. Recent work on active VLP [17] has shown that the selection of the distance metrics can have an impact on the localization accuracy of the WKNN algorithm. Consequently, the effect of distance metrics on the accuracy of the VLP system was investigated. The distance metrics are defined in Table III. The localization results for Database 1 are shown in Fig. 9 and Table IV. It can be observed that the Euclidean distance is not the most accurate metric. Two other distance metrics (*Manhattan* and *Canberra*) produce lower localization errors. Localization results for Database 2 and Database 3 show similar trend.

### E. Impact of the Number of Sensors

The impact of the number of sensors on the localization accuracy of the WoW was also investigated. Figure 10 shows the locations of the sensor nodes for these experiments. Figure 11 and Table V show the localization performance for various sensor numbers. As expected, the localization accuracy degraded when the number of deployed nodes was reduced. However, the 90-percentile error was below 50 cm even with only four sensors being employed.

F. Tracking a Mobile Target

In order to test how the developed system tracks a moving target, multiple paths were followed by the test subject. The target walked along a marked path at a constant speed of 0.2 m/s. The steps were synchronized to a metronome to ensure that the distance covered by each step and the walking speed remained constant. The deliberate walking speed allowed the ground truth (the actual location of the target) to be accurately estimated. Each path was recorded over 90 second period. Three different routes were employed to investigate the capability of the WoW at tracking a moving target. The results are shown in Fig. 12 and Table VI. It can be observed that the positioning error levels are similar for all three paths. In addition, the Euclidean distance performed worse among the three distances. The performances for the Canberra and Manhattan distances were nearly identical, and consequently, only the results for the Manhattan distance are shown against those of the Euclidean distance. When the subject walked around, the orientation of the subject varied leading to changes in the size of the shadow. The fingerprinting was performed with the subject in a single orientation (facing one of the walls). Larger errors were observed when the subject was facing significantly different directions compared to that when the fingerprints were taken. This is more noticeable in Fig. 12(a). The positioning error is smaller on the left and right sides. Here, the test subject walked towards or away from the sensors with the body orientation being similar to that observed during the fingerprinting. Whereas for the top and bottom parts of the trajectory, the positioning error is more pronounced. For these

TABLE V: LOCALIZATION ERRORS FOR DIFFERENT NUMBER OF LIGHT SENSORS

Error in cm (Sensor height 1.45 m, Database 1), Manhattan Distance									
14 diodes		8 diodes		6 diodes		4 diodes (middle)		4 diodes (end)	
Median	90-perc.	Median	90-perc.	Median	90-perc.	Median	90-perc.	Median	90-perc.
7	21	11	38	14.5	42	19	41	16	47
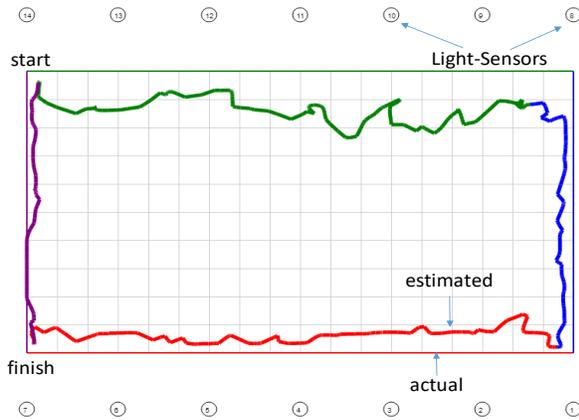
(a) Actual vs. estimated route – Path 1
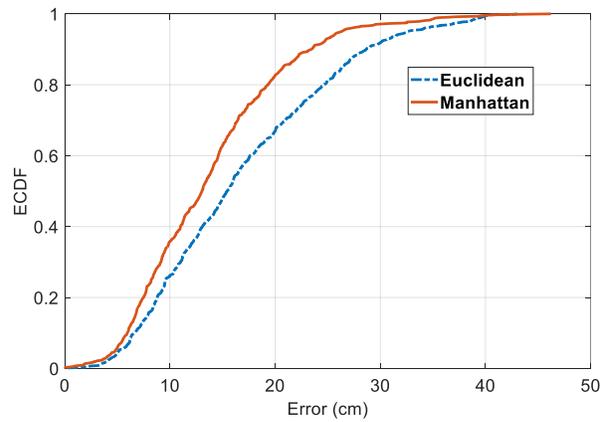
(b) ECDF of localization error - Path 1
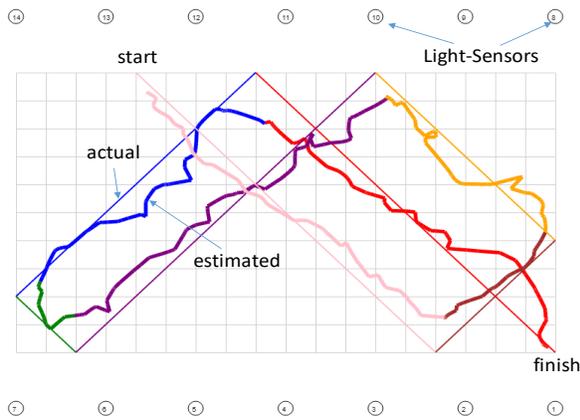
(c) Actual vs. estimated route - Path 2
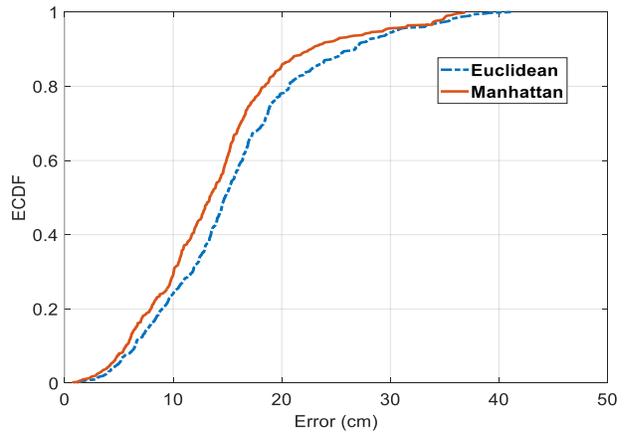
(d) ECDF of localization error - Path 2
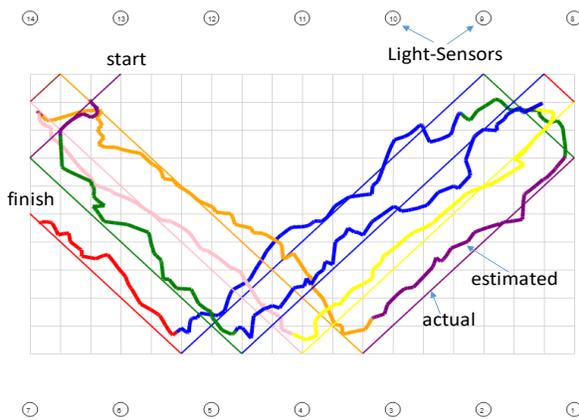
(e) Actual vs. estimated route - Path 3
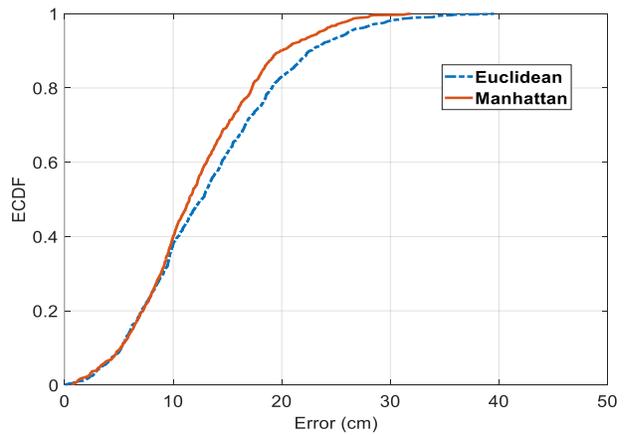
(f) ECDF of localization error - Path 3

Fig. 12. Tracking performance for three different target routes.

TABLE VI: LOCALIZATION ERRORS FOR VARIOUS TARGET TRAJECTORIES

Error in cm											
Path 1				Path 2				Path 3			
Euclidean		Manhattan		Euclidean		Manhattan		Euclidean		Manhattan	
Median	90-perc.	Median	90-perc.	Median	90-perc.	Median	90-perc.	Median	90-perc.	Median	90-perc.
16	29	13	23	15	27	13	23	13	23	12	21

trajectories, the orientation of the test subject is $\pm90^0$ turned compared to that observed during the fingerprinting.

The sampling rate of the VLP system is 10 Hz, which is the maximum sampling frequency the custom-designed light-sensors can handle at the 16-bit resolution. The conducted

experiments showed that the system could cope with a faster walking speed of up to 0.8 m/s. However, it was difficult to maintain a constant speed and achieve an accurate recording of the ground truth. Also, if the resolution is reduced, the sample rate of the current hardware can be increased allowing to track targets that are moving even faster. However, the loss of resolution would lead to a coarser estimation of RSS and could potentially lower the localization accuracy. In future work, the hardware design may need to be improved to increase the sampling rate without sacrificing the resolution in order to track faster targets.

IV. Conclusion & Future Works

This paper presents the development and implementation of a passive visible light-based indoor localization system that employs cost-effective components. The system was able to position a target with a median error of 7 cm in stationary and 12 cm in mobile scenarios using 14 wall-mounted light-sensors. The VLP system performed effectively using WKNN classifier and a fingerprint database consisting of 60 offline measurements within a 2 m x 3.6 m testbed. Weights computed using either Manhattan or Canberra distance provided better positioning accuracy than the traditional Euclidean distance for the WKNN classifier. The placement of the light-sensors within a range between 0.75 m and 1.45 m of height did not show any noticeable impact on the localization accuracy. Therefore a sensor placement height of 1.45 m is preferable to reduce the possibility of occlusion resulting from furniture or other paraphernalia. The localization accuracy degraded once the number of light-sensors was reduced. However, even with only 4 wall-mounted sensors, it was possible to attain a median positioning accuracy of 16 cm for a stationary target.

Further work will expand the test to a full room-scale with the light sensors embedded in all the room walls. In a larger room, enough shadows may not be cast by the target on the walls. In such a scenario, additional light sensors may need to be embedded in the ceiling and the floor, in particular, in the middle of the room.

The experiments were undertaken at night. Therefore changes in the level of ambient light were not investigated. The future work will study quantifying and mitigating the impact of changing the ambient light. Ambient light is measured as a DC signal at the output of the light-sensors. The proposed VLP system infers the location through monitoring the change a target causes to the DC levels at various light-sensors. Therefore, a change in the ambient light level could affect the localization accuracy of the proposed system. This can potentially be mitigated by using *Visible Light Communication* (VLC) -enabled luminaires that transmit modulated light. Under such circumstances, the sensors will monitor the RSS at a specific set of frequencies rather than the DC levels. If the ambient light level changes, the RSS at the modulating frequencies will not change. The change in RSS will be solely due to the occlusion, e.g., the presence of a target. Therefore, as long as the ambient light does not saturate the light-sensors, the accuracy of such a VLP system would not depend on variations in the ambient light. Another potential way to mitigate this could be to employ two separate RSS metrics measuring the long-term and the short-term levels of the ambient light. A

similar concept has been applied with RSS histograms for wireless DFL to offset the fluctuations of RSS occurring from the dynamic nature of the wireless channel [39].

The developed system was tested for a single target at a time, and as such, further investigation is planned to track multiple objects. Since generating the fingerprint database is a very time-consuming process, future research will look at the modelling of the RSS data, and their generation from a few strategically selected calibration points. Utilizing other types of classifiers (e.g., Neural Networks and Support Vector Machines) and comparing their performance with WKNN will be another direction for future research. Finally, the effect of the colour of the target's clothing on the system performance was also left for future investigation.

References

[1] J. Farrell, and M. Barth, *The global positioning system and inertial navigation*. Mcgraw-hill New York, 1999.

[2] M. B. Kjærgaard, H. Blunck, T. Godsk, T. Toftkjær, D. L. Christensen, and K. Grønbæk, "Indoor positioning using GPS revisited," in *International Conference on Pervasive Computing*, 2010, pp. 38-56.

[3] M. G. Wing, A. Eklund, and L. D. Kellogg, "Consumer-grade global positioning system (GPS) accuracy and reliability," *Journal of Forestry,* vol. 103, no. 4, pp. 169-173, 2005.

[4] C. Huang, L. Lee, C. C. Ho, L. Wu, and Z. Lai, "Real-Time RFID Indoor Positioning System Based on Kalman-Filter Drift Removal and Heron-Bilateration Location Estimation," *IEEE Transactions on Instrumentation and Measurement,* vol. 64, no. 3, pp. 728-739, 2015.

[5] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. D. Hamalainen, "Experiments on local positioning with Bluetooth," in *International Conference on Information Technology: Coding and Computing*, 2003, pp. 297-303.

[6] A. Makki, A. Siddig, M. Saad, J. R. Cavallaro, and C. J. Bleakley, "Indoor Localization Using 802.11 Time Differences of Arrival," *IEEE Transactions on Instrumentation and Measurement,* vol. 65, no. 3, pp. 614-623, 2016.

[7] V. Bianchi, P. Ciampolini, and I. De Munari, "RSSI-Based Indoor Localization and Identification for ZigBee Wireless Sensor Networks in Smart Homes," *IEEE Transactions on Instrumentation and Measurement,* no. 99, pp. 1-10, 2018.

[8] A. Cazzorla, G. De Angelis, A. Moschitta, M. Dionigi, F. Alimenti, and P. Carbone, "A 5.6-GHz UWB position measurement system," *IEEE Transactions on Instrumentation and Measurement,* vol. 62, no. 3, pp. 675-683, 2013.

[9] B. Kim and S.-H. Kong, "A novel indoor positioning technique using magnetic fingerprint difference," *IEEE Transactions on Instrumentation and Measurement,* vol. 65, no. 9, pp. 2035-2045, 2016.

[10] S. J. Kim and B. K. Kim, "Accurate hybrid global self-localization algorithm for indoor mobile robots with two-dimensional isotropic ultrasonic receivers," *IEEE Transactions on Instrumentation and Measurement,* vol. 60, no. 10, pp. 3391-3404, 2011.

[11] Y. Zhuang *et al.*, "A Survey of Positioning Systems Using Visible LED Lights," *IEEE Communications Surveys & Tutorials,* vol. 20, no. 3, pp. 1963-1988, 2018.

[12] T. Komine, and M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE Transactions on Consumer Electronics,* vol. 50, no. 1, pp. 100-107, 2004.

[13] W. Zhang and M. Kavehrad, "Comparison of VLC-based indoor positioning techniques," in *Broadband Access Communication Technologies VII*, 2013, vol. 8645, p. 86450M.

[14] B. Xie *et al.*, "LIPS: A light intensity--based positioning system for indoor environments," *ACM Transactions on Sensor Networks,* vol. 12, no. 4, p. 28, 2016.

[15] D. Konings, B. Parr, F. Alam, and E. M.-K. Lai, "Falcon: Fused Application of Light Based Positioning Coupled With Onboard Network Localization," *IEEE Access,* vol. 6, pp. 36155-36167, 2018.

[16] M. Yasir, S.-W. Ho, and B. N. Vellambi, "Indoor Position Tracking Using Multiple Optical Receivers," *Journal of Lightwave Technology,* vol. 34, no. 4, pp. 1166-1176, 2016.

[17] F. Alam, M. T. Chew, T. Wenge, and G. S. Gupta, "An Accurate Visible Light Positioning System Using Regenerated Fingerprint Database Based on Calibrated Propagation Model," *IEEE Transactions on Instrumentation and Measurement,* vol. 68, pp. 2714-2723, 2018.

[18] M. Youssef, M. Mah, and A. Agrawala, "Challenges: device-free passive localization for wireless environments," in *13th Annual ACM International Conference on Mobile Computing and Networking,* 2007, pp. 222-229.

[19] D. Konings, F. Alam, F. Noble, and E. M. Lai, "Device-Free Localization Systems Utilizing Wireless RSSI: A Comparative Practical Investigation," *IEEE Sensors,* vol. 19, no. 7, pp. 2747-2757, 2019.

[20] Y. Ma, G. Zhou, and S. Wang, "WiFi Sensing with Channel State Information: A Survey," *ACM Computing Surveys*, vol. 52(3), no 46, 2019.

[21] K. Qian, C. Wu, Y. Zhang, G. Zhang, Z. Yang, and Y. Liu, "Widar2. 0: Passive human tracking with a single wi-fi link," in *16th Annual International Conference on Mobile Systems, Applications, and Services*, 2018, pp. 350-361.

[22] M. Ibrahim, V. Nguyen, S. Rupavatharam, M. Jawahar, M. Gruteser, and R. Howard, "Visible light based activity sensing using ceiling photosensors," in *3rd Workshop on Visible Light Communication Systems*, 2016, pp. 43-48.

[23] V. Nguyen, M. Ibrahim, S. Rupavatharam, M. Jawahar, M. Gruteser, and R. Howard, "Eyelight: Light-and-Shadow-Based Occupancy Estimation and Room Activity Recognition," in *IEEE Conference on Computer Communications*, 2018, pp. 351-359.

[24] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, "Human Sensing Using Visible Light Communication," in *21st Annual International Conference on Mobile Computing and Networking*, 2015, pp. 331-344.

[25] T. Li, Q. Liu, and X. Zhou, "Practical Human Sensing in the Light," in *14th Annual International Conference on Mobile Systems, Applications, and Services*, 2016, pp. 71-84.

[26] S. Zhang, K. Liu, Y. Ma, X. Huang, X. Gong, and Y. Zhang, "An Accurate Geometrical Multi-Target Device-Free Localization Method Using Light Sensors," *IEEE Sensors,* vol. 18, no. 18, pp. 7619-7632, 2018.

[27] C. Zhang, J. Tabor, J. Zhang, and X. Zhang, "Extending Mobile Interaction Through Near-Field Visible Light Sensing," *21st Annual International Conference on Mobile Computing and Networking*, 2015, pp. 345-357.

[28] Y. Yang, J. Hao, J. Luo, and S. J. Pan, "CeilingSee: Device-free occupancy inference through lighting infrastructure based LED sensing," in *2017 IEEE International Conference on Pervasive Computing and Communications*, 2017, pp. 247-256.

[29] S. Hu, Q. Gao, C. Gong, and Z. Xu, "Efficient Visible Light Sensing in Eigenspace," *IEEE Communications Letters,* vol. 22, no. 5, pp. 994-997, 2018.

[30] K. Majeed and S. Hranilovic, "Passive Indoor Localization for Visible Light Communication Systems," in *2018 IEEE Global Communications Conference*, 2018, pp. 1-6.

[31] D. Demery, J. Parsons, and A. Turkmani, "Sounding Techniques for Wideband Mobile Radio Channels: A Review," *IEE Proceedings I (Communications, Speech and Vision),* vol. 138, no. 5, pp. 437-446, 1991.

[32] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Smart Wall: Passive Visible Light Positioning with Ambient Light Only," in *2019 IEEE International Instrumentation and Measurement Technology Conference*, 2019, pp. 1740-1745.

[33] (June 14, 2019). *Renesas, ISL29023 Integrated Digital Light Sensor with Interrupt, May 2014. Rev. 4.* Available: https://www.renesas.com/ja-jp/www/doc/datasheet/isl29023.pdf

[34] (March 05, 2019). *ESP8266 Technical Reference*. Available: https://www.espressif.com/sites/default/files/documentation/esp8266-technical_reference_en.pdf

[35] M. Brunato, and R. Battiti, "Statistical learning theory for location fingerprinting in wireless LANs," *Computer Networks,* vol. 47, no. 6, pp. 825-845, 2005.

[36] C. Laoudias, D. G. Eliades, P. Kemppi, C. G. Panayiotou, and M. M. Polycarpou, "Indoor localization using neural networks with location fingerprints," in *International Conference on Artificial Neural Networks*, 2009, pp. 954-963.

[37] S. A. Dudani, "The Distance-Weighted k-Nearest-Neighbor Rule," *IEEE Transactions on Systems, Man, and Cybernetics,* vol. SMC-6, no. 4, pp. 325-327, 1976.

[38] V. Prasath, H. A. A. Alfeilat, O. Lasassmeh, and A. Hassanat, "Distance and Similarity Measures Effect on the Performance of K-Nearest Neighbor Classifier-A Review," *arXiv preprint arXiv:1708.04321,* 2017.

[39] D. Konings, F. Alam, F. Noble, and E. M. Lai, "Improved Distance Metrics for Histogram-Based Device-Free Localization," *IEEE Sensors,* vol. 19, pp. 8940-8950, 2019.

Nathaniel Faulkner received his BE(Hons) from Massey University, Auckland, New Zealand in Electronic & Computer Engineering in 2016. He is currently pursuing Ph.D. at the same institute. His research interests include visible light positioning and internet of things.

Fakhrul Alam (M'17-SM'19) is a senior lecturer at the Department of Mechanical & Electrical Engineering, School of Food & Advanced Technology, Massey University, New Zealand. He received BSc (Hons) in Electrical & Electronic Engineering from BUET, Bangladesh and MS and Ph.D. in Electrical Engineering from Virginia Tech, USA. His research interest includes indoor localization, 5G & visible light communication, IoT & wireless sensor networks. He is a member of the IEEE I&M Society and IET.

Mathew Legg (M'19) Received his B.Sc, M.Sc, and Ph.D in Physics from the University of Auckland, New Zealand. He is currently a lecturer with the Department of Mechanical and Electrical Engineering, School of Food and Advanced Technology, Massey University. His research relates to the development of acoustic/ultrasonic measurement systems and techniques for acoustic imaging, non-destructive testing, and remote sensing.

Serge Demidenko (M'91-SM'94-F'04) is Professor and Dean of the School of Science and Technology, Sunway University, Malaysia. He is also associated with the Department of Mechanical & Electrical Engineering, School Food and Advanced Technology, Massey University, New Zealand. He graduated in Computer Engineering from the Belarusian State University of Informatics and RadioElectronics, and received Ph.D. from the Institute of Engineering Cybernetics of the Belarusian Academy of Sciences. His research interests include electronic design and test, signal processing, instrumentation and measurements. He is also Fellow of IET and UK Chartered Engineer.

CHAPTER 3 – CAPLOC

This chapter is republished in accordance with IEEE's copyright policy. This chapter is the accepted version of the published work, and as such may have stylistic differences from the final published article.

CapLoc: Capacitive Sensing Floor for Device-Free Localization and Fall Detection

Nathaniel Faulkner[1,2], Baden Parr[1], Fakhrul Alam[1], Senior Member, IEEE, Mathew Legg[1], Member, IEEE, and Serge Demidenko[1,2], Fellow, IEEE

[1]Department of Mechanical & Electrical Engineering, School of Food & Advanced Technology, Massey University, Auckland 0632, New Zealand
[2]School of Science & Technology, Sunway University, 47500 Selangor, Malaysia

Corresponding author: F. Alam (e-mail: f.alam@ massey.ac.nz).

ABSTRACT Passive indoor positioning, also known as *Device-Free Localization* (DFL), has applications such as occupancy sensing, human-computer interaction, fall detection, and many other location-based services in smart buildings. Vision-, infrared-, wireless-based DFL solutions have been widely explored in recent years. They are characterized by respective strengths and weaknesses in terms of the desired accuracy, feasibility in various real-world scenarios, etc. Passive positioning by tracking the footsteps on the floor has been put forward as one of the promising options. This article introduces CapLoc, a floor-based DFL solution that can localize a subject in real-time using capacitive sensing. Experimental results with three individuals walking 39 paths on the CapLoc show that it can detect and localize a single target's footsteps accurately with a median localization error of 0.026 m. The potential for fall detection is also shown with the outlines of various poses of the subject lying upon the floor.

INDEX TERMS Capacitive Sensing, Device-Free Localization, Electric Field Sensing, Fall Detection, Footstep Detection, Footstep Tracking, Human Sensing, Indoor localization, Indoor Positioning System (IPS), Passive Positioning.

I. INTRODUCTION

Passive indoor positioning is the key enabling technology for applications like *Ambient Assisted Living* (AAL) and *Human-Computer Interaction* (HCI). Unfortunately, even with the attention of researchers for over two decades, passive positioning or *Device-Free Localization* (DFL) remains a problem to be solved. Camera-based techniques can accurately locate and identify a tag-less target with reasonable accuracy [1]. However, they require good lighting conditions and are adversely impacted by occlusion. More importantly, privacy is a significant concern making such systems less acceptable in many applications, especially in a residential setting. Many accidents and falls happen in places such as bathrooms and bedrooms where cameras would be considered to be invasive. While efforts are underway to utilize privacy-preserving single-pixel cameras [2, 3], it is still early days for such a technique.

Passive localization using *Radio Frequency* (RF) sensing has been extensively researched in recent years [4, 5]. While RF-based localization has the advantage of potentially being able to repurpose the wireless networks within the built environment, there are some inherent disadvantages like

limited accuracy due to multipath reflections. Application of the *Channel State Information* (CSI) metric utilizing many Wi-Fi subcarriers can mitigate the multipath issue [6] to achieve much-improved accuracy [7-9] and even perform sophisticated tasks like activity recognition [10]. However, CSI is not available for the majority of the RF technologies (e.g., Bluetooth and ZigBee). In addition to this, most consumer-grade Wi-Fi hardware is yet to widely support the use of this metric thus limiting its practicality.

Passive *Visible Light Positioning* (VLP) [11, 12] is based on the principle that the presence of a subject alters optical channels. These changes can be detected by nearby light-sensors as variation in the *Received Signal Strength* (RSS) of the light level and subsequently used to estimate the subject's position. However, the majority of passive VLP techniques are vulnerable to change in ambient light levels. Also, they need good illumination conditions. *Infrared* (IR) sensing has been proposed as an alternative way for DFL by detecting the heat signature of a human target. *Passive IR* (PIR) sensors, commonly available as motion detectors, have been used for such localization [13-16]. However, PIR sensors require

TABLE 1. Comparison of CapLoc with other floor-based positioning systems.

Research	Sensing Method	Position Accuracy
Liau et al. [19]	Pressure	85-percentile error of 0.283 m
Andries et al. [20]	Pressure	Mean error of 0.13 m for a single person, 0.2 m for two people
Al-Naimi et al. [21]	Pressure	Mean error 0.0767 m
Murakita et al. [22]	Binary Pressure Sensors	Mean error of 0.2 m
Mirshekari et al. [23]	Vibration	Median localization error of 0.38 m
Alajlouni et al. [24]	Vibration	80-percentile error of 0.7 m
Poston et al. [25]	Vibration	RMSE of 0.6 m and 0.8 m in two separate environments
Smartcarpet [26]	Capacitive	MSE 0.0187 m (line) to 0.431 m (C-shape) for various trajectories
Rimmeinen et al. [30, 31]	Capacitive	Mean position error of 0.21 m
Capfloor [32]	Capacitive	"In the range of 50 cm"
Tiletrack [34, 35]	Capacitive	80-percentile error 0.1 m
CapLoc (This paper)	Capacitive	Median error 0.026 m, 90-percentile error 0.066 m

TABLE 2. Comparison of CapLoc with other passive positioning systems.

Research	Sensing Method	Position Accuracy
Watchers on the Wall [11]	Passive VLP	Median error 0.12 m
FieldLight [12]	Passive VLP	Median error 0.68 m to 1.2 m
D Yang et al. [14]	PIR	Mean error 0.21 m
B Yang et al. [15]	PIR	Mean error < 0.8 m
Liu et al. [16]	PIR	Mean error 0.47 m to 0.71 m
Tang et al. [42]	Passive EFS	Mean error 0.104 m to 0.272 m
P-Loc [43]	Passive EFS	Mean error 0.48 m
Zhao et al. [3]	Single pixel camera	Mean error 0.2 m
Tariq et al. [37]	Capacitive wall sensors	Mean error 0.307 m
Chen et al. [18]	Thermophile	RMSE of 0.19 m
Qu et al. [17]	Thermophile	Mean error 0.07 m
Zhang et al. [8]	Wireless CSI	Mean error 0.8 m
Shi et al. [9]	Wireless CSI	Mean error 0.63 m
SpringLoc [5]	Wireless RSSI	Median error 0.6 m to 1.57 m
CapLoc (This paper)	Capacitive	Median error 0.026 m, 90-percentile error 0.066 m

relative motion between the sensors and a target. Therefore, they are unable to deal with a stationary target. IR-sensing based positioning using thermopile sensors has been proposed [17, 18] to deal with both stationary and mobile targets. Unfortunately, such techniques are inherently vulnerable to changes in heat signatures resulting from clothing variations.

Humans spend much of their time in contact with the floor when they are inside a building. Therefore, the floor can be potentially repurposed as a large sensor for device free positioning of individuals. Table 1 summarizes the key works in the area of the floor-based DFL.

Pressure-sensitive floors [19-21] have been used for locating and identifying people. There are also systems using binary pressure-sensitive switches built into the floor [22]. Unfortunately, the pressure-sensitive floors appear to be complex to build. Besides, the pressure sensors (e.g., load cells) are also subject to wear and tear degradation, especially of the mechanical components.

Floor-based localization can also be accomplished by measuring footsteps-induced vibrations with a network of seismic sensors [23-25]. The footsteps (and hence the target) are located by exploiting the fact that vibration signals take different times to reach each sensor depending on the distance between the footstep and the sensors. This allows performing the localization using Time of Arrival (ToA) or Time Difference of Arrival (TDoA) techniques [25]. However, the floor is a complex heterogeneous medium. It varies significantly from one building to another. This makes the

calibration challenging thus complicating the transfer of a relevant system between different premises.

Capacitive sensing utilizing the change in capacitive coupling between a custom-designed floor and target can be an effective localization method. In this scenario, the floor and the target form (two plates of) a capacitor. The presence of the target alters the electric field, actively generated by a transmitter, manifesting as a measurable change in the capacitance. Smart Carpet [26] uses fabric into which conductive wires are sewn in serpentine patterns to form 0.15 m × 0.15 m panels. Similarly, SensFloor [27-29] uses conductive triangles embedded into a textile. Capacitive floor with metal squares was utilized in [30, 31]. CapFloor [32] uses two sets of parallel wires orthogonal to each other. A person walking above them changes the measured capacitance in these wires. Since a person is above at least one wire in each direction, an intersection point of these wires presents the person's estimated position.

In contrast to the aforementioned works that use the loading mode of the capacitive sensing [33], TileTrack [34, 35] employs the transmit mode. A square wave signal transmitted from the floor tiles is received by an additional electrode placed in the room as a receiver. The detected change in the signal amplitude caused by a person between the electrode and the floor tile helps infer the location. Capacitive sensing is also utilized in research [36, 37] where instead of using the floor-based solution, electrodes are set up on the walls.

When a person walks on a typical floor, a charge is built up due to the *Triboelectric Effect* [38]. The person can also be
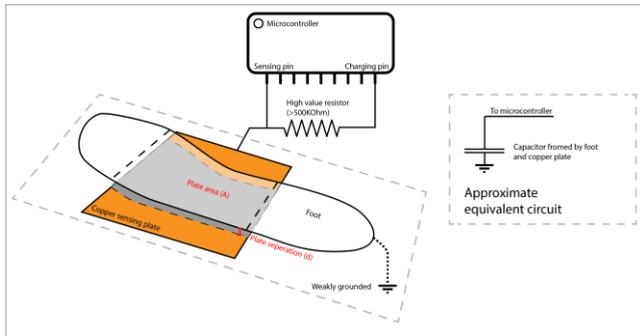
FIGURE 1. Loading mode capacitor formed by subject's foot on CapLoc, along with a simplified circuit diagram. *A* is the overlapping area of the two plates, and *d* is the separation between the two plates.

FIGURE 2. The structure of the floor. The floor can be topped with any non-conductive flooring material such as wood, vinyl or carpet.

considered as being an earthed conductor. Therefore, the ambient electric field created by the radiation from the AC powerlines (ever-present in buildings) is altered by the presence of a human target. This change can be measured with *Electric Potential Sensors* (EPS) and used for both identification of subjects [39] as well positioning of them [40-43]. Unfortunately, such opportunistic, passive electric sensing is vulnerable to ambient electrical field noise and interference. The relevant systems are mainly implemented using EPS units that are placed on the walls or ceiling of a room.

This paper proposes a new capacitive floor system named CapLoc for passive positioning. In a preliminary work [44], the authors presented how a static foot can be detected when a subject stands barefoot on a capacitive sensing panel. This paper utilizes that concept to develop CapLoc, a positioning system, for real time localization of a moving target accurately and potentially detect fall in an automated manner. It presents the following original contributions:

1. CapLoc can determine the position of a mobile target in real-time. It is not data-driven and therefore, requires minimal calibration for localization making it more invariant to changes in the setting. CapLoc is also robust and not vulnerable to factors that adversely affects other DFL systems like wireless multipath propagation (affects wireless DFL), illumination condition (impacts camera and passive VLP systems), clothing worn by the target (affects IR-based systems) etc.;

2. The experimental results showing the localization of a mobile target for multiple trajectories are presented. The median and 90 percentile localization errors while testing with three different subjects are found to be 0.026 m and 0.066 m, respectively. This makes CapLoc more accurate than most passive localization systems reported in the literature (see Tables 1 & 2). Also, the majority of the reported DFL systems were only tested for a handful of target trajectories. In contrast, CapLoc was tested for 39 different paths walked by multiple subjects. An accurate ground truth recording system was implemented using virtual reality technology (HTC Vive [45]) to ensure that the localization error is accurately measured. By utilizing

the procedure outlined in the article, other researchers will be able to record accurate ground truth in an automated manner, using an affordable consumer grade technology;

3. It is shown that the poses of a person lying on the floor can be captured easily. Potentially, this can be used for automated fall detection in a non-obtrusive manner.

The rest of the paper is organized as follows. Section II discusses the development of the CapLoc system. Section III presents the footstep detection process. Section IV demonstrates the localization performance. Pose capture for potential fall detection is shown in Section V. Section VI concludes the paper and discusses future research directions.

II. SYSTEM DEVELOPMENT
A. KEY CONCEPT

CapLoc is based on the formation and the subsequent sensing of loading mode capacitance [33, 46]. The concept is shown in Fig. 1 where the subject's foot and copper-foil tiles underneath the floor form the two plates of the capacitor. This capacitor can be modeled as:

$$C = \varepsilon \frac{A}{d} \, , \tag{1}$$

where C is the total capacitance, ε is the permittivity of the dielectric (assumed to be constant), A is the overlapping area of the two plates, and d is the separation between the two plates (details shown in Fig. 1). When the subject stands with a foot above the transmitting plate, the capacitance depends on two main factors: the proportion of the plate covered by the subject's foot (A), and the distance between the subject's foot and the plate (d). For a rigid floor type, the distance d remains fairly constant, whereas the area A changes as sensors could naturally be covered to a different extent.

B. PROTOTYPE HARDWARE DESIGN

A 0.6 m × 0.6 m sensing panel, with 25 individual copper-foil squares, is the basic building block of the CapLoc floor (Fig. 2 and 3). Each copper square is soldered to a wire that is connected along with 24 other wires to a microcontroller (100-pin ARM Cortex M3 [47]) where the capacitance is measured. The wires are routed within the gaps between the copper squares. The total component cost of a 0.6 m × 0.6 m sensing panel (excluding the cost of floorboards) is approximately $6.
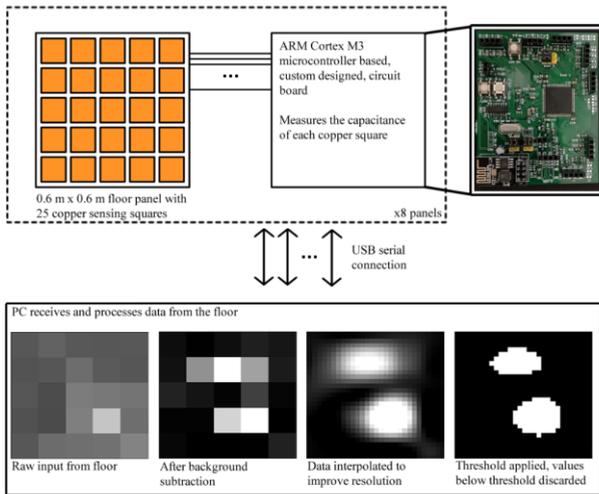
FIGURE 3. Block diagram of the CapFloor system architecture, with the custom designed hardware sampling the capacitance values which are sent to the PC app for the foot detection process. The foot detection is performed by adopting image processing techniques.

Therefore, the cost of implementing CapLoc, excluding labour, is less than $18/sqm while offering significant functionality. Also, the cost of the system is expected to decrease significantly with mass manufacture.

The capacitance is measured by evaluating the RC time constant of the equivalent capacitive circuit. The time taken to charge a capacitor to a set voltage V_0 is given by the well-known RC charging equation:

$$V(t) = V_0(1 - e^{-t/\tau}) \qquad (2)$$

where $\tau = RC$.

If the selected resistance value R is sufficiently high, it can be assumed to be reasonably constant and independent of the unknown resistance to the ground. Time taken by the capacitor to charge to a set value, therefore, depends solely on the capacitance. A microcontroller is used to charge the copper plate through a high value resistor by applying a voltage to the charging pin (Fig. 1). The time taken to reach a set voltage at the sensing pin is measured. When a subject's foot is near the copper plate, the effective capacitance is much greater than when there is no subject nearby. This leads to a significantly longer rise time of the signal. The raw capacitance measurements are sent from the microcontroller to an application running on PC over the USB serial communication
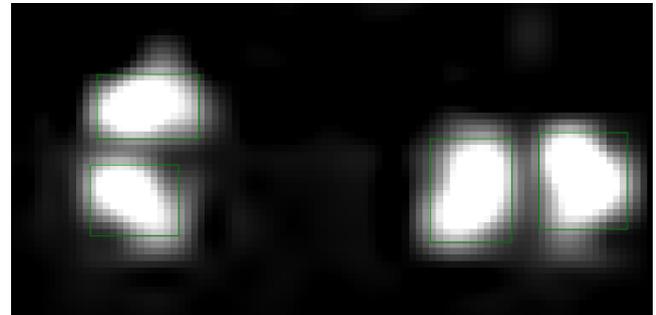
FIGURE 4. The simultaneous detection of multiple feet from multiple subjects (interpolated, before thresholding).
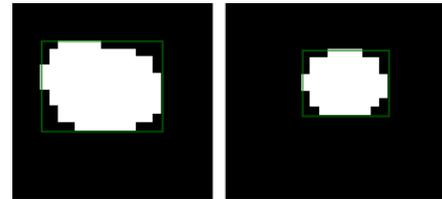
FIGURE 5. Foot after thresholding in socks (left) and in thick soled footwear (right). In thick footwear the foot is smaller in area after thresholding.

line. The PC app processes and displays the incoming data in real-time as well as saves the data for further analysis. The footstep detection algorithm takes less than 2 ms to run on a standard desktop PC running at 3.2 GHz. Trace drawing on the screen takes around 15-20 ms. The floor is sampled at around 10 Hz, giving the app plenty of time to process each frame whilst waiting for the next data frame from the sensors.

C. FOOT DETECTION

Figure 3 illustrates the foot detection process that is effectively an image processing algorithm where each capacitance value from the floor is represented as a single grayscale pixel. When CapLoc is first powered on, a number (currently set to 10 after many rounds of empirical testing) of capacitance readings are taken from the floor sensors as a background estimation. It is then subtracted from each subsequent capacitance measurement from the floor. Over time the background estimations can drift. To counteract this phenomenon, periodic CapLoc recalibration can be implemented by taking a new set of baseline capacitance readings when the floor is known to be vacant. Over a long period, the amount of time when a subject is standing on a

FIGURE 6. A sequence of footprints superimposed in time. Both pre (top) and post (bottom) thresholding. Estimated center of the foot marked with a cross.

(a) **Layout diagram of experimental setup**

(b) **Photo of experimental setup**

FIGURE 7. Layout of the floor and Vive calibration points.

square is small compared to that when the subject is not standing on it. Therefore, an alternate method is to take a long-term average of all capacitance readings taken whilst the system is in use and employ this long-term average as the baseline.

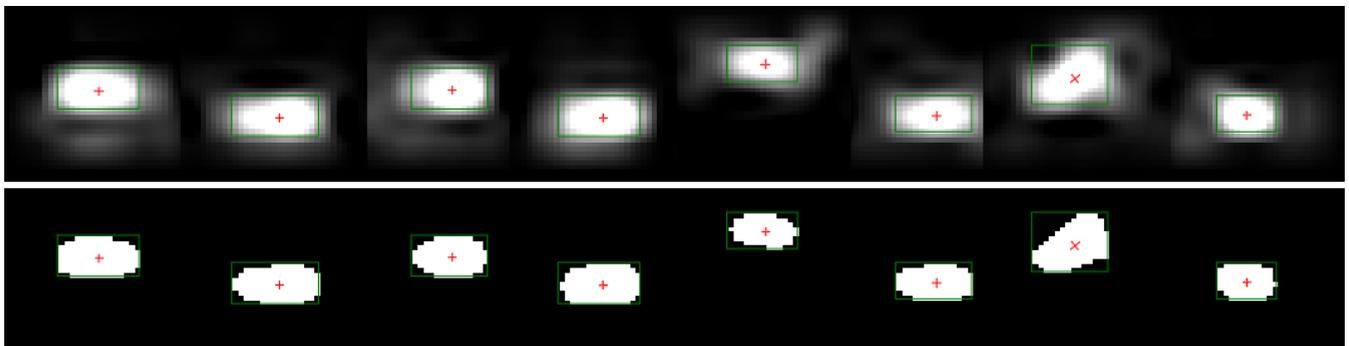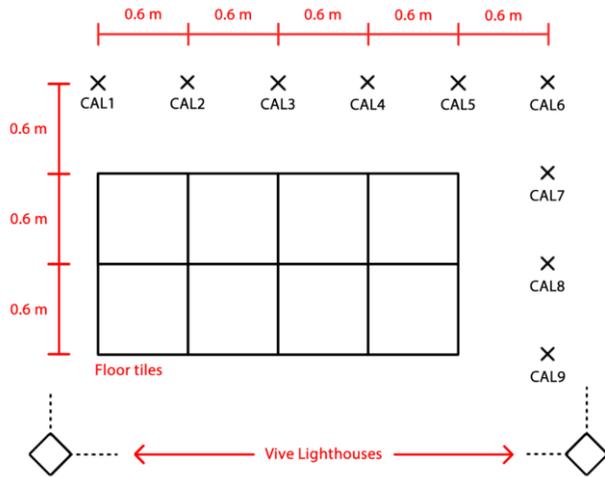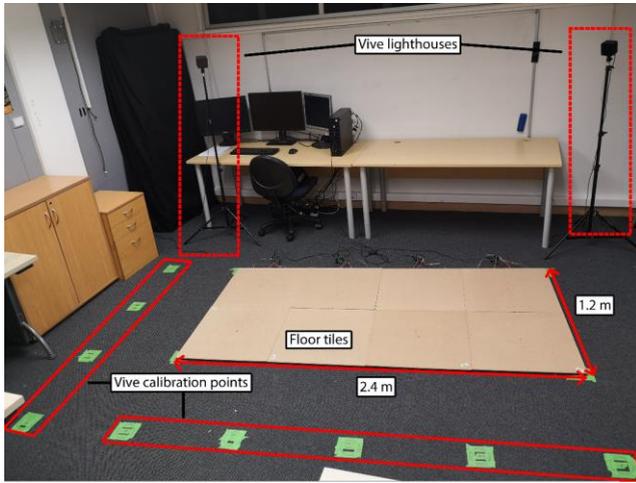In terms of image processing, the measured capacitance values form a very low-resolution image. Interpolation is applied to improve its quality. Several interpolation algorithms were tried. Cubic interpolation showed the best performance while enhancing 2×2 images to 7×7 interpolated ones.

A threshold is then applied to the data such that any capacitance values below the threshold are set to become "0" while those above the threshold are set to be "1". Once it is done, blob detection through connected component analysis [48] is applied whereby all connected squares are considered to be a part of the blob or cluster. Each blob corresponding to a single footprint can then be represented by a matrix **M** of 2 $\times$ N dimension, where N is the number of data points in the cluster. Each column of the matrix is a vector representing the position of a single data point in the cluster.

The center of the footprint $(\bar{x}, \bar{y})$ is estimated by averaging the position of each point in the $2 \times N$ cluster matrix **M** as:

$$
\begin{aligned}
\bar{x} &= \frac{\sum_{i=1}^{N} M_{1,i}}{N} \\
\bar{y} &= \frac{\sum_{i=1}^{N} M_{2,i}}{N}
\end{aligned}
\tag{3}
$$

The system was tested with multiple subjects. It detected the feet of several subjects concurrently given that they were sufficiently spaced apart. Fig. 4 shows two subjects' feet being detected individually. It was observed that feet on adjacent squares might be non-detectable as they merged into a larger blob. The copper sensing squares are spaced at 120 mm intervals thus providing that the feet separation is to be greater than around 200 mm to avoid the aliasing. This is because the partial occlusion of feet at the very edge of adjacent squares does not put them over the threshold. Initial testing, as reported in [44], found that the position of the subject's foot in a static situation could be measured accurately.

When the target is barefoot, the separation between the target's foot and the copper-foil (d of Fig. 1 and Equation 1) is the smallest. This results in a larger value of the capacitance compared to the case when a subject is wearing a footwear. Therefore, CapLoc enjoys the highest SNR when the subject is barefoot which is quite common in a home setting. The impact of footwear type on foot detection accuracy was thus investigated. It was found that the type of footwear had quite a minimal effect. Figure 5 demonstrates the cases where a subject stands on the floor wearing socks and a pair sneakers with thick soles. Whilst one can see the image for the foot in the sneaker is slightly smaller (due to it being further from the sensing squares), it is still detectable with its position being relatively unaffected.

III. FOOTSTEP LOCALIZATION

A test floor was set up using eight sensing panels to create 1.2 m × 2.4 m area. Data from the system were sampled at 10 Hz making it possible to track a person moving around the floor. Firstly, individual footprints were detected, and the center of each footprint was stored. The footprint centers were then clustered in time and space to determine if they come from the same footstep. The path of the subject was then estimated by taking the midpoints of the successive footsteps. Figure 6 shows the detected footprints from a subject walking on CapLoc (in 0.6 m × 4.8 m configuration).

Implementation of an accurate ground truth system to compare the estimated path with the actual one is a challenging task. Several approaches were reported in the literature. While motion capture can provide an extremely accurate ground truth [14], it is not cost-effective. The use of the Xbox Kinect was reported in the study [40]. A custom-designed solution was reported in [31] whereby a hat on the subject's head was connected via wires to pulleys with attached encoders.

In this work, the HTC Vive [45] was used as a ground truth system due to its low cost, availability, and sufficient
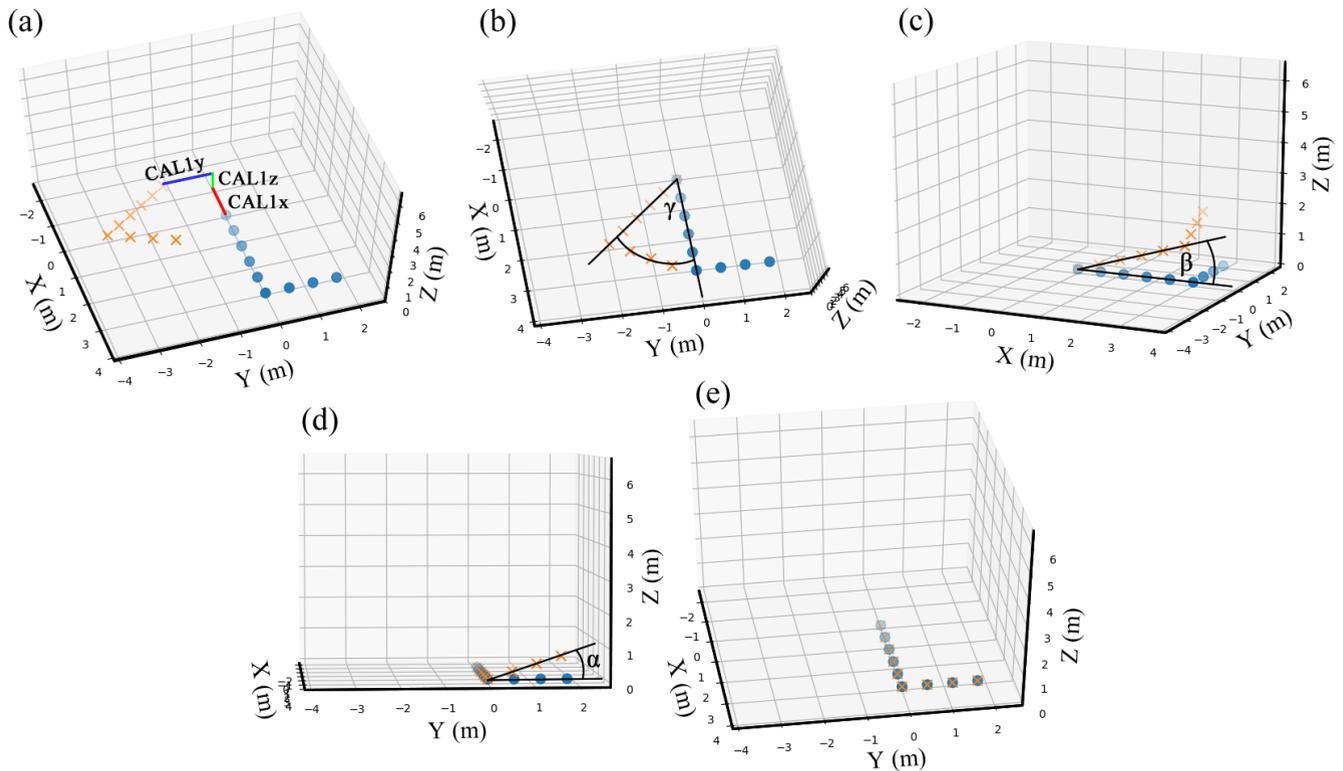
(a)

(b)

(c)

(d)

(e)

FIGURE 8. The process of aligning the Vive's calibration points with the floor. The orange crosses represent the calibration points in the Vive's frame of reference, the blue circles in the floor's frame of reference. (a): Translating the points so that the origins are aligned. (b): Rotating about the Z axis; (c): Rotating about the Y axis; (d): rotating about the X axis; (e): the final outcome with the two sets of points aligned. Note that, the angles (especially α and β) have been exaggerated for clarity. In reality, the translation and the rotation γ were usually enough for the ICP algorithm to align the points correctly.

FIGURE 9. The Vive tracker affixed atop a subject's head.

accuracy. It uses two base stations (called Lighthouses) to track a small device called Tracker.

In pre-experimental testing, the accuracy of the system was evaluated using an x-y CNC plotter with max deviation of 0.025 mm. Vive was found to be accurate within 10 mm. The positions reported by the Vive are relative to the primary lighthouse. To reconcile this coordinate system to that of the floor, a calibration process needs to be undertaken. This also means that positions of the lighthouses do not need to be carefully measured thus eliminating a potential source of error.

First, the ground truth system was calibrated using nine points around the edge of the floor (Fig. 7). The calibration points were used to align the Vive's reference plane with the floor as well as to align point CAL1 with the origin of the

floor. The calibration points were used to generate a transformation matrix ($\mathbf{R}$) that was then applied to all positions measured using the Vive.

$$\mathbf{x'} = \mathbf{R}\mathbf{x}, \tag{4}$$

Where [49]

$$\mathbf{R} = \mathbf{T_v} \cdot \mathbf{R_z} \cdot \mathbf{R_y} \cdot \mathbf{R_x} \tag{5}$$

and

$$\mathbf{T_v} = \begin{bmatrix} 1 & 0 & 0 & -CAL1_x \\ 0 & 1 & 0 & -CAL1_y \\ 0 & 0 & 1 & -CAL1_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{6}$$

$$\mathbf{R_z}(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 & 0 \\ \sin\gamma & \cos\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{7}$$

$$\mathbf{R_y}(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{8}$$

$$\mathbf{R_x}(\alpha) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{9}$$

Here $\mathbf{x}$ is a position from the Vive to be transformed, and $\mathbf{x'}$ is the transformed position relative to the floor. The values α, β,
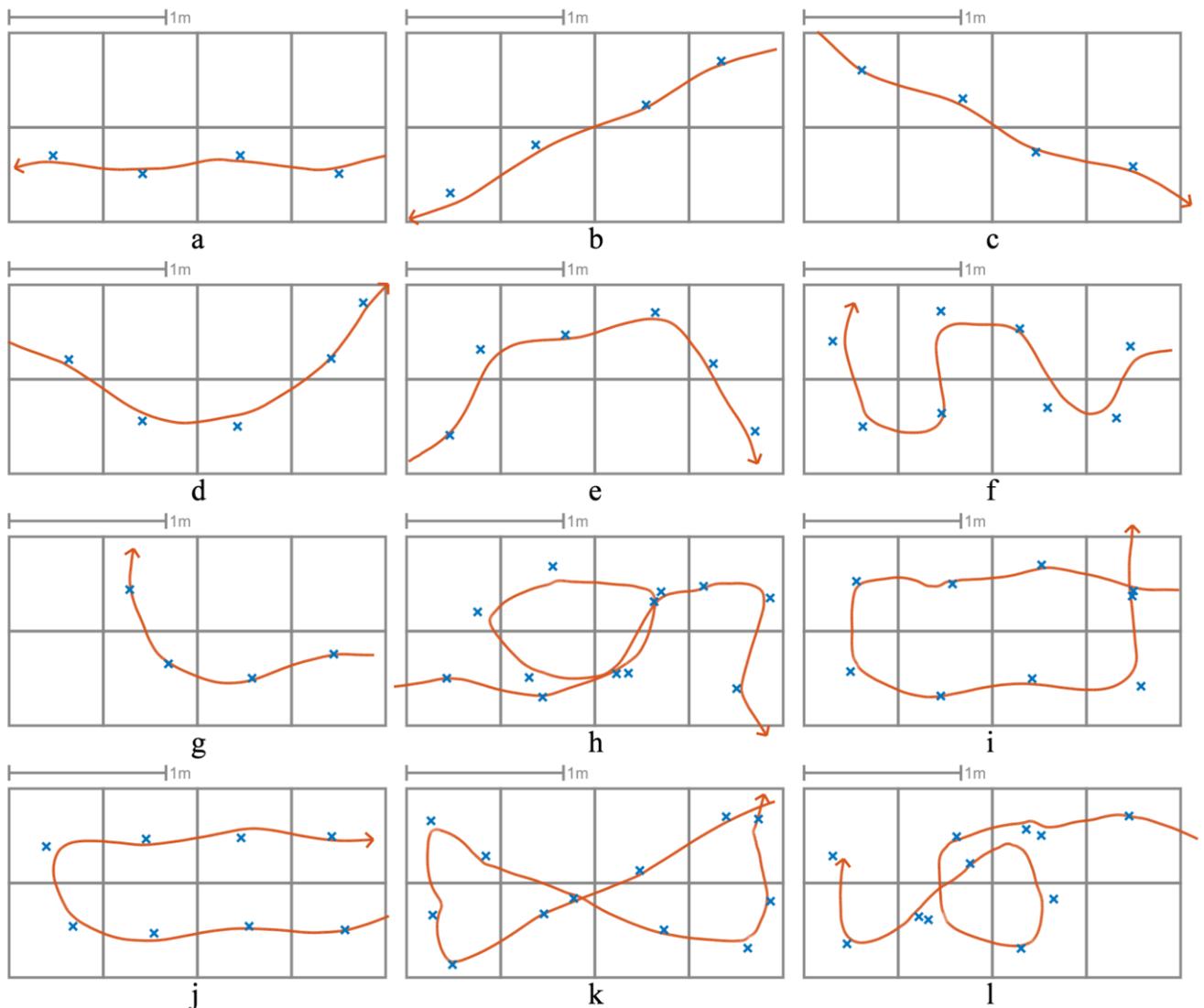
FIGURE 10. Twelve paths walked by Subject 1 on CapLoc: crosses represent the estimated foot positions and the lines show the ground truth (Vive tracker).
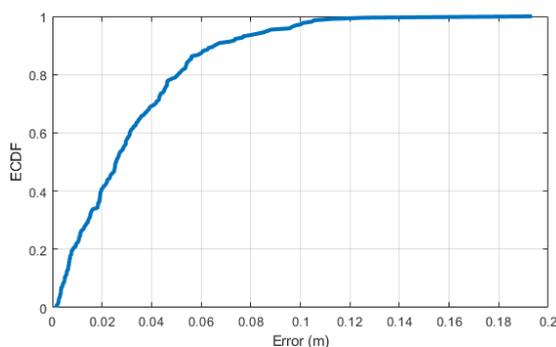
FIGURE 11. ECDF of localization error for 219 footsteps across 39 different paths. The median error was found to be 0.026 m and the 90 percentile error 0.066 m.

and γ are the *pitch, yaw,* and *roll* between the Vive's reference plane and the floor. Figure 8 illustrates the aforementioned process.

It was then further refined by employing the *Iterative Closest Point* (ICP) algorithm [50] to generate a

transformation matrix aiming to minimize the Euclidean error between the measured and actual positions of all nine calibration points. The combination of the two transformations was then used to transform the position data from the Vive.

Literature reports [51-53] suggest that tracking could be lost when a line of sight is absent between the lighthouses or between the tracker and the lighthouses. The tracker was therefore attached to the top of the subject's head (Fig. 9) to maintain a constant line of sight with the two lighthouses that were mounted at approximately 2 m above the ground, one on each side of the testbed.

Thirty-nine different paths, split between three subjects - two males (subjects 1 and 2) and one female (subject 3), were walked across CapLoc with the ground truth of the subject's head being recorded by the Vive. Fig. 10 shows the footsteps estimated by CapLoc and the position of the subject's head tracked by the Vive for 12 of the total 39 paths. It can be seen that the footsteps very closely match the ground truth. Localization errors were computed by considering the position
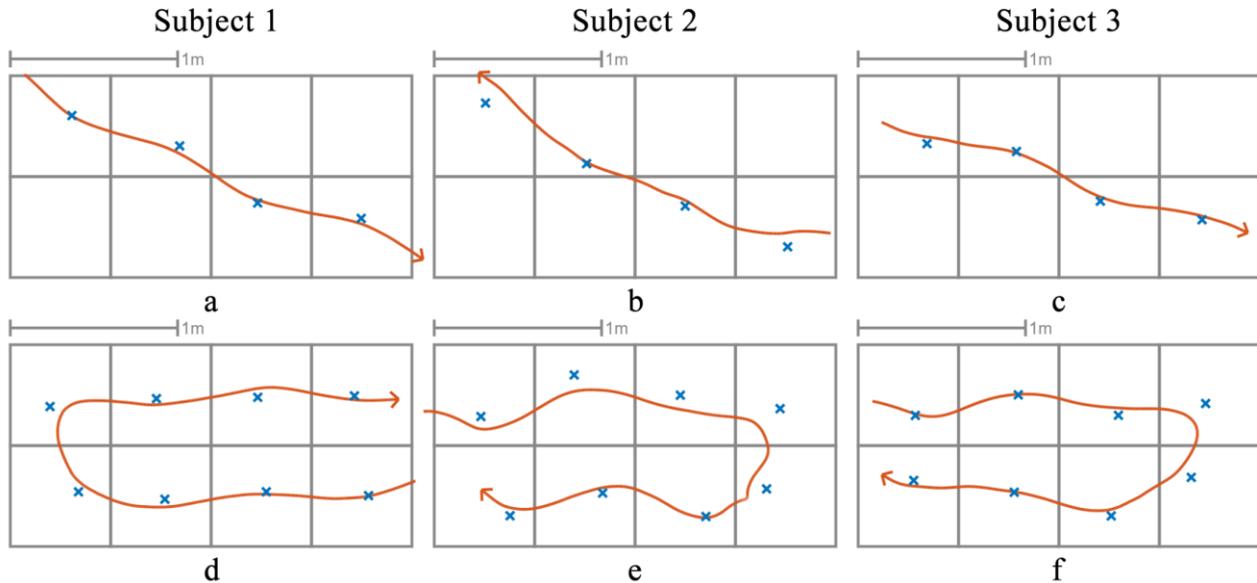
FIGURE 12. Paths walked by different subjects without the need for calibration in between.
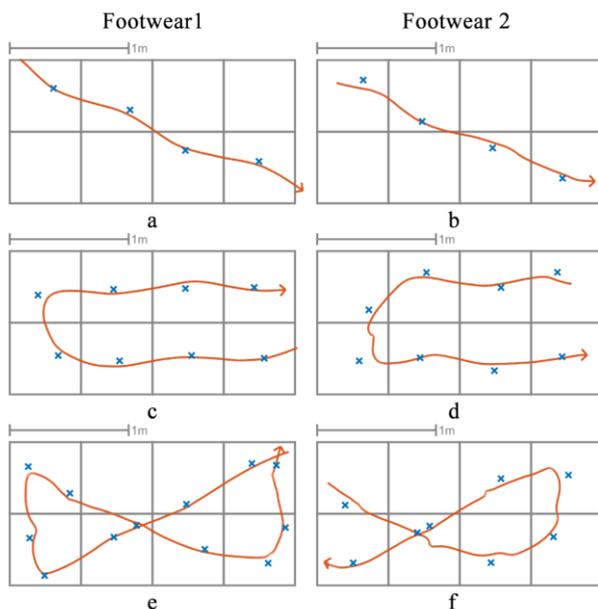
FIGURE 13. Paths walked in thin (1) and thick-soled footwear (2).

TABLE 3. Comparison of path tracking error for different subjects.

	Median (m)	90 Percentile (m)
Subject 1	0.025	0.056
Subject 2	0.039	0.097
Subject 3	0.026	0.069
Subject 1 - thick shoes	0.031	0.082

was done to verify that the floor was able to locate different subjects without the need for calibration in between. Figure 12 shows two paths for each subject. Table 3 shows the median and 90 percentile errors for each of the subjects.

Five of the paths were walked by subject 1 in a pair of sneakers having a thick sole. Other than that, the three subjects had similar footwear, considerably thinner than the sneaker. The results are shown in Fig. 13 and Table 3. For subject 1, the median and 90 percentile errors are slightly worse for the thick-soled footwear as the measured capacitance was lower (due to higher separation from copper-foil plates, please see Section II for more details), and therefore it was more affected by noise.

The results support the assertion that the floor can be used for human tracking without any foreknowledge of the subject or environment. The only requirement being that the floor must be vacant for several seconds after the initial powering on to measure the background capacitance.

Potentially, the error could be further reduced by employing a more sophisticated path estimation algorithm. Also, accurate tracking is complicated by the impossibility to define the subject (person) as a single point object. The top of the head is approximately in the center of the subject when viewed from a top-down perspective. However, when people walk, they tend to sway from side to side. This was noticed to be even more prevalent when a subject walked along pre-marked paths. Besides, the amount of the head movements is normally somewhat higher than that of the center of mass of the body, thus causing additional errors. This can be seen in the paths and error statistics for Subject 2, which are worse than those for the other two subjects. The U-shaped path in particular shows this subjects' propensity to move their heads as they walk. The head movements resulting from the subject's walking pattern may have as much or even more effect

of the subject to be the midpoints between the successive footprints and then comparing them to the relevant points of the Vive's reported path. *Empirical Cumulative Distribution Function* (ECDF) for the 219 footsteps corresponding to all 39 trajectories is shown in Fig. 11.

Both U-shaped and diagonal trajectories were walked by all three subjects, due to those being easily repeatable paths. That
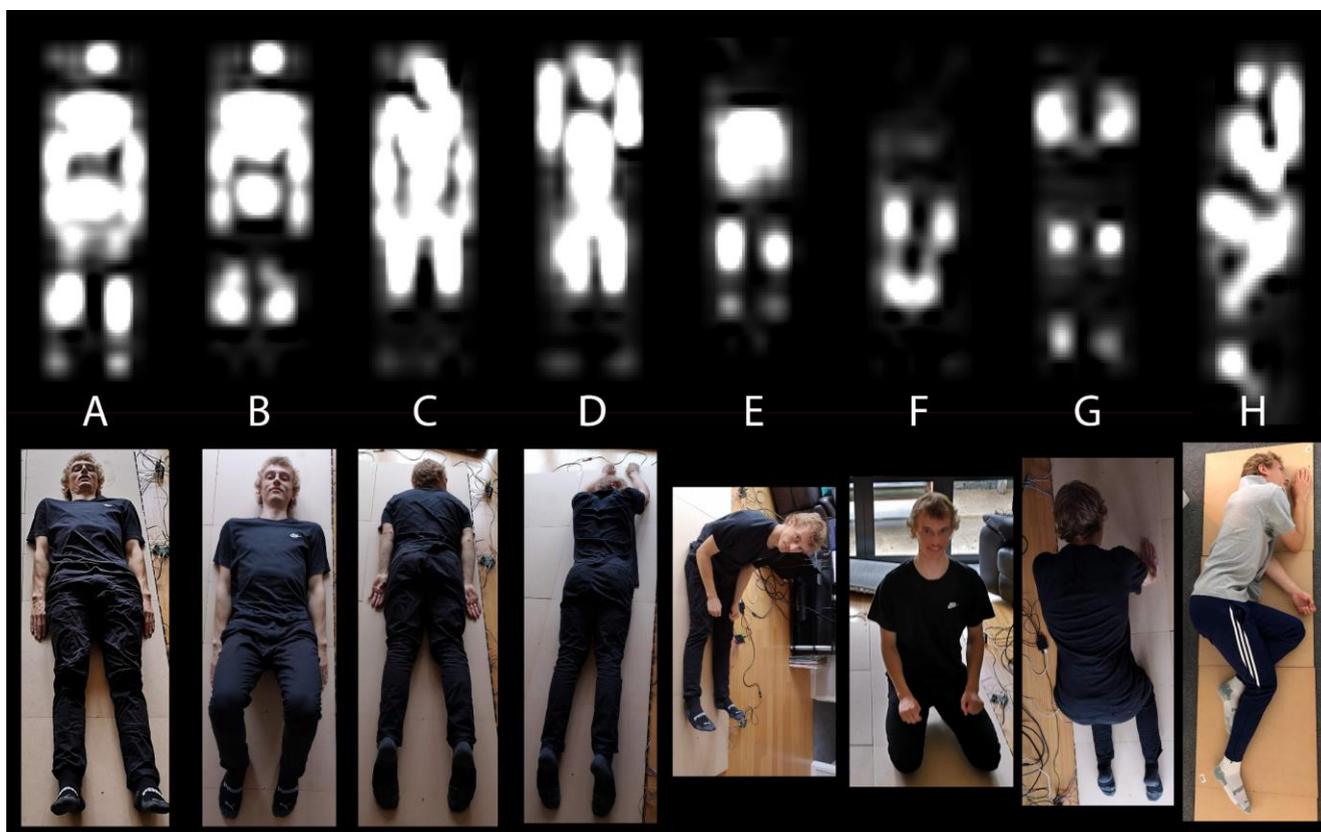
FIGURE 14. A subject in a variety of poses upon the floor.

compared to the thickness of the footwear. As can be observed, the localization error for subject 2 with thinner footwear is higher than that for subject 1 with thick shoes.

Tables 1 & 2 compare the localization accuracy of CapLoc against the state-of-the-art floor-based and other DFL systems. As can be seen, the proposed system is more accurate than other systems reported in the literature. CapLoc's accuracy is likely to be even higher than that which is being reported if the ground truth of the foot could be more reliably established. The problem with placing the tracker on the foot is that it can lose line of sight with the light houses. In such a scenario, the ground truth recording system loses calibration (as discussed earlier), reporting incorrect positions. Therefore, a practical compromise was made. It should be noted that, if a person is not in contact with the floor, they are not visible to CapLoc. However, in a real-life setting, people can only enter and exit a room at defined points. They can be tracked around the room and if they remove themselves from contact with the floor (e.g. by sitting on a chair) they can be assumed to be in that location until they are seen again (i.e. they stand up from the chair).

IV. POSES CAPTURED BY CAPLOC FOR FALL DETECTION

Fall is a major health risk for the elderly, negatively affecting their health and quality of lives. It poses also significant burden on the healthcare and elderly-care institutions. For someone living alone, timely and accurate fall detection is needed to initiate swift medical assistance.

Personal Alarm System (PAS) can be worn by an elderly person. In case of any problems (e.g., a fall), it enables the alarm activation by just pressing a button. Unfortunately, if the victim loses consciousness or is in a confused or panicked state, the button may not be pressed [54].

Wearable sensors, utilizing primarily accelerometers (e.g., presented in [55]) have been proposed for automated fall detection. However, they rely on the subject to wear a sensor at all times. Such a wearable device can be forgotten or misplaced or get damaged. It also requires charging or battery replacement that again can be missed. There may also be a reluctance from a person to wear the sensor. Smartphone-based fall detectors (e.g., discussed in [56]) are also associated with similar issues. Camera- [57] and sound- [58] based fall detection approaches are perceived to be invasive to privacy. Wireless- [59] and IR- [60] based systems rely on anomalous activity detection. They utilize the signatures for a fall that are not immediately obvious to the naked eye [61]. Large amounts of data are generally required to train a model to detect falls. However, the falls are rare events. Besides, it is very difficult to simulate them with human participants. All of this makes it hard to collect enough data to train a robust classifier for fall detection [62].

When using CapLoc, a simple and more naive algorithm potentially could be used for fall detection. For example, a sudden increase in the area of contact with the floor could suggest that a person has gone from a standing to a prone position. By combining it with pose capture and temporal

changes in the pose, it could be possible to detect an event such as a fall. Rather than trying to detect a rare, anomalous event, CapLoc can support a fall detection approach identifying the immediate aftermath of the fall, i.e., the subject lying on the floor.

A. LYING SUBJECT POSE CAPTURE

An investigation was undertaken to determine if different poses can be observed by using CapLoc. A subject laid on the floor in eight different poses, with the system output being recorded. The following poses were tried (Fig. 14): A – the subject was lying face up with the arms by the sides and legs flat; B – the subject was in the same pose except with the knees were in the air and the feet whilst still on the floor wre close to the body; C – the subject was lying face down with arms by the sides; D – the subject was lying face down with arms stretched above the head; E – the subject was sitting upright with the legs outstretched in front; F – the subject was kneeling; G – the subject was crawling on the hands and knees; H – the subject was lying in the fetal position. It can be seen that the poses were captured reasonably distinctively by the CapLoc.

This suggests that once sufficient data are available, not only fall detection but also fall pose recognition could be achieved while employing relevant classification models (e.g. applying histogram distances [63]).

B. POSE AREA ESTIMATION

Parts of the foot detection algorithm can also be used to estimate the contact area of a subject with the floor. Each individual capacitance reading (represented as a single pixel) is subject to background subtraction, cubic interpolation, and, finally, binary thresholding as discussed before. Each pixel then represents an area of the floor defined by the size of each copper-foil sensor and the interpolation factor. The number of pixels above the threshold then approximates the area of the contact.

Each of the poses in Fig. 14 had their areas estimated by the system to demonstrate the CapLoc potential for fall detection. It can be seen from Table 4 that the poses of the lying on the floor have much larger contact areas compared to a footprint, thus supporting the suggestion that the floor contact area could potentially be used for fall detection.

Certain poses (e.g., G) could be confused for multiple sets of footprints. However, if fall detection is combined with occupancy tracking, it could distinguish the fall from the case of three people standing near each other. People only enter and exit the room at defined points and hence they can be tracked around the room with reasonable accuracy. Therefore, if there is only one person in a room (or in a certain area of it), and an image of a potentially dangerous pose arrives, the system would be able to trigger the fall alarm. A body on the floor will have a significantly larger estimated contact area than a footprint regardless of the size of the body. An abrupt increase in area suggests that a fall may have occurred. Therefore, the

TABLE 4. Comparison of the area of different poses.

Pose	Area (m²)
A – lying on back	0.64
B – lying on back with knees up	0.54
C – lying on front, hands by side	0.64
D – lying on front, hands above head	0.59
E – Sitting with legs outstretched	0.28
F – Kneeling	0.16
G – Crawling on hands and knees	0.19
H – lying in fetal position	0.58
Single foot area	0.05

difference in body size should not impact the fall detection performance. Also, with large amount of data collected for people of varying body size, sophisticated image recognition techniques (e.g. a deep neural network classifier [64]) could be used in the future to recognize a fall event rather than just using the contact area.

V. CONCLUSIONS AND FUTURE WORK

The developed capacitive floor, CapLoc, can identify the position of a subject's feet and track a single individual while walking upon it. The median and 90 percentile error of CapLoc for a wide-range of trajectories were found to be 0.026 m and 0.066 m. The sample rate used by the prototype hardware was at 10 Hz per individual copper square. A new version of the hardwar e is currently undergoing development. It will offer higher sensitivity and a much-improved sample rate whilst still being compatible with the current flooring tiles as well as signal and data processing techniques. Further work will also help to reduce the stray capacitance by potentially using shielded cabling and to improve the background capacitance measurement.

The localization experiments were performed with a single person on the floor. However, it was demonstrated that the system was capable of detecting multiple targets simultaneously. For ambient signal based DFL techniques (e.g. wireless or IR), each subject adds interference and lowers the SNR leading to poor performance. In contrast, subjects on CapLoc that are spatially separated do not interfere with each other. Therefore, by dividing the floor into sperate smaller areas, it is possible to track targets within those spaces using the algorithm outlined in this paper. It can be further improved by incorporating a particle filter or some similar techniques. However, tracking multiple targets in a crossover scenario, where targets come together and then diverge, will require user identification. It was found that CapLoc systematically overestimates the foot area. However, such overestimation occurs uniformly around the foot perimeters. As such, it did not affect the position of the center of the foot. Unfortunately, the overestimation phenomena means that it would not be achievable at this stage to accurately identify individuals based on their estimated footprint area. However, it is possible to discern the different phases of a subject's footstep on CapLoc from the initial heel strike, through the midstance to the toe-off. During this sequence of events, the center of contact of the foot moves from the heel to the toe. With the improved hardware, in combination with other features (e.g., stride

CapLoc

33

length and foot angle) future work will also explore the identification of individuals using their gait patterns. In order to achieve this, significant data needs to be collected to train a machine learning algorithm [65].

Only flat footwear was employed during the experimental investigations while showing good results. Future investigations will also include performance evaluation of the proposed technique on a variety of footwear types (e.g., footwear with raised heels).

Finally, poses of a subject lying on the floor subject can be clearly captured for a variety of positions. Therefore, the proposed technique has the potential to be applied to develop an accurate yet noninvasive fall detection system. Future work will involve collecting sufficient pose data from multiple subjects of varying body size. These data can then be used to train a classifier to detect poses and subsequently identify the fall occurrence.

REFERENCES

[1] R. Poppe, "A survey on vision-based human action recognition," *Image and Vision Computing,* vol. 28, no. 6, pp. 976-990, 2010/06 2010, doi: 10.1016/j.imavis.2009.11.014.

[2] D. Roeper, J. Chen, J. Konrad, and P. Ishwar, "Privacy-preserving, indoor occupant localization using a network of single-pixel sensors," in *2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2016: IEEE, pp. 214-220.

[3] J. Zhao, N. Frumkin, P. Ishwar, and J. Konrad, "CNN-Based Indoor Occupant Localization via Active Scene Illumination," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019: IEEE, pp. 2636-2640.

[4] S. Palipana, B. Pietropaoli, and D. Pesch, "Recent advances in RF-based passive device-free localisation for indoor applications," *Ad Hoc Networks,* vol. 64, pp. 80-98, 2017/09 2017, doi: 10.1016/j.adhoc.2017.06.007.

[5] D. Konings, F. Alam, F. Noble, and E. M. Lai, "SpringLoc: A device-free localization technique for indoor positioning and tracking using adaptive RSSI spring relaxation," *Ieee Access,* vol. 7, pp. 56960-56973, 2019.

[6] M. Zhou, Y. Wang, Z. Tian, Y. Lian, Y. Wang, and B. Wang, "Calibrated data simplification for energy-efficient location sensing in internet of things," *IEEE Internet of Things Journal,* vol. 6, no. 4, pp. 6125-6133, 2018.

[7] Z. Yang, Z. Zhou, and Y. Liu, "From RSSI to CSI," *ACM Computing Surveys,* vol. 46, no. 2, pp. 1-32, 2013/11 2013, doi: 10.1145/2543581.2543592.

[8] Y. Zhang, C. Qu, and Y. Wang, "An Indoor Positioning Method Based on CSI by Using Features Optimization Mechanism With LSTM," *IEEE Sensors Journal,* vol. 20, no. 9, pp. 4868-4878, 2020.

[9] S. Shi, S. Sigg, L. Chen, and Y. Ji, "Accurate location tracking from CSI-based passive device-free probabilistic fingerprinting," *IEEE Transactions on Vehicular Technology,* vol. 67, no. 6, pp. 5217-5230, 2018.

[10] X. Yang, R. Cao, M. Zhou, and L. Xie, "Temporal-Frequency Attention-Based Human Activity Recognition Using Commercial WiFi Devices," *IEEE Access,* vol. 8, pp. 137758-137769, 2020.

[11] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Watchers on the Wall: Passive Visible Light-Based Positioning and Tracking With Embedded Light-Sensors on the Wall," *IEEE Transactions on Instrumentation and Measurement,* vol. 69, no. 5, pp. 2522-2532, 2020/05 2020, doi: 10.1109/tim.2019.2953373.

[12] D. Konings, N. Faulkner, F. Alam, E. M.-K. Lai, and S. Demidenko, "FieldLight: Device-Free Indoor Human Localization Using Passive Visible Light Positioning and Artificial Potential Fields," *IEEE Sensors Journal,* vol. 20, no. 2, pp. 1054-1066, 2019.

[13] Q. Hao, F. Hu, and Y. Xiao, "Multiple human tracking and identification with wireless distributed pyroelectric sensor systems," *IEEE Systems Journal,* vol. 3, no. 4, pp. 428-439, 2009.

[14] D. Yang, B. Xu, K. Rao, and W. Sheng, "Passive infrared (PIR)-based indoor position tracking for smart homes using accessibility maps and a-star algorithm," *Sensors,* vol. 18, no. 2, p. 332, 2018.

[15] B. Yang, Q. Wei, and L. Yuan, "Location ambiguity resolution and tracking method of human targets in wireless infrared sensor network," *Infrared Physics & Technology,* vol. 96, pp. 174-183, 2019.

[16] X. Liu, T. Yang, S. Tang, P. Guo, and J. Niu, "From relative azimuth to absolute location: Pushing the limit of pir sensor based localization," in *Proceedings of the 26th Annual International Conference on Mobile Computing and Networking,* 2020, pp. 1-14.

[17] 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.

[18] 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.

[19] W.-H. Liau, C.-L. Wu, and L.-C. Fu, "Inhabitants tracking system in a cluttered home environment via floor load sensors," *IEEE Transactions on Automation Science and Engineering,* vol. 5, no. 1, pp. 10-20, 2008.

[20] M. Andries, O. Simonin, and F. Charpillet, "Localization of humans, objects, and robots interacting on load-sensing floors," *IEEE Sensors Journal,* vol. 16, no. 4, pp. 1026-1037, 2015.

[21] I. Al-Naimi and C. B. Wong, "Indoor human detection and tracking using advanced smart floor," in *2017 8th International Conference on Information and Communication Systems (ICICS)*, 2017: IEEE, pp. 34-39.

[22] T. Murakita, T. Ikeda, and H. Ishiguro, "Human tracking using floor sensors based on the Markov chain Monte Carlo method," presented at the Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004., 2004. [Online]. Available: http://dx.doi.org/10.1109/icpr.2004.1333922.

[23] M. Mirshekari, S. Pan, J. Fagert, E. M. Schooler, P. Zhang, and H. Y. Noh, "Occupant localization using footstep-induced structural vibration," *Mechanical Systems and Signal Processing,* vol. 112, pp. 77-97, 2018.

[24] S. e. Alajlouni and P. Tarazaga, "A new fast and calibration-free method for footstep impact localization in an instrumented floor," *Journal of Vibration and Control,* vol. 25, no. 10, pp. 1629-1638, 2019.

[25] J. D. Poston, R. M. Buehrer, and P. A. Tarazaga, "Indoor footstep localization from structural dynamics instrumentation," *Mechanical Systems and Signal Processing,* vol. 88, pp. 224-239, 2017.

[26] D. Savio and T. Ludwig, "Smart Carpet: A Footstep Tracking Interface," presented at the 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07), 2007/05, 2007. [Online]. Available: http://dx.doi.org/10.1109/ainaw.2007.338.

[27] C. Lauterbach, A. Steinhage, and A. Techmer, "Large-area wireless sensor system based on smart textiles," presented at the International Multi-Conference on Systems, Sygnals & Devices, 2012/03, 2012. [Online]. Available: http://dx.doi.org/10.1109/ssd.2012.6198101.

[28] A. Steinhage, C. Lauterbach, and A. Techmer, "Large-Area Wireless Sensor System for Ambient Assisted Living," presented at the AMA Conference Proceedings SENSOR 2013, 2013.

[29] M. Sousa, A. Techmer, A. Steinhage, C. Lauterbach, and P. Lukowicz, "Human tracking and identification using a sensitive floor and wearable accelerometers," presented at the 2013 IEEE International Conference on Pervasive Computing and Communications (PerCom), 2013/03, 2013. [Online]. Available: http://dx.doi.org/10.1109/percom.2013.6526728.

[30] H. Rimminen, M. Linnavuo, and R. Sepponen, "Human tracking using near field imaging," presented at the 2008 Second International Conference on Pervasive Computing Technologies for Healthcare, 2008/01, 2008. [Online]. Available: http://dx.doi.org/10.1109/pcthealth.2008.4571055.

[31] H. Rimminen, J. Lindström, and R. Sepponen, "Positioning accuracy and multi-target separation with a human tracking system using near

field imaging," *International Journal on Smart Sensing and Intelligent Systems,* vol. 2, no. 1, pp. 156-175, 2009.

[32] A. Braun, H. Heggen, and R. Wichert, "CapFloor–a flexible capacitive indoor localization system," in *International Competition on Evaluating AAL Systems through Competitive Benchmarking*, 2011: Springer, pp. 26-35.

[33] J. Smith, T. White, C. Dodge, J. Paradiso, N. Gershenfeld, and D. Allport, "Electric field sensing for graphical interfaces," *IEEE Computer Graphics and Applications,* vol. 18, no. 3, pp. 54-60, 1998, doi: 10.1109/38.674972.

[34] M. Valtonen, J. Maentausta, and J. Vanhala, "TileTrack: Capacitive human tracking using floor tiles," presented at the 2009 IEEE International Conference on Pervasive Computing and Communications, 2009/03, 2009. [Online]. Available: http://dx.doi.org/10.1109/percom.2009.4912749.

[35] M. Valtonen, T. Vuorela, L. Kaila, and J. Vanhala, "Capacitive indoor positioning and contact sensing for activity recognition in smart homes," *Journal of Ambient Intelligence and Smart Environments,* vol. 4, no. 4, pp. 305-334, 2012, doi: 10.3233/ais-2012-0158.

[36] P. Lindahl, A.-T. Avestruz, W. Thompson, E. George, B. R. Sennett, and S. B. Leeb, "A Transmitter—Receiver System for Long-Range Capacitive Sensing Applications," *IEEE Transactions on Instrumentation and Measurement,* vol. 65, no. 10, pp. 2412-2423, 2016/10 2016, doi: 10.1109/tim.2016.2575338.

[37] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural Networks for Indoor Human Activity Reconstructions," *IEEE Sensors Journal,* 2020.

[38] T. Ficker, "Electrification of human body by walking," *Journal of Electrostatics,* vol. 64, no. 1, pp. 10-16, 2006/01 2006, doi: 10.1016/j.elstat.2005.04.002.

[39] K. Kurita, "Human Identification from Walking Signal based on Measurement of Current Generated by Electrostatic Induction," *Kansei Engineering International Journal,* vol. 11, no. 4, pp. 183-189, 2012, doi: 10.5057/kei.11.183.

[40] T. Grosse-Puppendahl *et al.*, "Platypus: Indoor localization and identification through sensing of electric potential changes in human bodies," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, 2016, pp. 17-30.

[41] B. Fu, F. Kirchbuchner, J. von Wilmsdorff, T. Grosse-Puppendahl, A. Braun, and A. Kuijper, "Performing indoor localization with electric potential sensing," *Journal of Ambient Intelligence and Humanized Computing,* vol. 10, no. 2, pp. 731-746, 2018/06/05 2018, doi: 10.1007/s12652-018-0879-z.

[42] X. Tang and S. Mandal, "Indoor Occupancy Awareness and Localization Using Passive Electric Field Sensing," *IEEE Transactions on Instrumentation and Measurement,* vol. 68, no. 11, pp. 4535-4549, 2019/11 2019, doi: 10.1109/tim.2018.2890319.

[43] T. Zhou *et al.*, "P-Loc: a device-free indoor localization system utilizing building power-line network," in *Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers*, 2019, pp. 611-615.

[44] N. Faulkner, B. Parr, F. Alam, M. Legg, and S. Demidenko, "Device Free Localization with Capacitive Sensing Floor," presented at the 2020 IEEE Sensors Application Symposium (SAS), Kuala Lumpur, Malaysia, 9-11 March, 2020.

[45] A. Borrego, J. Latorre, M. Alcaniz, and R. Llorens, "Comparison of Oculus Rift and HTC Vive: feasibility for virtual reality-based exploration, navigation, exergaming, and rehabilitation," *Games for health journal,* vol. 7, no. 3, pp. 151-156, 2018.

[46] T. G. Zimmerman, J. R. Smith, J. A. Paradiso, D. Allport, and N. Gershenfeld, "Applying electric field sensing to human-computer interfaces," presented at the Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '95, 1995. [Online]. Available: http://dx.doi.org/10.1145/223904.223940.

[47] ST Microelectronics, "STM32F103xC, STM32F103xD, STM32F103xE Datasheet." [Online]. Available: https://www.st.com/resource/en/datasheet/cd00191185.pdf

[48] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[49] G. B. Arfken and H. J. Weber, *Mathematical methods for physicists*. American Association of Physics Teachers, 1999.

[50] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Sensor fusion IV: control paradigms and data structures*, 1992, vol. 1611: International Society for Optics and Photonics, pp. 586-606.

[51] W. Jansen, D. Laurijssen, W. Daems, and J. Steckel, "Automatic Calibration of a Six-Degrees-of-Freedom Pose Estimation System," *IEEE Sensors Journal,* vol. 19, no. 19, pp. 8824-8831, 2019.

[52] D. C. Niehorster, L. Li, and M. Lappe, "The accuracy and precision of position and orientation tracking in the HTC vive virtual reality system for scientific research," *i-Perception,* vol. 8, no. 3, p. 2041669517708205, 2017.

[53] M. Borges, A. Symington, B. Coltin, T. Smith, and R. Ventura, "HTC Vive: analysis and accuracy improvement," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2018: IEEE, pp. 2610-2615.

[54] P. Vallabh and R. Malekian, "Fall detection monitoring systems: a comprehensive review," *Journal of Ambient Intelligence and Humanized Computing,* vol. 9, no. 6, pp. 1809-1833, 2018.

[55] S. B. Khojasteh, J. R. Villar, C. Chira, V. M. González, and E. De la Cal, "Improving fall detection using an on-wrist wearable accelerometer," *Sensors,* vol. 18, no. 5, p. 1350, 2018.

[56] B. Andò, S. Baglio, C. O. Lombardo, and V. Marletta, "An event polarized paradigm for ADL detection in AAL context," *IEEE Transactions on instrumentation and Measurement,* vol. 64, no. 7, pp. 1814-1825, 2015.

[57] Z. Zhang, C. Conly, and V. Athitsos, "A survey on vision-based fall detection," presented at the Proceedings of the 8th ACM International Conference on PErvasive Technologies Related to Assistive Environments - PETRA '15, 2015. [Online]. Available: http://dx.doi.org/10.1145/2769493.2769540.

[58] M. Cheffena, "Fall detection using smartphone audio features," *IEEE journal of biomedical and health informatics,* vol. 20, no. 4, pp. 1073-1080, 2015.

[59] Y. Wang, K. Wu, and L. M. Ni, "Wifall: Device-free fall detection by wireless networks," *IEEE Transactions on Mobile Computing,* vol. 16, no. 2, pp. 581-594, 2016.

[60] Q. Guan, C. Li, X. Guo, and B. Shen, "Infrared signal based elderly fall detection for in-home monitoring," in *2017 9th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2017, vol. 1: IEEE, pp. 373-376.

[61] J. Liu, H. Liu, Y. Chen, Y. Wang, and C. Wang, "Wireless Sensing for Human Activity: A Survey," *IEEE Communications Surveys & Tutorials,* pp. 1-1, 2020, doi: 10.1109/comst.2019.2934489.

[62] S. S. Khan and J. Hoey, "Review of fall detection techniques: A data availability perspective," *Medical Engineering & Physics,* vol. 39, pp. 12-22, 2017/01 2017, doi: 10.1016/j.medengphy.2016.10.014.

[63] G. Feng, J. Mai, Z. Ban, X. Guo, and G. Wang, "Floor pressure imaging for fall detection with fiber-optic sensors," *IEEE Pervasive Computing,* vol. 15, no. 2, pp. 40-47, 2016.

[64] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.

[65] C. Wan, L. Wang, and V. V. Phoha, "A survey on gait recognition," *ACM Computing Surveys (CSUR),* vol. 51, no. 5, pp. 1-35, 2018.

NATHANIEL FAULKNER received his B.E. (Hons.) from Massey University, New Zealand in Electronics & Computer Engineering in 2016. He is currently pursuing a Ph.D. at the same institution. In 2020 he was also appointed as a Visiting Research Fellow at the School of Science and Technology, Sunway University, Malaysia. His research interests include indoor positioning, embedded systems design, and the Internet of Things.

BADEN PARR Received his B.E. degree (Hons.) in Electronics and Computer Engineering from Massey University, Auckland, New Zealand in 2017, where he is currently pursuing a Ph.D. degree. His doctoral research focuses on automating grape yield estimation in vineyards using 3D camera technology. In addition to these studies, he is also involved in research on indoor localization using visible light, acoustic imaging using active and passive arrays, dense urban IoT sensor networks for air quality monitoring, and robot design and locomotion strategies.

FAKHRUL ALAM (M'17-SM'19) is an Associate Professor at the Department of Mechanical & Electrical Engineering, School of Food & Advanced Technology, Massey University, New Zealand. He received BSc (Hons) in Electrical & Electronic Engineering from BUET, Bangladesh, and MS and Ph.D. in Electrical Engineering from Virginia Tech, USA. His research interest includes indoor localization, 5G & visible light communication, IoT & wireless sensor networks. He is a member of IET.

MATHEW LEGG (M'19) Received his B.Sc, M.Sc, and Ph.D. in Physics from the University of Auckland, New Zealand. He is currently a senior lecturer with the Department of Mechanical and Electrical Engineering, School of Food and Advanced Technology, Massey University. His research relates to the development of acoustic/ultrasonic measurement systems and techniques for acoustic imaging, non-destructive testing, and remote sensing.

SERGE DEMIDENKO (M'91-SM'94-F'04) is Professor and Dean of the School of Science and Technology, Sunway University, Malaysia. He is also associated with the Department of Mechanical & Electrical Engineering, School Food and Advanced Technology, Massey University, New Zealand. He graduated in Computer Engineering from the Belarusian State University of Informatics and Radio Electronics and received Ph.D. from the Institute of Engineering Cybernetics of the Belarusian Academy of Sciences. His research interests include electronic design and test, signal processing, instrumentation and measurements. He is also Fellow of IET and UK Chartered Engineer.

CHAPTER 4 – DFL USING WALL-MOUNTED THERMOPILES

This chapter is republished in accordance with IEEE's copyright policy. This chapter is the accepted version of the published work, and as such may have stylistic differences from the final published article.

Device-Free Localization Using Privacy-Preserving Infrared Signatures Acquired from Thermopiles and Machine Learning

Nathaniel Faulkner[1,2], Fakhrul Alam[1,2], Senior Member, IEEE, Mathew Legg[1], Member, IEEE, and Serge Demidenko[1,2], Fellow, IEEE

[1]Department of Mechanical & Electrical Engineering, School of Food & Advanced Technology, Massey University, Auckland 0632, New Zealand
[2]School of Engineering & Technology, Sunway University, Bandar Sunway, 47500 Selangor, Malaysia

Corresponding author: F. Alam (e-mail: f.alam@ massey.ac.nz).

ABSTRACT The development of an accurate passive localization system utilizing thermopile sensing and artificial intelligence is discussed in this paper. Several machine learning techniques are explored to create robust angular and radius coordinate models for a localization target with respect to thermopile sensors. These models are leveraged to develop a reconfigurable passive localization system that can use a varying number of thermopiles without the need for retraining. The proposed robust system achieves high localization accuracy (with the median error between 0.13 m and 0.2 m) while being trained using a single human subject and tested against multiple other subjects. It is shown that the proposed system does not experience any significant performance deterioration when localizing a subject at different ambient temperatures or with different configurations of the thermopile sensors placement.

INDEX TERMS Device-free localization (DFL), human sensing, indoor positioning system (IPS), infrared sensing, machine learning, passive localization, thermopile.

I. INTRODUCTION

Smart cities [1] and smart homes [2] are radically changing how we live by offering, among other things, location based services [3] and ambient assisted living [4] requiring reliable positioning systems. Two recent decades have seen intensive research activities associated with the development of Indoor Positioning System (IPS) solutions [5].

IPS can be of active and passive types. Active or device-based solutions use a network of static nodes (often termed as anchors) to localize a transceiver carried by a human target. Given the immense popularity of mobile phones, many solutions propose to locate individuals by tracking their phones. These techniques leverage the large number of onboard sensors (e.g. camera, Inertial Measurement Unit (IMU), light-sensors) and communication capabilities (cellular, Wi-Fi, Bluetooth) of the phones [6]. The passive or Device-Free Localization (DFL) systems [7] do not require the tracked entity to carry a transceiver. Passive positioning can be achieved by using regular camera vision techniques. However, there is a privacy issue that has to be considered here: people are normally quite reluctant to have such imaging

devices, particularly in private areas of their residences. Camera-based techniques are also impacted by the illumination conditions.

Various alternative sensing techniques, based on the use of the Radio Frequency (RF) Received Signal Strength Indicator [8], Wi-Fi Channel State Information [9], visible light [10, 11], and electric field [12, 13] were proposed for DFL. Localization using pressure-sensitive [14] and capacitive [15] floors were also investigated. There were reports on techniques that applied single-pixel cameras [16], ultrasonic [17], and acoustic [18] sensing. Footstep induced vibrations captured by seismic sensors were also proposed for the localization [19]. Whilst considerable progress has been achieved in the DFL-associated research, the area is still of significant and on-going interest amongst researchers aiming to improve existing techniques and develop new solutions.

Human subjects can be detected from their Infrared (IR) emission. In most indoor surroundings, a person having a higher temperature than the environment can be distinguished from the background. Two popular devices used for IR localization are Passive IR (PIR) and Thermopile sensors. PIR

sensors are commonly used as motion detectors in security systems. However, PIR-based techniques require a relatively large number of sensors. They may also need significant sensor modifications (e.g., [20]), making commercial off-the-shelf motion sensors unusable. Besides, they are inherently incapable of localizing a stationary target as PIR sensors require relative motion between them and the target. Rotating sensors [21] or shutters [22] could overcome that issue. Regrettably, such sensors are more complex, expensive and are characterized by increased power consumption.

A. LOCALIZATION USING THERMOPILE SENSORS

Low-resolution thermopile sensors (e.g., AMG8833 Grid-EYE[1]) are effectively thermal cameras that can detect both stationary and moving targets. At the same time, due to their lower image resolution, such sensors do not compromise the privacy of subjects. Figures 1(a) and 1(b) illustrate the difference between an image acquired from a standard camera and a typical output of a thermopile sensor, taken at its maximum 8×8 pixel resolution. As well as being privacy-preserving, thermopiles are invariant to changes in illumination.

Shetty et al. proposed tracking a subject using the foreground regions from the thermopile images [23]. Unfortunately, the authors did not report any accuracy values. Using a similar method, Kuki et al. were able to obtain an accuracy ranging between 0.15 m and 0.35 m (depending upon the activity) in a 2.56 m^2 area using a 4×4 pixel sensor [24]. The same authors then extended their work to achieve multi-person detection [25]. Qu et al. performed multi-target localization using a ceiling-mounted sensor [26]. They were able to distinguish between subjects even for crossover events. They also investigated the sensor lens distortion and performed distortion correction. Ng et al. were able to locate subjects with an accuracy of approximately 0.5 m in a 12.5 m^2 area using five sensors [27]. Kowalski et al. [28] were able to localize a subject to a 0.5 m grid square with a 73% probability in a 3.75 m^2 area. Whilst the accuracy of that setup was lower than results in some other reported works, the authors extended their sensors to have a 180-degree Field-of-View (FOV) by collocating three GridEye sensors directed 60 degrees from each other whilst also covering a much larger area. Tariq et al. used a low-resolution 16-pixel thermopile sensor with a variety of neural networks to achieve 0.096 m Root Mean Squared Error (RMSE) in a 9 m^2 space [29]. Narayana et al. [30] were able to locate a subject with a median 0.22 m accuracy within a 72 m^2 area using a higher resolution 32×24-pixel sensor. It should be noted that their proposed system required an additional calibrated PIR sensor for depth estimation.

Singh et al. [31] compared the application of various Machine Learning (ML) classifiers for the detection and
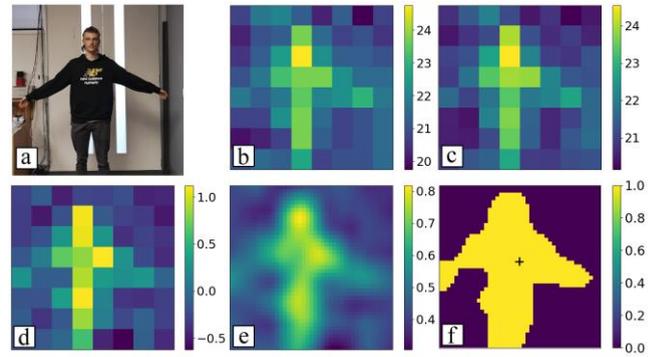

FIGURE 1. The sensor data pre-processing and parametric model process used to find the angle of the subject. Photo (a) shows a standard smartphone camera image of the subject taken simultaneously from the same position as that obtained by the thermopile sensor. Plot (b) shows the raw input from the sensor. The remaining plots illustrate (c) the Gaussian denoising application, (d) SVD background removal and data normalization, (e) interpolation, and (f) thresholding and averaging of pixel positions above the threshold to find the center of the largest blob.

activity recognition of multiple human subjects using thermopile sensors. Similarly, Tateno et al. [32] and Tao et al. [33] used deep learning networks for fall detection and activity recognition respectively by utilizing ceiling-mounted sensors. Gochoo et al. [34] used a deep learning network to classify 26 separate yoga poses.

B. KEY CONCEPT AND CONTRIBUTION

The model-based localization techniques proposed in the literature (e.g., [23], [27]) appear to be unable to accurately capture the complex relationship between IR data and the relative position of the target with respect to the sensors. The reported ML techniques (e.g., [28], [29]) largely adopt the fingerprinting approach. They employ a single dataset, collected with one test subject, that is split for training, validation and testing. Therefore, it is difficult to ascertain whether these systems generalize well to different environments or subjects.

This paper proposes a new approach for target (i.e., a human subject) localization based on training models of thermopile sensors with the application of ML techniques. These models provide accurate estimation of an Angular Coordinate (AC) and Radius Coordinate (RC) of the target with respect to the center of the sensor, which are the direction and range to the target. Therefore, in the proposed approach, the subject can be localized by using just a single sensor (Fig. 2(a)). If the subject is in FOV of at least two sensors of known positions, two ACs can be estimated and used for positioning in a manner that is similar to the Angle of Arrival (AoA) [35] method (Fig. 2(b)). Similarly, if there are three or more sensors, one can use the lateration technique [35] to find the target position, using the distances between the subject and the sensors (Fig. 2(c)).

In the proposed approach, the sensor models for AC and RC need to be trained just once for one sensor. They can be then transferred to other sensors without measurable compromise

---

[1]https://na.industrial.panasonic.com/products/sensors/sensors-automotive-industrial-applications/lineup/grid-eye-infrared-array-sensor
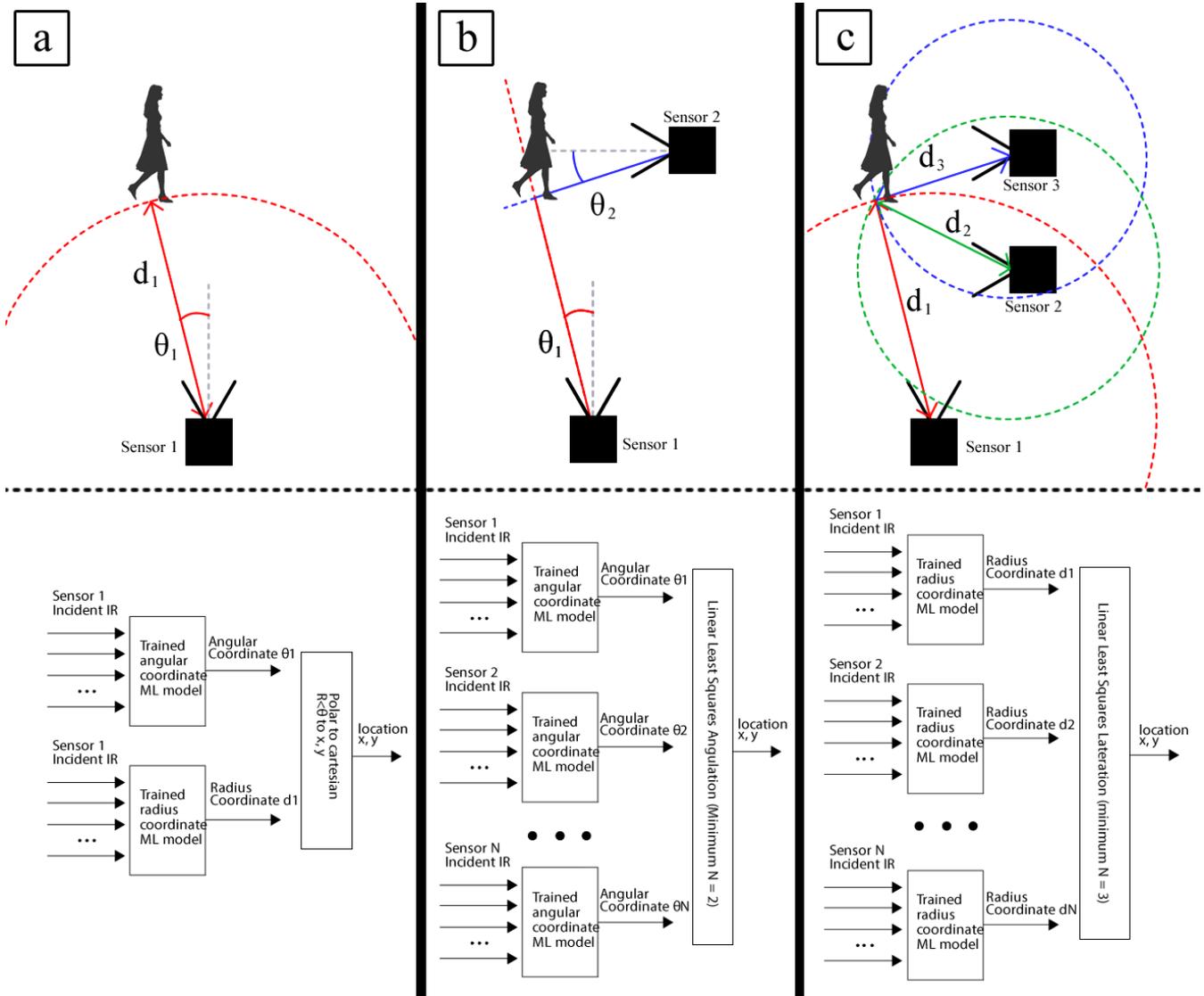
FIGURE 2. Three methods of positioning a subject where the sensor positions and orientations are known a priori. Diagram (a) shows the positioning using a combination of a single AC and RC/range. Diagram (b) shows the angulation using a minimum of two ACs. Diagram (c) shows the lateration using a minimum of three ranges.

in terms of localization accuracy. This leads to a robust and reconfigurable indoor positioning system that does not need to be retrained when deployed outside the training environment. The performance of the proposed system was tested with several different subjects. Each subject walked arbitrary paths for several minutes. The system was able to localize a subject, which it was not trained upon, with a median error of less than 0.2 m. The results show that the proposed approach is largely invariant to the subjects, system configuration, and deployment environments.

The remainder of the paper is structured as follows. Section II discusses data acquisition, ground truth estimation, and data preprocessing methods. Section III discusses the training, tuning, and evaluation of the various ML models used for the

estimation of the angular and radius coordinates of a subject. Section IV demonstrates how these models can be used for positioning, and it also investigates the positioning performance of the system. Section V concludes the paper and discusses the limitations and future work to address these.

II. SYSTEM DEVELOPMENT
A. DATA ACQUISITION
The thermopile sensor used in this work is the Grid-EYE AMG8833. An interface to connect the AMG8833 to a computer was designed and constructed. It uses the STM32f103[2] microcontroller and a USB to serial adapter (Fig. 3). An arbitrary number of sensors may then be connected to a computer via USB cables whilst using a simple script for

---

[2]https://www.st.com/content/st_com/en/products/microcontrollers-microprocessors/stm32-32-bit-arm-cortex-mcus/stm32-mainstream-mcus/stm32f1-series/stm32f103/stm32f103cb.html

logging incoming sensor data, a corresponding device ID, and the timestamp of the data to a text file.

Several different datasets were recorded. Firstly, a 1.8 m tall male subject walked around a test area for approximately one hour. Two sensors were affixed to the walls at a height of 1.4 m (chosen to be above the height of most furniture) as seen in Fig. 4(a). The ambient temperature was measured to be at 24 °C. This dataset is henceforth referred to as *Dataset 1*. The second dataset (*Dataset 2*) was taken at a later date in the same test area, using five different subjects, henceforth known as *Subjects 1-5*. The subjects were between 1.65 and 1.85 m tall (one female and four males). *Subject 1* was the subject used to collect *Dataset 1*. Each subject (including *Subject 1*) walked around the test area (having the same layout as the one used for the *Dataset 1* collection) for approximately 5-7 minutes each. The ambient temperature was measured to be at 22 °C. The third dataset (*Dataset 3*) was taken at another date, in a different room, with a three-sensor setup as seen in Fig. 4(b). It only featured *Subject 1* walking around the test area for approximately 10 minutes. The ambient temperature was at 26°C. The final dataset (*Dataset 4*) was also taken in this room, on another date, with *Subject 1* moving within the test area for approximately 7 minutes. The positions of the three sensors (see Fig. 4(c)) were different from those chosen for *Dataset 3*. The ambient temperature was again at 22 °C.

These multiple datasets were taken for specific purposes. *Dataset 1* was used to train various sensor models discussed later in this paper. *Datasets 2-4* were employed to investigate the generalizability of the proposed approach with respect to different environments or configurations and with different subjects from whence the system was trained.

B. GROUND TRUTH ESTIMATION

Accurate ground truth is very important when designing and evaluating a positioning system. In order to train and test a robust model, a large amount of labelled data is required. It is possible to mark out predefined paths and have a subject walking whilst following them at a set pace. However, such an arrangement is not ideal as it requires a high level of concentration from the subject. Besides, it can potentially force the subject into an unnatural gait. Ideally, a ground truth system should accurately track subjects as they naturally walk within the testing area. For this reason, the HTC Vive[3] was used as the ground truth tool. In previous works, it was found to be accurate to within several mm for extended time periods [15, 36]. The "tracking puck" needs to be kept within the line of sight of the "lighthouses". Therefore, the puck was attached to the subject's head, to approximate the subject's position in two dimensions.

C. DATA PRE-PROCESSING

Data received from the thermopile sensors should be preprocessed to make them resistant to changes in ambient
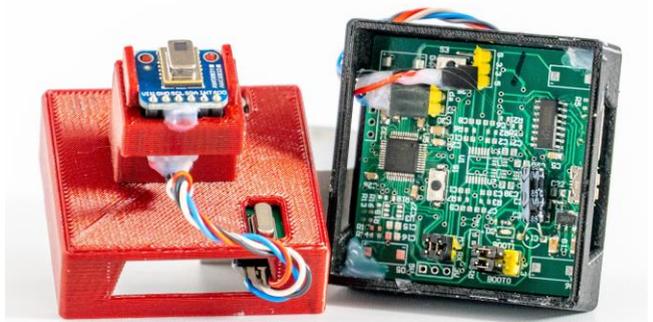
FIGURE 3. The thermopile sensor unit, housing both the thermopile module and the microcontroller used to communicate with a computer.

conditions. After that, the data can be used to train or test various ML models.

1) GAUSSIAN DENOISING

The temperature data produced by the thermopiles are noisy and as such, a single pixel fluctuates between frames randomly. This could cause the misdetection of subjects. To address it, each pixel is taken as a single time-series element and a one-dimensional Gaussian kernel is then applied along with the time-series data for each pixel (Fig. 1(c)). Each 8×8 frame is flattened into a single 64×1 vector:

$$f = \begin{bmatrix} p_1 \\ p_2 \\ \vdots \\ p_{64} \end{bmatrix}. \tag{1}$$

Such vectors form columns of a matrix:

$$\mathbf{A} = [f_1 \quad f_2 \quad \cdots \quad f_N], \tag{2}$$

which is a time-series of length N samples with each row representing a single pixel (p_n) with respect to time. Each row of $\mathbf{A}$ is convolved with a 1-dimensional Gaussian kernel G. A kernel with a sigma value of 3 was empirically chosen with the following 5 samples defining the function:

$$G = [0.1784 \quad 0.2104 \quad 0.2223 \quad 0.2104 \quad 0.1784] \tag{3}$$

2) BACKGROUND REMOVAL

The background temperature of a room is prone to change over time. However, such temperature changes are highly correlated between pixels in an empty scene. Moreover, the difference between two given background pixels at the same point in time appears to stay constant over time. This is illustrated in the first panel of Fig. 5. Each vertical slice represents a single frame from the thermopile. The first set of approximately 1800 samples in the figure corresponds to the room with no subject being present for approximately 3 minutes. The fluctuations in the background with time as the room temperature changes can be seen by the vertical strata in
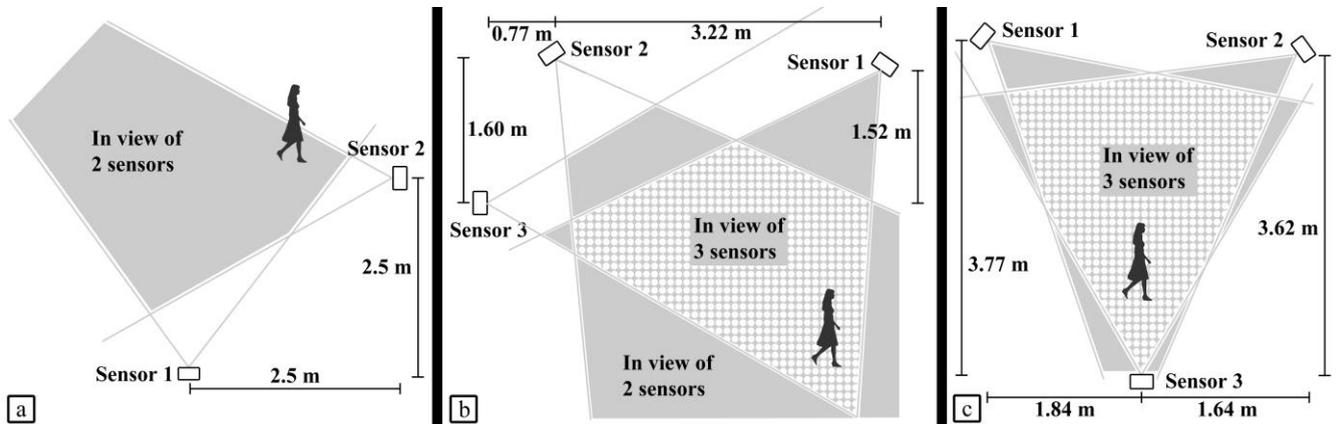
FIGURE 4. Room layouts for *Dataset 1* and *Dataset 2* (a), *Dataset 3* (b), and *Dataset 4* (c).

the data, whereas the constant offset between pixels can be seen in the horizontal stratification. Singular Value Decomposition (SVD) was employed for dimension reduction (such an approach has been used in computer vision for separating the background and foreground in videos [37]). SVD factorizes [38] the M × N matrix **A** (in this case M is 64 and N is the number of samples in the dataset):

$$\mathbf{A} = U\mathbf{L}V^{T}. \qquad (4)$$

Here, the columns of U are the left singular vectors and the columns of V are the right singular vectors; **L** contains the singular values of **A** in a diagonal matrix, arranged in descending order; and T represents the transpose operation. The matrix **A** can be reassembled by multiplying the matrices together as given in Equation (4). However, it is possible to modify values of **L** before reassembling, to various effects. The singular values effectively represent how strongly each singular vector contributes to the matrix. The background data are highly correlated across the dataset, both in time and across the frame. Therefore, a good representation of the background can be found by reassembling the data using only the first singular value (the most dominant dimension) and zeroing out the others (see the bottom panel of Fig. 5). The foreground can therefore be found by doing the opposite – zeroing the first singular value and reassembling the matrix, as shown in the middle panel of Fig. 5.

3) DATA NORMALIZATION
After the background subtraction, there is still some variation in mean and standard deviation between different datasets. Furthermore, it is often advantageous to have input data for machine learning ranging between 0 and 1. Both Min-Max scaling and standardization were used in different circumstances. In a permanent, real-world deployment of the system, this would be done using a predetermined number of previous samples (e.g., several minutes' worth). When the system is first powered on, it would require an initial self-calibration period until a sufficient number of samples is captured.

III. SENSOR MODELS
The authors propose to create transferable sensor models that are trained once and then can be used for any sensor of the same make (e.g., the GridEye sensors). Locations of the sensors could be set arbitrarily. The same model could be used for all sensors regardless of their positions in a room without any retraining. In essence, the sensors are calibrated to produce the angular and/or radius coordinates of a target in two dimensions.

Multiple sensor models were trained and validated on *Dataset 1*, which was the largest one with 32,000 data points (frames). It was randomly split into 80-10-10 training-validation-test segments. The split was the same for each trained model (i.e., the same segments of data were used for training, validation, and test for each model). Each model (including the outlined below parametric model that did not require training data) was then tested against the data from five subjects taken at a different time (*Dataset 2*) as well as the test split from *Dataset 1*. This was done to investigate the generalizability of the models between subjects.

A. ANGULAR COORDINATE MODELS
These models take the temperature data from the sensor (after they were preprocessed as outlined in Section II D) as an input. They then output the angular coordinates of the subject with respect to the sensor. The ground truth angle and distance are computed using the subject's ground truth position and the sensor position (see θ_1 and d_1 respectively in Fig. 2(a)). A parametric method was used, as well as several ML methods (such as Multi-Layer Perceptron (MLP), random forest, Weighted K-Nearest Neighbor (WKNN)) to develop the AC models. The advantage of the parametric model is that it does not require any prior training. At the same time, the studied ML methods require a one-off training of the model that then ideally could be generalized to any subject or room.

1) PARAMETRIC MODEL
The devised parametric model utilizes an approach similar to the thermopile positioning [23] and capacitive floor footprints
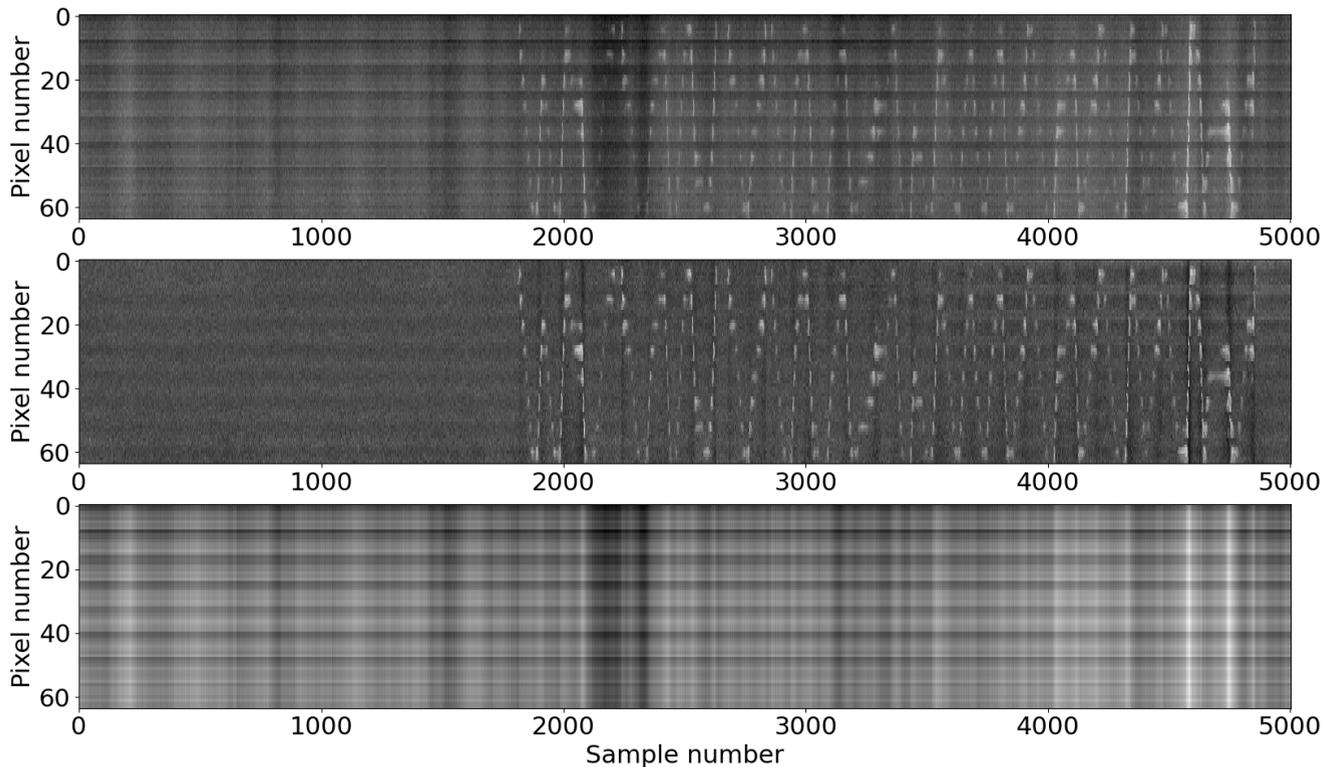
FIGURE 5. Plots demonstrating how the thermopile temperature matrix *A* (see Equation 2) is processed to remove background thermal effects. Each column is a single flattened frame, *f* from the sensor, with the horizontal axis representing time. The first approximately 1800 samples show an empty room, with the rest showing a subject moving around. The top panel is before background removal, the middle panel is after background removal and the bottom panel the removed background.

detection reported in [15]. After the temperature data are pre-processed (as described above in Section II D), it is reshaped to an 8×8 frame. Then it undergoes bicubic interpolation to a 55×55 matrix (see Fig. 1(e)) with each element of the matrix being referred to as a single pixel. Binary thresholding is then applied to select only the foreground objects. After that, the connected component analysis is employed for blob detection. If the number of blobs is more than one, the largest of them is assumed to be the subject. The center of the blob is found by taking the mean *x* and *y* positions of all the pixels in the blob. Fig. 1(f) illustrates this process. The position of the blob is then converted into an angular form:

$$\theta = \left(\frac{x}{FW} - 0.5\right) \times FOV. \tag{5}$$

Here θ is the AC of the subject with respect to the sensor, x is the position of the center of the blob pertaining to the frame (in pixels), FW is the frame width in pixels (55 pixels after the interpolation), and FOV is the width of the horizontal field of view of the sensor (60 degrees for GridEye). A zero value of θ indicates that a subject is on a line perpendicular to the sensor (i.e., in the middle of the field of view). Negative values of θ correspond to the left-hand side while positive - to the right-hand side plane.

2) MULTI-LAYER PERCEPTRON MODEL

An MLP model was trained with 64×1 input vectors of preprocessed temperature data from the sensor while the outputs were ACs of the subject with respect to the sensor. The input layer of the MLP had 64 perceptions fully connected to the first hidden layer. The output layer was a single perceptron, fully connected to the final hidden layer. A grid search within a wide range of hyperparameters was used to tune the MLP. Each candidate model, defined by a unique combination of hyperparameters, was trained for 1000 epochs and tested against the validation data after each epoch. Early stopping was used to avoid overtraining by observing the validation RMSE curve over these 1000 epochs. The final models for each candidate model were then sorted by validation RMSE.

It was found that 2 hidden layers of 500 perceptrons with ReLU activation [39] on the hidden layers and sigmoid activation on the output layer gave the best performance (see Table 1). Several other parameter combinations gave similar performance results whilst enabling trading of the performance (RMSE) for a simpler model if needed. Larger models take significantly longer time and more computing resources to both train and run. Whilst training is a one-off event and can be generalized to multiple environments or subjects, running the network has potential processing considerations. This is because the models could be run on the

TABLE 1. Final hyperparameter values selected for each AC model after hyperparameter tuning.

Model	Hyperparameter	Value
MLP	Number of hidden layers	2
	Hidden layer size	500
	Hidden layer activation	ReLU
	Output layer activation	Sigmoid
Random Forest	Number of estimators	500
	Minimum sample per leaf node	1
	Minimum samples to split leaf	2
WKNN	K-value	2
	Distance metric	Euclidean

sensors themselves as opposed to on a PC. Such sensor-based solutions are resource-constrained. Thus, a simpler model is preferred as it would run faster and with lower power consumption.

3) RANDOM FOREST MODEL
A random forest regressor was trained in a manner that was similar to the MLP case by searching through a range of parameters to find optimal hyperparameters. It was observed that limiting the maximum tree depth or maximum leaf nodes had very little impact on the accuracy of the models. The best model had 500 estimators, a minimum of 2 samples to split a leaf node and a minimum of 1 sample per leaf node (see Table 1). However, the RMSE difference between 100 and 500 estimators was less than 1%, thus suggesting an opportunity for utilizing a simpler regressor.

4) WEIGHTED K-NEAREST-NEIGHBOR MODEL
WKNN regressor was optimized for the number of neighbors (K) and the distance metrics. It was found that a K-value of 2 with either Euclidean or Canberra distance metric provided the best performance.

B. PERFORMANCE OF ANGULAR COORDINATE MODELS
The performance of each model for estimating the AC for *Dataset 2* can be seen in Fig. 6. The parametric model was significantly outperformed by the three ML models. Also, there were no significant differences between these three machine learning models. At the same time, MLP might be somewhat preferable over WKNN in cases where a large database is needed. The model was processed on a PC with the raw data arriving from the device. However, it might be preferred in the future to move the model onto the microcontroller where storage concerns could preclude the use of a large WKNN database. The results showed that while the accuracy of the AC estimation was decreased compared to the test set split from *Dataset 1*, it was reasonably subject-invariant. *Subject 1* was used for training the model. The obtained results were on par with those for the other four subjects (whilst worse than the training set). It suggested that the main difference was in the environments (the training set was taken on a different day, with the temperature of the room being approximately 2°C lower).
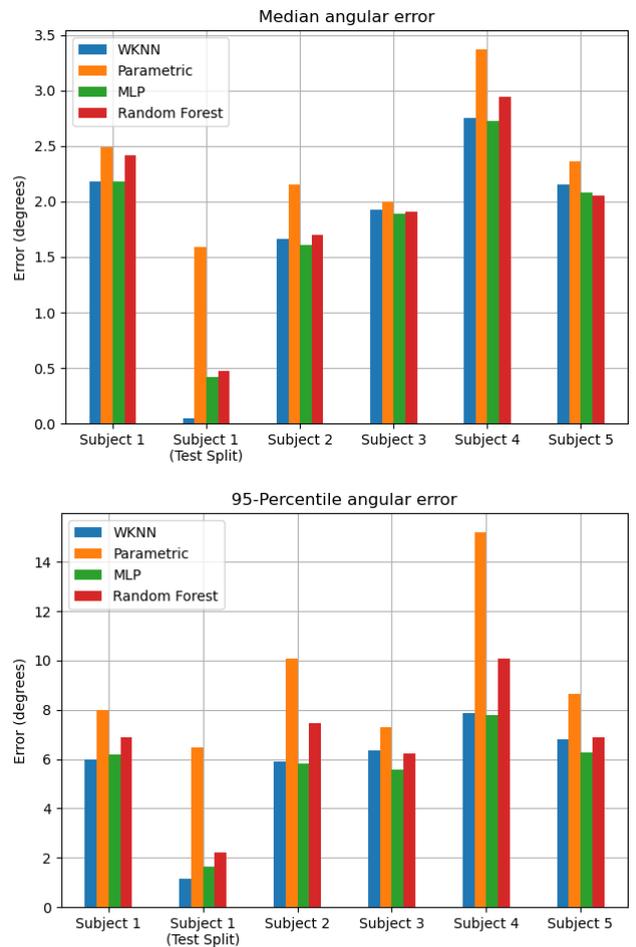

FIGURE 6. Median and 95-percentile angular coordinate error for the four angular coordinate models. Each model was tested on all five subjects (*Dataset 2*), plus the 10% test split from the training dataset (*Dataset 1*).

C. RADIUS COORDINATE OR RANGE MODELS
The range models estimate the distance between the subject and the sensor. In a similar manner as for the AC models, a flattened 8×8 frame from the sensor was inputted as a 64×1 vector. The output of the model was the range. The range can be taken as the radius coordinate and then it can be combined with the angular coordinate to perform single-sensor based positioning (Fig. 2(a)). Also, the distances from multiple sensors can potentially be used for the lateration (Fig. 2(c)). Three regressors (MLP, random forest and WKNN) were trained. The hyperparameters are listed in Table 2.

D. PERFORMANCE OF RANGE MODELS
The three different range estimation models were compared for *Dataset 2*, i.e., for five subjects (Fig. 7). The MLP outperformed the other two models quite significantly. It was more robust to variations of the environment and different subjects than the other methods. Interestingly, the WKNN and random forest methods struggled most with *Subject 1* upon whose data the models were trained albeit with data collected on a different day.

TABLE 2. Final hyperparameter values selected for each range model after hyperparameter tuning.

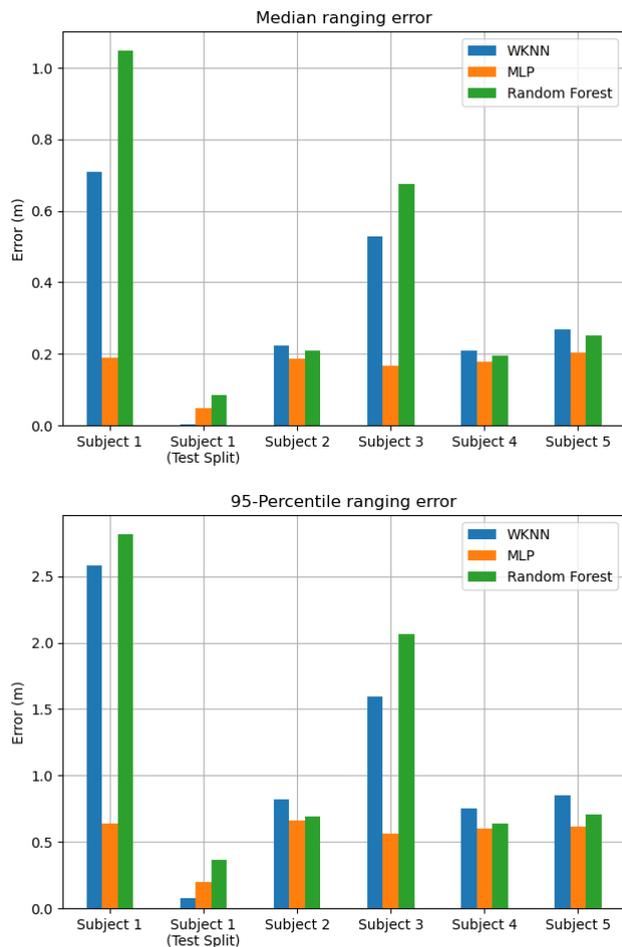| Model	Hyperparameter	Value
MLP	Number of hidden layers	3
	Hidden layer size	500
	Hidden layer activation	ReLU
	Output layer activation	Sigmoid
Random Forest	Number of estimators	100
	Minimum sample per leaf node	1
	Minimum samples to split leaf	2
WKNN	K-value	2
	Distance metric	Canberra

FIGURE 7. Median and 95-percentile radius coordinate error for the three range models. Each model was tested on all five subjects (*Dataset 2*), plus the 10% test split from the training dataset (*Dataset 1*).

IV. POSITIONING METHODS

Several methods of positioning a subject can now be developed using the sensor models proposed in Section III. The use of the AC and range with a single sensor or just the AC-based model on multiple sensors allows for the ease of configuring the system. Sensors can be set up at suitable locations in a room. Additional sensors can be incorporated for extended coverage or higher accuracy without the need for retraining. This provides a significant advantage over the multiple sensor fingerprint-based method that relies on the

sensor number and the geometry remaining consistent. For such fingerprint-based methods, retraining would be required if sensors are to be spaced at different distances, or additional sensors need to be added.

A. MODEL-BASED POSITIONING

1) SINGLE-SENSOR BASED POSITIONING USING THE AC AND RANGE

Positioning can be performed using a single sensor using AC and range data. Two MLP models are employed for subject positioning, with one of them outputting the AC whilst the other, the RC - the distance between the subject and sensor. Simple geometry is then used to calculate the position of the subject relative to the sensor, see Fig 2(a). It should be noted that once the range is estimated from three or more sensors, it is possible to perform the lateration-based localization. However, the reported research did not pursue that approach as it would require extra sensors while preliminary results did not show noticeable performance benefits.

2) MULTIPLE SENSOR BASED POSITIONING USING AC

This approach uses the positions of multiple sensors and the angular coordinates of the subject with regard to the sensors. ACs are estimated using the MLP model (outlined in Section III A). The position of the subject is found similarly to a standard AoA technique. Fig. 2(b) shows an example where two sensors are used. If more than two sensors are employed, the system is over defined and linear least squared estimation can be used [35]. The problem can be formulated as

$$Ax + q = b \qquad (6)$$

where:

$$A = \begin{bmatrix} \sin(\theta_1) & -\cos(\theta_1) \\ \sin(\theta_2) & -\cos(\theta_2) \\ \vdots & \vdots \\ \sin(\theta_n) & -\cos(\theta_n) \end{bmatrix}, \qquad (7)$$

$$b = \begin{bmatrix} \sin(\theta_1)x_1 - \cos(\theta_1)y_1 \\ \sin(\theta_2)x_2 - \cos(\theta_2)y_2 \\ \vdots \\ \sin(\theta_n)x_n - \cos(\theta_n)y_n \end{bmatrix}, \qquad (8)$$

and q is a measure of the noise. The estimate of x, the 2x1 position vector is:

$$\hat{x} = (A^T A)^{-1} A^T b \qquad (9)$$

B. FINGERPRINT-BASED POSITIONING

The proposed model-based positioning techniques were benchmarked against fingerprint-based techniques. Fingerprinting is commonly used for ML-based positioning as reported in the literature. A single-sensor based fingerprint positioning used an MLP with its input being a flattened array of the pixels from a single sensor. The output presents the x and y coordinates of the subject relative to the sensor. Training, validation, and testing were done following the

process described in Section III A. A dual-sensor fingerprint-based positioning was also implemented where an MLP was trained to take a flattened array of 128 pixels (64 from each sensor) to estimate the position of the subject.

C. POSITIONING RESULTS EVALUATION

The four different positioning methods were evaluated using each of the five subjects (corresponding to *Dataset 2*). It can be seen from Fig. 8 that two-sensor-based positioning provided higher accuracy than the single-sensor ones. The proposed dual-sensor AC-based positioning was the most accurate. However, the field of view of the dual-sensor models (being the intersection of the individual FOVs) was smaller than the total area (Fig. 4(a)). The single sensor models offered a larger coverage, at the cost of lower accuracy.

One solution could be to use the dual-sensor configuration where there is coverage and utilize the single-sensor-based solution only where there is coverage by a single sensor. Another solution could be to use a higher density of sensors or employ sensors with a wider FOV (e.g., [28]).

In order to make a fair comparison, the single-sensor based positioning algorithms were also run a second time whilst using only the data that corresponded to the combined FOVs of both the sensors in the dual-sensor configuration. While there was a modest improvement for the single-sensor systems, the dual sensor configurations were still more accurate.

It was observed that there was not much variation in position errors between the subjects. As each subject walked about randomly, they entered and exited the FOV of the sensors. Time intervals between a subject entering and exiting sensor FOVs were saved as for individual paths. For each subject, that equated to between 25 to 35 paths of varying lengths and trajectories. Some examples of paths can be seen in Fig 9.

D. PERFORMANCE IN DIFFERENT CONFIGURATIONS

A three-sensor system layout (Fig. 4(b)) was implemented to carry out the experiment (corresponding to the case of *Dataset 3* outlined in Section II B). The AC-based position evaluations were performed using all combinations of sensor pairs (1-2), (2-3), (3-1) as well as for the three-sensors case. The AC MLP model used was trained on *Dataset 1* that was acquired at an earlier date and in a different room. Fig. 10 and Table 3 show the localization results. When localization was carried out with the pairs of sensors, the obtained positioning accuracy was similar for all of them (the median varied between 0.13 m and 0.14 m; the 95-percentile varied between 0.27 m and 0.34 m for each of the three pairs). This was on par with what was observed for *Dataset 2* even though the relative sensors' positionings were markedly different. It clearly demonstrates the robust nature of the AC-based localization system. As seen in Fig. 11, the positioning accuracy could be improved by using the ACs from all three sensors. Such flexibility is not readily available with the
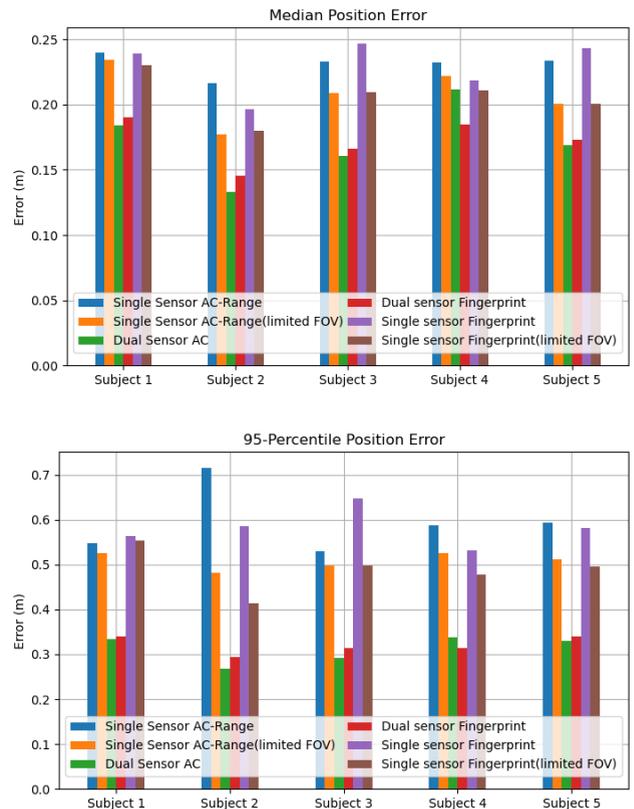

FIGURE 8. Median and 95-percentile position error for the four different positioning methods, plus the two single sensor methods being limited to the same FOV as the dual sensor models (the intersection of the two FOVs).

other positioning techniques.

The performance of the single-sensor based positioning methods (AC plus range and fingerprinting) were not impacted by changes in the configuration. The accuracy was consistently close between *Dataset 2* and *Dataset 3* while less accurate than when employing the multiple-sensor based methods with AC estimation. In contrast, the dual sensor fingerprinting method experienced significant accuracy degradation. Closer inspection revealed that the accuracy of the (1-2) pair was nearer to that achieved with *Dataset 2* (see Fig. 12). The two sensors of the (1-2) pair were placed in a relatively similar configuration to those employed for acquiring *Dataset 2*. However, the accuracy levels achieved with (2-3) and, especially, (3-1) sensor pairs were considerably poorer due to the sensor positions being significantly different from the *Dataset 2* acquisition case.

While the AC-based method is shown to be robust and flexible, the selection of sensor locations still needs to be done judiciously. This is evident from the AC-based positioning performance shown in Fig. 13 for *Dataset 4*. The sensor layout is given in Fig 4(c) whilst the dataset details are outlined in Section II B. The accuracy for the sensor pair (1-2) appears to be similar to what was observed for *Dataset 2* and *Dataset 3*. However, the accuracy achieved by the sensor pairs (1-3) and (2-3) shows significant degradation. In all previous cases, the sensors were out of FOV of each other,
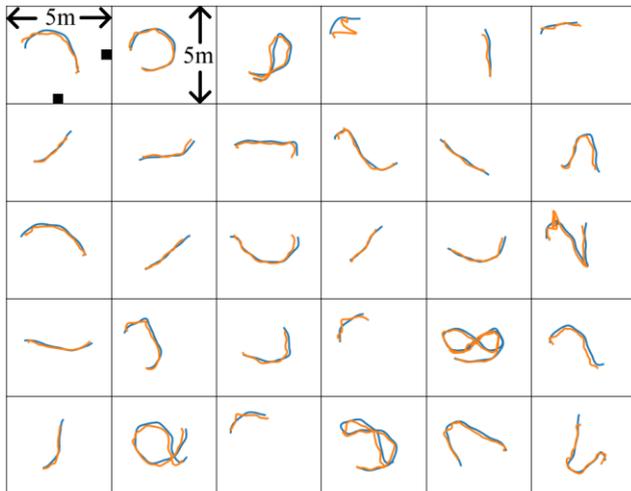
FIGURE 9. The paths walked by *Subject 1* in *Dataset 2*. The HTC Vive ground truth is shown in blue, along with the estimated positions using the dual camera AC-based method in orange. The black squares in the top left panel represent the sensor positions (shown only for the first panel)
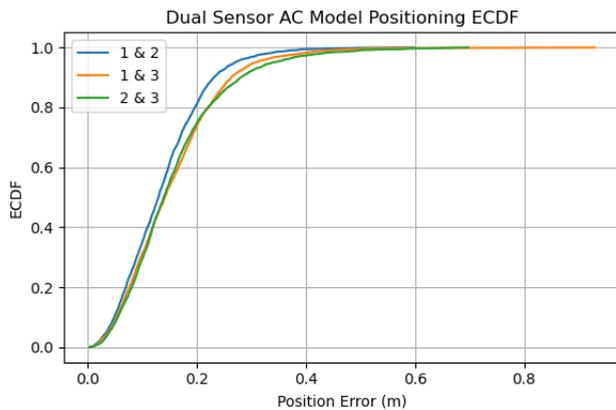

FIGURE 10. Dual sensor AC model positioning accuracy for pairs of sensors for *Dataset 3*.

whereas in the configuration under discussion, *Sensor 3* was within the fields of view of both *Sensor 1* and *Sensor 2*. Thus, if a subject stands directly between *Sensor 3* and one of the other sensors, the two angular coordinates would produce near-parallel lines. Therefore, even a small error in the AC estimations can cause a very large error in the position estimation. This is only a problem when the subject is in the FOV of two sensors, the sensors are within the FOV of each other, and the subject is directly between them. This can be mitigated through continuous tracking (as long as the target does not move along the direct path from one sensor to the other). However, the sensors could be positioned to ensure that such a scenario is unlikely to happen any often. Also, since the single-sensor based positioning is not impacted, the system can switch to a single sensor operation mode in such a scenario.

V. CONCLUSIONS AND FUTURE WORK

ML model-based systems showed great promise in performing accurate localization using single and multiple thermopile sensors. Multiple-sensor based positioning was shown to be

TABLE 3. Median and 95-percentile errors for the various positioning models used on *Dataset 3*.

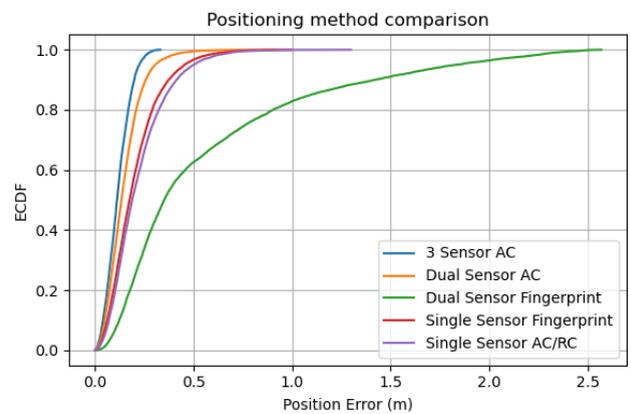| Model	Sensor id	Median error (m)	95-Percentile error (m)
Three-sensor AC model	1, 2 & 3	0.11	0.23
Dual-sensor AC model	1 & 2	0.13	0.27
	2 & 3	0.14	0.34
	3 & 1	0.14	0.31
Dual-sensor fingerprint	1 & 2	0.22	0.41
	2 & 3	0.35	0.98
	3 & 1	0.95	2.24
Single-sensor fingerprint	1	0.19	0.45
	2	0.17	0.39
	3	0.16	0.52
Single-sensor AC/range	1	0.17	0.40
	2	0.18	0.51
	3	0.22	0.56

FIGURE 11. Positioning accuracy for the different positioning models on *Dataset 3*. For the dual sensor models, the errors were combined from each of the three pairwise combinations. For the single sensor models the errors were combined from the three individual sensors.

more accurate than the single-sensor based one. However, the single-sensor based positioning offered an important advantage of larger area coverage.

The ML regressors were trained with one human subject and tested with other subjects as well as in different environments as opposed to only training and testing on the same dataset. The model-based techniques generalized well.

The fingerprinting-based positioning also appeared to be able to cope with a change of subjects and environments when only a single sensor was used. However, the fingerprinting-based positioning with multiple sensors was essentially limited by the configuration (e.g., the number of sensors and their relative positions) that was used for training.

The most apparent limitation of the proposed system is that it is only capable of localizing a single subject. This could be addressed in the future by applying a stacked approach. The authors were able to train an accurate classifier (over 95%) to detect the presence of a subject in a frame. This can be further extended to count the number of subjects. Several different models can then be trained for a varying number of subjects. With multiple sensors, it should be possible to track the subjects without losing their identity. In the case of crossover events where one subject is occluded by another, it is unlikely that subjects are occluded from both sensors. It may also be
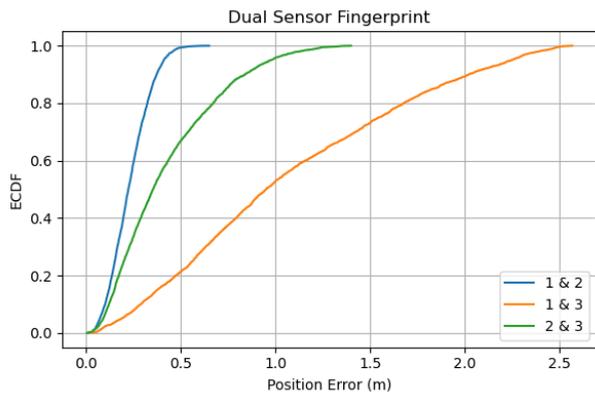
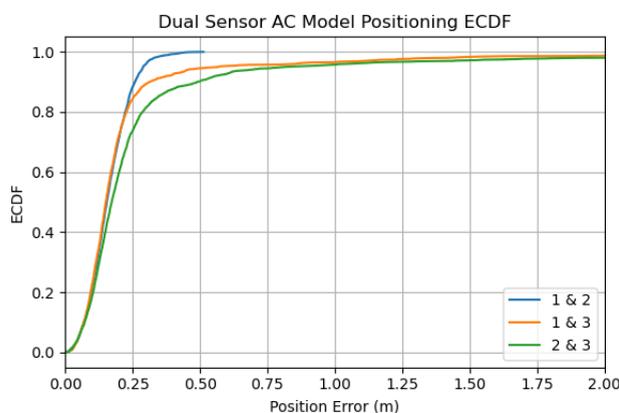**FIGURE 12. Dual sensor Fingerprint-based positioning accuracy for pairs of sensors for *Dataset 3*.**

FIGURE 13. Dual sensor AC model positioning accuracy for pairs of sensors for *Dataset 4*. Note the graph has been truncated at 2 m along the x axis for clarity due to the very large worse case errors for 1-3 and 2-3 (8.5 m and 97.5 m respectively).

possible to assign a short-term identity to a person based on their thermal profile [26] whilst they are within the FOV. The impact of changing the heights of the sensors have also not been investigated and can be a topic of further studies.

Thermopile sensors rely on the subject emitting IR, which is influenced by the clothes the subject is wearing. For example, while wearing a very heavy coat, a subject may not be visible to the sensor except at very close proximity. The experiments were undertaken in standard office wear that is appropriate for the ambient temperature. An exploratory investigation was undertaken where *Subject 1* wore a thick winter jacket. Positioning nearer the sensors appeared to be relatively unaffected. However, the performance was degraded at further distances as the jacket reduced the effective range of the sensor. It should be noted that the subject was much warmer than comfortable and would not have worn such apparel in a climate-controlled room. This would be an interesting area for a future investigation.

The performance of additional ML techniques (e.g., recurrent neural networks) and the impact of hyperparameters on ML techniques have been also identified for a future study.

REFERENCES
[1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things journal,* vol. 1, no. 1, pp. 22-32, 2014.

[2] B. L. R. Stojkoska and K. V. Trivodaliev, "A review of Internet of Things for smart home: Challenges and solutions," *Journal of Cleaner Production,* vol. 140, pp. 1454-1464, 2017.

[3] I. A. Junglas and R. T. Watson, "Location-based services," *Communications of the ACM,* vol. 51, no. 3, pp. 65-69, 2008.

[4] P. Rashidi and A. Mihailidis, "A survey on ambient-assisted living tools for older adults," *IEEE journal of biomedical and health informatics,* vol. 17, no. 3, pp. 579-590, 2012.

[5] F. Khelifi, A. Bradai, A. Benslimane, P. Rawat, and M. Atri, "A survey of localization systems in internet of things," *Mobile Networks and Applications,* vol. 24, no. 3, pp. 761-785, 2019.

[6] A. Yassin *et al.*, "Recent advances in indoor localization: A survey on theoretical approaches and applications," *IEEE Communications Surveys & Tutorials,* vol. 19, no. 2, pp. 1327-1346, 2016.

[7] F. Alam, N. Faulkner, and B. Parr, "Device Free Localization: A Review of Non-RF Techniques for Unobtrusive Indoor Positioning," *IEEE Iot Journal* 2020.

[8] D. Konings, F. Alam, F. Noble, and E. M. Lai, "Device-free localization systems utilizing wireless RSSI: A comparative practical investigation," *IEEE Sensors Journal,* vol. 19, no. 7, pp. 2747-2757, 2018.

[9] S. Shi, S. Sigg, L. Chen, and Y. Ji, "Accurate location tracking from CSI-based passive device-free probabilistic fingerprinting," *IEEE Transactions on Vehicular Technology,* vol. 67, no. 6, pp. 5217-5230, 2018.

[10] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Watchers on the Wall: Passive Visible Light-Based Positioning and Tracking with Embedded Light-Sensors on the Wall," *IEEE Transactions on Instrumentation and Measurement,* 2019.

[11] D. Konings, N. Faulkner, F. Alam, E. M.-K. Lai, and S. Demidenko, "FieldLight: Device-Free Indoor Human Localization Using Passive Visible Light Positioning and Artificial Potential Fields," *IEEE Sensors Journal,* vol. 20, no. 2, pp. 1054-1066, 2019.

[12] A. Ramezani Akhmareh, M. T. Lazarescu, O. Bin Tariq, and L. Lavagno, "A tagless indoor localization system based on capacitive sensing technology," *Sensors,* vol. 16, no. 9, p. 1448, 2016.

[13] X. Tang and S. Mandal, "Indoor occupancy awareness and localization using passive electric field sensing," *IEEE Transactions on Instrumentation and Measurement,* vol. 68, no. 11, pp. 4535-4549, 2019.

[14] M. Andries, O. Simonin, and F. Charpillet, "Localization of humans, objects, and robots interacting on load-sensing floors," *IEEE Sensors Journal,* vol. 16, no. 4, pp. 1026-1037, 2015.

[15] N. Faulkner, B. Parr, F. Alam, M. Legg, and S. Demidenko, "CapLoc: Capacitive Sensing Floor for Device-Free Localization and Fall Detection," *IEEE Access,* vol. 8, pp. 187353-187364, 2020.

[16] J. Zhao, N. Frumkin, P. Ishwar, and J. Konrad, "CNN-Based Indoor Occupant Localization via Active Scene Illumination," in *2019 IEEE International Conference on Image Processing (ICIP)*, 2019: IEEE, pp. 2636-2640.

[17] E. A. Wan and A. S. Paul, "A tag-free solution to unobtrusive indoor tracking using wall-mounted ultrasonic transducers," in *2010 International Conference on Indoor Positioning and Indoor Navigation*, 2010: IEEE, pp. 1-10.

[18] Y. Guo and M. Hazas, "Localising speech, footsteps and other sounds using resource-constrained devices," in *Proceedings of the 10th ACM/IEEE International Conference on Information Processing in Sensor Networks*, 2011: IEEE, pp. 330-341.

[19] M. Mirshekari, S. Pan, J. Fagert, E. M. Schooler, P. Zhang, and H. Y. Noh, "Occupant localization using footstep-induced structural

vibration," *Mechanical Systems and Signal Processing,* vol. 112, pp. 77-97, 2018.

[20] S. Narayana, R. V. Prasad, V. S. Rao, T. V. Prabhakar, S. S. Kowshik, and M. S. Iyer, "PIR sensors: Characterization and novel localization technique," in *Proceedings of the 14th international conference on information processing in sensor networks*, 2015, pp. 142-153.

[21] Y. Li, D. Li, Y. Cheng, G. Liu, J. Niu, and L. Su, "A novel human tracking and localization system based on pyroelectric infrared sensors: demonstration abstract," in *Proceedings of the 15th international conference on information processing in sensor networks*, 2016, pp. 1-2.

[22] L. Wu, Y. Wang, and H. Liu, "Occupancy detection and localization by monitoring nonlinear energy flow of a shuttered passive infrared sensor," *IEEE Sensors Journal,* vol. 18, no. 21, pp. 8656-8666, 2018.

[23] A. D. Shetty, B. Shubha, and K. Suryanarayana, "Detection and tracking of a human using the infrared thermopile array sensor—"Grid-EYE"," in *2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*, 2017: IEEE, pp. 1490-1495.

[24] 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)*, 2012: IEEE, pp. 2042-2047.

[25] M. Kuki, H. Nakajima, N. Tsuchiya, K. Kuramoto, S. Kobashi, and Y. Hata, "Mining multi human locations using thermopile array sensors," in *2013 IEEE 43rd International Symposium on Multiple-Valued Logic*, 2013: IEEE, pp. 59-64.

[26] 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.

[27] H. M. Ng, "Human localization and activity detection using thermopile sensors," in *2013 ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2013: IEEE, pp. 337-338.

[28] C. Kowalski, K. Blohm, S. Weiss, M. Pfingsthorn, P. Gliesche, and A. Hein, "Multi Low-resolution Infrared Sensor Setup for Privacy-preserving Unobtrusive Indoor Localization," in *Proceedings of the 5th International Conference on Information and Communication Technologies for Ageing Well and e-Health - Volume 1: ICT4AWE*, Heraklion, Crete, Greece, 2019, pp. 183-188.

[29] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural Networks for Indoor Person Tracking With Infrared Sensors," *IEEE Sensors Letters,* vol. 5, no. 1, pp. 1-4, 2021.

[30] S. Narayana *et al.*, "LOCI: Privacy-aware, Device-free, Low-power Localization of Multiple Persons using IR Sensors," in *2020 19th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*, 2020: IEEE, pp. 121-132.

[31] S. Singh and B. Aksanli, "Non-Intrusive Presence Detection and Position Tracking for Multiple People Using Low-Resolution Thermal Sensors," *Journal of Sensor and Actuator Networks,* vol. 8, no. 3, p. 40, 2019.

[32] 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.

[33] L. Tao, T. Volonakis, B. Tan, Z. Zhang, and Y. Jing, "3D convolutional neural network for home monitoring using low resolution thermal-sensor array," in *3rd IET International Conference on Technologies for Active and Assisted Living (TechAAL 2019)*, London, 2019: IET, pp. 1-6.

[34] M. Gochoo *et al.*, "Novel IoT-based privacy-preserving yoga posture recognition system using low-resolution infrared sensors and deep learning," *IEEE Internet of Things Journal,* vol. 6, no. 4, pp. 7192-7200, 2019.

[35] R. Zekavat and R. M. Buehrer, *Handbook of position location: Theory, practice and advances*. John Wiley & Sons, 2011.

[36] A. H. A. Bakar, T. Glass, H. Y. Tee, F. Alam, and M. Legg, "Accurate visible light positioning using multiple photodiode receiver and machine learning," *IEEE Transactions on Instrumentation and Measurement,* 2020.

[37] D. Chetverikov, S. Fazekas, and M. Haindl, "Dynamic texture as foreground and background," *Machine vision and Applications,* vol. 22, no. 5, pp. 741-750, 2011.

[38] S. J. Prince, *Computer vision: models, learning, and inference*. Cambridge University Press, 2012.

[39] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375,* 2018.

NATHANIEL FAULKNER received his B.E. (Hons.) from Massey University, New Zealand in Electronics & Computer Engineering in 2016. He is currently pursuing a Ph.D. at the same institution. In 2020 he was also appointed as a Visiting Research Fellow at the School of Engineering and Technology, Sunway University, Malaysia. His research interests include indoor positioning, embedded systems design, and the Internet of Things.

FAKHRUL ALAM (M'17-SM'19) is an Associate Professor at the Department of Mechanical & Electrical Engineering, School of Food & Advanced Technology, Massey University, New Zealand. He has been appointed as Adjunct Professor with the School of Engineering and Technology of Sunway University, Malaysia for 2021-22. He received BSc (Hons) in Electrical & Electronic Engineering from BUET, Bangladesh, and MS and Ph.D. in Electrical Engineering from Virginia Tech, USA. His research interest includes indoor localization, 5G & visible light communication, IoT & wireless sensor networks.

MATHEW LEGG (M'19) Received his B.Sc, M.Sc, and Ph.D. in Physics from the University of Auckland, New Zealand. He is currently a senior lecturer with the Department of Mechanical and Electrical Engineering, School of Food and Advanced Technology, Massey University. His research relates to the development of acoustic/ultrasonic measurement systems and techniques for acoustic imaging, non-destructive testing, and remote sensing.

SERGE DEMIDENKO (M'91-SM'94-F'04) is Professor and Dean of the School of Engineering and Technology, Sunway University, Malaysia. He is also associated with the Department of Mechanical & Electrical Engineering, School Food and Advanced Technology, Massey University, New Zealand. He graduated in Computer Engineering from the Belarusian State University of Informatics and Radio Electronics and received Ph.D. from the Institute of Engineering Cybernetics of the Belarusian Academy of Sciences. His research interests include electronic design and test, signal processing, instrumentation and measurements. He is also Fellow of IET and UK Chartered Engineer.

CHAPTER 5 – DFL USING CEILING-MOUNTED THERMOPILES

This chapter is republished in accordance with IEEE's copyright policy. This chapter is the accepted version of the published work, and as such may have stylistic differences from the final published article.

Machine Learning Techniques for Device-Free Localization Using Low-Resolution Thermopiles

Nathaniel Faulkner, Daniel Konings, *Member, IEEE*, Fakhrul Alam, *Senior Member, IEEE*, Mathew Legg, *Member, IEEE* and Serge Demidenko, *Fellow, IEEE*

Abstract—**Indoor Device-Free Localization (DFL) has many uses including aged care, location-based services, ambient assisted living, and fire safety management. In recent publications, thermopile sensors (very low-resolution infrared cameras) have been shown as being able to localize individuals whilst preserving their privacy. This paper reports the performance evaluation of a large number of supervised machine learning techniques for the localization of a target using a ceiling-mounted thermopile. The algorithms were trained and validated using a large dataset constructed from an individual walking arbitrary paths with the accurate ground truth provided by a virtual reality system. For robust performance evaluation, the algorithms were tested with datasets collected on a different day with several other subjects. A 2D Convolutional Neural Network exploiting spatial correlation and several Recurrent Neural Network structures exploiting temporal correlation among the captured data provided the most accurate localization performance. Several datasets, constructed from the thermopile's readings for four individual targets, were made available online for other researchers to use.**

Index Terms—**Device-Free Localization (DFL), human sensing, Indoor Positioning System (IPS), passive localization, thermopile, infrared sensing, machine learning, neural network, supervised learning, LSTM, CNN**

I. INTRODUCTION

PASSIVE indoor localization, otherwise known as Device-Free Localization (DFL) [1] has uses in security, assisted living for older adults, consumer habit tracking in commercial venues amongst others. There is currently no *de facto* solution for DFL, with it being an active area of research during recent years. This is also driven by the rapid emergence of the Internet of Things (IoT) [2]. Standard video cameras, coupled with state-of-the-art computer vision advancements [3], could provide a large volume of high-quality information streams. However, this comes at the cost of significant processing power requirements. It is also associated with potential privacy issues, especially when employed in a residential setting. Finally, it normally requires adequate lighting. Less invasive DFL solutions have been widely investigated using a range of technologies split between wireless Radio Frequency (RF) and non-wireless approaches (please see review articles [4] and [5] respectively). DFL solutions typically rely on detecting the changes a subject makes to an environment (e.g., shadowing RF links, causing vibrations when walking, etc.). It allows the subject to be localized without the need to carry any special devices. Whilst there are many examples of the relevant solutions reported in the literature, there is still no singular universally adopted technology for indoor DFL. A robust practical solution for indoor DFL is likely to be multi-modal (i.e., utilizing multiple technologies). This will likely include fusing RF sensing with other non-wireless technologies. Wireless DFL is a mature, well-investigated technique. Non-wireless approaches have not yet been adequately explored thus leading to the need to study thermopile and other non-RF localization techniques.

Human subjects produce Infrared (IR) radiation, as their body is normally at a higher temperature than the ambient environment. There are two main ways this IR radiation can be detected: by using a Passive Infrared (PIR) sensor [6] or by employing thermopiles [7]. An advantage of PIR sensors is their affordability. However, they are only able to sense a temperature change. Therefore, they can detect only moving subjects unless significant enhancements are made (e.g., rotating sensors by using motors [8] or adding electromechanical shutters [9]).

In contrast to PIR sensors, thermopiles have the advantage of being able to acquire absolute temperature values and thus to

N. Faulkner, D. Konings, M. Legg are all with the Department of Mechanical & Electrical Engineering (MEE), School of Food & Advanced Technology (SF&AT), Massey University, Auckland 0632, New Zealand (n.faulkner@massey.ac.nz, d.konings@massey.ac.nz, m.legg@massey.ac.nz).

F. Alam is with the Department of Mechanical & Electrical Engineering (MEE), School of Food & Advanced Technology (SF&AT), Massey University, Auckland 0632, New Zealand and as Adjunct Professor with the School of Engineering & Technology, Sunway University, Bandar Sunway 47500, Malaysia (f.alam@massey.ac.nz)

S. Demidenko is with the School of Engineering & Technology, Sunway University, Bandar Sunway 47500, Malaysia, and (as a co-supervisor) with the Department of MEE of SF&AT, Massey University, Auckland 0632, New Zealand (sdemidenko@sunway.edu.my, s.demidenko@massey.ac.nz).

inherently detect a stationary target. Thermopile sensors, such as the commonly used Panasonic AMG8833[10] can detect a human target at a distance of up to 7 m in ideal conditions. This led to their use in a variety of human sensing applications including fall detection [11], pose recognition [12], occupancy prediction [13], and DFL. Research works on human target localization using thermopiles utilized both wall-mounted [14, 15] and ceiling-mounted [16-19] sensors. A large proportion of the reported works utilized ceiling-mounted sensors. Such an arrangement is not adversely impacted by furniture and other obstacles present in the environment. It allows the sensor to establish a clear line of sight with a target. Due to these advantages, the ceiling-mounted sensor configuration is employed in this reported research.

The basic premise for positioning is as follows: the sensor is sampled at regular intervals, with the data then being fed through a preprocessing pipeline for noise reduction, background removal, and normalization. The data itself represents the temperature measurements within the coverage area of each pixel. Therefore, a human subject is manifested by a higher temperature area in the data. The challenge is then to map the position of such hotspots within the data frame to the real-world position of the target. This is illustrated in Fig. 1.

Thermopile based localization is accurate compared to other localization techniques [20-24] as can be seen in Table I. Thermopile based localization can complement the RF employing DFL. The RF solution can provide localization estimates that work through walls. At the same time, the thermopiles can provide more accurate positioning for line-of-sight targets. A future fused approach utilizing both these two technologies would improve the robustness of localization systems compared to a solution built on the use of just a single one. Another advantage of the thermopile based solution is its modular aspect: just a single thermopile is sufficient to localize/track a subject within (albeit small) an area. The coverage area can be extended by adding additional thermopiles as required. With RF-based and many other localization techniques, multiple sensors are required for lateration, angulation or for creating enough "links" for the target to affect regardless of how small the intended coverage area is. Considering the coverage of a single thermopile and the possibility of scaling up coverage in a modular fashion, the thermopile based localization allows for a sparser sensor density compared to other techniques (as highlighted in Table I).

A. Related Works

Kuki et al. [16] were able to localize a subject to a mean accuracy of 0.215 m with a sensor placed 2.73 m above the ground. Connected component blobs detection [25] was used to find the position of probable human subject in the frame. The positions of the centroids of these blobs within the frame were then mapped to the position of the subject within the room. Qu et al. [17] were able to achieve 0.07 m mean accuracy, with a sensor placed at 3.5 m above the ground. In a fashion similar [16], the centroids of the connected components were used, with a Kalman filter added for tracking/smoothing.
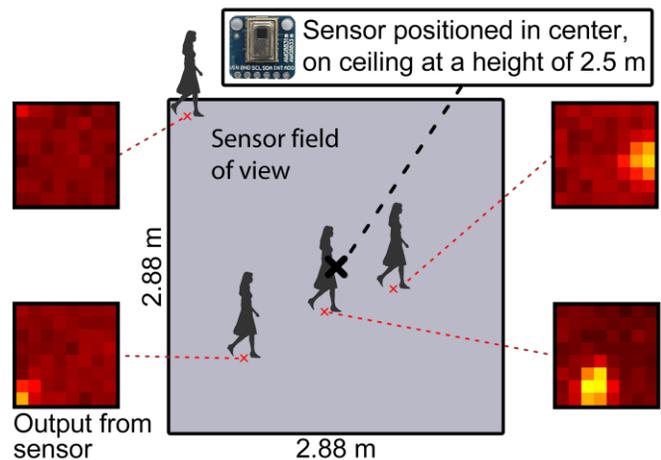

Fig. 1. Top-down view of test area demonstrating the readings from the sensor when the subject is at different positions.

Unfortunately, the authors did not specify the number of samples collected or the time period over which the samples were acquired. Gu et al. [18] used a higher resolution sensor (24 x 32 pixels) to achieve a Root Mean Square Error (RMSE) level of 0.05–0.2 m. They tracked the trajectory of connected components. Two subjects were used for multi-person positioning. However, it was not specified in the report whether both subjects were used for the single person position testing. Besides, the number of trials and lengths of the movement were not given. Tariq et al. [19] had a subject walking about in a natural fashion for 30 minutes to collect 9000 data points. These data were split into three parts: 60% - for training, 20% - for validation, and the remaining 20% - for testing. Four neural networks were employed and compared, achieving an RMSE of 0.096 m. However, the localization accuracy was only evaluated on the test split of the initial dataset, and not tested on other subjects or using data taken at a later date. Thus, the real-world performance of the system, likely to suffer degradation, was not presented and objectively confirmed.

Accurate ground truth is vital in the evaluation of a positioning system. To quantify the localization error of the system, one first needs to have an accurate knowledge of the subject's position before computing the error of the subject's estimated position. A small error in ground truth measurement can cause significant inaccuracy in positioning evaluation [26]. Walking along a predetermined path is a common method of establishing the ground truth [17, 18]. However, this can lead to a less natural walking style. It also restricts the coverage of the area where a subject can walk. Kuki et al. [16] used a camera to record the video of their experiments. The authors then manually extracted the ground truth from the video. However, such an approach would appear to be time-consuming and could lead to manual processing errors. These limitations in ground truth recording restrict the ease with which a large dataset can be collected, leading to small dataset sizes. Consequently, such systems are trained/calibrated and tested on a small set of paths, lowering the robustness of the models and the analysis of such models. Tariq et al. [19] used a commercial-off-the-shelf (COTS) ultrasonic solution for recording ground truth. It

TABLE I: Comparison Between The Proposed Thermopile Based Localization And Approaches Presented In The Literature.

Method	Median Accuracy	Coverage Per Sensor	Remarks
WoW (Visible Light Positioning) [20]	0.12 m	0.72 sqm/sensor (14 sensors for 2.8m x 3.6m)	Requires higher sensor density than in the proposed solution.
SpringLoc (Wireless RSSI) [21]	~0.6 m	1.25 sqm/sensor (20 Zigbee sensors to cover 5m x 5m)	Requires higher sensor density than in the proposed solution. Accuracy is worse than in the proposed solution.
Shi et al. (Wireless CSI) [22]	~0.6 m	3.1-11.9 sqm/sensor (4 transmitting and 3 receiving nodes for areas from 1.8m x 12m to 11.6m x 7.2m)	Accuracy is worse than in the proposed solution.
Mirshekari et al. (Vibration) [23]	0.38 m	2.2 sqm/sensor (9 geophones for 4m x 5m area)	Requires higher sensor density than in the proposed solution. Accuracy is worse than proposed.
Yang et al. (PIR) [24]	0.21 m (mean)	8.6 sqm/sensor (10 sensors for 10m x 7.2m)	Comparable sensor density to the proposed solution. Accuracy is worse than in the proposed solution.
Proposed solution	0.027 m for the same dataset 0.16 m for other datasets	8.3 sqm with just one sensor	Coverage can potentially be extended in a modular manner by adding more sensors.

allowed for the natural movement of subjects and larger datasets. However, the dataset was hindered by the limited accuracy of the ground truth system (average - 39 mm, maximal - 64 mm). The accuracy of the ground truth system employed in the research, being of the same order as the results reported, made the reported error statistics less robust. Walking along a predetermined path, as reported in many articles, can also produce unwanted artifacts, especially where the model uses a sequence of readings. This is because the model can unintentionally learn the pattern of the predetermined path, and as such will match a subject to the path very well. It can have poor real-world performance unless many paths are tested. Ground truth systems allowing subjects to walk about a test area in a natural manner while not following any particular path would ensure that the model does not learn any specific path geometry.

B. Contribution

In light of the discussion presented in the previous section, the state of the art of thermopile-based localization have the following deficiencies:

1. **Inadequate ground truth**: walking along a predetermined path forces subjects to walk in an unnatural manner. In addition to this, there is potential for the positioning models to use the geometry of the path as a feature, which would not be representative of a real-world scenario.
2. **Insufficient description of the data collection process:** the number of participants, the number of samples are not always disclosed, making it impossible to perform fair benchmarks against existing works.
3. **Lack of robust performance evaluation:** using a single participant in a single environment with a single dataset split for both training and testing. In reality, we have found that small changes to the environment lead to reduced localization accuracy. Using multiple datasets for testing helps to evaluate and justify the real-world performance of a system.
4. **Lack of data:** there is no robust dataset available for researchers to train and benchmark against.

To address the deficiencies, this paper utilizes experimental data taken across multiple environments while employing four different test subjects. Accurate ground truth is automatically recorded by using an HTC Vive Virtual Reality (VR) system[1]. This also allows the subjects to retain natural walking motion whilst roaming the test area in an arbitrary manner (i.e., without the requirement to follow a small number of predefined trajectories). The investigation carried out and reported is, therefore, underpinned by a rigorous data collection and structured measurement methodology. A wide range of supervised Machine Learning (ML) techniques [27] have been trained and tested upon the dataset. The novelty and the contributions of the work can be summarized as follows.

1. **Devising and offering the first thermopile dataset thus addressing the lack of such datasets in the public domain.** To the best authors' knowledge, no dataset is currently available for thermopile-based localization. Datasets are commonly provided as a novel contribution as they allow researchers to train and evaluate localization algorithms using experimental data without developing bespoke hardware and implementing their own testbeds. Such datasets also allow fair, "apple-to-apple" comparison and benchmarking between different solutions. The comprehensive dataset, collected with a ceiling-mounted thermopile for multiple subjects, on three different occasions has been made available for a wide research audience.
2. **Presenting one of the first comprehensive studies of the performance of Machine Learning algorithms for**

---

[1] https://developer.vive.com/eu/vive-tracker-for-developer/

thermopile localization. The only published work investigating the performance of machine learning for thermopiles known to the authors is by Tariq et al. [19]. That study used a much smaller dataset compared to the one developed and employed in the reported research – it was collected using a single subject with the walking of approximately 30 minutes. Unfortunately, that dataset has not been made publicly available. Besides, the investigation was somewhat limited to the performance of just four neural networks. In contrast, the presented work covers a large number of neural networks and other supervised ML techniques. The results demonstrate that ML approaches capable of leveraging spatial and temporal correlation provide the highest accuracy.

3. **Outlining the first thermopile based localization work to utilize multiple subjects for performance evaluation.** To the best knowledge of the authors, this article presents the first thermopile localization work to include different subjects that were used for training for performance evaluation. Additionally, tests were performed using both light and heavy apparel. These measurements allowed the generalizability of the algorithms to be verified and provided insights into issues that can occur in real-world scenarios. The presented results show that there is a clear performance degradation when testing while employing different subjects and doing it with prolonged time breaks between the sessions. Such effects have not been investigated in the reported literature on thermopile based localization and were rather sparingly covered within the localization discipline.

The remainder of this paper is structured as follows: Section II discusses the system overview and methodology; Section III presents the results of the different machine learning methods on the datasets; Section IV provides conclusions and recommendations for future research.

II. SYSTEM OVERVIEW

The system utilizes the 64-pixel Panasonic AMG8833 infrared array sensor combined with a custom-designed interface to a logging computer. The sensor unit is mounted on the ceiling (2.5 m above the floor level), facing directly downwards. The sensor has an onboard thermistor with a resolution of 0.0625 °C. The sensor provides a viewing angle of 60° and a Noise Equivalent Temperature Difference (NETD) of 0.16 °C. The sensor is sampled at 10 Hz.

The ground truth is provided by the HTC Vive tracker unit attached to the top of the subject's head. The 2D X,Y ground truth is aligned with the participant's center of mass. The Vive ground truth data is collected at 60 Hz. The data are then resampled to 10 Hz to match the infrared sensor rate. Glass et al. [28] showed that the HTC Vive can provide high-quality ground truth recordings with a precision of 0.65 mm and an accuracy of 5 mm. The accuracy and reliability of the Vive system were also highlighted in [29].

A. Data Collection

The test area is a 5 m by 5 m square, where subjects are able to roam freely during the data collection. Of this, a 2.88 m by 2.88 m square is within the sensor's Field of View (FOV) with the sensor being affixed to the ceiling above the center of the area (Fig. 2). It should be noted that the detection area can be further extended by increasing the deployment height for the sensor (e.g., see Qu et al. [17]). Data is collected from the sensor as a subject moves around the test area whilst positions of the subject are logged from the HTC Vive. The data, once preprocessed, can then be used to train and evaluate ML-based regression models. The inputs to the model are the processed temperature values from each of the 64 pixels of the sensor. The output is the estimated X,Y coordinates of the subject within the room. The model can be evaluated by comparing the mean-squared error of the Euclidean distance between the estimated position from the model to the actual position measured from the HTC Vive. Models can be compared by using error statistics such as the median, 95-percentile and RMSE of the position errors as well as by using Empirical Cumulative Density Function (ECDF) plots of the position errors. This process is illustrated in Fig. 3. Multiple error metrics were chosen for benchmarking as per recommendations by ISO/IEC 18305 [30], EvAAL framework [31], and the standard practices given in the literature. A detailed discussion on performance metrics for DFL can be found in [32].

Six datasets were collected. The first (Dataset 1) was the largest. Its aim was to be used as a training set. To collect it, a male subject (henceforth referred to as Subject 1) walked around the test area for 50 minutes, equating to approximately 30,000 data points (of which around 20,000 were within the sensor's FOV). The samples associated with positions outside the sensors FOV were removed. Then the dataset was split into three smaller sets. The first subset (80% of the data) was used for training, the second (10%) - for validation, while the final (10%) - for testing. The split was done in contiguous time blocks, so that relationship between sequential frames of data could be leveraged by memory dependent supervised machine learning techniques. That also ensured reasonable temporal separation among the three subsets.

To investigate how the trained models perform on different days and with different subjects, four additional datasets (named Datasets 2-5) were collected approximately two weeks later. They were shorter (each of them corresponded to 5-10 minutes of a subject's movement). They featured the same Subject 1 (Dataset 2), as well as three new subjects: male Subject 2 (Dataset 3), male Subject 3 (Dataset 4), and female Subject 4 (Dataset 5). By using them as test sets for the models trained on Dataset 1, it became possible to evaluate how well each trained model generalizes to changes in the environment and previously unseen subjects. The final dataset featured Subject 1 in a heavy coat was taken on the same afternoon as Dataset 1. It is referred to as Dataset 6. The Datasets 1-6 are available online[2] along with instructions on how to load and use them.

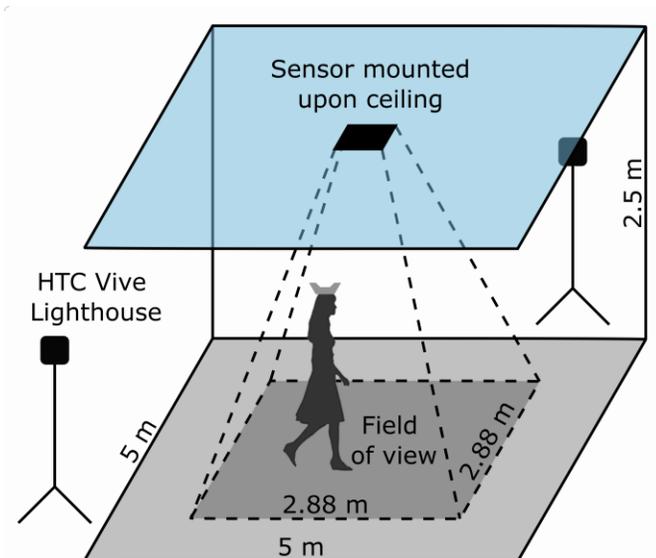[2] https://github.com/natfaulk/open-thermopile-dataset

Fig. 2. Room layout, showing the subject with the HTC Vive tracking puck upon their head. Also demonstrated is the thermopile sensor's field of view within the larger test area.
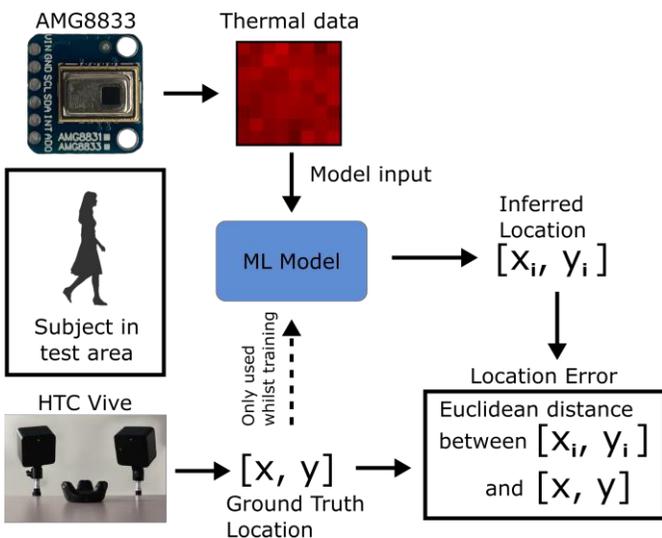

Fig. 3. System overview: Thermopile and ground truth data are captured for training the ML models and evaluating the localization error. The models are trained to infer the (x,y) location of the target. The Euclidean error between the inferred location and ground truth is used as the localization error during evaluation. The mean-squared Euclidean error between the inferred location and ground truth is used for model hyperparameter tuning.

For each machine learning approach, a Bayesian optimizer was used to tune the hyperparameters to attain the best performance. The two-dimensional (X,Y) Euclidean error of the validation set was used as the Bayesian objective function's minimization target. This was chosen as it maintains separation between the train and test sets, while also minimizing the chance of overtraining on the training set. An overview of the training approach is presented in Fig. 4. Table II shows the final hyperparameter values.

B. Data Preprocessing

After collecting the datasets, Dataset 1 was preprocessed before being used to train the machine learning approaches. Data for each of the 64 individual pixels can be taken as an independent time series for that pixel. There is a reasonable
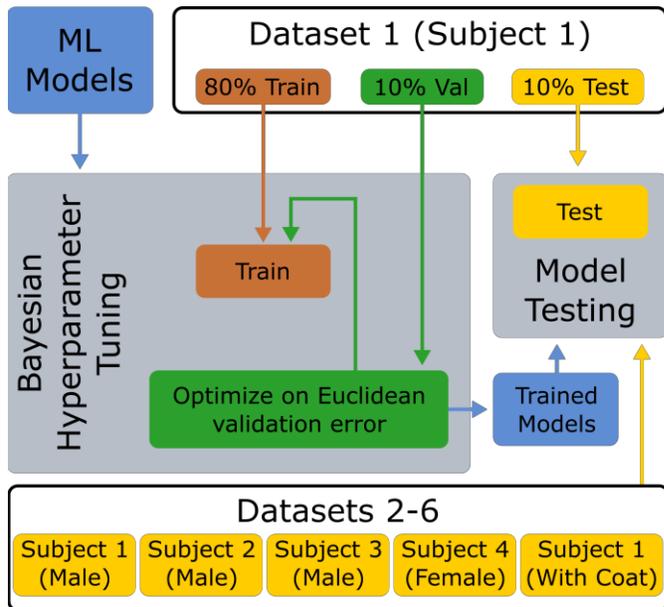

Fig. 4. Datasets and hyperparameter tuning.

amount of noise from sample to sample. Hence, denoising filters were tested – such as median filters of 3-6 sample lengths alongside a linear Gaussian filter with several different kernel sizes. Through empirical testing, it was found that a four-sample median filter displayed the best performance. Since this filter uses one sample ahead of the current time at each timestep, it is non-causal. However, with a 10 Hz sampling rate, this only introduces a 0.1 second delay if used in a real-time system with live data. This remains realistic for real-time human location reporting [33]. Single Value Decomposition (SVD) [34] is used to reduce the influence of background noise by zeroing the first component of the diagonal matrix before the data are reconstructed and normalized between zero and one. Fig. 5 shows a block diagram of the data preprocessing pipeline (it is utilized before using each dataset). A more detailed explanation of this process can be found in the earlier work [15].

C. Tested Supervised Learning Techniques

Machine learning techniques have been used for various localization works for over a decade [35] with the success across multiple technologies including Wi-Fi RSS [36], Wi-Fi CSI [37, 38], Bluetooth RSS [39], and visible light RSS [40]. There are many supervised learning techniques, and the ones chosen were broadly divided into three main groups: classical regression, ensemble learning, and neural networks. The classical regression category included: Linear, Ridge (linear with L2 penalty), Lasso (linear with L1 penalty), Elastic Net (linear with L1 + L2 penalty), Support Vector Regression (SVR), and K-Nearest Neighbors (KNN) [41]. In the ensemble learning category, the following techniques were investigated: Random Forest, Bagged Ensemble, and Boosted Ensemble [42]. Finally, the Neural Network category included the Multilayer Perceptron (MLP), three implementations of Convolutional Neural Network (CNN), and several Recurrent Neural Network (RNN) structures, e.g., Long Short Term Memory (LSTM), Bidirectional LSTM (BLSTM), and Gated Recurrent Unit (GRU) [43]. A cascade structure of a CNN

Algorithm	Hyperparameter Range	Final Hyperparameter Values
1D-CNN	Section depth: [1, 3] Filter size: [2, 7] Pooling: [1, 5] Number of filters: 2^[3, 8]	Section depth: 3 Filter size: 3 Pooling: 5 (1x5) Number of filters: 128
2D-CNN	Section depth: [1, 3] Filter size: [2, 7] Pooling: [1, 5] Number of filters: 2^[3, 8]	Section depth: 3 Filter size: 7 (7x7) Pooling: 1 (No pooling) Number of filters: 16
3D-CNN	Section depth: [1, 4] Filter size: [2, 7] Filter depth: [1, 7] Pooling: [1, 5] Number of filters: 2^[3, 8]	Section depth: 4 Filter size: 4 Depth: 5 (4 x 4 x 5) Pooling: 3 (3x3x3) Number of filters: 16
MLP	Section depth: [2, 4] Neurons: 2^[6, 10] Dropout: [0, 0.3]	Section depth: 3 Neurons: 1024 Dropout: 0.025
CNN-LSTM	Section depth: [1, 3] Filtersize: [2, 7] Dropout: [0, 0.5] Number of filters: 2^[3, 8] Number of hidden units (1): 2^[4, 9] Number of hidden units (2): 2^[4, 8]	Section depth: 3 Filtersize: 2 Dropout: 0 Number of filters: 8 Number of hidden units (1): 64 Number of hidden units (2): 16
LSTM	Number of hidden units (1): 2^[4, 9] Number of hidden units (2): 2^[4, 8] Dropout: [0, 0.5]	Number of hidden units (1): 256 Number of hidden units (2): 32 Dropout: 0
BLSTM	Number of hidden units (1): 2^[4, 9] Number of hidden units (2): 2^[4, 8] Dropout: [0, 0.5]	Number of hidden units (1): 256 Number of hidden units (2): 32 Dropout: 0
GRU	Number of hidden units (1): 2^[4, 9] Number of hidden units (2): 2^[4, 8] Dropout: [0, 0.5]	Number of hidden units (1): 256 Number of hidden units (2): 64 Dropout: 0
Bagged Ensemble	Number of learning cycles: [10, 1000] Minimum leaf size: [1, (NumberObservations/2)] Maximum number of splits: [1, (NumberObservations -1)] Number of variables to sample: [1, 64]	Number of learning cycles: 939 Minimum leaf size: 1 Maximum number of splits: 7973 Number of variables to sample: 44
Boosted Ensemble	Learnrate: [0.001, 1] Minimum leaf size: [1, (NumberObservations/2)] Maximum number of splits: [1, (NumberObservations -1)] Number of Learning cycles: [10, 1000]	Learnrate: 0.10434 Minimum leaf size: 1 Maximum number of splits: 13567 Number of Learning cycles: 494
Random Forest	Minimum leaf size: [1, (NumberObservations/2)] Maximum number of splits: [1, (NumberObservations -1)] Number of trees: [10, 100]	13 1509 40
Ridge	Alpha: [0.001, 10]	5.5
Lasso	Alpha: [0.0001, 1]	0.00013
ElasticNet	Alpha: [0.0001, 1] L1 Ratio: [0.1, 0.9]	0.00026 0.4
Weighted KNN	K: [1, 30] Weighting Metric: [None, Euclidean, Manhatten, Chebyshev, Canberra]	22 Chebyshev
SVR	Kernel: [Radial Basis Function, Linear, Sigmoid] C: [0.01, 100] Epsilon: [0.001, 10]	Radial Basis Function 10 0.001

followed by an LSTM, CNN-LSTM, was also considered. The structures of the neural networks are shown in Fig. 6 -13.

D. Neural Network Models

All the neural networks used an output structure consisting of two fully connected neurons, followed by a regression layer computing the mean-squared-error loss. The 2D-CNN follows a standard structure. It convolves across the 8 pixels by 8 pixels input image while extracting spatial features. The 1D- and 3D-CNN convolve across time by taking a 1 second (10 samples) consecutive input and predicting the location corresponding to the 5th sample. The 3D-CNN can obtain both spatial and temporal features (as it convolves over both the 8 pixels by 8 pixels input frames and consecutive samples in time). In contrast to its 2D counterpart, the 1D-CNN only obtains temporal features. It treats each pixel separately and is convolving over 10 samples. This means that technically the 1D- and 3D-CNN are non-causal as they use future information. However, in practice, it is realized as a 5 sample (0.5 second) delay in the output which remains practical for most indoor human movement tracking. The CNN and MLP based models contain section depth hyperparameters relating to the number of 'blocks' the network contains. For example, a CNN block is made up of a convolutional layer, a batch normalization layer, and a ReLU layer [44].

III. RESULTS

There is a large disparity in the complexity of the models. Some (especially, more classical regression models) require
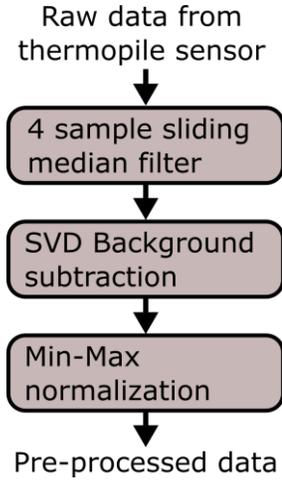
Fig. 5. Block diagram of the data preprocessing pipeline.

Fig. 6. 1D CNN layout.

Fig. 7. 2D CNN layout.

Fig. 8. 3D CNN layout.

Fig. 9. MLP layout.

both a very low training effort and very low complexity for inference of new data. The neural networks whilst offering better performance, require a significantly more expensive training effort. When comparing the models, the test split from Dataset 1 has been compared along with Datasets 2-5 that have been grouped together for easier comparison. The grouping was achieved by applying the model to each dataset and then combining the errors from each into one large set of errors. The performance metrics (median, 95-percentile) were then computed from this large, combined error set. The results for each model have been summarized in Table III.

A. Classical Regression models

The ECDF curves for the classical regression models can be seen in Fig. 14. Aside from the KNN, the classical regression models did not perform well, having high position errors. When tested against the test split of Dataset 1, the linear regression-based models (Linear, Ridge, Lasso, and ElasticNet) had almost identical performance with a median position error of ~0.60 m and a 95-Percentile position error of ~1.26-1.27 m. However, a big difference was found when testing against the other datasets (the more general data). The accuracy degraded drastically resulting in very large errors in position estimates. The linear models with regularization performed slightly worse against the other datasets than against the test split: a median position error of ~0.65-0.67 m and a 95-percentile error of ~1.53-1.54 m. This shows that whilst the positioning accuracy attained by the linear models was not very good, as long as there was some regularization to avoid overfitting, they generalized reasonably well. At the same time, since their performance was very poor, the generalizability of the models was somewhat immaterial. The weighted KNN model performed much better - achieving a median position error of 0.09 m and 95-percentile of 0.28 m on the test split of Dataset 1. This error increased when tested against the other datasets: to 0.26 m median and 1.27 m 95-percentile. SVR performed better than the linear regression models. However, it was worse than the KNN on the test split of Dataset 1 and on the remaining test sets. A median positioning error of 0.23 m and 95-percentile of 0.84 m was achieved for the test split of Dataset 1, and a median of 0.45 m
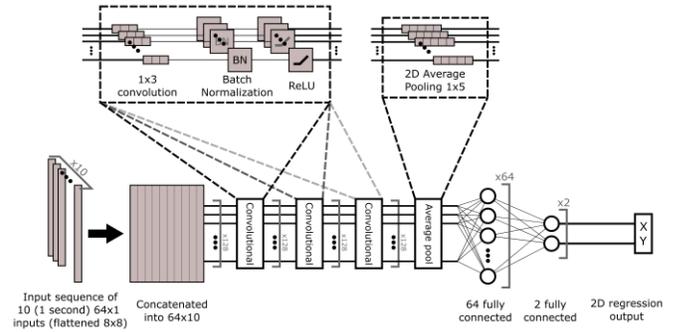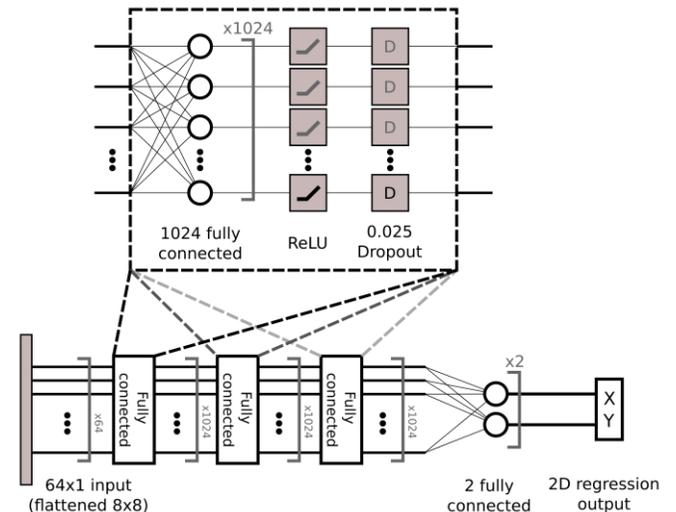
Fig. 10. LSTM layout.
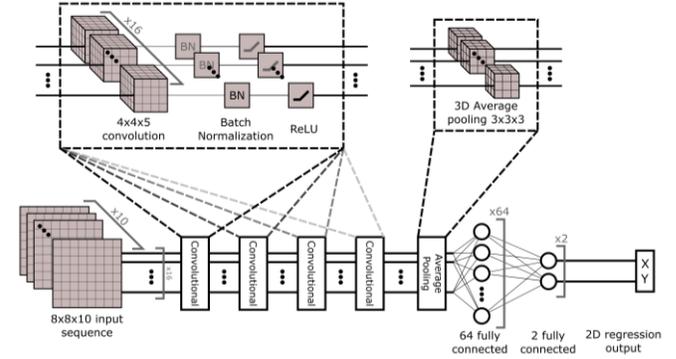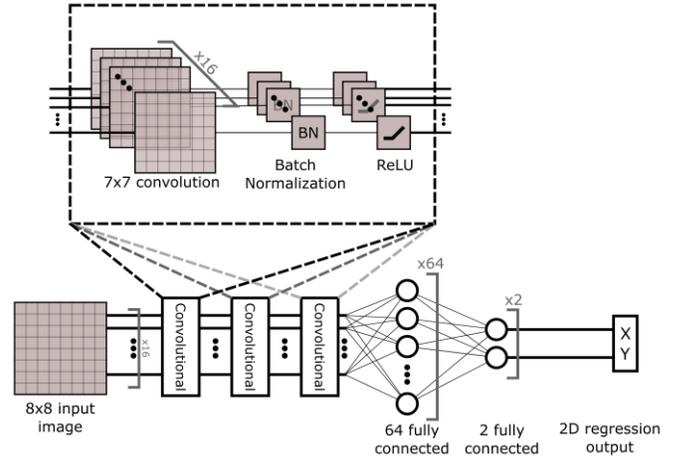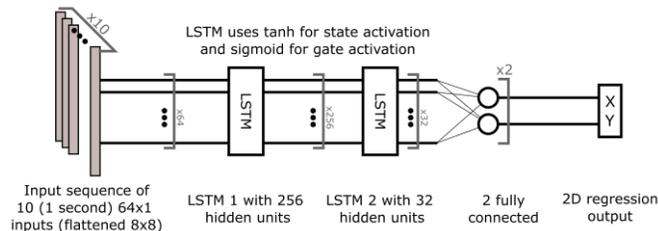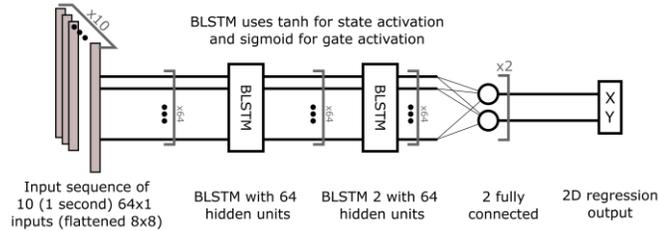

Fig. 11. BLSTM layout.
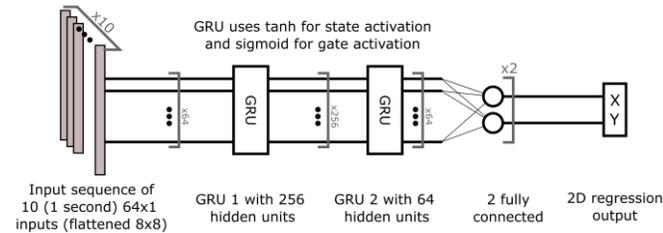
Fig. 12. GRU layout.

Fig. 13. CNN LSTM layout.

Fig. 14. ECDF plots for the classical regression models, both for (top) the Dataset 1 test split and (bottom) for Datasets 2-5 combined together. Note: For clarity, linear regression is omitted from the second plot due to the extraordinarily large errors.

Fig. 15. ECDF plots for the ensemble models, both for (top) the Dataset 1 test split and (bottom) for Datasets 2-5 combined together.
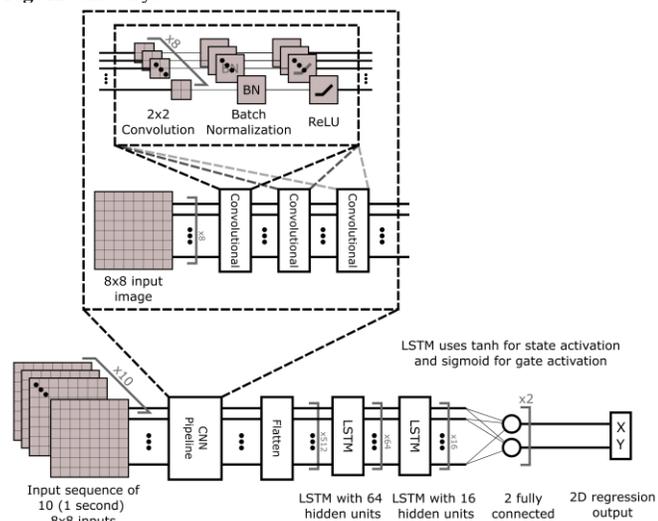
and a 95-percentile of 1.60 m - for the remaining datasets. The linear models are very fast to train and run. Similarly, KNN is very fast to run. At the same time, KNN requires a larger amount of memory due to its large database. However, this is unlikely to ever be an issue except in the most extreme memory-constrained environments.

B. Ensemble Learning

Three methods of ensemble learning were used: Random Forest, Bagged Trees, and Boosted Trees. The ECDF curves for these methods are shown in Fig. 15. Using Random Forest, a median position error of 0.14 m and 95-percentile one of 0.69 m were achieved when testing against the test split of Dataset 1. When tested against the other datasets, a median position error of 0.37 m and a 95-percentile error of 2.41 m were observed. Bagged Ensemble performed slightly better, with a median positioning error of 0.09 m and 95-percentile of 0.50 m against the test split of Dataset 1. Against the other datasets, the median position error was 0.35 m and the 95-percentile error was 1.85 m. The Boosted Ensemble performed slightly worse than the Bagged Ensemble (while being also slightly better than the Random Forest) with a median position error of 0.10 m and 95-percentile error of 0.48 m against the test split of Dataset 1. Against the other datasets, a median error of 0.36 m and a 95-percentile error of 2.15 m were observed.

Overall, the differences between the ensemble methods were quite small. Compared to the regularized linear regression models, the ensemble learning models had lower median

	Dataset 1 Test Split			Datasets 2-5			Dataset 6 (Heavy Clothing)		
	Median (m)	95-Percentile (m)	RMSE (m)	Median (m)	95-Percentile (m)	RMSE (m)	Median (m)	95-Percentile (m)	RMSE (m)
Classical Regressions									
Linear Regression	0.601	1.254	0.732	9×10^{11}	2×10^{12}	5×10^{12}	5×10^{12}	6×10^{12}	5×10^{12}
Ridge Regression	0.600	1.267	0.732	0.652	1.533	0.847	0.716	1.453	0.850
Lasso Regression	0.601	1.261	0.732	0.663	1.536	0.851	0.719	1.440	0.837
Elastic Net	0.600	1.263	0.731	0.656	1.532	0.848	0.714	1.449	0.840
Weighted KNN	0.094	0.278	0.219	0.257	1.273	0.528	0.185	0.474	0.329
SVR	0.233	0.842	0.415	0.447	1.599	0.762	0.477	1.221	0.645
Ensemble									
Random Forest	0.145	0.690	0.373	0.368	2.412	0.998	0.388	2.294	0.950
Bagged	0.098	0.509	0.250	0.347	1.845	0.793	0.654	1.703	0.908
Boosted	0.107	0.481	0.300	0.359	2.147	0.901	0.724	2.349	1.165
Neural Networks									
MLP	0.082	0.220	0.159	0.276	2.105	0.749	0.204	0.631	0.388
1D-CNN	0.275	0.630	0.350	0.357	1.028	0.524	0.580	1.361	0.765
2D-CNN	0.029	0.068	0.038	0.188	0.833	0.420	0.209	0.825	0.401
3D-CNN	0.086	0.186	0.107	0.235	0.943	0.468	0.355	1.302	0.624
LSTM	0.055	0.148	0.079	0.197	1.798	0.649	0.141	0.357	0.222
BLSTM	0.050	0.139	0.072	0.188	1.394	0.596	0.148	0.359	0.260
GRU	0.061	0.169	0.088	0.184	1.639	0.615	0.154	0.384	0.227
CNN-LSTM	0.027	0.068	0.036	0.162	0.758	0.458	0.248	0.846	0.424

position errors but worse 95-percentile errors. Weighted KNN significantly outperformed them all.

C. Neural Networks

Initially, a standard MLP was tried, achieving median and 95-percentile positioning errors of 0.082 m and 0.22 m respectively for the test split of Dataset 1. When tested on the combined results from the other datasets (Subject 1 day 2, Subject 2, Subject 3, and Subject 4), the median and 95-percentile positioning errors were 0.28 m and 2.11 m respectively.

Adjacent pixels in the frame from the sensor have some correlation. This spatial correlation could potentially be used to improve the positioning. The MLP does not leverage the possible spatial correlation between nearby pixels. Therefore, a 2D-CNN was tried. The 2D-CNN performed very well with median and 95-percentile position errors of 0.028 m and 0.068 m respectively. For the remaining datasets, the median and 95-percentile errors were found to be 0.19 m and 0.83 m. That was a marked improvement upon the performance of the MLP on the test split of the training data and on the more generalized datasets.

As a human subject moves about slowly in comparison to the sample rate of the sensor, successive readings from the thermopile are very likely to have some correlation. A 1D-CNN was tested in an attempt to leverage the potential temporal correlation of pixels across several consecutive frames. Unfortunately, the performance was poor with a median and 95-percentile positioning errors of 0.28 m and 0.63 m for the test split of Dataset 1, and of 0.36 m and 1.03 m for the other datasets. In order to achieve an improvement, several Recurrent Neural Networks (RNN) were tried. The LSTM, BLSTM, and GRU methods all performed well: all achieving very similar results with median positioning errors in the range of 0.05 – 0.061 m and 95-percentile errors in the range of 0.13-0.17 m for the test split of Dataset 1. On the remaining datasets, the median errors were at 0.18-0.20 m and 95-percentile errors were: 1.80 m (LSTM), 1.39 m (BLSTM), and 1.64 m (GRU). If one was

to look solely at the test split data, the 2D-CNN vastly outperformed the RNNs. However, when the remaining datasets are taken into consideration, the RNNs performed very similarly to the 2D-CNN. This emphasizes the importance of testing upon multiple subjects on different days for realistic performance comparisons. In the above case, by using just the test split of Dataset 1, the performance of 2D-CNN would be vastly inflated compared to a more realistic scenario.

Finally, to leverage both the spatial and temporal features of the data, a 3D-CNN and CNN-LSTM were evaluated. Whilst the performance of the 3D-CNN was better than that of the 1D-CNN, it was still worse than the performance of the 2D-CNN. For the 3D-CNN, the median and 95-percentile errors on the test split of Dataset 1 were found to be of 0.086 m and 0.19 m respectively. For the remaining sets, the corresponding errors were 0.24 m and 0.94 m. The CNN-LSTM had a median and 95-percentile positioning errors of 0.027 m and 0.068 m respectively for the test split of Dataset 1. They then rose to 0.16 m and 0.76 m for the remaining datasets. The results show a slight improvement over the 2D-CNN and the other RNNs on the test split of Dataset 1 as well as on the combined Datasets 2-5. This is more evident when comparing the ECDFs in Fig. 16 as the median and 95-percentile error statistics do not fully encapsulate the performance. A comparison of the best-in-class models (i.e., KNN from the classical regression, Bagged Ensemble from the ensemble regression, and CNN-LSTM in the neural network regression) can be seen in Fig. 17.

From these results, it is apparent that the models perform far better on the test split of Dataset 1. This is important to emphasize, as, without the additional datasets, the models' accuracy would be greatly overstated (i.e., the CNN-LSTM achieving 0.027 m position error on the test split of Dataset 1 compared to 0.18 m on the combined Datasets 2-5 – the differences can be clearly seen for three of the models in Fig. 18). The presented results show that the real-world performance of a system cannot be determined directly from the results of the test split. Models need to be evaluated against more diverse
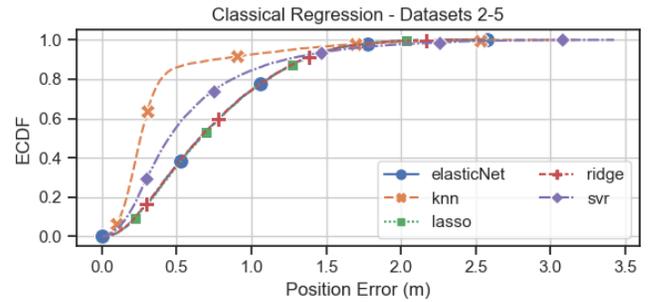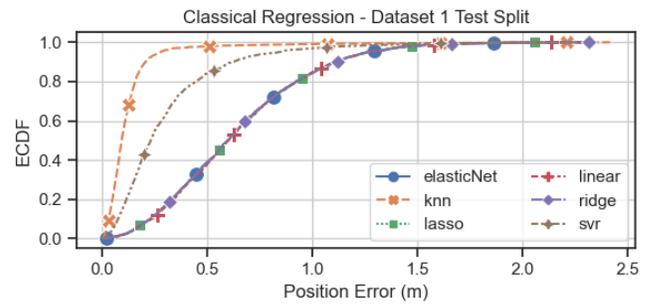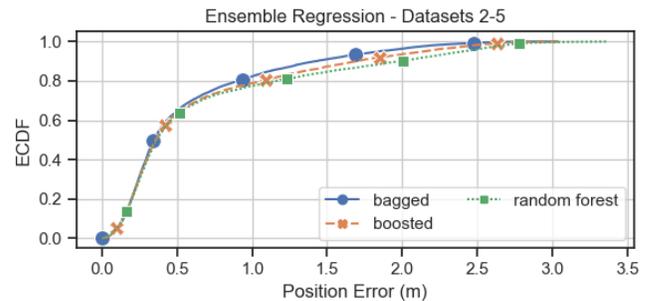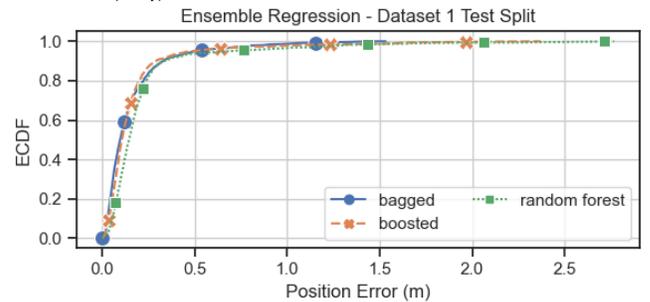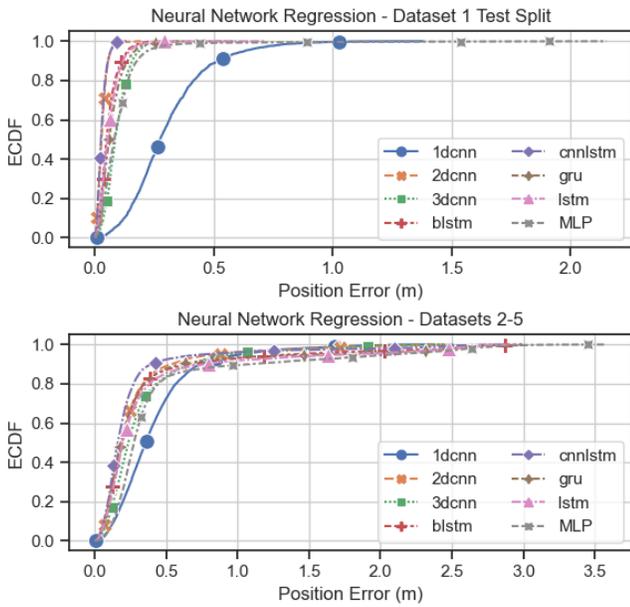
Fig. 16. ECDF plots for the neural network models, both for (top) the Dataset 1 test split and (bottom) for Datasets 2-5 combined together.
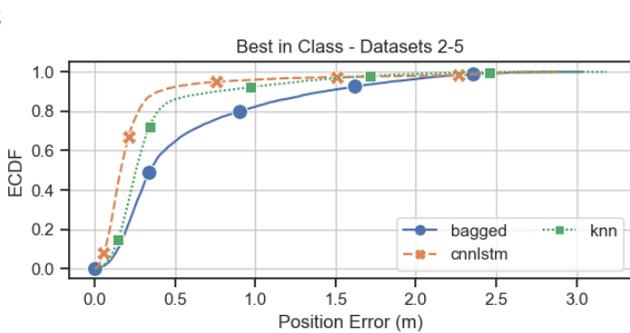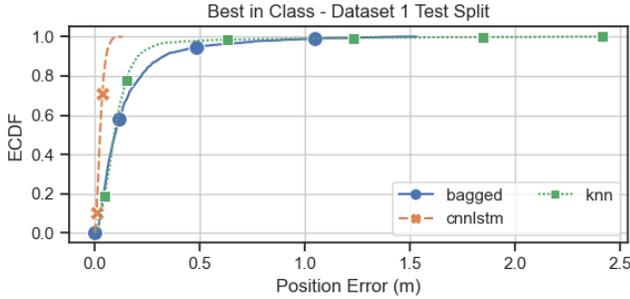

2

Fig. 17. ECDF plots for the best-in-class models – CNN-LSTM, KNN, and bagged ensemble. Both for (top) the Dataset 1 test split and (bottom) for Datasets 2-5 combined together.

test sets for robust performance analysis. Further research should be done to improve the generalizability of the models, as potentially improved preprocessing could help remove some of the differences. A larger, more diverse training dataset could also help train more generalized models. However, none of this can be achieved without first having multiple test sets to benchmark against.

D. Person to Person Generalizability

An important criterion is that the models can position subjects that they have not *seen* before. This has not been reported upon in the existing literature on thermopile-based localization. However, it is a very important factor to consider for real-world practical applications where a system is very
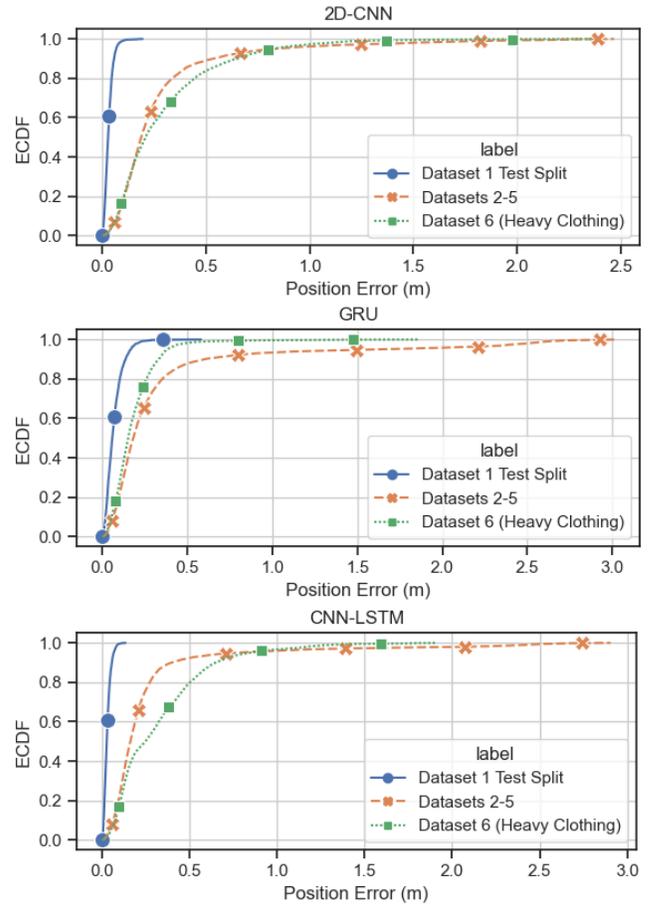

Fig. 18. ECDFs comparing the performance between Dataset 1 test split, Datasets 2-5, and Dataset 6. Shown are the 2D-CNN model (top), GRU model (middle), and the CNN-LSTM model (bottom).

likely to operate primarily on subjects that it has not been trained upon.

In this work, the models were all trained using the large dataset obtained with Subject 1. They were all tested both on the testing split of this data and on all the other remaining datasets. Fig. 19 shows the ECDF curves for three models, four subjects, and the test split of Dataset 1. As can be seen here, there are no significant variations in accuracy between the subjects. At the same time, some differences are apparent between the models. The differences between the subjects are very small for the two CNN-based models. Therefore, there is very good generalizability to subjects the models have not been trained upon. However, it should be noted that the results for all the subjects are significantly worse than for the case with the test split of the training data.

E. The Impact of Heavy Clothing

Dataset 6 was collected on the same afternoon as Dataset 1. Subject 1 walked around the test space for approximately 12 minutes while wearing heavy clothing (long pants and a thick coat). Heavy clothing traps more heat. Thus, such a subject is less visible to the sensor. Across all the models, the positioning error of the subject in Dataset 6 was significantly worse than with the test split of Dataset 1. When compared to the combined results of the subjects in Datasets 2-5, the results of Dataset 6 were reasonably similar while still dependent upon a model.
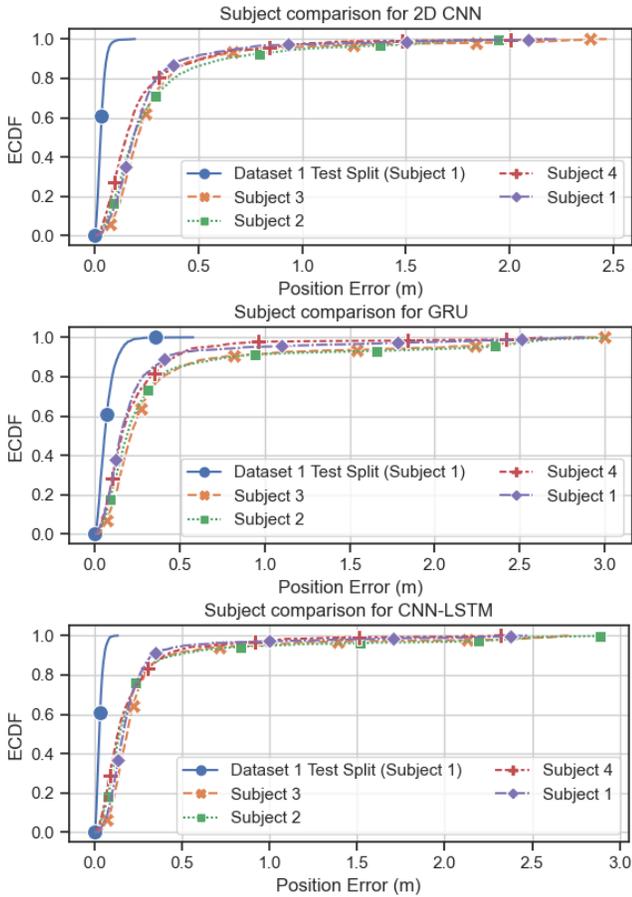
Fig. 19. ECDFs comparing the for each subject for selected models. Shown are the 2D-CNN model (top), GRU model (middle), and the CNN-LSTM model (bottom).

Fig. 18 shows the ECDFs for a selection of the models (2D-CNN, GRU, and CNN-LSTM). It can be seen that the 2D-CNN somewhat struggled in the heavy clothing case. This was also true for the 1D- and 3D-CNN (see also Table III). The RNN-based methods were better at positioning the subject in heavy clothing (Fig. 18, the GRU model). The CNN-LSTM performed worse than both the 2D-CNN and the RNN-based methods.

Whilst heavy clothing increases the positioning error compared to the test split, the models still generalize well. It suggests that they can handle subjects in a wide range of clothing. In addition to this, the subject was uncomfortably warm in such clothing and would not realistically have dressed in a such a manner. If the air temperature was lower, such clothing might be worn, but in such a case there would be a larger difference between the temperature of the subject and the environment, potentially reducing the impact of the heavy clothing. There is still further investigation that could be done to further understand the impact of clothing on the location accuracy.

F. Spatial Distribution of Positioning Errors

When the datasets were collected, the subjects walked around in an arbitrary fashion in a 5 m by 5 m space. However, due to the sensor's field of view and the height of the room, the visible floor area for the sensor was just 2.88 m by 2.88 m (Fig. 2). Furthermore, when subjects were at the very edge of that space,

only their feet were within the field of view of the sensor. Most of the subject bodies were out of the field of view due to the FOV's pyramidal shape. Significantly lower signal-to-noise ratios were present at the edges of the field of view (Fig. 1). In turn, it made it much harder to locate subjects in those areas. This is illustrated by the heatmap in Fig. 20. The diagram presents positioning errors (from the 2D CNN model for Datasets 2-5) plotted at the relevant positions given by the ground truth. The points are sorted by error magnitude: the largest errors are overlaid on top, with the colors of the points corresponding to the error magnitudes. It can be seen clearly that the worst error magnitudes occurred at the edges. This suggests further limiting the field of view (and thus the usable floor area) in cases where the location accuracy is of importance and should be increased. Alternatively, multiple sensors could be employed in future research to cover a larger area while providing the required positioning capability. The extent of the sensor fields of view overlapping could be an important factor to study.

IV. Conclusion

To the best knowledge of the authors, this work offers the first open dataset for thermopile-based indoor localization. Given the volume and quality of the data, it will enable researchers to develop, train and evaluate machine learning and other localization algorithms in a robust manner. The dataset will also enable researchers to benchmark their algorithms against each other upon a common dataset.

The presented comprehensive study shows that the majority of the recurrent and convolutional neural networks greatly outperformed the ensemble and classical methods. However, it comes at the cost of increased complexity. Therefore, a careful trade-off between these two parameters is required especially in resource-constrained environments. Of the other methods, the weighted KNN was the only one comparable in performance to the best performing neural networks and could potentially be used where computational power is limited. The 2D-CNN and CNN-LSTM networks performed exceptionally well on the test split of Dataset 1 (the large training dataset) and were able to achieve the median positioning error of less than 0.03 m. However, there was a clear degradation in performance when models were trained on the data obtained on one day while tested on the data obtained on another day. It highlighted the importance of the testing performed upon more diverse data.

The majority of the convolutional and recurrent neural networks were able to maintain a median accuracy of less than 0.2 m (compared to less than 0.07 m on the testing split of the training dataset). This shows that thermopiles can offer reliable sub-meter localization accuracy, even in the presence of environmental and subject changes. This has not been explored in the literature.

Across the ML algorithms, the degradation in performance between subjects was much lower than the degradation in performance between the sessions taken at different days/times. In real terms, this suggests that across the majority of the tested ML algorithms there was generalizability to other subjects which were not part of the training corpus.
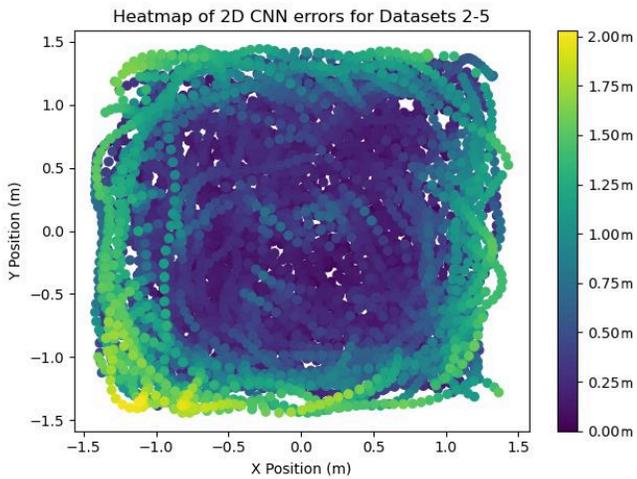
Fig. 20. A plot showing the spatial distribution of the positioning errors for the 2D CNN model. Each point is plotted at the ground truth location, with the color of the point representing the magnitude of the error. The points are first sorted by error magnitude so larger errors are overlaid on top.

It was found that the CNN-based methods performed very well on Datasets 1-5 (light clothing). However, the performance significantly worsened on Dataset 6 (heavy clothing). In contrast, the RNN based models had slightly inferior performance compared to the CNN methods on Datasets 1-5. At the same time, they outperformed the CNN methods on Dataset 6. This suggests that the RNN based methods (e.g., LSTM) with their inherent tracking abilities, may be better suited to deal with subjects wearing a wider range of apparel types.

The disparity in accuracy between the initial network training and the testing on data collected at a later date also offers several considerations for future work. Namely, when testing localization systems, it is important to check system degradation over a time to ensure that the proposed method can generalize to realistic environmental changes. Data augmentation could also be explored to see whether temperature fluctuations due to seasonal variation, air conditioning, or varying human heat signatures can be modelled and included as supplementary training samples. When testing the subjects with Datasets 2-5, the trained models were left unchanged. If data collection remains possible after the initial training period, transfer learning [45] could be explored. This could be achieved using a supervised approach using a small volume of data. Alternatively, it can be achieved for a continuous unsupervised system to maintain localization accuracy as the environment changed.

There may also be scope for reducing the environmental differences between the datasets with more advanced preprocessing, specifically the background subtraction and normalization. Using a ceiling mounted thermopile, Trofimova et al. were able to increase the accuracy of human detection from 70% to 97% by using more advanced background tracking and subtraction [46]. A similar approach applied to positioning of subjects could be investigated to potentially reduce the difference in performance between datasets. Recent literature show that block-sparse coding based ML [47] and convolutional autoencoder [48] have led to accurate and robust

localization for wireless-based DFL at low Signal to Noise Ratio (SNR) conditions. Such approaches should be explored in the future to further improve the discussed thermopile-based localization.

In this work, the variations between each subject's speed was relatively small. All walked at a standard adult's pace , not exceeding 1.5 m/s. The thermopile sampling rate of 10 Hz was adequate to capture the data. Future work can explore what minimum sampling rate is required for tracking given subject speeds and ascertain how the system's performance degrades as subjects reach or exceed the maximum supported speed.

The current method is applicable for a single subject positioning. Further work is required to allow for the detection and tracking of multiple subjects.

The field of view of the employed sensor was not large enough to cover a whole room. Therefore, the use of multiple sensors is planned to be studied. In particular, finding an ideal number of sensors, their parameters, extent of their field of view overlaps, the optimal height of sensor placing over a floor, etc., would be of value. This would also lead to the subject tracking across multiple sensors and subject handoff between the sensors.

REFERENCES

[1] J. Zhang, W. Xiao, and Y. Li, "Data and knowledge twin driven integration for large-scale device-free localization," *IEEE Internet of Things Journal,* vol. 8, no. 1, pp. 320-331, 2020.
[2] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of things for smart cities," *IEEE Internet of Things journal,* vol. 1, no. 1, pp. 22-32, 2014.
[3] A. Morar *et al.*, "A comprehensive survey of indoor localization methods based on computer vision," *Sensors,* vol. 20, no. 9, p. 2641, 2020.
[4] R. C. Shit *et al.*, "Ubiquitous localization (UbiLoc): A survey and taxonomy on device free localization for smart world," *IEEE Communications Surveys & Tutorials,* vol. 21, no. 4, pp. 3532-3564, 2019.
[5] S. Ali, T. Glass, B. Parr, J. Potgieter, and F. Alam, "Low Cost Sensor with IoT LoRaWAN Connectivity and Machine Learning Based Calibration for Air Pollution Monitoring," *IEEE Transactions on Instrumentation and Measurement,* 2020.
[6] J. Yun and J. Woo, "A comparative analysis of deep learning and machine learning on detecting movement directions using PIR sensors," *IEEE Internet of Things Journal,* vol. 7, no. 4, pp. 2855-2868, 2019.
[7] M. Gochoo *et al.*, "Novel IoT-Based Privacy-Preserving Yoga Posture Recognition System Using Low-Resolution Infrared Sensors and Deep Learning," *IEEE Internet of Things Journal,* vol. 6, no. 4, pp. 7192-7200, 2019.
[8] Y. Li, D. Li, Y. Cheng, G. Liu, J. Niu, and L. Su, "A novel human tracking and localization system based on pyroelectric infrared sensors: demonstration abstract," in *Proceedings of the 15th international conference on information processing in sensor networks*, 2016, pp. 1-2.
[9] L. Wu, Y. Wang, and H. Liu, "Occupancy detection and localization by monitoring nonlinear energy flow of a shuttered passive infrared sensor," *IEEE Sensors Journal,* vol. 18, no. 21, pp. 8656-8666, 2018.
[10] P. I. Devices, "AMG8833 | Panasonic Industrial Devices." [Online]. Available: https://na.industrial.panasonic.com/products/sensors/sensors-automotive-industrial-applications/lineup/grid-eye-infrared-array-sensor/series/70496/model/72453
[11] 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.
[12] S. Singh and B. Aksanli, "Non-intrusive presence detection and position tracking for multiple people using low-resolution thermal sensors," *Journal of Sensor and Actuator Networks,* vol. 8, no. 3, p. 40, 2019.
[13] B. Sirmacek and M. Riveiro, "Occupancy prediction using low-cost and low-resolution heat sensors for smart offices," *Sensors,* vol. 20, no. 19, p. 5497, 2020.

[14] C. Kowalski, K. Blohm, S. Weiss, M. Pfingsthorn, P. Gliesche, and A. Hein, "Multi low-resolution infrared sensor setup for privacy-preserving unobtrusive indoor localization," in *ICT4AWE*, 2019, pp. 183-188.

[15] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Device-Free Localization Using Privacy-Preserving Infrared Signatures Acquired from Thermopiles and Machine Learning," *IEEE Access,* 2021.

[16] 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)*, 2012: IEEE, pp. 2042-2047.

[17] 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.

[18] N. Gu, B. Yang, and T. Li, "High-resolution thermopile array sensor-based system for human detection and tracking in indoor environment," in *2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2020: IEEE, pp. 1926-1931.

[19] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural networks for indoor person tracking with infrared sensors," *IEEE Sensors Letters,* vol. 5, no. 1, pp. 1-4, 2021.

[20] N. Faulkner, F. Alam, M. Legg, and S. Demidenko, "Watchers on the Wall: Passive Visible Light-Based Positioning and Tracking With Embedded Light-Sensors on the Wall," *IEEE Transactions on Instrumentation and Measurement,* vol. 69, no. 5, pp. 2522-2532, 2019.

[21] D. Konings, F. Alam, F. Noble, and E. M. Lai, "SpringLoc: A device-free localization technique for indoor positioning and tracking using adaptive RSSI spring relaxation," *Ieee Access,* vol. 7, pp. 56960-56973, 2019.

[22] S. Shi, S. Sigg, L. Chen, and Y. Ji, "Accurate location tracking from CSI-based passive device-free probabilistic fingerprinting," *IEEE Transactions on Vehicular Technology,* vol. 67, no. 6, pp. 5217-5230, 2018.

[23] M. Mirshekari, S. Pan, J. Fagert, E. M. Schooler, P. Zhang, and H. Y. Noh, "Occupant localization using footstep-induced structural vibration," *Mechanical Systems and Signal Processing,* vol. 112, pp. 77-97, 2018.

[24] D. Yang, B. Xu, K. Rao, and W. Sheng, "Passive infrared (PIR)-based indoor position tracking for smart homes using accessibility maps and a-star algorithm," *Sensors,* vol. 18, no. 2, p. 332, 2018.

[25] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.

[26] A. Popleteev, "HIPS: Human-based indoor positioning system," in *2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2016: IEEE, pp. 1-7.

[27] P. Roy and C. Chowdhury, "A Survey of Machine Learning Techniques for Indoor Localization and Navigation Systems," *Journal of Intelligent & Robotic Systems,* vol. 101, no. 3, pp. 1-34, 2021.

[28] T. Glass, F. Alam, M. Legg, and F. Noble, "Autonomous Fingerprinting and Large Experimental Data Set for Visible Light Positioning," *Sensors,* vol. 21, no. 9, p. 3256, 2021.

[29] A. Paijens, L. Huang, and A. Al-Jumaily, "Measurement of the positions of light sensors on a mobile frame being tracked with a lighthouse localization system," *Sensors and Actuators A: Physical,* vol. 331, p. 112979, 2021.

[30] ISO/IEC JTC 1/SC 31, *Information technology–real time locating systems–test. and evaluation of localization and tracking systems (ISO/IEC 18305: 2016)*. International Organization for Standardization Geneva, Switzerland, 2016.

[31] F. Potortì *et al.*, "Comparing the performance of indoor localization systems through the EvAAL framework," *Sensors,* vol. 17, no. 10, p. 2327, 2017.

[32] D. Konings, F. Alam, F. Noble, and E. M. Lai, "Device-Free Localization Systems Utilizing Wireless RSSI: A Comparative Practical Investigation," *IEEE Sensors Journal,* vol. 19, no. 7, pp. 2747-2757, 2018.

[33] O. B. Tariq, M. T. Lazarescu, and L. Lavagno, "Neural networks for indoor human activity reconstructions," *IEEE Sensors Journal,* vol. 20, no. 22, pp. 13571-13584, 2020.

[34] D. Chetverikov, S. Fazekas, and M. Haindl, "Dynamic texture as foreground and background," *Machine vision and Applications,* vol. 22, no. 5, pp. 741-750, 2011.

[35] Y. Li *et al.*, "Toward location-enabled IoT (LE-IoT): IoT positioning techniques, error sources, and error mitigation," *IEEE Internet of Things Journal,* vol. 8, no. 6, pp. 4035-4062, 2020.

[36] X. Guo, S. Zhu, L. Li, F. Hu, and N. Ansari, "Accurate WiFi localization by unsupervised fusion of extended candidate location set," *IEEE Internet of Things Journal,* vol. 6, no. 2, pp. 2476-2485, 2018.

[37] X. Wang, L. Gao, and S. Mao, "CSI phase fingerprinting for indoor localization with a deep learning approach," *IEEE Internet of Things Journal,* vol. 3, no. 6, pp. 1113-1123, 2016.

[38] D. Konings, R. Grace, and F. Alam, "A stacked neural network-based machine learning framework to detect activities and falls within multiple indoor environments using Wi-Fi CSI," *IEEE Sensors Letters,* vol. 5, no. 5, pp. 1-4, 2021.

[39] M. Mohammadi, A. Al-Fuqaha, M. Guizani, and J.-S. Oh, "Semisupervised deep reinforcement learning in support of IoT and smart city services," *IEEE Internet of Things Journal,* vol. 5, no. 2, pp. 624-635, 2017.

[40] X.-X. Du, Y. Mu, Z.-W. Ye, and Y.-J. Zhu, "A passive target recognition method based on LED lighting for industrial internet of things," *IEEE Photonics Journal,* vol. 13, no. 4, pp. 1-8, 2021.

[41] A. J. Izenman, *Modern multivariate statistical techniques* (Regression, classification and manifold learning). New York, NY: Springer, 2008.

[42] C. Zhang and Y. Ma, *Ensemble machine learning: methods and applications*. New York: Springer-Verlag, 2012.

[43] C. C. Aggarwal, *Neural networks and deep learning*. Cham: Springer, 2018, pp. 978-3.

[44] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv preprint arXiv:1803.08375,* 2018.

[45] H. Zou, Y. Zhou, H. Jiang, B. Huang, L. Xie, and C. Spanos, "Adaptive localization in dynamic indoor environments by transfer kernel learning," in *2017 IEEE wireless communications and networking conference (WCNC)*, 2017: IEEE, pp. 1-6.

[46] A. A. Trofimova, A. Masciadri, F. Veronese, and F. Salice, "Indoor human detection based on thermal array sensor data and adaptive background estimation," *Journal of Computer and Communications,* vol. 5, no. 4, pp. 16-28, 2017.

[47] L. Zhao *et al.*, "Block-Sparse Coding-Based Machine Learning Approach for Dependable Device-Free Localization in IoT Environment," *IEEE Internet of Things Journal,* vol. 8, no. 5, pp. 3211-3223, 2020.

[48] L. Zhao, H. Huang, X. Li, S. Ding, H. Zhao, and Z. Han, "An accurate and robust approach of device-free localization with convolutional autoencoder," *IEEE Internet of Things Journal,* vol. 6, no. 3, pp. 5825-5840, 2019.

Nathaniel Faulkner received his B.E. (Hons.) from Massey University, New Zealand in Electronics & Computer Engineering in 2016. He is currently pursuing a Ph.D. at the same institution. In 2020, he was also appointed as a Visiting Research Fellow at the School of Science and Technology, Sunway University, Malaysia. His research interests include indoor positioning, embedded systems design, and the Internet of Things.

Daniel Konings (M'17) received his BE(Hons) and Ph.D. in Computer Systems and Electronics Engineering from Massey University, New Zealand in 2014 and 2019 respectively. He is currently a Lecturer with the Department of Mechanical and Electrical Engineering, School of Food and Advanced Technology, Massey University, New Zealand. His research interests include indoor localization, machine learning, IoT, sensor networks, and engineering education.

Fakhrul Alam (M'17–SM'19) received the B.Sc. degree (Hons.) in electrical and electronic engineering from BUET, Bangladesh, and the M.S. and Ph.D. degrees in electrical engineering from Virginia Tech, USA. He is currently an Associate Professor with the Department of Mechanical and Electrical Engineering, School of Food and Advanced Technology, Massey University, New Zealand. He has been appointed as an Adjunct Professor with the School of Engineering and Technology, Sunway University, Malaysia, since 2021. His research interests include indoor localization, 5G and visible light communication, the IoT, and wireless sensor networks.

Mathew Legg (M'19) received his B.Sc, M.Sc, and Ph.D. in Physics from the University of Auckland, New Zealand. He is currently a Senior Lecturer with the Department of Mechanical and Electrical Engineering, School of Food and Advanced Technology, Massey University. His research relates to the development of acoustic/ultrasonic measurement systems and techniques for acoustic imaging, non-destructive testing, and remote sensing.

Serge Demidenko (M'91-SM'94-F'04) is Professor and Dean of the School of Science and Technology, Sunway University, Malaysia. He is also associated with the Department of Mechanical & Electrical Engineering, School Food and Advanced Technology, Massey University, New Zealand. He graduated in Computer Engineering from the Belarusian State University of Informatics and Radio Electronics and received Ph.D. from the Institute of Engineering Cybernetics of the Belarusian Academy of Sciences. His research interests include electronic design and test, signal processing, instrumentation and measurements. He is also Fellow of IET and UK Chartered Engineer.

CHAPTER 6 – CONCLUSION

This PhD study has resulted in the publication of three top-tier (Q1) journal articles, with a fourth pending a review. In addition, two peer-reviewed conference papers were published. Furthermore, the research led to the publication of several other papers associated with the use of the developed hardware. These results cohesively fit together to provide a range of indoor localisation solutions that can potentially be combined through sensor fusion to produce a robust and accurate localisation system.

The median localisation error for a moving subject for all three systems is under 0.15 m (Passive VLP: 0.12 m; Capacitive floor: 0.03 m; Wall-mounted thermopiles using three sensors: 0.11 m; Ceiling-mounted thermopile using CNN-LSTM: 0.182 m). The capacitive floor and thermopile systems appear to be generalisable to new environments with very little calibration. The passive VLP system and thermopile systems also do not require large amounts of additional infrastructure to be installed. In particular, the thermopile system can operate using only a single thermopile sensor with an option for extra sensors to be added for increased localisation accuracy.

The infrastructure complexity of the floor-sensing is its main drawback. However, it can detect a person lying on the floor and capture the poses of such a person. The high resolution of the data from the floor sensors makes it a much stronger candidate for subject identification and automated fall detection in an unobtrusive manner than a number of other solutions. This could offset the relatively higher cost.

Throughout the PhD research, there was a large improvement in the data collection methodology for the localisation of a subject. It progressed from using predetermined paths (Chapter 2) to employing the HTC Vive for automated data collection and labelling (Chapters 3, 4, and 5). In addition to the better data collection, this much-improved approach also allowed subjects to walk around the testing area in a natural fashion. Chapters 3, 4, and 5 also presents the results of investigating the generalisability of the various localisation methods across multiple subjects and environments. It was found that changes in the environment and the use of different subjects can cause significant performance degradation - an aspect that has been widely overlooked in the earlier research presented in the literature. To highlight it, the CapLoc system was tested for 39 different paths across multiple subjects while for the thermopile

system several hundred paths were walked by the subjects in several different environments (Chapter 4). Data collection at that scale and accuracy would not have been possible without employing the HTC Vive ground truth data collection.

6.1 Future Works

Several passive indoor localisation solutions having distinctive strengths and weaknesses were developed over the duration of the reported PhD study. It is likely that the next generation of indoor localisation systems will be of a multi-modal type combining data from multiple sensors of different types to improve the overall location estimation. Each of the investigations reported in this thesis used a singular sensing modality. It would therefore be a natural next step to attempt to build a system fusing several reported solutions and, perhaps, enhancing it with entirely new ones.

Localising multiple subjects simultaneously is an important goal for a real-world positioning system. Earlier developed systems were only suitable for tracking a single subject at a time. It will be of value to extend the study towards simultaneously determining locations of multiple subjects. A classifier could be used to count the number of the subjects. A separate model would then be chosen and trained for each subject number. For a larger deployment, each system could be spatially partitioned into smaller individual sub-systems. There is a maximum number of subjects that can realistically fit into any set area at any given moment of time. This, therefore, limits the maximum number of subjects that are needed to be positioned simultaneously. For such a system, the research would also need to be undertaken to investigate the *handoff* of subjects at the boundaries of adjacent sub-systems.

There are privacy considerations when carrying out subject localisation. In the real-world scenario, subjects' location information is sensitive, and it should not be disclosed without their explicit permission. Therefore, the security of a positioning system needs to be considered prior to its deployment. The security aspects of the localisation systems have not been considered in the reported study. There is scope for future investigation in this area. Furthermore, as no authentication of subjects was offered in the reported research, there is the possibility of spoofing the system by a malicious entity. Thus, enhancements covering subject authentication or identity verification would be of importance.

APPENDICES

Appendix 1

This chapter is republished in accordance with IEEE's copyright policy. This chapter is the accepted version of the published work, and as such may have stylistic differences from the final published article.

Smart Wall: Passive Visible Light Positioning with Ambient Light Only

Nathaniel Faulkner*, Fakhrul Alam*, Mathew Legg* and Serge Demidenko†*

*School of Engineering & Advanced Technology, Massey University, Private Bag 102904, Auckland, New Zealand
†School of Science & Technology, Sunway University, No.5 Jalan Universiti, Bandar Sunway, 47500 Selangor, Malaysia
Email: *{N.Faulkner, F.Alam, M.Legg, S.Demidenko}@Massey.ac.nz, †SDemidenko@Sunway.edu.my

Abstract—**This paper reports the experimental results from a novel visible light positioning (VLP) system. The developed VLP is completely passive as it does not require the target to carry any active device or tag and, at the same time, it does not require any modification to the lighting infrastructure. The developed VLP localizes a target based on measuring the change it creates in the received signal strength (RSS) of the ambient light recorded at an array of photodiodes embedded in the wall. Experimental results from a prototype system show that a median error of 7.9 cm can be achieved.**

Index Terms—**Indoor localization, Indoor Positioning System (IPS), Visible Light Positioning (VLP), Device Free Localization (DFL), Passive VLP**

I. Introduction

Indoor positioning has been a burgeoning area of research over the past decade. In terms of outdoor positioning, GPS [1] is the de facto solution, due to it being both ubiquitous and free to use. However, it has limitations, especially in built up areas or indoors [2]. The signal is negatively impacted by multipath reflections and struggles to penetrate walls. Furthermore, the offered accuracy of several metres [3] is not good enough for indoor applications. For these reasons, other methods have been proposed using infrared signals [4], RFID [5], Bluetooth [6] and Wi-Fi [7], [8]. Whilst most of these represent an improvement over GPS for indoor localization, the majority of them still do not meet the desired levels of accuracy, reliability or simplicity. With Light Emitting Diodes (LEDs) steadily replacing traditional lighting sources, a new method of positioning has come to the fore – Visible Light Positioning (VLP). Visible light has the benefits of being far less susceptible to multipath interference and flat fading due to its vastly higher frequency than RF [9]. LED lighting can also perform multiple roles – illumination, communication and positioning. Active VLP has been well researched and relies on a mobile object having a receiver containing either a photodiode or image sensor [10]. There are several active VLP methods that have been implemented on indoor testbeds, with the main approaches being Received Signal Strength (RSS) lateration [11]–[13], Angle of Arrival (AOA) angulation [14], [15], and fingerprint matching [16].

Passive VLP allows for the detection of people and objects without the need for the tracked object to have an attached

receiving device. It is highly desirable to be able to track passively rather than relying on a wristband or other smart device which must be consciously put on. There are several existing works for passive VLP. In [17] the authors used co-located LED luminaires and photodiodes to passively detect humans. The light from the LED luminaires was multiplexed using Time-Division Multiple Access (TDMA) to identify the source of incoming light at each photodiode. In the aforementioned paper, the data was primarily used to detect whether a door was open or closed. The authors further extended this work in [18] to also track human movement and detect room occupancy. The authors were able to achieve 93.7% occupancy count accuracy and 0.89 m median error positioning accuracy in a 45 m^2 room. In [19], the floor is inlaid with 324 photodiodes, with 5 LED luminaires placed on the ceiling above. This setup is then used to detect the position of a human's body and limbs from the shadows cast onto the floor. The authors were able to achieve a mean angular accuracy of 10 degrees for the 5 main body joints. The work was further extended in [20] using only 20 photodiodes, albeit with a much larger number of LED panels on the ceiling. This simplifies the infrastructure at the cost of slightly decreased accuracy – 13.6 degrees mean angular error instead of 10. Similarly, the authors in [21] also use a grid of photodiodes embedded into the floor. LED luminaires on the roof cast shadows from test subjects onto said photodiodes. However, this paper reports results based on mostly simulation, with the only experimental part being a single point to point LED to photodiode link to gather parameters for the larger scale simulation. In simulations, the authors were able to achieve median error of 8cm in an 8m x 8m x 4m room with 4 LED luminaires, and photodiodes uniformly spaced at 0.5m in the floor. In [22], the authors use a passive VLP approach for mobile device input using an LED and two photodiodes to detect a user's finger. The LED improves the reliability in the presence of changing ambient light. The authors were able to position a user's finger in a 9x7 cm grid with 0.7 cm median error. CeilingSee [23] uses reverse biased LED luminaires as photodiodes for occupancy sensing. However, the authors did not use the system for positioning of test subjects or objects and therefore, do not have a position accuracy.

This paper, focuses on achieving accurate positioning of an object in ambient light conditions without the need of any modification to the existing lighting, unlike the majority of
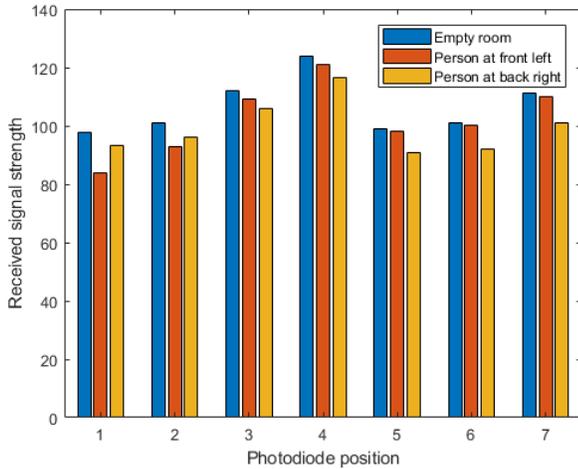
Fig. 1. Received power at each photodiode for three scenarios - empty test bed, test subject at left hand side close to the wall with the photodiodes affixed, and right hand side away from wall with photodiodes affixed. Corridor environment

VLP solutions.

II. SYSTEM OVERVIEW

In a room, there are generally multiple light sources – several interior lights and in many cases, windows as well. In addition to this, walls are generally light coloured and, therefore, cause a portion of the light to be reflected. A person moving around a room produces several shadows of different intensities to be projected onto the floor and walls. The main shadows comes from blocking the direct path from the ambient light sources. However, many other shadows are generated because the reflected components from the light sources are also blocked. These shadows can be detected by photodiodes placed around a room and then used to locate objects. This can be observed in Fig. 1. The blue bars are the RSS at seven different photodiodes placed along a wall when the test area is free from obstructions. The red and orange bars present a case when a person is standing at the front left (close to the wall) and the back right (further from the wall) respectively. This causes RSS to drop compared to the empty room, with there being a greater drop at photodiodes closer to the test subject. For example, when the test subject is in the front left position, the RSS drop is more significant in photodiodes 1 and 2 and there is very little drop in photodiodes 6 and 7. When the person is at the back right, the opposite is true - photodiodes 6 and 7 are affected to a much larger degree than 1 and 2.

The testbed makes use of seven ISL29023 [24] integrated digital light sensors placed on a board (wall) at a height of 1.05 meters from the floor. The light sensors are comprised of a photodiode, transimpedance amplifier, and analog-to-digital converter (ADC) located on the same package. Each light sensor is connected to a low-cost Wi-Fi microchip (ESP8266 [25]) as shown in Fig. 2. The ambient light manifests as DC at the output of the transimpedance amplifier. The DC level is
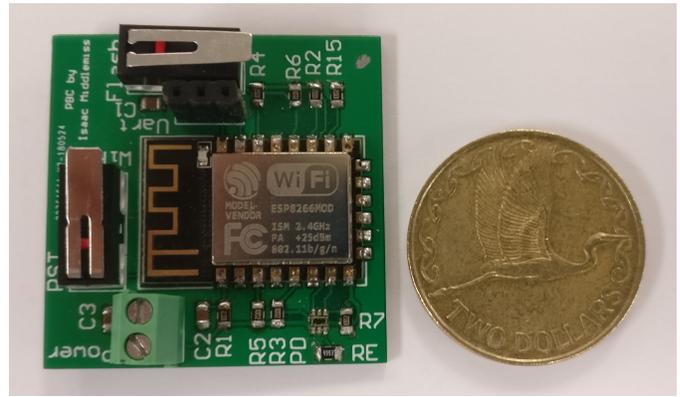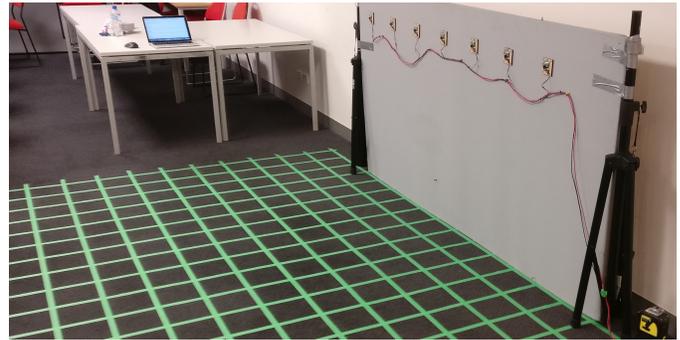

Fig. 2. Photodiode receiver.

Fig. 3. Smart wall with the embedded photodiode receivers, open room environment.

a measure of the RSS of the ambient light and is sampled by the embedded ADC and is retrieved by the microcontroller in the ESP8266. The latest 100 samples are stored in the memory until they are retrieved over Wi-Fi. The data can then be requested in 100 value packets from a laptop and saved to the hard drive.

A 3.4m x 2.2m grid with 20cm squares was marked out using masking tape and a laser straight edge.

Two experiments were performed, one with the photodiodes along the wall facing into the room – henceforth known as the open room environment (see Fig. 3). For the second experiment, the photodiodes were positioned along the side of the grid furthest from the wall, with the photodiodes pointing back towards the wall – henceforth known as the corridor environment (see Fig. 4). Immediately before and after each test, the background ambient light measured and recorded. This was then used to normalise the data at each photodiode and verify the ambient light levels remained constant - a factor which this work is reliant on. Changes in ambient light would introduce extra uncertainty and consequently decreased positioning accuracy. Each measurement consists of 100 RSS readings over 10 seconds at each photodiode. This could potentially be reduced (or the data sample rate increased) in later works to hasten the data acquisition process.

Data were collected at each grid intersection for a total of 198 locations, with the data being split into two parts with
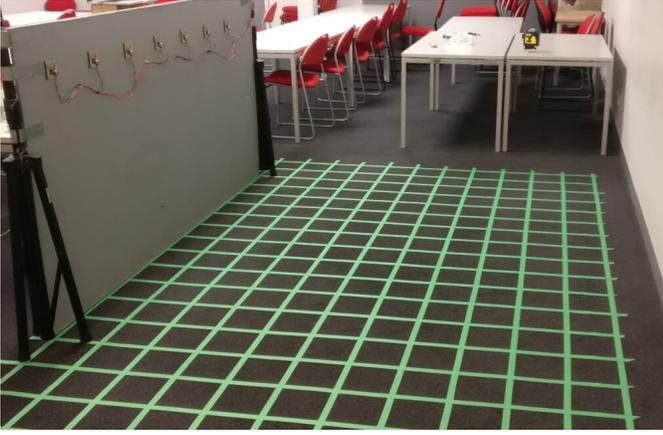
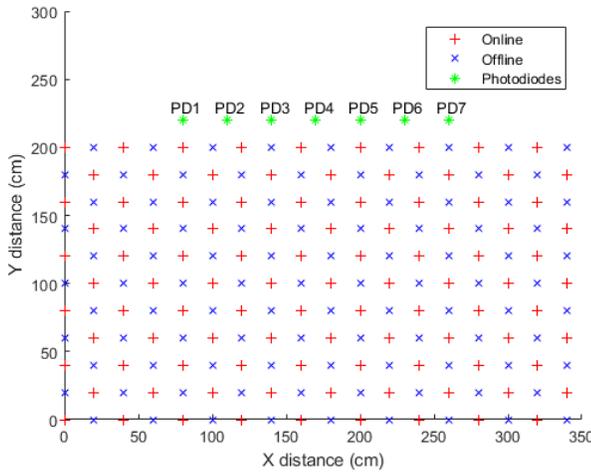Fig. 4. Smart wall with the embedded photodiodes, corridor environment.

Fig. 5. Online vs offline data points.

half forming the offline fingerprint database, and the other half, the online RSS measurements. The online and the offline locations are shown in Fig. 5. The measured RSS values at each photodiode are shown in Fig. 6. These plots show the change in RSS with a test subject (180 cm in height) standing at each individual point on the grid - combining both the online and offline points. A very large dip can be seen on the top left edge of each plot where the test subject stood immediately in front of the photodiode causing a strong shadow. Taking the RSS value from the same location on each plot gives the fingerprint value for that position. Weighted K Nearest Neighbours (WKNN) [26] was employed to classify the online readings using the offline fingerprints. Euclidean distance was used to measure the distance between the online reading and each entry on the fingerprint database.

A. WKNN algorithm

Weighted K nearest neighbours is an extension of the K nearest neighbours algorithm [27]. The algorithm takes a live reading R_{live} which is a vector of M RSS readings - one from

each photodiode. This is compared to the offline fingerprint database R which stores a vector R for each point that has been mapped. The Euclidean distance d_i between R_{live} and an entry in the database $R_{i,j}$ is taken as follows:

$$d_i = \sqrt{\sum_{j=1}^{M}(R_{i,j}^2 - R_{live}^2)} \qquad (1)$$

The K smallest distance entries in the database are taken and used to estimate the weights for each of the K database readings as:

$$W_k = \frac{1}{d_k} \qquad (2)$$

These are then used to weight each of the locations before they are combined. This is so that the database readings closest to the live reading have a greater influence on the final position estimation. The final position is found as follows:

$$\widetilde{x}_j = \frac{\sum_{k=1}^{K} W_{j,k} \times x_k}{\sum_{k=1}^{K} W_{j,k}} \qquad (3)$$

$$\widetilde{y}_j = \frac{\sum_{k=1}^{K} W_{j,k} \times y_k}{\sum_{k=1}^{K} W_{j,k}} \qquad (4)$$

Where $[\widetilde{x}_j, \widetilde{y}_j]$ is the estimated position, $W_{j,k}$ are the weights calculated in equation (2) and $[x_k, y_k]$ is the associated coordinate for that weight. The position error is then calculated by finding the Euclidean distance between the estimated location and the actual location where the live RSS values were taken.

III. EXPERIMENTAL RESULTS

A K value of 3 was experimentally chosen for the WKNN algorithm, as it provides a good balance between optimising both the median and maximum error for both environments. This can be clearly seen in Fig. 7.

TABLE I
POSITION ERROR FOR K = 3 FOR BOTH ENVIRONMENTS

	Corridor	Open room
Median error (cm)	7.9	12.3
Max error (cm)	97	357
Standard deviation (cm)	14.3	40.8

Table I shows the position errors for both the experiments. In Fig. 8, the estimated positions for the corridor are plotted in relation to their actual locations to show the spatial error distribution. One can see that the errors are concentrated at the boundaries of the testbed. In part, this is due to the positions being further from the photodiodes and in part, due to having less fingerprints around the position. Fig. 9 shows the localization errors for the open room environment. As expected from the Cumulative Distribution Function (CDF) plot in Fig. 10, one can see that the position estimation is
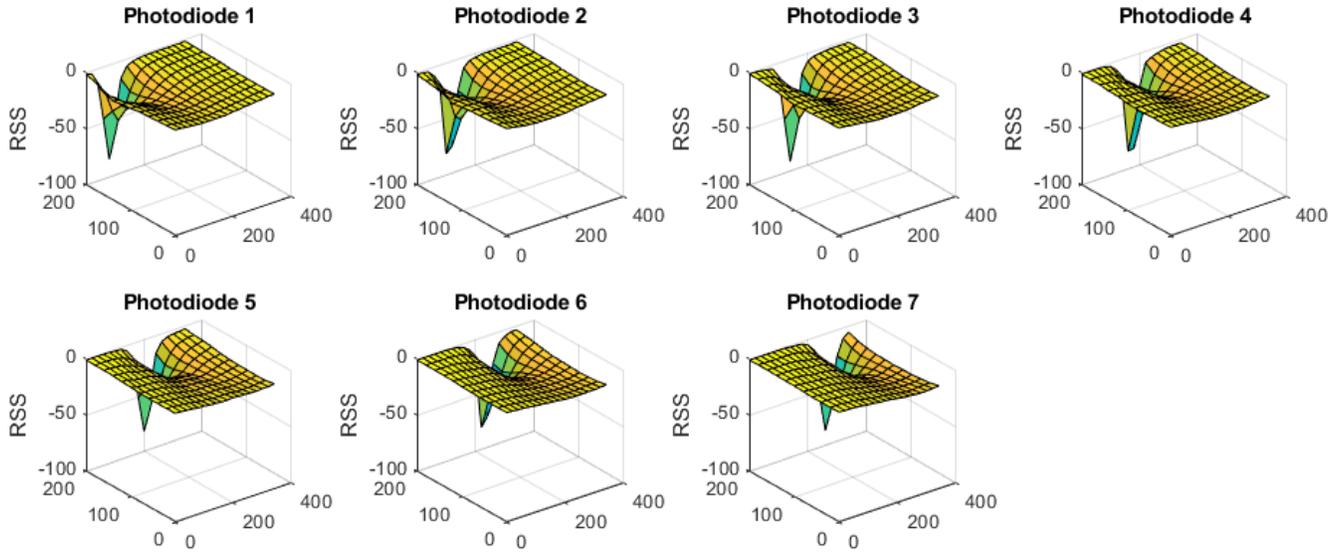
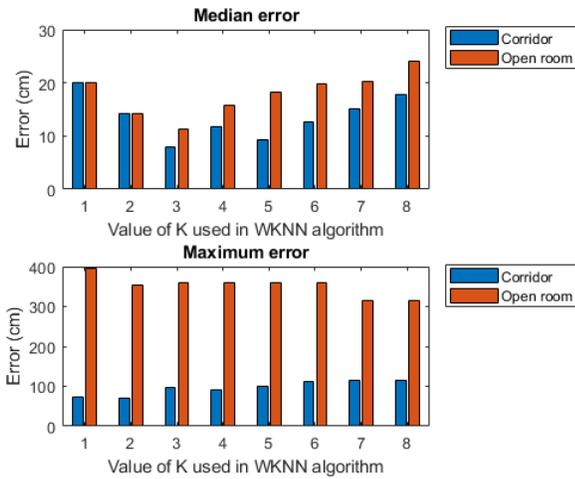Fig. 6. RSS fingerprints for each photodiode. Corridor environment.

Fig. 7. The impact of the value of K on both the median and maximum error for both the environments (corridor and open room).

Fig. 8. Actual vs estimated positions in the corridor environment. Green asterisks denote the photodiode positions, red crosses the actual positions, the blue circles the estimated positions and the black lines the magnitude of the error between the actual and estimated positions.

more accurate for the corridor environment. It is more accurate due to the light reflecting off the white wall behind causing more distinct shadows. The worst case errors in the open room scenario are at the very edges of the testbed. In particular, the two corners closest to the photodiode wall which are at a very acute angle to the majority of the photodiodes and, therefore, do not experience a discernible shadow. This can be seen in Fig. 11 where the two corner RSS plots are compared to the background RSS reading and also a position with low error. The RSS readings at the two corner plots are both very similar to the background readings. They are most different at photodiode 1 for the left hand reading and photodiode 7 for the right hand reading as these are the least acute angles.

As the RSS readings are so close to the background readings, small amounts of noise can cause erroneous identification of neighbours leading to large errors in the position estimate. This can be addressed by extending the row of the photodiodes further along the wall.

IV. CONCLUSION

The authors believe that this is the first reported passive VLP reported in the literature that uses only the ambient light. The system is able to position an object with a median error of 7.9cm in a corridor environment. In an open room scenario, this increases to 11.4 cm median error. Further work should
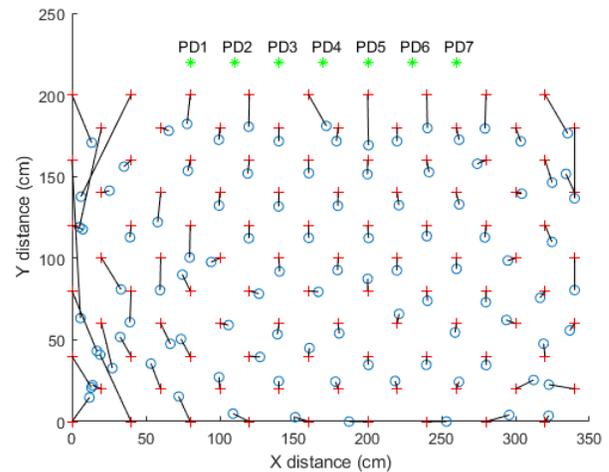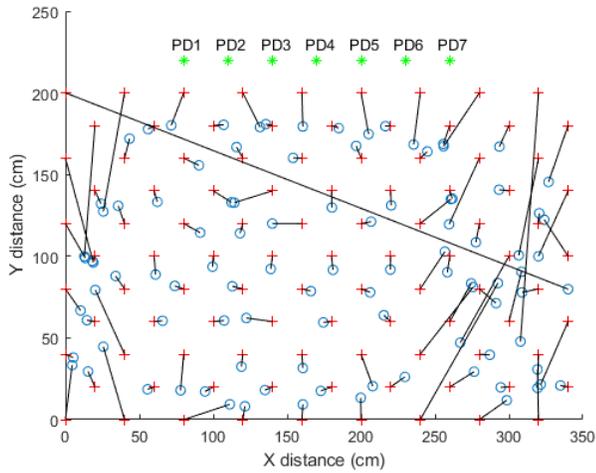
Appendices

Fig. 9. Actual vs estimated positions in the open room environment. Green asterisks denote the photodiode positions, red crosses the actual positions, the blue circles the estimated positions and the black lines the magnitude of the error between the actual and estimated positions.

Fig. 11. The worst errors are found at (0, 200), (320, 200). This is compared to a location with a much lower error at (160, 160) and the readings when no test subject is present. Open room environment
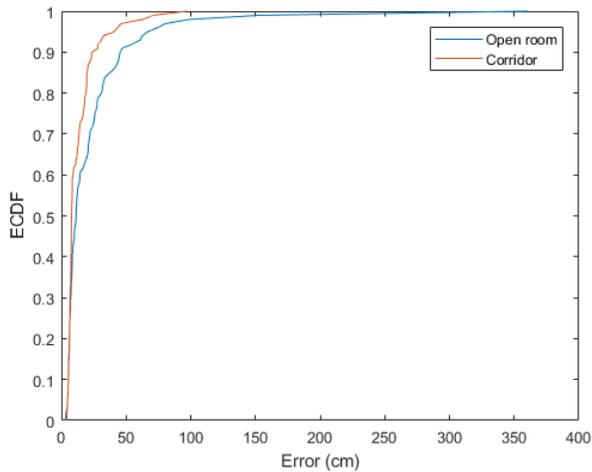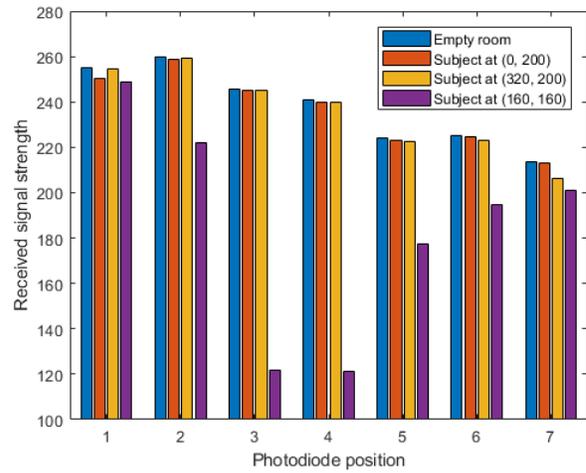
Fig. 10. Localization precision as CDF of error in both environments.

expand the test to a full room scale. The experiments were undertaken at night and, therefore changes in the level of ambient light was not investigated. This is the area for future investigations to quantify and potentially mitigate the impact it may have. Modulated light could potentially be used from LED ceiling luminaires to mitigate the effect of ambient light. Currently the system has been tested for a single object at a time and as such, further investigation is needed to detect multiple objects. With fingerprint based systems, generating the fingerprint database is a very time consuming process. In the future, the authors will investigate how to model these data and generate them from a few strategically selected calibration points.

REFERENCES

[1] J. Farrell and M. Barth, *The global positioning system and inertial navigation*, vol. 61. Mcgraw-hill New York, 1999.

[2] M. B. Kjærgaard, H. Blunck, T. Godsk, T. Toftkjær, D. L. Christensen, and K. Grønbæk, "Indoor positioning using GPS revisited," in *International conference on pervasive computing*, Springer, 2010.

[3] M. G. Wing, A. Eklund, and L. D. Kellogg, "Consumer-grade global positioning system (gps) accuracy and reliability," *Journal of forestry*, vol. 103, no. 4, p. 169, 2005.

[4] R. Want, A. Hopper, V. Falcao, and J. Gibbons, "The active badge location system," *ACM Transactions on Information Systems (TOIS)*, vol. 10, pp. 91–102, 1992.

[5] L. M. Ni, Y. Liu, Y. C. Lau, and A. P. Patil, "LANDMARC: indoor location sensing using active RFID," in *Proceedings of the First IEEE International Conference on Pervasive Computing and Communications*, IEEE, 2003.

[6] A. Kotanen, M. Hannikainen, H. Leppakoski, and T. D. Hamalainen, "Experiments on local positioning with Bluetooth," in *Information technology: Coding and computing [computers and communications], 2003. Proceedings. ITCC 2003. International conference on*, IEEE, 2003.

[7] M. Youssef and A. Agrawala, "The Horus WLAN location determination system," in *Proceedings of the 3rd international conference on Mobile systems, applications, and services*, ACM, 2005.

[8] P. Bahl and V. N. Padmanabhan, "RADAR: An in-building RF-based user location and tracking system," in *INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, Ieee, 2000.

[9] T. Komine and M. Nakagawa, "Fundamental analysis for visible-light communication system using LED lights," *IEEE transactions on Consumer Electronics*, vol. 50, pp. 100–107, 2004.

[10] W. Zhang and M. Kavehrad, "Comparison of VLC-based indoor positioning techniques," in *Broadband Access Communication Technologies VII*, International Society for Optics and Photonics, 2013.

[11] P. Hu, L. Li, C. Peng, G. Shen, and F. Zhao, "Pharos: Enable physical analytics through visible light based indoor localization," in *Proceedings of the Twelfth ACM Workshop on Hot Topics in Networks*, ACM, 2013.

[12] B. Xie, K. Chen, G. Tan, M. Lu, Y. Liu, J. Wu, and T. He, "LIPS: A light intensity–based positioning system for indoor environments," *ACM Transactions on Sensor Networks (TOSN)*, vol. 12, p. 28, 2016.

Appendices

[13] D. Konings, B. Parr, F. Alam, and E. M. Lai, "Falcon: Fused application of light based positioning coupled with onboard network localization," *IEEE Access*, vol. 6, pp. 36155–36167, 2018.

[14] Y.-S. Kuo, P. Pannuto, K.-J. Hsiao, and P. Dutta, "Luxapose: Indoor positioning with mobile phones and visible light," in *Proceedings of the 20th annual international conference on Mobile computing and networking*, ACM, 2014.

[15] M. Yasir, S.-W. Ho, and B. N. Vellambi, "Indoor position tracking using multiple optical receivers," *Journal of Lightwave Technology*, vol. 34, no. 4, pp. 1166–1176, 2016.

[16] F. Alam, M. T. Chew, T. Wenge, and G. S. Gupta, "An accurate visible light positioning system using regenerated fingerprint database based on calibrated propagation model," *IEEE Transactions on Instrumentation and Measurement*, 2018.

[17] M. Ibrahim, V. Nguyen, S. Rupavatharam, M. Jawahar, M. Gruteser, and R. Howard, "Visible light based activity sensing using ceiling photosensors," in *Proceedings of the 3rd Workshop on Visible Light Communication Systems*, ACM, 2016.

[18] V. Nguyen, M. Ibrahim, S. Rupavatharam, M. Jawahar, M. Gruteser, and R. Howard, "Eyelight: Light-and-shadow-based occupancy estimation and room activity recognition," in *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, IEEE, 2018.

[19] T. Li, C. An, Z. Tian, A. T. Campbell, and X. Zhou, "Human sensing using visible light communication," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ACM, 2015.

[20] T. Li, Q. Liu, and X. Zhou, "Practical human sensing in the light," in *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services*, ACM, 2016.

[21] S. Zhang, K. Liu, Y. Ma, X. Huang, X. Gong, and Y. Zhang, "An accurate geometrical multi-target device-free localization method using light sensors," *IEEE Sensors Journal*, vol. 18, no. 18, pp. 7619–7632, 2018.

[22] C. Zhang, J. Tabor, J. Zhang, and X. Zhang, "Extending mobile interaction through near-field visible light sensing," in *Proceedings of the 21st Annual International Conference on Mobile Computing and Networking*, ACM, 2015.

[23] Y. Yang, J. Hao, J. Luo, and S. J. Pan, "CeilingSee: Device-free occupancy inference through lighting infrastructure based LED sensing," in *IEEE International Conference on Pervasive Computing and Communications (PerCom), 2017*, IEEE, 2017.

[24] Renesas, *ISL29023 Integrated Digital Light Sensor with Interrupt*, May 2014. Rev. 4.

[25] Espressif Systems, *ESP8266EX Datasheet*, November 2018. Rev. 6.

[26] S. A. Dudani, "The distance-weighted k-nearest-neighbor rule," *IEEE Transactions on Systems, Man, and Cybernetics*, no. 4, pp. 325–327, 1976.

[27] N. S. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, vol. 46, no. 3, pp. 175–185, 1992.

Appendices

Appendix 2

This chapter is republished in accordance with IEEE's copyright policy. This chapter is the accepted version of the published work, and as such may have stylistic differences from the final published article.

Device Free Localization with Capacitive Sensing Floor

Nathaniel Faulkner*, Baden Parr*, Fakhrul Alam*, Mathew Legg* and Serge Demidenko†*

*Department of Mechanical & Electrical Engineering (MEE), School of Food & Advanced Technology (SF&AT), Massey University,
Auckland 0632, New Zealand

†School of Science & Technology, Sunway University, 47500 Selangor, Malaysia

Email: *{N.Faulkner, B.Parr, F.Alam, M.Legg, S.Demidenko}@Massey.ac.nz, †SDemidenko@Sunway.edu.my

Abstract— **Passive indoor positioning has many applications including intrusion detection, fall detection of the elderly, and occupancy sensing to name a few. However, current Device Free Localization (DFL) solutions fall short of the desired accuracy requirements and are difficult to implement in a real-world scenario. This research investigates the use of a capacitive floor-based sensing solution, which can simultaneously detect multiple footsteps of a subject. The developed sensing floor prototype achieved a median positioning error of 13.5 mm and a median angular accuracy of 10.4°.**

Keywords— **Indoor localization, Device Free Localization (DFL), Capacitive Flooring, Footprint Detection**

I. INTRODUCTION

In an increasingly technologically connected world, passive indoor localization service is still a problem to be solved. Image processing or computer vision-based techniques can accurately localize and identify an un-tagged target with reasonable accuracy [1]. However, privacy is a significant concern and thus has limited utility in many applications. Whilst people may be accepting of cameras in public spaces, most people would find cameras inside their house invasive to their privacy. Many accidents, especially amongst the elderly, happen in areas where cameras would not be welcome such as in bathrooms and bedrooms. Passive localization using wireless technology has seen extensive research effort in the recent years. Wireless-based localization has the advantage of potentially being able to localize using existing infrastructure by leveraging the ubiquitous presence of wireless networks within the built environment. A survey of wireless Device Free Location (DFL) indicates a saturated research field [2]. In addition, there are some inherent disadvantages with RF wireless technology such as the limited accuracy due to multipath reflections. A more recent development has been the use of the Channel State Information (CSI) metric which uses all of the many Wi-Fi subcarriers for much improved accuracy [3]. However, commercial hardware has yet to widely support the use of this metric limiting it use to experimental setups. Passive VLP [4] works around the principle that as subjects move around a room, they cast shadows which can be detected by a light sensitive device. These shadows can then be used to estimate the subject's position. However, the passive VLP techniques are often vulnerable to change in ambient light level.

When inside a building, humans spend much of their time in contact with the floor. The major exceptions being when one is in bed, sitting with one's feet off the ground or in the bath. This therefore lends the floor a new potential purpose; namely, becoming a large sensor for both positioning and identifying people indoors. There are several works that have attempted to do this with varying degrees of success.

There have been several works that have investigated the use of pressure sensitive floors for locating and identifying people [5, 6, 7]. Pressure sensitive floors have the advantage that they are able to sense the force at which a subject's foot hits the ground, however this is offset by the generally worse spatial resolution they offer. As such pressure sensitive floors are very good for identifying people, however they appear to be complex and expensive to build and cannot handle multiple occupants in close proximity. Capacitive sensing has mainly been used for positioning and fall detection as it is much harder to utilize it for user identification. However, it has the advantage of being easier to extend for use with multiple occupants, which is an area most pressure sensitive solutions have been unable to solve.

One of the earliest capacitive systems is Smart Carpet [8], which uses fabric into which conductive wires are sewn in serpentine patterns to form 150 mm by 150 mm panels. The system is used to estimate subject's paths through a room. Similarly SensFloor [9, 10] uses conductive triangles embedded into a textile. It was able to identify individuals when used in conjunction with a hip mounted accelerometer [11]. The authors were also able to track multiple subjects, however the details are sparse. SensFloor has since been made into a commercial product [12] and more details are not available in the published literature. Rimminen et al. [13] were similarly able to track occupants in a room using metal squares of 0.3 m by 0.3 m embedded in the floor, however the authors did not investigate the localization accuracy. The authors also demonstrated that the pattern seen from the floor is different when a person is lying on it versus standing on it, however, the authors did not provide any quantitative figures. This work was significantly improved upon in [14] where the authors used a room of 4 m by 4.5 m with sensor panels of size 0.25 m by 0.5 m. The authors tested the capacitive floor on moving subjects and found that a mean positioning error of 210 mm could be achieved. Multiple target tracking was employed using Rao-Blackwellized Monte Carlo Data Association. Two subjects could be individually separated with 90% accuracy if they were more than 0.8 m apart and with 99% accuracy if they were more than 1.1 m apart. The authors also implemented fall detection in [15] which used the previous works for tracking people and then classified poses based upon their area amongst other metrics. However, the methodology is very brief and there is very limited discussion of the results. Arshad et al. describe a similar system with a very basic proof of concept showing that a change in capacitance can be detected at a metallic electrode by a microcontroller [16]. The authors discuss how this could be used for fall detection with some very limited proof of concept testing [17]. CapFloor [18] uses two sets of parallel wires orthogonal to each other. A person walking above these changes the measured capacitance in any wires that they are above. As there are two sets of wires in orthogonal directions, a person will be above at least one wire in each direction, with the intersection point of these wires being the person's

estimated position. The position error is given as "in the range of 50cm".

All the previously mentioned works have used loading mode [19] capacitive sensing. TileTrack [20] instead uses transmit mode where an additional electrode is placed in the room as a receiver. A square wave is transmitted from the floor tiles and received by the said electrode. The change in amplitude that is caused in this square wave by a person between the electrode and the floor tile can be detected. The squares are 0.6 m by 0.6 m and a frequency of 32 KHz used for the square wave. The system can position a stationary person to within 143 mm worst case and within 100 mm in 80% of cases. Several paths were tested, and the maximum error was found to be 407 mm. This work was further extended in [21]. A whole apartment of 69 m² had the floor fitted with either 0.3 m by 0.3 m squares or 0.6 m by 0.6 m squares depending on the room. The position accuracy was found to be 70 mm when standing on the on 0.3 m squares and 110 mm when standing on the 0.6 m tiles. For walking the accuracy was found to be 170 mm on the 0.3 m squares and 330 mm on the 0.6 m tiles. These accuracy values are with 90% confidence i.e. the accuracy of the 90th percentile of the data.

II. SYSTEM DEVELOPMENT

A. Key Concept

There are three main sensing modes for capacitive electric field sensing as discovered by Zimmerman et al. [22] and Smith et al. [19]: transmit mode, shunt mode and loading mode. In transmit mode the signal from the transmitter is coupled by the subject's body, which then becomes an electric field emitter. This only occurs when the subject is very close to the transmitter and the body effectively becomes an extension of the transmitter. In shunt mode, the subject's body conducts a portion of the signal to ground. The remainder of the signal which is not blocked by the subject can then be measured at the receiver. This happens when the subject is not close to either electrode. In loading mode, there is no receiver and the environment effectively forms the second plate of the capacitor to ground (Fig. 1).

A parallel plate capacitor can be modelled using the following equation:

$$C = \frac{\varepsilon_0 \varepsilon_r A}{d} \tag{1}$$

Where C is the total capacitance, ε_0 is the electric constant (8.854 × 10⁻¹² Fm⁻¹), ε_r is the relative permittivity of the dielectric (which is assumed to be constant), A is the overlapping area of the two plates, and d is the separation between the two plates. In the case of a flooring solution, a subject stands with their foot above the transmitting plate. The capacitance then depends on two main factors – the proportion of the plate covered by the subject's foot (A) and the distance between the subject's foot and the plate (d). The distance will remain fairly constant between footsteps and between users with the main factor being their footwear. Whereas the area will change very often as sensors are usually only partially covered.
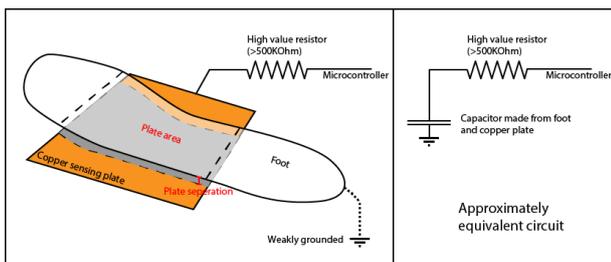
B. Hardware Design

Squares of copper are affixed to the underside of a sheet of 6 mm MDF sheet which may be seen in Fig. 2. The copper squares are 90 mm by 90 mm in size and are spaced 10 mm apart. The current testbed hardware is made up of four 0.6 m by 0.6 m panels adjacent to each other, with each panel having 36 individual copper squares (Fig. 3). Each copper square is soldered to a wire which is connected along with 35 other wires to a microcontroller where the capacitance is measured. The wires are routed along the gaps between the copper squares.

There are several ways to measure the capacitance. One can use a low frequency signal into the plate using a 30 – 100 KHz sine wave, as suggested by Smith et al. [19]. One can then measure the current of this signal using either a transimpedance amplifier or, more simply, a shunt resistor. Another method is to use the RC time constant of a capacitive circuit.

The time taken to charge a capacitor to a set voltage V_0 is given by the well-known RC charging equation:

$$V(t) = V_0(1 - e^{-t/\tau}) \tag{2}$$

Here τ is the RC time constant given by multiplying the circuit resistance by the circuit capacitance. If a high value resistance is used, the resistance can be assumed to be reasonably constant and independent of the unknown resistance to ground. The time taken for the capacitor to charge to a set value will therefore depend solely on the capacitance.

This information can be used to measure the capacitance with a microcontroller. A microcontroller's digital logic pins are set so that they have a threshold for the voltage that constitutes a digital zero and a digital one. Two digital pins are connected by a high value resistor (in the range of 1-5 M ohm). The pins shall be called the sender (S pin) and receiver (R pin). The resistor connects the two and the copper plate is attached

Fig. 2: Copper floor sensing tiles. Each tile is connected to a microcontroller with a wire.
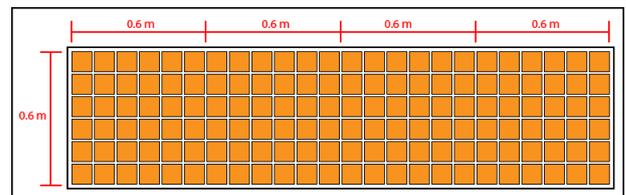

Fig. 3: Testbed copper plate layout

Fig. 1: Loading mode capacitor formed by a user's foot on the sensing floor

to the R pin. The S pin is used as an output in push-pull configuration whilst the R pin is used as an input. The S pin is set to output a logic low (GND) and a certain amount of time is waited so that the R pin has time to settle. A timer is started, and the S pin is set to output a logic high (3.3 v). The timer is stopped as soon as the R pin registers a digital 1 read. This process is repeated multiple times to reduce the measurement noise through averaging. When a subject is near the copper plate the effective capacitance is much greater than when there is no subject nearby. This leads to the pin having a much longer rise time when there is a subject in close proximity.

The circuit board is equipped with a 100 pin ARM Cortex M3 from ST Microelectronics [23]. The majority of the aforementioned pins are used for the capacitance measurements. An ESP8266 Wi-Fi module [24] is used for communications. The circuit board is shown in Fig. 4 and a block diagram of the entire system is shown in Fig. 5. The PC app is written primarily in JavaScript (nodeJS) using the electron wrapper to package it as a GUI desktop app. It hosts an http server which the devices POST data to. The IP of the server is currently hardcoded into the firmware of the ESP8266, however a small wireless router is used to create a subnet onto which all the devices and the PC running the app are connected. The IP address of the PC can then be configured through the router's DHCP server. The app has several main functions. The first is that it displays a live real time feed directly from the floor with the option to display the output of the foot detection algorithms overlaid on top. One can also use the app to configure parameters for the algorithm in real time. It also allows for recording of incoming data to a file for replaying and later analysis.
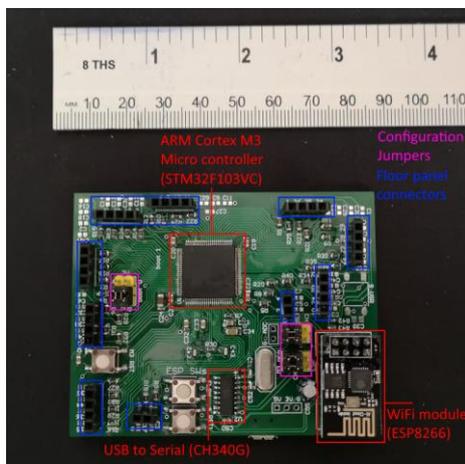

Fig. 4: The custom designed electronic hardware used to measure the capacitance of the sensing floor tiles.
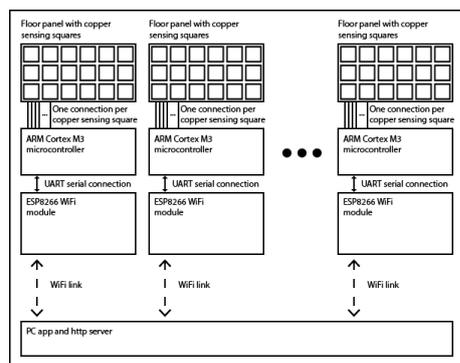
Fig. 5: Block diagram of the sensing floor

When the system is first powered on, a number (currently set to 10 after limited empirical testing) of capacitance readings are taken from the floor. These readings are then used as initial baseline readings which are then subtracted from each subsequent capacitance measurement from the floor. Over time these baseline readings tend to drift, so to counteract this several measures can be taken. Firstly, one can manually recalibrate the system by taking a new set of baseline capacitance readings periodically when the system is known to be empty. A more automated method is to take a long-term average of all capacitance readings taken whilst the system is in use and use this long-term average as the baseline. The assumption being that over a long period of time the amount of time in which a subject is standing on a square is small compared to that in which a subject is not standing on a square. However, this does mean that if a person stands still for a very long period of time they will eventually be lost by the system.

C. Foot Detection

Foot detection is done using the following algorithm. Firstly, the capacitance values from the floor are interpolated to improve the resolution. Several interpolations have been tried, with cubic interpolation performing the best. A threshold is then applied to the data, such that any capacitance values below the threshold value are set to zero and any capacitance values above the threshold value set to one. Once this has occurred, cluster detection is applied whereby all connected squares are considered to be a cluster. In the future, a more sophisticated clustering method can be used. Each cluster, representing a single footprint, can then be represented by a 2 x N matrix, M where N is the number of data points in the cluster. Each column of the matrix is a vector representing the position of a single data point in the cluster. Figure 6 shows this process. The centre of the footprint $(\bar{x}, \bar{y})$ is currently taken by averaging the position of each point in the 2 x N cluster matrix M as follows:

$$\bar{x} = \frac{\sum_{i=1}^{N} M_{1,i}}{N}$$
$$\bar{y} = \frac{\sum_{i=1}^{N} M_{2,i}}{N} \tag{3}$$

The orientation of the footprint is then found by using Principal Component Analysis (PCA) [25]. The covariance of two vectors can be calculated as follows:

$$cov(a, b) = \frac{\sum_{i=1}^{N}(a_i - \bar{a})(b_i - \bar{b})}{N} \tag{4}$$

Fig. 6: Foot detection process

A covariance matrix M_{cov} can be formed by taking the top row of M to be the vector, a and the bottom row the vector, b. The vector a is a vector containing all th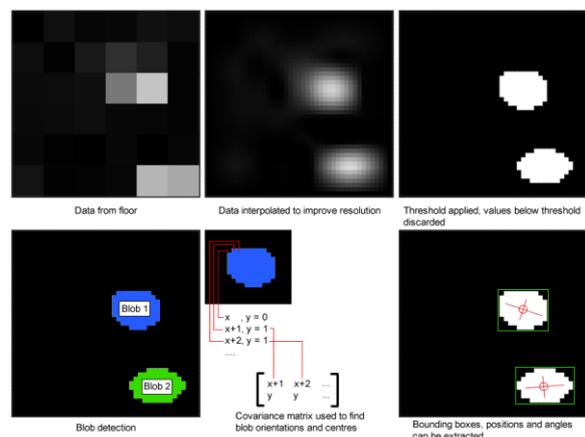e x positions of each point in the cluster and the vector b is a vector containing all the y positions of each point in the cluster.

$$M_{cov} = \begin{bmatrix} cov(a,a) & cov(a,b) \\ cov(b,a) & cov(b,b) \end{bmatrix} \qquad (5)$$

The eigenvectors of this matrix can then be used to calculate the vectors of the orientation of the foot. The bounding box of the foot can be found by taking the maximum and minimum x and y positions of the points in the cluster.

III. SYSTEM PERFORMANCE

A. Position Accuracy

To investigate the position accuracy of a subject's footprint on the floor, fifteen locations were chosen. The position of the subject's foot was measured using a ruler and measurements were taken from the floor itself. The ruler was used as the ground truth to verify the position estimates from the sensing floor against. The outline of the subject's right foot was drawn onto a sheet of cardboard and cut out. A square corner was left protruding from the top left to measure against. The distance was measured from both the top and right edges of the sensing floor to this protruding corner of the footprint. This is shown in Fig. 7. An attempt was made to keep the foot's orientation constant between measurements but was only done by visual estimation and therefore, the orientation varied by approximately ±5° between measurements.

After performing the test at the 15 locations on the sensing floor, the median position error was found to be 13.5 mm and the maximum position error was found to be 25.6 mm. Figure 8 shows the positions of each location and the error at each location.

It should be noted that the ground truth was measured to the top right corner of the foot outline, whereas the floor is estimating the position of the centre of the foot. Therefore, a variation in the orientation of the foot causes the error to increase due to the offset. Hence the measured errors are likely to be in part due to the methodology and it is believed that the error could potentially be lower, with a more accurate ground truth. Also, the ground truth and estimated values are at different positions on the foot, the estimated results must be translated so that they match up. This translation is applied uniformly to all the estimated results. However, the calculation of the translation assumes that the errors are evenly distributed in all directions and therefore the translation is taken to be the median of the error on each axis.
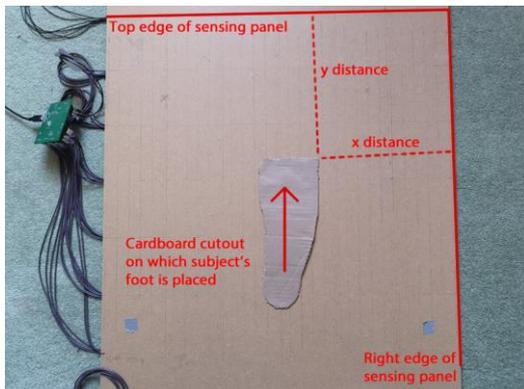
Fig. 7: Experimental setup for sensing floor position accuracy testing.
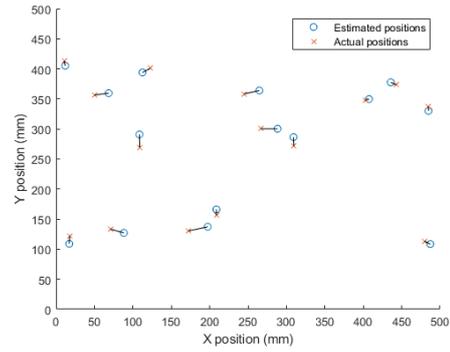
Fig. 8: Estimated vs actual positions of a test subject's foot on the sensing floor

B. Angular Accuracy

A test was undertaken to investigate the accuracy of estimating the subject's foot angle. A similar setup was used to that in the position accuracy testing. Using a protractor, lines were marked out at 10-degree increments from 0° to 90° and an extra line at 45°. The same cardboard cut-out was used to locate the foot with minor modifications. A slit was added down the centre of the foot cut out so that the lines can be seen underneath as well as the origin point about which the rotation was done. Capacitance samples were taken over a period of 5 seconds at each angle and from this the foot angle is estimated. The setup can be seen in Fig. 9.

The median angular error was found to be 10.4° and the maximum angular error was found to be 18.8°. Figure 10 shows the error for different angular orientations of the subject's foot. The error shows strong signs of non-linearity which means it may be possible to correct for this in the detection algorithm. Further investigation is required as this may depend on the location of the foot with respect to the copper squares underneath the floor. Therefore, this same test needs to be performed at different locations on the floor. Improving the accuracy is desirable as the foot angle is a possible metric to be investigated for gait identification.
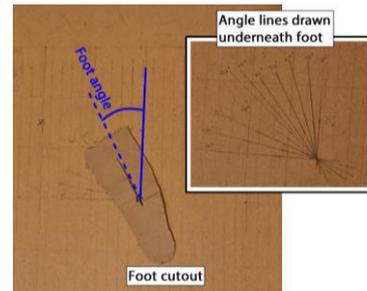
Fig. 9: Angular accuracy testing setup
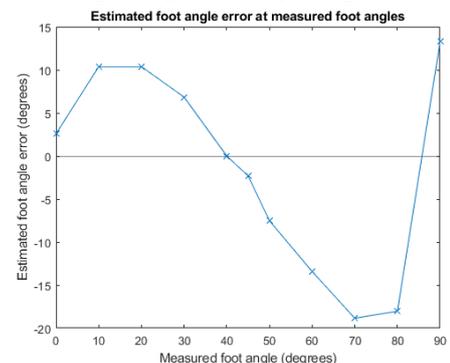
Fig. 10: Estimated foot angle error with respect to actual foot angle.

Appendices

C. Foot Detection

Currently, whilst detection of multiple feet has been implemented, no quantifiable data on the performance has been collected. The system has been tested with multiple subjects and can detect the feet of several subjects concurrently given that they are sufficiently spaced apart. Figure 11 shows both of a subject's feet being detected individually. The position of each foot is marked by the intersection of the red lines, with the longer line corresponding to the orientation. The orientation is only valid in a 180-degree hemisphere. This means that assumptions must be made about the direction the foot is facing. One can assume that people generally walk forwards rather than backwards, so over the course of several footsteps, one can deduce the direction of the foot. It has been observed that feet on adjacent squares can get lost as they merge with each other into a larger blob. As the copper sensing squares are 100 mm wide, providing the feet are greater than around 150 mm apart, they do not appear to alias. This is because the partial occlusion of feet at the very edge of adjacent squares does not put them over the threshold.

IV. CONCLUSION AND FUTURE WORKS

The developed capacitive floor can position a subject's foot with a median position error of 13.5 mm and a median angular error of 10.4°. It has the potential to be an accurate, yet noninvasive passive localization system. The current solution is still in early stages of development with scope for future improvements. Whilst multiple footprints can be simultaneously located, estimation of a subject's body position from a set of successive footprints has not been implemented. Further work is needed to identify an individual from their gait pattern and develop a classification model to detect poses of people lying on the floor. This could then be used to monitor for falls and if necessary, alert caregivers or emergency services to such an event.

Fig. 11: Multiple foot detection as seen in the PC app.

REFERENCES

[1] R. Poppe, "A survey on vision-based human action recognition," Image and vision computing, vol. 28, no. 6, pp. 976-990, 2010.

[2] S. Palipana, B. Pietropaoli and D. Pesch, "Recent advances in RF-based passive device-free localisation for indoor applications," Ad Hoc Networks, vol. 64, pp. 80-98, 2017.

[3] Z. Yang, Z. Zhou and Y. Liu, "From RSSI to CSI: Indoor localization via channel response," ACM Computing Surveys (CSUR), vol. 46, no. 2, p. 25, 2013.

[4] N. Faulkner, F. Alam, M. Legg and S. Demidenko, "Smart Wall: Passive Visible Light Positioning with Ambient Light Only," in 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2019.

[5] M. Addlesee, A. Jones, F. Livesey and F. Samaria, "The ORL active floor," IEEE Personal Communications, vol. 4, no. 5, pp. 35-41, 1997.

[6] S. Pirttikangas, J. Suutala, J. Riekki and J. Röning, "Footstep identification from pressure signals using Hidden Markov Models," in Proc. Finnish Signal Processing Symposium (FINSIG'03), 2003.

[7] J.-W. Jung, T. Sato and Z. Bien, "Dynamic footprint-based person recognition method using a hidden markov model and a neural network," International journal of intelligent systems, vol. 19, no. 11, pp. 1127-1141, 2004.

[8] D. Savio and T. Ludwig, "Smart carpet: A footstep tracking interface," in 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07), 2007.

[9] C. Lauterbach, A. Steinhage and A. Techmer, "Large-area wireless sensor system based on smart textiles," in International Multi-Conference on Systems, Signals & Devices, 2012.

[10] A. Steinhage, C. Lauterbach and A. Techmer, "Large-Area Wireless Sensor System for Ambient Assisted Living," in AMA Conferences Proceedings SENSOR 2013, Nürnberg, 2013.

[11] M. Sousa, A. Techmer, A. Steinhage, C. Lauterbach and P. Lukowicz, "Human tracking and identification using a sensitive floor and wearable accelerometers," in 2013 IEEE International Conference on Pervasive Computing and Communications (PerCom), 2013.

[12] Future-Shape GmbH, "Solutions - Future-Shape EN," [Online]. Available: https://future-shape.com/en. [Accessed October 2019].

[13] H. Rimminen, M. Linnavuo and R. Sepponen, "Human tracking using near field imaging," in 2008 Second International Conference on Pervasive Computing Technologies for Healthcare, 2008.

[14] H. Rimminen, J. Lindström and R. Sepponen, "Positioning accuracy and multi-target separation with a human tracking system using near field imaging," International Journal on Smart Sensing and Intelligent Systems, vol. 2, no. 1, pp. 156-175, 2009.

[15] H. Rimminen, J. Lindström, M. Linnavuo and R. Sepponen, "Detection of falls among the elderly by a floor sensor using the electric near field," IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 6, pp. 1475-1476, 2010.

[16] A. Arshad, S. Khan, A. Z. Alam, R. Tasnim, T. S. Gunawan, R. Ahmad and C. Nataraj, "An activity monitoring system for senior citizens living independently using capacitive sensing technique," in 2016 IEEE International Instrumentation and Measurement Technology Conference Proceedings, 2016.

[17] A. Arshad, S. Khan, A. Z. Alam, A. F. Ismail and R. Tasnim, "Capacitive proximity floor sensing system for elderly tracking and fall detection," in 2017 IEEE 4th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA), 2017.

[18] A. Braun, H. Heggen and R. Wichert, "CapFloor–a flexible capacitive indoor localization system.," in International Competition on Evaluating AAL Systems through Competitive Benchmarking, Springer, Berlin, Heidelberg, 2011.

[19] J. Smith, T. White, C. Dodge, J. Paradiso, N. Gershenfeld and D. Allport, "Electric field sensing for graphical interfaces," IEEE Computer Graphics and Applications, vol. 18, no. 3, pp. 54-60, 1998.

[20] M. Valtonen, J. Maentausta and J. Vanhala, "Tiletrack: Capacitive human tracking using floor tiles," in 2009 IEEE international conference on pervasive computing and communications, 2009.

[21] M. Valtonen, T. Vuorela, L. Kaila and J. Vanhala, "Capacitive indoor positioning and contact sensing for activity recognition in smart homes," Journal of Ambient Intelligence and Smart Environments, vol. 4, no. 4, pp. 305-334, 2012.

[22] T. G. Zimmerman, J. R. Smith, J. A. Paradiso, D. Allport and N. Gershenfeld, "Applying electric field sensing to human-computer interfaces," in Proceedings of CHI'95; ACM Conference on Human Factors in Computing, 1995.

[23] STMicroelectronics, "STM32F103xC, STM32F103xD, STM32F103xE Datasheet," July 2018. [Online]. Available: https://www.st.com/resource/en/datasheet/cd00191185.pdf. [Accessed October 2019].

[24] Espressif, "ESP8266 Technical Reference Version 1.4," August 2019. [Online]. Available: https://www.espressif.com/sites/default/files/documentation/esp8266-technical_reference_en.pdf. [Accessed October 2019].

[25] I. Jolliffe, Principal Component Analysis, Second Edition, New York: Springer-Verlag, 2002.

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Nathaniel Faulkner
Name/title of Primary Supervisor:	Associate Professor Fakhrul Alam

In which chapter is the manuscript /published work: Chapter 2

Please select one of the following three options: .

◉   The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

N. Faulkner, F. Alam, M. Legg and S. Demidenko, "Watchers on the Wall: Passive Visible Light-Based Positioning and Tracking With Embedded Light-Sensors on the Wall," in IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 5, pp. 2522-2532, May 2020, doi: 10.1109/TIM.2019.2953373

◯   The manuscript is currently under review for publication – please indicate:

- The name of the journal:

- The percentage of the manuscript/published work that was contributed by the candidate: 90.00

- Describe the contribution that the candidate has made to the manuscript/published work:

The candidate developed the hardware, performed the data collection and analysis, and produced the first draft of the article.

◯   It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Nathaniel Faulkner *Digitally signed by Nathaniel Faulkner Date: 2021.11.22 17:31:43 Z*
Date:	22-Nov-2021
Primary Supervisor's Signature:	Fakhrul Alam *Digitally signed by Fakhrul Alam DN: cn=Fakhrul Alam, c=NZ, o=Massey University, ou=SEAT, email=f.alam@massey.ac.nz Date: 2021.11.25 08:48:25 +13'00'*
Date:	25-Nov-2021

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Nathaniel Faulkner
Name/title of Primary Supervisor:	Associate Professor Fakhrul Alam

In which chapter is the manuscript /published work:	Chapter 3

Please select one of the following three options:

◉ The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

N. Faulkner, B. Parr, F. Alam, M. Legg and S. Demidenko, "CapLoc: Capacitive Sensing Floor for Device-Free Localization and Fall Detection," in IEEE Access, vol. 8, pp. 187353-187364, 2020, doi: 10.1109/ACCESS.2020.3029971

○ The manuscript is currently under review for publication – please indicate:

- The name of the journal:

- The percentage of the manuscript/published work that was contributed by the candidate: 90.00

- Describe the contribution that the candidate has made to the manuscript/published work:

The candidate developed the hardware, performed the data collection and analysis, and produced the first draft of the article.

○ It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Nathaniel Faulkner	Digitally signed by Nathaniel Faulkner Date: 2021.11.22 17:29:00 Z
Date:	22-Nov-2021	
Primary Supervisor's Signature:	Fakhrul Alam	Digitally signed by Fakhrul Alam DN: cn=Fakhrul Alam, c=NZ, o=Massey University, ou=SEAT, email=f.alam@massey.ac.nz Date: 2021.11.25 08:50:42 +13'00'
Date:	25-Nov-2021	

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Nathaniel Faulkner
Name/title of Primary Supervisor:	Associate Professor Fakhrul Alam

In which chapter is the manuscript /published work:　　Chapter 4

Please select one of the following three options: .

⦿   The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

N. Faulkner, F. Alam, M. Legg and S. Demidenko, "Device-Free Localization Using Privacy-Preserving Infrared Signatures Acquired From Thermopiles and Machine Learning," in IEEE Access, vol. 9, pp. 81786-81797, 2021, doi: 10.1109/ACCESS.2021.3086431

◯   The manuscript is currently under review for publication – please indicate:

- The name of the journal:

- The percentage of the manuscript/published work that was contributed by the candidate:　　90.00

- Describe the contribution that the candidate has made to the manuscript/published work:

The candidate developed the hardware, performed the data collection and analysis, developed the localisation algorithms, and produced the first draft of the article.

◯   It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Nathaniel Faulkner　Digitally signed by Nathaniel Faulkner Date: 2021.11.22 17:30:44 Z
Date:	22-Nov-2021
Primary Supervisor's Signature:	Fakhrul Alam　Digitally signed by Fakhrul Alam DN: cn=Fakhrul Alam, c=NZ, o=Massey University, ou=SEAT, email=f.alam@massey.ac.nz Date: 2021.11.25 08:49:19 +13'00'
Date:	25-Nov-2021

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Nathaniel Faulkner
Name/title of Primary Supervisor:	Associate Professor Fakhrul Alam

In which chapter is the manuscript /published work:	Chapter 5

Please select one of the following three options: .

◉   The manuscript/published work is published or in press

   •   Please provide the full reference of the Research Output:
   N. Faulkner, D. Konings, F. Alam, M. Legg and S. Demidenko, "Machine Learning Techniques for Device-Free Localization Using Low-Resolution Thermopiles," in IEEE Internet of Things Journal, doi: 10.1109/JIOT.2022.3161646.

◯   The manuscript is currently under review for publication – please indicate:

   •   The name of the journal:

   •   The percentage of the manuscript/published work that   90.00
        was contributed by the candidate:

   •   Describe the contribution that the candidate has made to the manuscript/published work:
   The candidate developed the hardware, performed the data collection and analysis, implemented the machine learning algorithms with the assistance of the second author, and produced the first draft of the article.

◯   It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Nathaniel Faulkner Digitally signed by Nathaniel Faulkner Date: 2022.06.08 21:04:06 +01'00'
Date:	08-Jun-2022
Primary Supervisor's Signature:	Fakhrul Alam Digitally signed by Fakhrul Alam DN: cn=Fakhrul Alam, c=NZ, o=Massey University, ou=SF&AT, email=f.alam@massey.ac.nz Date: 2022.06.09 08:08:06 +12'00'
Date:	8-Jun-2022

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

GRADUATE RESEARCH SCHOOL

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA
UNIVERSITY OF NEW ZEALAND

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Nathaniel Faulkner
Name/title of Primary Supervisor:	Associate Professor Fakhrul Alam

In which chapter is the manuscript /published work:	Appendix 1

Please select one of the following three options:

◉ The manuscript/published work is published or in press

  • Please provide the full reference of the Research Output:

  N. Faulkner, F. Alam, M. Legg and S. Demidenko, "Smart Wall: Passive Visible Light Positioning with Ambient Light Only," 2019 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), 2019, pp. 1-6, doi: 10.1109/I2MTC.2019.8826875

◯ The manuscript is currently under review for publication – please indicate:

  • The name of the journal:

  • The percentage of the manuscript/published work that was contributed by the candidate: 90.00

  • Describe the contribution that the candidate has made to the manuscript/published work:

  The candidate developed the hardware, performed the data collection and analysis, and produced the first draft of the article.

◯ It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Nathaniel Faulkner Digitally signed by Nathaniel Faulkner Date: 2021.11.22 17:30:26 Z
Date:	22-Nov-2021
Primary Supervisor's Signature:	Fakhrul Alam Digitally signed by Fakhrul Alam DN: cn=Fakhrul Alam, c=NZ, o=Massey University, ou=SEAT, email=f.alam@massey.ac.nz Date: 2021.11.25 08:49:37 +13'00'
Date:	25-Nov-2021

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.

STATEMENT OF CONTRIBUTION
DOCTORATE WITH PUBLICATIONS/MANUSCRIPTS

We, the candidate and the candidate's Primary Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of candidate:	Nathaniel Faulkner
Name/title of Primary Supervisor:	Associate Professor Fakhrul Alam

In which chapter is the manuscript /published work: Appendix 2

Please select one of the following three options:

.

◉ The manuscript/published work is published or in press

- Please provide the full reference of the Research Output:

N. Faulkner, B. Parr, F. Alam, M. Legg and S. Demidenko, "Device Free Localization with Capacitive Sensing Floor," 2020 IEEE Sensors Applications Symposium (SAS), 2020, pp. 1-5, doi: 10.1109/SAS48726.2020.9220042

○ The manuscript is currently under review for publication – please indicate:

- The name of the journal:

- The percentage of the manuscript/published work that was contributed by the candidate: 90.00

- Describe the contribution that the candidate has made to the manuscript/published work:

The candidate developed the hardware, performed the data collection and analysis, and produced the first draft of the article.

○ It is intended that the manuscript will be published, but it has not yet been submitted to a journal

Candidate's Signature:	Nathaniel Faulkner Digitally signed by Nathaniel Faulkner Date: 2021.11.22 17:29:45 Z
Date:	22-Nov-2021
Primary Supervisor's Signature:	Fakhrul Alam Digitally signed by Fakhrul Alam DN: cn=Fakhrul Alam, c=NZ, o=Massey University, ou=SEAT, email=f.alam@massey.ac.nz Date: 2021.11.25 08:50:18 +13'00'
Date:	25-Nov-2021

This form should appear at the end of each thesis chapter/section/appendix submitted as a manuscript/ publication or collected as an appendix at the end of the thesis.