A Long-Term Study Of mmWave Sensing In An Outdoor Urban Scenario

Wang, Weizheng; Vaidya, Girish; Bhattacharjee, Anup; Fioranelli, Francesco ; Zuniga, Marco

DOI
10.1109/DCOSS-IoT58021.2023.00049

Publication date
2023

Document Version
Final published version

Published in
Proceedings of the 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)

Citation (APA)
Wang, W., Vaidya, G., Bhattacharjee, A., Fioranelli, F., & Zuniga, M. (2023). A Long-Term Study Of mmWave Sensing In An Outdoor Urban Scenario. In L. O'Conner (Ed.), *Proceedings of the 2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)* (pp. 240-247). IEEE. https://doi.org/10.1109/DCOSS-IoT58021.2023.00049

2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT)

2023 19th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT) | 979-8-3503-4649-7/23/$31.00 ©2023 IEEE | DOI: 10.1109/DCOSS-IOT58021.2023.00049

A Long-Term Study Of mmWave Sensing In An Outdoor Urban Scenario

Weizheng Wang*, Girish Vaidya*†, Anup Bhattacharjee*, Francesco Fioranelli*, Marco Zuniga*
* *Delft University of Technology, Delft, Netherlands*
† *Amsterdam Institute for Advanced Metropolitan Solutions, Amsterdam, Netherlands*
Email: {w.wang-14, g.vaidya, a.k.bhattacharjee, f.fioranelli, m.a.zunigazamalloa}@tudelft.nl

***Abstract*—Sensing people with mmWave radars is gaining significant attention. This growing interest is due to two factors: radar monitoring provides more privacy than camera-based alternatives, and radio waves are not as easily blocked as light waves. Most mmWave studies, however, have three common characteristics. They are done *indoors*, without protecting the sensor (*no casing*), and the evaluation is performed for *short periods* of time. To assess the suitability of mmWave sensing in realistic outdoor scenarios, we deploy two nodes to track the flow of pedestrians over a period of three months. This long-term deployment provides three main contributions. First, we follow a detailed process to design a casing that can protect the sensors from harsh environmental conditions. Second, we install our nodes close to a set of cameras that were already deployed in the area. To compare the performance of both types of sensors, we propose a framework that considers the different coverage patterns of cameras and radars. Third, the time frame of our evaluation considers various types of weather, from sunny days to rainy and windy. Our results indicate that mmWave sensors need to be explored further outside the comfort zone of indoor spaces. To the best of our knowledge, this is the first long-term study assessing the reliability of radar sensors in the 60 GHz ISM band.**

I. Introduction

A. Motivation

The modernization of the urban infrastructure promises efficient and safe delivery of services in public places. Many of these services require sensing people's flow. For instance, pedestrian monitoring can avoid unsafe crowding situations. Similarly, through active crowd management at tourist destinations, it is possible for civic authorities to intervene in a timely manner [1].

There are two popular mechanisms for sensing people-traffic, viz., device-based and vision-based. The device-based mechanisms count the number of devices, such as mobile phones, to estimate the traffic in a particular area [2]. The main limitation of this method is that it assumes that every individual carries one device. On the other hand, the affordability of cameras and the simultaneous advancement in vision processing algorithms have made vision-based systems a popular choice for monitoring people-movement. The city of Amsterdam alone has more than 1000 registered cameras deployed for crowd sensing [3].

While the deployment of sensors is necessary for monitoring the crowd flow, the increasing number of cameras in public spaces can be problematic. Cities such as London, with more than six hundred thousand surveillance cameras, have triggered concerns amongst privacy advocates [4]. To circumvent this concern, a few privacy-friendly cameras are commercially available. Some cameras blur the faces, while others can be configured not to share any images and instead pass only the people count for further analysis. However, such software mechanisms may still be prone to hacking. More importantly, there is the issue of *perceived privacy*, which cause many citizens to have strong reservations even with privacy-aware cameras [5].

B. Challenges in mmWave sensing

Millimeter wave (mmWave) radar is emerging as a promising alternative for privacy-preserving crowd monitoring. mmWave radar detects people and objects as point clouds, making it difficult to collect personal information. The mmWave radars are available as compact single-chip solutions, and this improved availability and affordability has encouraged their wider acceptance and deployment.

While recent works have explored the possibility of mmWave radar deployment for sensing people's flow [6]–[8], they have several limitations. First, most deployments have been done under indoor conditions. For urban applications, the mmWave sensor must be deployed outdoors. This implies that the radar must be enclosed in a radome[1]. The design of the radome –which comprises its material, thickness, and placement of radar– significantly influences the performance of the system. Secondly, the studies have been conducted for a short duration, thus not capturing the impact of sustained operations under varying weather conditions. Thirdly, most of the current literature captures the performance of the radar under a controlled flow of people. Through this work, we design a suitable radome and investigate the performance of mmWave sensing with a flow of real-world pedestrians and bikers in the license-free band (60 GHz).

C. Our contribution

To overcome the above limitations, we perform a thorough evaluation of mmWave radar using the IWR6843ISK platform [9]. In particular, our contribution is three-fold.

- *A careful radome design [Section III]*. We perform a careful set of experiments to design an appropriate radome. Our design considers important parameters such as the width of the material and the exact location of the radar sensor inside the radome. We compare the performance

[1]A radome is a structure to protect radar equipment, and it has to be made from material that is transparent to radio waves

2325-2944/23/$31.00 ©2023 IEEE
DOI 10.1109/DCOSS-IoT58021.2023.00049

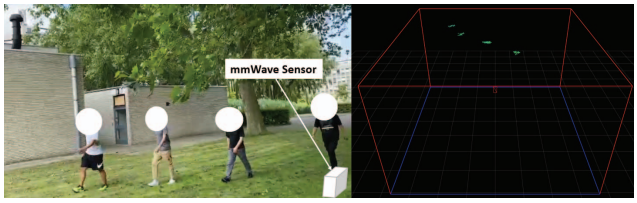


Fig. 1: Point Cloud Tracking

of our prototype with a commercial alternative and show that we perform better in terms of coverage.

- *A real-world deployment [Section IV].* We deploy two mmWave radar sensors to monitor real-world people's movement and compare their outputs with those from commercial camera-based sensors. The location is regularly used by cyclists and pedestrians. We gather per-minute data over a period of three months. The time frame considers sunny, foggy, rainy, and windy days.
- *A detailed evaluation [Sections V and VI].* We utilize precision, recall, and correlation metrics with different levels of granularity to quantify the performance of mmWave sensing. Our results show that, compared to the high accuracy of indoor scenarios, radar systems outdoors still need to improve. Detecting empty spaces has an accuracy of 85% and estimating the flow of people has correlation coefficients between 60% and 85%, depending on the weather condition. We believe that, compared to the error rate provided by people-counting cameras (below 10%), the current performance of radars still needs to improve.

II. Background

A. Properties of mmWave sensing

Frequency Modulated Continuous Wave (FMCW) radars are a special class of mmWave radars, that are commercially available in a miniaturized form-factor for wide-scale deployment. FMCW radars have a carefully designed array of transmitter and receiver antennas. The radar sends chirp signals, which are reflected by the environment and processed by the system to provide point clouds for all the *moving* objects, as shown in Figure 1. The point clouds provide angular information of the objects, as well as the range and velocity [10].

The fact that people, or any dynamic element, are represented as points is the key strength of mmWave sensors to maintain privacy. In the GHz bands, the wavelengths are in the order of a few millimeters, which are too long to capture the fine granularity present in human faces. Thus, even if the radar platform is hacked to obtain the raw signals, the attacker will be fundamentally limited regarding the depth of information that can be attained.

To capture mobile elements as cloud points, the mmWave radar has algorithms that remove *static* clutter [11]. Since indoor scenarios offer a *stable* static background, mobile elements reflect clear signals. Furthermore, indoor evaluations do not require adding a case, and hence, do not need to consider the attenuation and distortion that is inherent to mechanical enclosures. These two properties, static background and no need for casings, allow mmWave studies to deliver indoor applications that not only provide tracking [12] but also 3D skeleton reconstruction [13] and vital sign monitoring [14]. Outdoors, the radars need casing and the background is not static. Animals can be under the field of view and weather conditions can affect the quality of the reflected signal (rain or fog). Furthermore, wind also has a pernicious effect because it blows leaves, plants, grass and other objects that appear in the point cloud measurements. It is, hence, more challenging to perform mmWave sensing in outdoor scenarios.

B. Frequency ranges for mmWave sensing

Another element that is often overlooked in mmWave studies is the operational band. mmWave radars are available in three bands: 24GHz, 60GHz and 77GHz. The 77GHz band is generally reserved for automotive applications. Initially, we wanted to use a sensor on this band (to complement the traffic information gathered by cars), but we consulted a municipality that did not approve of its use. The 24GHz band has two main components, an Ultra-Wide Band (UWB) and a NarrowBand (NB). Spectrum regulations and standards developed by the European Telecommunications Standards Institute (ETSI) and Federal Communications Commission (FCC) prohibit new industrial products from using the 24GHz UWB. The NB offers a limited bandwidth of 250MHz, adversely affecting people-detection in this spectrum. On the contrary, the 60GHz sensor operates in the free ISM band and offers a wide bandwidth of 4GHz, providing richer point clouds and thus better object identification and tracking [15]. In our deployment, we use the IWR6843ISK platform from Texas Instruments (TI) in the 60GHz band.

III. Radome design

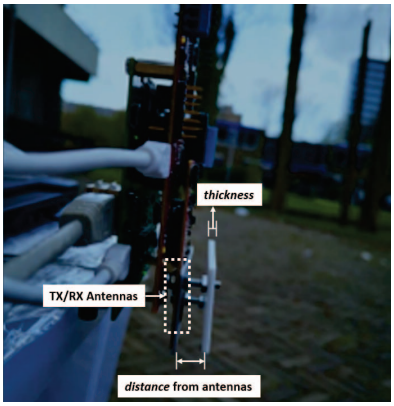
An important consideration for any outdoor radar deployment is the design of the casing to protect the hardware from weather-related phenomena. This enclosure is called the radome and it needs to be transparent to radio waves.

We considered using a commercial product that already has a radome but each node costs 600 $ and the node is designed for indoor scenarios [16]. Given that the price of the mmWave chip is 120 $, significant savings can be achieved if we design our own casing. Later in this section, we compare our design with the commercial product and show that our prototype not only saves costs but also attains a better performance.

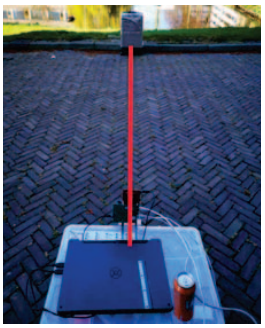
A. Parameter analysis

To design a radome, one needs to take into account two main aspects. The first one is simple: To consider the mechanical and electronic requirements for placing the radar chip, the microcontroller, and the power circuitry. The second aspect is more complex: To analyze electromagnetic effects in order to minimize distortions to the emitted and received signals.

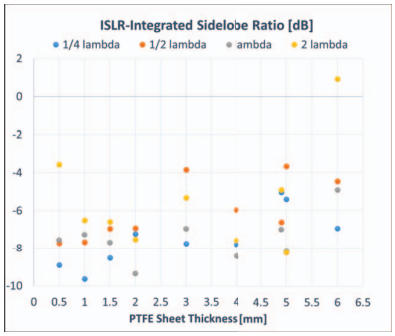
A thorough design of a radome requires full electromagnetic simulations [17], [18]. Our objective is not to make an optimal design from an electromagnetic perspective, but to build an inexpensive casing that does not hinder performance. With that goal in mind, we focus on three key parameters: the type of material, its thickness, and the distance of the radome to the radar's antennas. These parameters are captured in Figure 2a.



(a) Key parameters for a radome's design.



(b) Experimental setup for radome design, the radar has a front aluminum plate.



(c) Results of the tests with PTFE (Teflon) sheets of different thicknesses and distances from the radar antennas.

Fig. 2: The experiments for radome design.

Material. Following the suggestion from the manufacturer (Texas Instruments), we use PTFE (Teflon) as the material for our casing. Teflon has a relatively low cost and it is easy to cut (manipulate) into the desired shape. Furthermore, there are off-the-shelf sheets available with different *thicknesses*, which is important because the thickness is another critical parameter.

Thickness. In mmWave systems, the signal's bandwidth usually spans several GHz. This means that the wavelength value will vary across such bandwidth and thus the optimal radome thickness will vary as well. Thus, in our evaluation we test sheets with nine thickness values, equal to 0.5, 1.0, 1.5, 2.0, 2.9($\sim$ 3.0), 4.0, 4.9($\sim$ 5.0), 5.1, and 6.0 mm.

Radome-antenna distance. The *radome-antenna distance* is determined by the signal's wavelength. Theoretically, it can be demonstrated that the optimal distance to minimize reflections caused by the radome is an integer multiple of half the wavelength in air [17], [18]. Using the central frequency as the reference point, our evaluation considers four distance values: 1.25 mm, 2.5 mm, 5 mm, and 10 mm, corresponding in terms of wavelength to $\lambda/4$, $\lambda/2$, λ and 2λ.

Shape. For imaging applications, the best shape is a spherical casing to ensure that the waves' propagation distance traveled within the dielectric radome is constant with respect to the angle of the radar's field of view [17], [18]. Similar to the commercial product we use as a benchmark (explained next), we use a flat surface due to its simplicity and the fact that we are not focused on exact imaging applications, but on simple tracking.

B. Evaluation.

To obtain the final radome's design, we test Teflon sheets considering all 36 combinations of thicknesses (9 values), and antenna-teflon distances (4 values).

To test the different configurations we use a flat aluminum plate as a target (due to its strong reflection). This plate is placed in front of the radar at the boresight, at a distance of approximately 6m. Figure 2b captures the setup used for the radome design.

As a quantitative metric to compare the different configurations, we use the Integrated Side Lobe Ratio (ISLR). This is the ratio between the energy of the sidelobes in the radar range profile and the energy of the main lobe associated with the target response. Lower values indicate better performance in the sense that the target lobe will be more easily detectable.

Figure 2c shows the ISLR results. It can be seen that the best values are obtained for PTFE sheets of thickness equal to 1 mm and a distance from the antennas of $\lambda/4$, i.e. 1.25 mm. These are the values used for the final fabrication of the radome, shown in Figure 4. It is important to highlight the importance of the empirical evaluation because the theoretical values for thickness and distance are not optimal.

TABLE I: Costs of Materials

Component	Price($)/piece(excl. Taxes)
mmWave sensor: IWR6843ISK	120.87
Processor: RaspberryPi 4B	49.77
Casing: Front Radome (Teflon)	101.12
Casing: Back (Plastic)	12.74
Other (adapters, cables, etc.)	≈26.50
Total Price($) (excl. taxes)	**311**

C. Comparison with commercial product

We bought one unit of a commercial product, viz., WAYV Air from Ainstein, to benchmark the performance of our system. It uses the same mmWave chipset we use in the 60 GHz band. The brochure states that the sensor is designed for indoor scenarios with a maximum range of 6 m. Considering that the sensor is not designed for outdoor use (the casing does not appear to be weatherproof), our evaluation is done on a clear day in the same setup as our platform, c.f. Figure 3. In this outdoor evaluation, the range was around 2.5 m, and the system had trouble detecting more than three people. The cost of the Ainstein sensor is 600 $, while the total cost of our system is 310 $. The itemized costs are presented in Table I. Besides reducing our costs by half, we also triple the range. In the next section, we describe the configuration used to achieve a range around 8 m. We cannot make conclusive remarks about the reasons for the differences in the design of our systems since all the design details of the Ainstein sensor are not available.

Fig. 3: Scenario. The gray path is for pedestrians and the orange path is for bikes. The scenario has two cameras (outside our control) and two of our mmWave sensors.

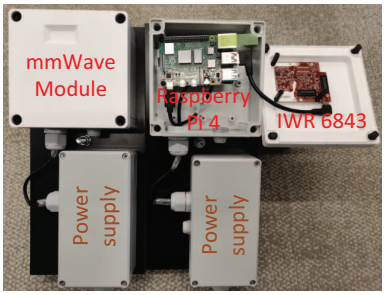


Fig. 4: The mmWave modules installed on poles.

IV. Experiment Setup

Two mmWave sensors are deployed on our university campus. The data analyzed for this submission corresponds to three months: from January 2022 until April 2022. During this time, the sensors experienced diverse weather conditions. We collected the temperature, humidity, wind, precipitation and fog parameters through a weather station located in proximity to our deployment.

A. Installation setup

Figure 3 shows our deployment. The radome-enclosed sensors are placed on lamp posts adjacent to a road frequently used by pedestrians and cyclists. The nodes are installed close to two people-counting cameras that were present in the scenario to benchmark the performance of the mmWave sensors. The nodes are placed at a height of 3.54 m to avoid any pilferage. The orientations have an elevation angle of $25°$, and azimuth angles of $30°$ and $150°$. The casings have rubber gaskets, cable glands and pressure valves to ensure the radomes are resilient to water leakage.

B. System configuration

Sensor node: Figure 4 shows two sensor nodes. The left sensor is closed and ready to be deployed, and the right sensor is open. Each node consists of a TI IWR6843ISK as the mmWave chipset and a Raspberry Pi (RPi) 4B as the computational unit. The mmWave sensor computes the point clouds and counts the number of passing objects. This data is shared with the RPi through a UART link, and the RPi uploads the data to the cloud via WiFi. The power supply unit is connected to the AC mains and generates a regulated DC voltage suitable for the mmWave module and RPi.

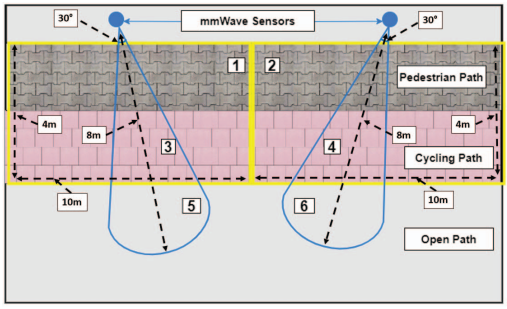


Fig. 5: Eagle-eye view of deployment

The computation of point clouds, static-noise removal, and overall tracking is done by the firmware provided by the manufacturer (Texas Instruments). The firmware allows selecting different configurations, through empirical evaluation, we choose the configuration for 'Sense and Direct HVAC Control'. The configuration generates a 2-dimensional point cloud and covers the range and velocity parameters for our use-case. The range and velocity are 10.35m and 6.065 m/s respectively, while their resolutions are 0.089 m and 0.047 m/s.

C. An indirect measurement of accuracy

As stated before, we co-locate the mmWave sensors with people counting cameras. The cameras, however, are not under our control. These cameras only report the number of people in the Field-of-View (FoV), but their FoV differ in shape and size from the mmWave's FoV. Figure 5 shows that while the FoV of the cameras is rectangular (two yellow boxes), the mmWave sensor has a conical FoV (blue cones). Thus, the objects seen by the cameras and mmWave sensors differ. For instance, while objects 3 and 4 are detected by mmWave sensors as well as the corresponding cameras; only the cameras will observe objects 1 and 2; and only the mmWave sensors will observe objects 5 and 6. Overall, due to their bigger coverage, the cameras will see more people on an average.

One approach to overcome this disagreement would be to modify the orientation of the cameras and utilize basic signal processing to bound the cameras' coverage to match the radars' coverage. However, due to stringent privacy regulations, we were not allowed to access the camera's images or modify their FoV in any way[2].

The inability to change the parameters or configuration of infrastructure that is already deployed is an important and frequent challenge in real urban setups. These constraints are in place to avoid jeopardizing in any way the safety, security or privacy of citizens.

To alleviate the disagreement between the different coverages, in the next section we propose a framework with different metrics to compare the performance of these two systems.

[2]The Ethics Review Board gave us permission to deploy the mmWave nodes.

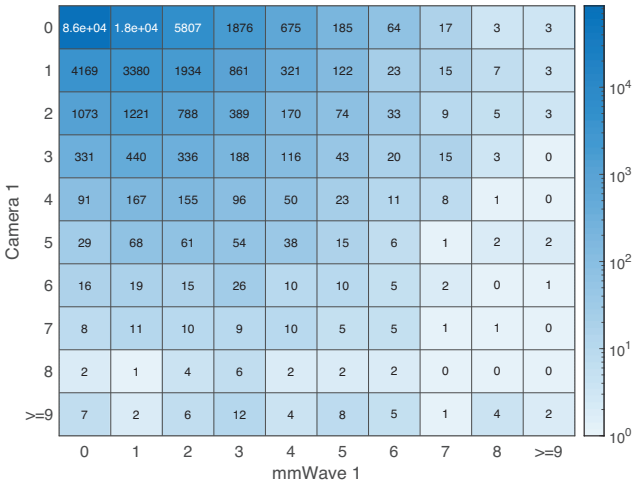


Fig. 6: Confusion matrix. The heatmap follows a log scale.

V. Evaluation Framework

Considering that the cameras and mmWave sensors have different coverages, we first need to identify the appropriate metrics, periods and quantization levels to make sure that we do not distort the comparisons further. We start by formally defining the metrics, and then, we use those metrics to analyze the effect of different monitoring periods and quantization levels. Our analysis first focuses on empty areas and then on scenarios with one or more people.

A. Analysis of empty areas

Our sensors have a sampling rate of one image per minute. Thus, during the three-month period, we gather around 130K samples for each sensor. Figure 6 shows the confusion matrix for one of the cameras and its corresponding mmWave sensor. Given that the coverage of the sensors is different, this matrix has to be considered carefully because the camera does *not* provide the ground truth, only an approximation of it. In spite of these differences, we are able to obtain one important insight.

Insight 1: Outdoors, mmWave sensors have an error of 15% detecting empty scenarios. Despite the differences in coverage, the fact that the camera's view is greater than the radar's, allows us to state –with high probability– that if the camera does not detect any person, the radar should not detect any either. Thus, using the confusion matrix in Figure 6, we could use precision and recall to quantify the performance of mmWave sensors in detecting if an area is indeed empty. The precision of the system is 94%, that is, if the mmWave sensor indicates that the area is empty, very likely it is. The recall, on the other hand, is 76%, which means that almost a quarter of the empty scenarios are deemed as busy, i.e. the mmWave sensor indicates that there are one or more people when there are none in fact. The low recall is an artifact of the high background noise present in outdoor scenarios, any mobile element –an animal or a moving object – may be confused as a person. Overall, combining the precision and recall of the mmWave sensor into the F1 metric gives an accuracy of 84%, which is lower than the high accuracy reported for indoor scenarios (error ≤ 1 person for 97.8% [19]).

B. Analysis of occupied areas (one or more people)

The analysis done for empty areas cannot be applied to busy areas (one or more people). Given that the camera has a bigger coverage, it can detect the presence of persons that are not under the radar's coverage. Partly due to this smaller coverage, the radar's recall for presence detection drops to 67%, that is, around 30% of the time when the camera reports a person (or persons), the radar reports an empty scenario. The confusion matrix in Figure 6 also captures this underestimating behavior due to the larger coverage of the camera: the lower triangle has more points than the upper triangle. Next, we propose using correlation metrics to overcome the different coverages for scenarios with one or more people.

To analyze occupied periods, we first filter out all the samples where the camera indicates zero presence, and then, use correlation coefficients to analyze the count-similarity between the sensors. The use of a correlation metric builds upon the assumption that, over time, the flow of people observed by the mmWave sensor is proportional to the flow observed by the camera.

1) Correlation coefficient: The Pearson correlation coefficient is a widely accepted measure to evaluate the correlation between two variables, X and Y. The coefficient has a range between [-1,+1]. Values around ± 0.8 are considered to have a high correlation, values around ± 0.4 are deemed as a medium correlation and 0 implies that there is no dependency between the variables. Equation 1 gives the mathematical definition.

$$\rho_{X,Y} = \frac{\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{N}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{N}(y_i - \bar{y})^2}} \quad (1)$$

where, N is the sample size in X and Y. $x_i,\ y_i$ are the i^{th} individual sample points of X and Y. $\bar{x},\ \bar{y}$ are sample means for X and Y, i.e., $\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$.

2) People-Flow: Most crowd-sensing applications are not interested in an instantaneous view, but the flow over a period of time. Thus, instead of correlating the individual samples, we correlate the flow of people. The people's flow is the total number of people detected over a certain period of time. Formally, if we denote a_j as the number of people sampled at time j, the flow (x_i) for a period τ is given by: $x_{i=t} = \sum_{j=t}^{t+\tau} a_j$.

In our analysis, we denote x_i^τ and y_i^τ as the flows measured at the mmWave sensor and camera, respectively, for a period τ. These flows are inserted in Equation 1 to obtain the correlation between the two types of sensors.

Figure 7 shows the comparison of correlation coefficients for different durations, viz., 15 minutes, 30 minutes, 1 hour and 2 hours. We see that the correlation coefficients increase as the period increases, but the difference is not significant. Thus, for simplicity, we use only the hourly flow for the remainder of our evaluation.

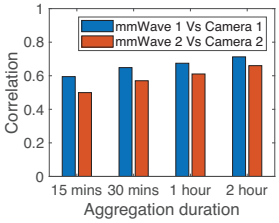


Fig. 7: Correlation under different aggregation durations.

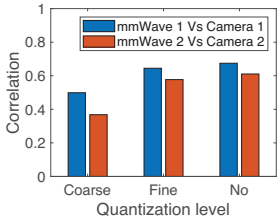


Fig. 8: Correlation under different quantization levels.

3) Quantization levels: Often, crowd-monitoring applications do not require the exact number of people but labels such as 'Sparse' or 'Busy'. Thus, we compare the correlation of three granularity levels, viz., (i) No quantization, where the original x_i^τ and y_i^τ values are used in the correlation; (ii) Fine-quantization, where the crowd levels are divided into five categories and only the crowd levels are correlated; and (iii) Coarse-quantization, with three crowd levels. Tables II and III define the thresholds for the fine and coarse quantization, respectively.

Figure 8 shows the correlation coefficients for the two mmWave radars and their corresponding cameras at different quantization levels. We observe that the correlation increases for more fine-granular levels. The coarse quantization shows a correlation of around 0.4 while the fine- and no-quantization are around 0.6. This difference is not due to limitations of the hardware or the algorithms to measure low granularity levels, but it is rather an artifact of Pearson's equation, which penalizes the inaccuracy introduced by reducing the maximum range of the coarse quantization, capped at 31 people in Table III. To avoid this artificially generated error, we consider the *no-quantization* in our next results.

VI. Weather Observations

Until now we have analyzed the aggregated data. In this section, we evaluate the impact of individual weather conditions. This evaluation is imperative for outdoor deployment. Using the information from a weather station close to our university, we divide the samples based on different weather parameters: temperature, humidity, precipitation, solar radiation, wind and fog. There were also a few days with snow but the samples were too few to analyze that condition. It is important to highlight that the weather station only provides data every six hours. Hence for true representation, we label only the samples close to the reporting time ($\pm$ 1 hour). In effect, this means that over a period of six hours, we only have two samples for the hourly flow, instead of six.

To measure the correlation between both sensors, we first filter out the samples where the camera indicates zero presence.

TABLE II: Threshold levels for fine-quantization

Flow Levels	Empty	Sparse	Normal	Busy	Crowded
Hourly people flow	0	1-12	13-30	31-60	$\geqslant 61$

TABLE III: Threshold levels for coarse-quantization

Flow Levels	Empty	Normal	Crowded
Hourly people flow	0	1-30	$\geqslant 31$

This is done for two reasons. First, 85% of the samples show no people, and hence, we do not want to skew the analysis. Second, the assessment of basic presence has been already done with precision and recall. After filtering the data, we use the hourly flow of people considering no quantization. The unquantized data for the radar and camera are then evaluated with the correlation coefficient.

To assess the impact of each weather condition, we segment them into bins. While the temperature, humidity, solar radiation and wind speed are divided based on different intervals, precipitation and fog are binned depending upon their presence. For example, Figure 10b shows four bins for temperature, but Figure 10d only shows two bins for precipitation: rain and no rain.

There are two important aspects that need to be considered for correlation analysis. First, we need a sufficient number of samples because, statistically, having few samples may incorrectly represent the impact. Second, the range of the values needs to be similar, otherwise, the Pearson equation adversely skews the correlation for lower ranges (as we saw during the analysis of different quantization levels, c.f. Figure 8).

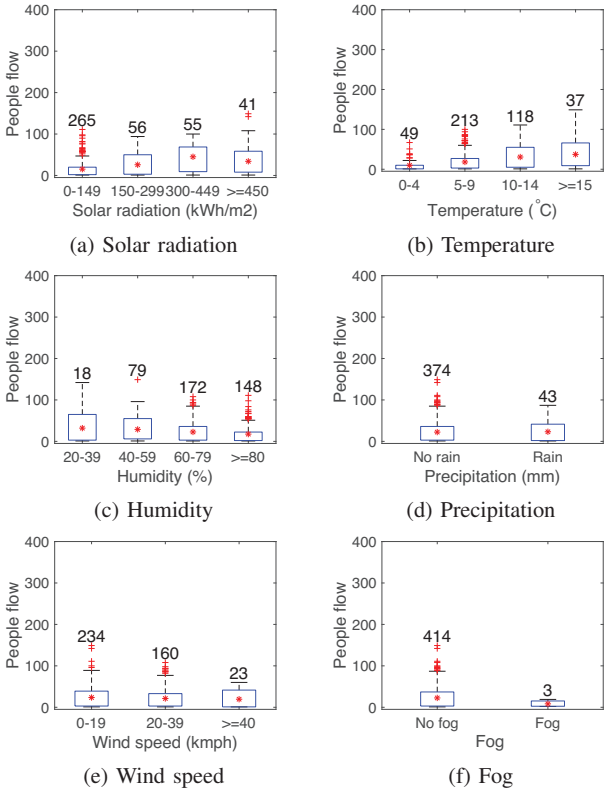
To obtain the number of samples and ranges for all weather conditions (and their bins), we use the data from camera-1. Figure 9 show the box-plots for temperature, humidity, solar radiation, precipitation, wind speed and fog; and Figure 10 plot their respective correlation coefficients. The boxplots capture the quantiles and mean (red asterisk), and at the top of the plot, we state the number of samples. The correlations for the two sensors are represented in different colors: blue for radar/camera 1 and red for radar/camera 2 in Figure 10.

We divide the results into two main groups: benign and harsh conditions. Benign conditions are those where we do not expect the sensor to be impacted much, such as changes in temperature, humidity or solar radiation. Harsh conditions capture more challenging parameters, such as precipitation wind and fog.

A. Benign conditions

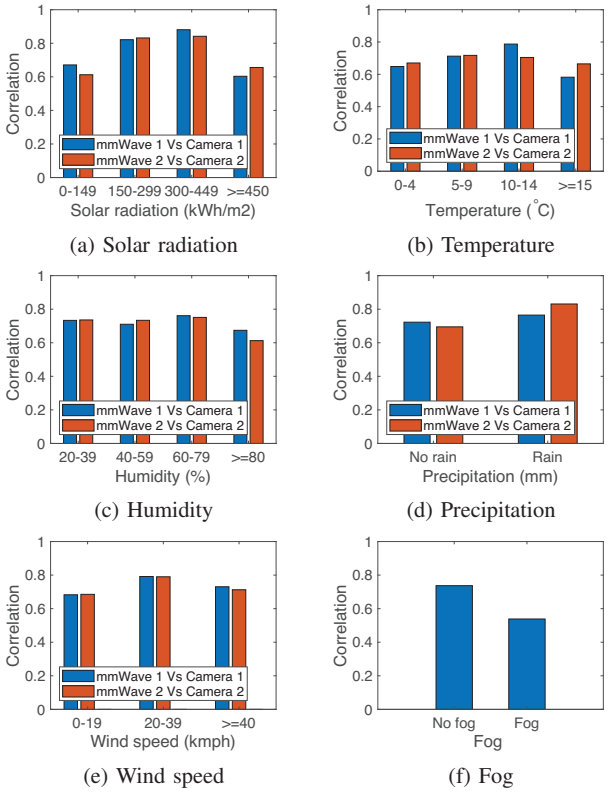
For temperature, humidity and solar radiation, the correlation values in Figure 10 follow the mean (red asterisks) in Figure 9. For example, for solar radiation the correlation increases for the first three bins and then drops (Figure 10a), following the same pattern of the mean in Figure 9a. There are some cases where the correlation coefficient does not follow the mean, but this effect can be explained by the lower number of samples. For example, for temperature, the correlation coefficients of the first three bins (Figure 10b) follow the mean of Figure 9b, but the last point drops because the number of samples decreases significantly from 118 to 37. The trends for humidity are more stable, the correlation coefficients (Figure 10c) capture the slow decreasing trend of the mean (Figure 9c) but without major differences. Overall, the above observations provide a second important insight.

Insight 2: Under benign weather conditions, radar sensors provide correlation coefficients that lay mainly between 60% and 85%. Correlation coefficients cannot be used strictly as



(a) Solar radiation (b) Temperature (c) Humidity (d) Precipitation (e) Wind speed (f) Fog

Fig. 9: People Flow



(a) Solar radiation (b) Temperature (c) Humidity (d) Precipitation (e) Wind speed (f) Fog

Fig. 10: Correlation coefficients

metrics for accuracy, i.e. we cannot state that radars are 60%-85% accurate, but they indicate that even though radar sensors cannot track the people-flow in a fine-grained manner (as they do indoors), they can provide coarse-grained flow information.

B. Harsh conditions

We now focus our attention on more challenging weather conditions: precipitation, wind and fog. A positive result is that precipitation does not seem to have a negative impact on the radar sensor. Even though we have more sample points for the no-rain case, there is still a sufficient number of samples for the rain case (Figure 9d), and the correlation coefficients remain stable across the board, between 70% and 85% (Figure 10d).

The wind speed has a slightly higher impact than precipitation but the effect is not so detrimental. All wind categories have a similar mean (Figure 9e), and in spite of the differences in the number of samples, the correlation coefficients remain between 70% - 80% across the range (Figure 10e).

For the case of fog, the situation is more complex, when there is no fog, the correlation reaches a value higher than the aggregated data, around 70% (Figure 10f) compared to the 60% with no quantization in Figure 8. This better correlation makes sense because without fog all measurements are clearer. The problem we faced is that with fog we have limited data (only seven samples) and the pair of sensors-2 (red) were malfunctioning during some of this time. For sensors-1 (blue) the correlation is around 50%, but no conclusions can be made because of the limited amount of data and because cameras are less accurate on foggy days as well. Overall, the evaluation under harsher conditions provides a third important insight.

Insight 3: Precipitation and wind speed do not seem to affect the performance further, compared to the benign scenarios. Rain seems to have little to no effect, compared to clear days. Wind can cause false positives with radars, and thus, an important research problem for the community is the removal of background clutter with mmWave sensors, but the correlation coefficient during busy periods remains stable. Fog, on the other hand, is the situation where cameras are more vulnerable and a longer study with accurate ground truth data is required to assess the performance of radar sensors.

All in all, our outdoor evaluation provides a fourth overarching insight.

Insight 4: mmWave radars in the ISM band (60 GHz) are not yet ready to provide accurate information in outdoor setups. The research community could balance the effort put into analyzing radar sensors between indoor and outdoor setups. The tracking accuracy of radars indoors is high (above 95%) and hence there is no major gap compared to using cameras. Outdoors, however, the gap between cameras and radars increases to levels that prevent mmWave being used for fine-granular analysis.

VII. Related Work

To the best of our knowledge, there is no study evaluating the performance of mmWave radar for crowd tracking in the

ISM band (60 GHz). Thus, we position our work within the following related areas.

Outdoor people counting: The majority of outdoor people counting solutions are implemented with video-based methods. Sacchi et. al. exploit novel image processing methods for real-time estimation of tourist-flows [20]. Under different weather conditions, they report a counting error of about 10% for up to 300 tourists. Hou et. al. estimate the number of people in complicated outdoor scenarios (up to 100 people) [21], with a best average error of around 10%. Overall, camera-based solutions are a reliable alternative for people counting but they raise up serious privacy concerns.

Another alternative for people counting is based on characteristics of Wifi signals. Depatla et. al. develop an analytic model mapping the probability distribution of the received signal amplitude to people count [22]. Their accuracy with count error less than 1 is 92% for up to nine people. This WiFi solution preserves privacy but requires dedicated antennas and extensive training data.

mmWave for indoor applications: mmWave radars are finding a wide range of applications including vital sign monitoring [14], activity recognition [23], people identification [6] and 3D pose reconstruction [13]. All these applications however are performed in *controlled* indoor setups with minimal background clutter and specific instructions to the users. Through digital beamforming, multi-target detection and a robust clustering technique, Wu et. al. can count people (up to 5) with an error ≤ 1 person for 97.8% of the time [19]. Weiss et. al. use an adaptive OS-CFAR peak detection algorithm and a vital sign verification algorithm to detect and check targets. They can achieve 85.4% accuracy for indoor counting for 0-4 people [24]. Similarly, the study by Gross et. al. [25] demonstrates an accuracy of nearly 97% for up to 5 people in indoor scenarios. All these indoor counting studies provide high accuracy but cannot be directly mapped to outdoor environment.

Automotive radar for outdoor people tracking: The 77-81 GHz frequency band is licensed for automotive applications, limiting the amount of work for people detection. For example, Scheiner et. al. propose a method to detect and track an object/person around corners with an automotive radar [26]. Through an image formation model for Doppler radar non-line-of-sight (NLOS) measurements, they can derive the position and velocity of people via reflections. Our work complements the studies done in the automotive bands, as we target a long-term urban scenario providing relevant information for traffic management.

VIII. Conclusion

We analyze the performance of mmWave sensing by deploying them outdoors for a period of three months and under different weather conditions. We provide a step-by-step design for a radome to enable outdoor deployment. Our study shows that the performance of mmWave sensors drops in the outdoor deployment as compared to the indoor case. In general, our study indicates that radar sensors need to be further studied outside controlled indoor setups. We plan to continue our investigations by gaining access to camera images, comparing them with mmWave point clouds, and developing advanced algorithms for improving the accuracy, thus paving the way for the deployment of mmWave radars in urban infrastructure.

Acknowledgement

The authors would like to thank Sascha Hoogendoorn for providing permission to install the mmWave sensor; Paul van Gent from the Mobility Innovation Centre for providing access to the camera counts; Peter van Oossanen from Civil Engineering for his assistance during the deployment of the sensors; Fernando Kuipers for his support throughout the project; and R. Gündel, S. Yuan, X. Li, and L. Ren from the MS3 group for their contribution to the enclosure design. The research was partially supported by a grant from the Amsterdam Institute for Advanced Metropolitan Solutions (AMS).

References

[1] K. Still *et al.*, "Place crowd safety, crowd science? case studies and application," *Journal of Place Management and Development*, 2020.
[2] A. Thiagarajan *et al.*, "Accurate, low-energy trajectory mapping for mobile devices," 2011.
[3] "Sensorenregister," https://sensorenregister.amsterdam.nl/.
[4] "Big brother watch," https://bigbrotherwatch.org.uk//.
[5] R. K. Chellappa, "Consumers' trust in electronic commerce transactions: the role of perceived privacy and perceived security," *Citeseer*, 2008.
[6] P. Zhao *et al.*, "mid: Tracking and identifying people with millimeter wave radar," in *IEEE DCOSS*, 2019.
[7] T. Gu *et al.*, "Mmsense: Multi-person detection and identification via mmwave sensing," in *mmNets*, 2019.
[8] S. Li *et al.*, "A field people counting test using millimeter wave radar in the restaurant," 2021.
[9] T. Instruments, *60GHz mmWave Sensor EVMs*, May 2022.
[10] C. Iovescu *et al.*, "The fundamentals of millimeter wave sensors," *Texas Instruments*, 2017.
[11] K. Garcia, "Bringing intelligent autonomy to fine motion detection and people counting with timmwave sensors," *Texas Instruments*, 2018.
[12] X. Huang *et al.*, "Indoor detection and tracking of people using mmwave sensor," *Journal of Sensors*, 2021.
[13] A. Sengupta *et al.*, "mm-Pose: Real-time human skeletal posture estimation using mmWave radars and CNNs," *IEEE Sensors Journal*, 2020.
[14] Y. Wang *et al.*, "Remote monitoring of human vital signs based on 77-ghz mm-wave fmcw radar," *Sensors*, 2020.
[15] "Choosing 60-GHz mmWave sensors over 24-GHz to enable smarter industrial applications," https://www.ti.com/lit/wp/spry328/spry328.pdf.
[16] "WAYV Air: Short-Range IoT Radar Sensor," https://sensing.ai/products/wayv-air.
[17] "Mm-wave radar radome design guide," https://www.ti.com/lit/an/swra705/swra705.pdf.
[18] "White paper on radar wave propagation through materials: Walls and radomes," https://www.infineon.com/dgdl/Infineon-Radar_wave_propagation_through_materials-Whitepaper-v01_00-EN.pdf?fileId=5546d462766cbe8601768a120c6d36cf&da=t.
[19] C. Wu *et al.*, "mmtrack: Passive multi-person localization using commodity millimeter wave radio," in *INFOCOM*, 2020.
[20] C. Sacchi *et al.*, "Advanced image-processing tools for counting people in tourist site-monitoring applications," *Signal Processing*, 2001.
[21] Y.-l. Hou *et al.*, "Automated people counting at a mass site," in *IEEE ICAL*, 2008.
[22] S. Depatla *et al.*, "Occupancy estimation using only wifi power measurements," *IEEE JSAC*, 2015.
[23] Y. Wang *et al.*, "m-activity: Accurate and real-time human activity recognition via millimeter wave radar," in *ICASSP*, 2021.
[24] J. Weiß *et al.*, "Improved people counting algorithm for indoor environments using 60 ghz fmcw radar," in *RadarConf 20*, 2020.
[25] C. Groß *et al.*, "Towards an occupancy count functionality for smart buildings-an industrial perspective," in *IESES*, 2020.
[26] N. Scheiner *et al.*, "Seeing around street corners: Non-line-of-sight detection and tracking in-the-wild using doppler radar," in *CVPR*, 2020.