This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Robotic Exploration for Learning Human Motion Patterns

Sergi Molina, Grzegorz Cielniak, and Tom Duckett

***Abstract*—Understanding how people are likely to move is key to efficient and safe robot navigation in human environments. However, mobile robots can only observe a fraction of the environment at a time, while the activity patterns of people may also change at different times. This article introduces a new methodology for mobile robot exploration to maximize the knowledge of human activity patterns by deciding where and when to collect observations. We introduce an exploration policy driven by the entropy levels in a spatio-temporal map of pedestrian flows, and compare multiple spatio-temporal exploration strategies including both informed and uninformed approaches. The evaluation is performed by simulating mobile robot exploration using real sensory data from three long-term pedestrian datasets. The results show that for certain scenarios the models built with proposed exploration system can better predict the flow patterns than uninformed strategies, allowing the robot to move in a more socially compliant way, and that the exploration ratio is a key factor when it comes to the model prediction accuracy.**

***Index Terms*—Human motion model, long-term data, mobile robots, spatio-temporal exploration.**

I. Introduction

THE INTRODUCTION of robots into environments shared with humans, such as hotels, warehouses, or shopping centres, creates new challenges. As pointed out in [1] and [2], navigation around people is one of the key aspects for long-term acceptance of mobile robots in human-populated environments. One of the current barriers for this acceptance is that many navigation methods are based only on reactive behavior, where the robot tracks the people in its vicinity and replans its trajectory accordingly. The work in [3] reported that the errors of the state-of-the-art methods for prediction of human motion exceed 0.4 m for predictions horizons of 1 s, which can lead to frequent replanning, causing an erratic or aggressive robot behavior. Instead, giving the robot the ability to model and predict in advance how the people are likely to behave from a long-term perspective is an important enabler for safe and efficient navigation [4]. For example, a robot may use the human motion (or "flow awareness") knowledge to go with the main direction of flow to minimize the possibility of collisions or trajectory replanning [5], [6]. Alternately a robot might choose to avoid areas likely to contain human traffic in order to maximize safety in industrial applications. However, obtaining an applicable model of the environment is not a straightforward task due to several sources of noise in the collected data, e.g., visibility, occlusions, and time constraints. The robot's view is constrained by sensor limitations and occlusions, and the robot usually has to perform other tasks, limiting the time available for data gathering.

Hence, a mobile robot requires an exploration strategy to provide a planned sequence of data observations with the primary aim of building a model of an environment to overcome the aforementioned constraints. Traditional exploration strategies assume a static scenario, i.e., the states measured do not change over time. Their aim is to build a spatial model that covers the operational environment as quickly as possible. If the flow of people remains constant (static in time), this would mean visiting the locations which have never been visited before. However, when modeling human motion, the environment tends to change over time, e.g., with different flow patterns at different times of the day [8]. Thus, apart from deciding where to explore, the exploration strategy must also consider the temporal dimension (i.e., when to explore) in order to obtain the maximum information from each observation. By learning and refining its knowledge of human motion patterns over time, the robot is able to operate more reliably in challenging environments shared with humans.

In this article we present, as the main contributions the following.

1) A new methodology for mobile robot exploration to maximize the knowledge of human activity patterns, by deciding where and when to collect observations.
2) A study comparing multiple spatio-temporal robotic exploration policies, evaluating their performance in building and refining a time-dependent probabilistic model of human activity patterns (spatio-temporal flow map, or STeF-Map) introduced in [9].
3) The introduction of an exploration policy driven by the entropy levels of the cells in a STeF-map, computed from the previous observations.
4) A new long-term pedestrian dataset
containing a total of 22 consecutive days of observed pedestrian motions. This is publicly available and can be found in [10].

Manuscript received December 16, 2020; accepted June 28, 2021. This work was supported within H2020-ICT by the EC under Grant 732737 (ILIAD). This article was recommended for publication by Associate Editor A. M. Zanchettin and Editor W. Burgard upon evaluation of the reviewers comments. *(Corresponding author: Sergi Molina.)*

The authors are with the Lincoln Centre for Autonomous Systems, University of Lincoln, LN5 8AA Lincoln, U.K. (e-mail: sergimolina91@gmail.com; gcielniak@lincoln.ac.uk; tom.d.duckett@gmail.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TRO.2021.3101358.

Digital Object Identifier 10.1109/TRO.2021.3101358

1552-3098 © 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

As the name implies, each spatio-temporal exploration strategy is composed of both a spatial policy and a temporal policy, which define, respectively, the set of rules the robot must follow to decide the location and time to collect data. In the experiments we treat both independently, so we can see the impact of all the possible combinations. For each domain, we test the entropy-based policy, plus two more uninformed policies (random and round-robin).

The evaluation is performed using three long-term term pedestrian datasets providing complex time-varying human motion patterns. Using the prerecorded datasets together with the provided ground truth data, we simulate an environment where a mobile platform can schedule observations at particular locations and times, following the assigned spatio-temporal exploration strategy.

The results show that for the dataset with a large environment containing many people detections, the entropy-based strategy learns the flow patterns and increases model prediction quality better than the uninformed strategies. We also determine that the time devoted to exploration is the factor with the biggest impact on model quality, obtaining a similar correlation between model quality and the exploration ratio for all three datasets.

II. Related Work

A. Robot Exploration

Classical exploration methods are aimed at guaranteeing the completeness of the robot's environment model. One of the earliest approaches is frontier-based exploration [11], [12], which drives the exploration toward the boundaries between the known and unknown until the map is complete. The scalability of this approach due to its ability to distribute the frontiers among multiple robots is one of the main advantages, but the method does not consider the model quality. Because the approach does not refine or update the map over time, this can lead to inaccuracies in the representation, which could cause robot localization and navigation failures.

Another approach is adopted by next-best-view strategies, which focus on building the initial map after the deployment stage, based on optimizing different criteria. For example, Amigoni and Caglioti [13] present a system that optimizes the estimated time to reach a given location and the amount of information expected to be gathered there. Vallvé and Andrade-Cetto [14] propose a method of calculating the entropy decrease in the robot configuration space and then use these estimates to evaluate different exploratory trajectories for robot mapping. Fentanes *et al.* [15] utilize the Poisson uncertainty to drive the observation and mapping of soil moisture by counting neutrons using a cosmic-ray sensor. Stachniss and Burgard [16] presented an information-gain-based exploration framework that integrates not only uncertainties of the map but also the uncertainties of the robot's localization.

However, the above approaches do not attempt to maintain the environment models after their acquisition, meaning that the model will lose accuracy as the environment changes. To deal with this problem, another branch of algorithms aim to create models of the environment that allow them to predict where and when to make observations of specific phenomena, reasoning about the best times and locations to explore. In Marchand and Ramos [17] and Singh *et al.* [18], for example, the decision-making for environmental surveillance and monitoring is based on Gaussian processes, which allow the robot to learn the temporal patterns in the environment.

Other approaches [19], [20] are instead based on the assumption that some of the environmental variations observed are caused by people's daily routines. Santos *et al.* [19] present a method for life-long spatio-temporal exploration of dynamic environments, using the entropy of binary state predictions in an occupancy map to create a scheduler that determines the areas and times to explore for each day. These methods rely on building a metric representation of the world structure, estimating the probability of a (binary) state of a cell to contain a person or not. However, reasoning about human motion requires more complex model representations, and binary states are not sufficient to describe the pedestrian flows covered in this article.

B. Human Motion Modeling

Several authors present approaches to model human motion in space [21]. Some of them focus on learning human affordance maps, which define the probability of finding people in certain regions of the environment. For example, Limosani *et al.* [22] create affordance maps by overlapping the data from 2D laser scans from multiple sessions. Tipaldi and Arras [23] model the probability of activity events by a Poisson process, which is exploited by a motion planning algorithm to find people as quickly as possible. Other authors focus more on modeling the typical directions of pedestrians in an environment. Senayake and Ramos [24] present a probabilistic directional grid map to represent the long-term angular motion of the environment dynamics. In similar work, Kucner *et al.* [25] further take the observed speed of the people into account as well as their orientation. Asahara *et al.* [26] and Wang *et al.* [27] use Markov models to learn the transition probabilities of people moving from one cell to another, and Han *et al.* [28] present an approach for pedestrian motion modeling that includes both low-level trajectory patterns and high-level discrete transitions.

Despite being focused on representing human motion, the aforementioned methods do not take into account temporal variations in the model, resulting in less accurate predictions. Thus, Jovan *et al.* [29] use periodic Poisson processes to characterize the behavior in time of different rooms of a building, which helps to better capture the nature of human activities. Zhi *et al.* [30] use an Long Short-Term Memory (LSTM) network to model a multimodal probability density function over the possible directions in which an object can move over time. Krajnik *et al.* [20] present a model for introducing time into discrete and continuous spatial representations by modeling long-term, pseudoperiodic variations caused by human activities or natural processes, wrapping time in several dimensions.

With mobile robots, the models need to be built from data that is sparse in space and time, with incremental online updates as the robot gathers new information. However, most of the existing models are built offline, as model creation and updating can be

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

computationally expensive and tedious. Wada *et al.* [31] and Kucner *et al.* [25] aim to solve this issue in the spatial domain, while Chenellato *et al.* [32] address the incremental building of a von Mises mixture model from streaming data to represent multimodal distributions of human activity.

The spatio-temporal exploration strategies compared in the experimental part of this article aim to build the STeF-Map model introduced in [8], which creates a time-dependent probabilistic map able to model and predict patterns of people in indoor environments from sparse data. This representation is compared with other similar maps of human dynamics in the literature, where it achieves favorable performance not only in terms of prediction accuracy using the Chi-squared metric [33] but also, it allows a more human compliant path planning and scheduling of robot activities in human-populated environments [34]. However, introducing the temporal dimension to robot mapping requires novel exploration strategies to build and maintain the spatio-temporal models during the robot's deployment in a given environment [35].

The work presented here extends the methodology first presented in [36]. However, the prior work assumed a fixed 24-h period for human activities, which is not always the case. In the current work, we relax this assumption and instead perform spectral analysis in the time domain for each modeled cell in the environment model. Moreover, in the experiments carried out here, we use a clustering algorithm to segment the environment into regions, the number of periodicities for each cell are learnt independently, and we increase the fidelity of the simulation by reproducing robot navigation in the environment. We also introduce an additional long-term pedestrian dataset, and present further analysis and discussion of the results, including their general applicability to a wide range of domains in environments shared with humans.

III. Problem Definition

This chapter addresses the problem of creating a spatio-temporal pedestrian flow model as accurately as possible from sparse observations of nontrivial environments, taking into account the limited sensory capabilities and time constraints of a mobile robot.

Assuming an indoor environment with a known spatial layout, which is divided into a set of C square cells forming a grid [see, e.g., the small squares in Fig. 1 (bottom)], the main aim of the robot is to learn a spatio-temporal model (STeF-Map [8]) of each cell that best represents the human motion patterns, i.e., minimizing the error between the internal model and the true human motion distributions over time.

In order to simulate a robot's limited visibility, let us define a region R_r as a subset of $c_r \subset C$ cells that can be observed simultaneously, such that by visiting c_r during an interval of time $[t, t + \Delta t)$, observations of the human motion distributions of only the c_r cells that belong to R_r are obtained. The remaining $C - c_r$ cells in the environment during that time interval remain unseen. The observations are performed over a predefined interval of time because, in contrast to other environmental variables such as the state of an occupancy cell (*free* or *occupied*), a distribution over human motion cannot be obtained in a single instant of time, since the robot cannot count enough detections to build a meaningful distribution. In the experiments carried out in this article, each subset corresponding to each R_r is nonoverlapping with the other regions, such that $C = \sum_{r=1}^{\rho} c_r$, ρ is the total number of regions R into which the environment has been divided.

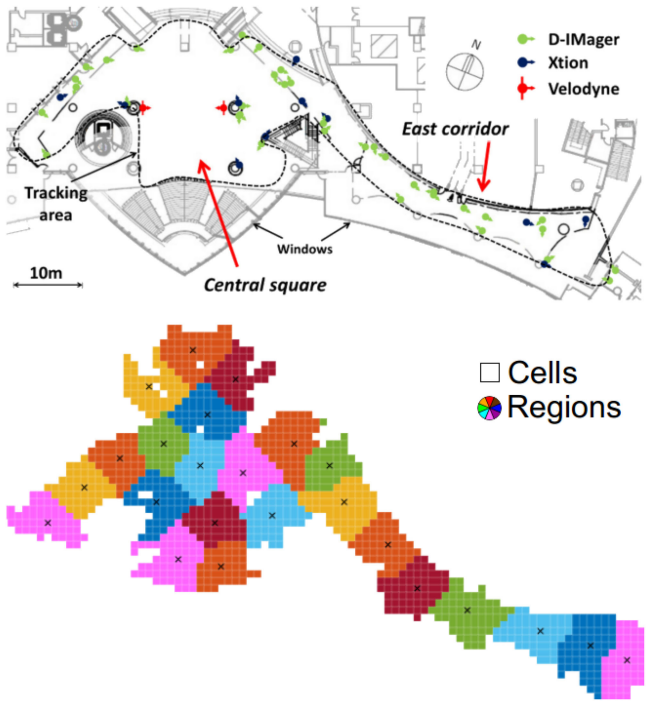


Fig. 1. Top: Shopping centre map with the location of all the sensors and the two main zones, taken from [7]. Bottom: Spatial division of the environment into 1 × 1-m cells, and the 24 regions available for the robot to explore.

Although an improvement in model accuracy can be achieved by visiting the regions as often as possible, the number of observations is typically limited and the robot can spend only a fraction of the total time on actual exploration. In the experiments this fraction is referred as the exploration ratio e, e.g., $e = 0.25$ means that the robot can spend only 25% of its operational time on gathering information. The rest of time, the robots keep a stationary position and remains "blind."

Therefore, given an exploration ratio e and a set of τ nonoverlapping time intervals $[t_j, t_j + \Delta t)$, it is the exploration strategy's job to define which regions to observe and in which time interval, complying with the e ratio defined, in order to improve the accuracy of the spatio-temporal human motion model as much as possible.

IV. STeF-Map Model

A. Spatial Representation

The underlying geometric space is represented by a grid, where each cell contains k temporal models, corresponding to k discretized orientations of people moving through the given cell over time. Since the total number of temporal models, which are of a fixed size, is $k \times C$, where C is the total number of cells,

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

the spatio-temporal model does not grow over time regardless of the duration of data collection. This makes the model not only memory efficient but also enables probabilistic predictions of the likely flow of people in a given direction for a given cell at any instant of time.

B. Temporal Framework—FreMEn

The temporal models, which can capture patterns of people movement, are based on the FreMEn framework [37]. FreMEn is a mathematical tool based on the Fourier transform, which considers the probability of a given state as a function of time and represents it by a combination of harmonic components. The model not only allows representation of environment dynamics over arbitrary timescales with constant memory requirements but also the prediction of future environment states based on the patterns learned. The idea is to treat a measured state as a signal, decompose it by means of the Fourier transform, and obtain a frequency spectrum with the corresponding amplitudes, frequencies, and phase shifts. Then, transferring the most prominent spectral components to the time domain provides an analytic expression representing the likelihood of that state at a given time in the past or future. Assuming that the observed direction of motion follows underlying patterns due to the rhythmic nature of human activities, we apply the FreMEn concept to discretized directions of people movement through a particular cell.

C. Building the Model

Our model assumes that it is provided with pedestrian detection data, comprising the person position, orientation and timestamp of the detection (x, y, α, t). At the beginning of the model construction, we associate each cell with k bins, corresponding to the discretized orientation of human motion, each with an associated temporal model. When building the model, the x, y positions are discretized and assigned to the corresponding cell, and the orientation α is assigned to one of the k bins, whose value is incremented by 1. In other words, we count the number of people detections occurring in each orientation bin of each cell. After a predefined interval of time, we normalize the bins and use the normalized values to update the spectra of the temporal models by the scheme described in [37]. Then, we reset the bin values to 0 and start the count again. Note that when building the model we do not use the absolute number of detections in each cell, only the relative number of occurrences among all k orientations.

D. Making Predictions

To predict the behavior of human movement through a cell at a future time t, we calculate the probability for each discretized orientation θ, ($\theta = i\frac{2\pi}{k}$ and $i \in \{0 \dots k-1\}$), associated with that cell as

$$p_{\theta_i}(t) = p_0 + \sum_{j=1}^{m} p_j \cos(\omega_j t + \varphi_j) \tag{1}$$

where p_0 is the stationary probability, m is the number of the most prominent spectral components, and p_j, ω_j, and φ_j are their amplitudes, periods, and phases. The spectral components ω_j are drawn from a set of ω_s that cover periodicities ranging from hours to 1 week with the following distribution:

$$\omega_s = \frac{7 \cdot 24 \cdot 3600}{1+s}, \quad s \in 0, 1, 2, 3, 4, \dots, 15. \tag{2}$$

The choice of m determines how many periodic processes are considered for prediction. Setting m too low could mean omitting other less prominent environment processes, while setting it too high might decrease the generalization capabilities of the model.

An example of a STeF-Map can be found in Fig. 2. The two images show the predictions obtained from a STeF-Map model created after some training in the environment at two different times of the day. In each cell, the arrow corresponds to the discretized orientation with the highest probability of human motion at that point in time. Looking at the arrows, we can see that the model is able to capture and learn the changes in the direction of pedestrian flows occurring in the central square and the East corridor.

V. Exploration

The representation shown in Fig. 2 was created assuming full observability in time and space (using the ATC dataset [7] with overhead cameras covering the entire scene). However, for the work presented in this article, we assume the robot is not able to see the entire environment at once, but only a single region at a time [the regions are shown by different colors in Fig. 1 (bottom)], and only a given percentage of the total time in a day (defined by the exploration ratio) is devoted to data gathering. Hence, the robot needs a strategy to explore (gather data) from the regions at different times to maximize the accuracy of the spatio-temporal flow model.

In this work, the physical uncertainty, also called the irreducible or aleatory uncertainty, is used to guide the robot's exploration activities. This uncertainty refers to the inherent variation in the physical system to be modeled, being an intrinsic property or variable of the system (a human motion distribution in this case). In order to measure this uncertainty, one of the alternative options is to use the variance. However, while the variance works well as a measure of uncertainty when the variable measured presents a unimodal distribution, the motion distributions obtained to create the STeF-map can be multimodal. Hence, the entropy is a better option as with multimodal outcome distributions, variance may incorrectly characterize changes in uncertainty ([38]).

In the following subsections, the definition of a cell's inherent entropy is explained, together with a description of the entropy-based heuristic policies for deciding where and when to explore.

A. Defining the Entropy of a Cell

As explained in Section IV-C on model building, STeF-map is based on counting occurrences of persons moving in different directions. Taking, for example, a single randomly chosen cell c with N detected persons, its probability distribution over the orientation of human motion is defined as $P(X) = \{x_1 =$

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

(a) (b)

Fig. 2. STeF-Map model predictions for all cells in the ATC dataset at two different points in time in a day: 11:00 and 18:00. The arrow shown in each cell corresponds to the discretized orientation with the highest predicted probability of human motion at that point in time.

$n_1/N, x_2 = n_2/N, \ldots, x_k = n_k/N\}$, where the count vector $n = \{n_1, n_2, \ldots, n_k\}$ accumulates the observed occurrences of each orientation bin k. The entropy associated with that given cell is

$$H_{\text{cell}}[P(X)] = -\sum_{i=1}^{k} P\left(\frac{n_i}{N}\right) \cdot \log_2\left(P\left(\frac{n_i}{N}\right)\right). \quad (3)$$

Equation (3) yields the correct answer as N/k tends to infinity, but in many practical cases, statistical fluctuations of small samples induce both statistical and systematic deviations in the entropy estimates [39]. To mitigate this issue, the first-order Miller and Madow correction [40] is applied

$$H_{\text{cell}}[P(X)] = -\sum_{i=1}^{k} P\left(\frac{n_i}{N}\right) \cdot \log_2\left(P\left(\frac{n_i}{N}\right)\right) + \frac{k-1}{2\,N} \cdot \log_2(e) \quad (4)$$

where e is Euler's number. Although this estimate still retains some bias when $N << k$ or $N \sim k$ [41], this is not the case for this application as the number of person detections tends to be greater than the number of bins. This correction adds more entropy to the cells with fewer detections, expressing the fact that having fewer data to define the motion models indicates that the distribution obtained can be trusted less.

B. Entropy-Based Exploration

It is reasonable to think that cells in diverse parts of the environment at different times could present different motion patterns, so the Shannon entropy of a cell is used to characterize those patterns. Lower entropy values indicate that persons tend to follow well-defined motions, while the higher the entropy becomes, the more randomly persons move across the boundaries of that cell.

In the experiments carried out, the exploration activities are not aimed at reducing the entropy levels in the environment, as the entropy describes the inherent uncertainties. So, the entropy is not used as the reducible epistemological uncertainty (as in the work by [19] or in [17], where the covariance is used as a measure of reducible uncertainty), but as a way to bias the exploration toward areas/intervals that are more suitable or worth visiting from a pattern-based spatio-temporal modeling perspective. Lower entropy levels mean better-defined motions, and hence more predictable patterns. On the other hand, it is assumed that high entropy levels describe a situation, where no matter how many times the robot gathers new data, the persons' patterns are likely to be unpredictable, or at least more difficult to predict than lower entropy values with sparse data, because their behavior is closer to being random. There are approaches such as Dempster–Shafer theory [42], [43] that allow to unify multiple sources of uncertainties, like the inherent uncertainty and the epistemological uncertainty which can appear due to the small finite number of data samples. However, this work considers that the uncertainty computed from the measurements only comes from the intrinsic "physical" characteristics that define the motions (besides the Miller and Madow entropy bias correction). This approach also aligns well with one of the key advantages of the STeF-map representations, i.e., the nonnecessity to store all past observations to update the model and obtain predictions for new data observations.

Therefore, in the following sections explaining the entropy-based policies for deciding when and where to perform new observations, the probabilities that define the chances of visiting particular regions, and time intervals are based on an heuristic that says those probabilities are inversely proportional to the entropy values. This tells us that, for example, after gathering some data in a region during a certain time interval, if the resulting entropy is low, we want that region to have higher chances of being visited in the future.

C. Defining When to Explore

This subsection explains how, given a set of observable τ intervals occurring in the future, the entropy-based temporal policy decides which ones to explore while complying with the exploration ratio e. The basic idea is to use the map entropy from previous observations to compute the entropy for the future τ time intervals, defining the chances of selecting each interval for exploration purposes. The steps involved are as follows: *1)* After each interval of time in which a data gathering action has been performed, the map entropy is calculated as the sum of all

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

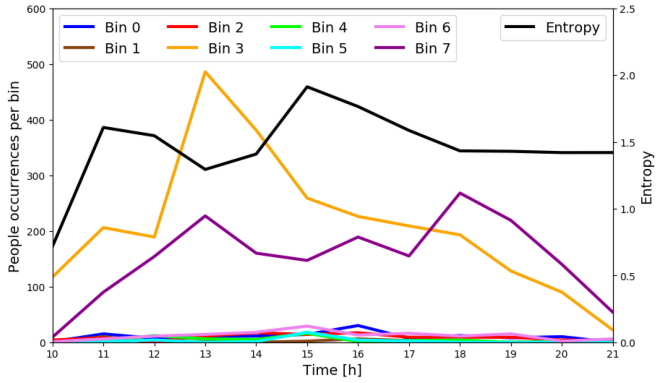


Fig. 3. Entropy calculated over 1 d using the distributions obtained in a cell in 1-h time intervals.

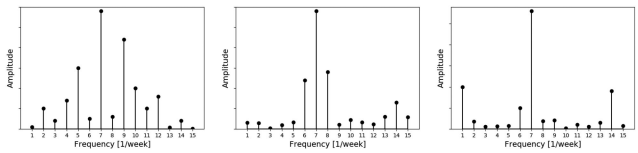


Fig. 4. Frequency spectra of the map entropy values for the three environments. (a) ATC. (b) Corridor. (c) Office.

cell entropies

$$H_{\text{map}}([t, t+\Delta t)) = \sum_{i=1}^{C} H_{\text{cell}_i}[P(X|[t, t+\Delta t))] \quad (5)$$

where $H_{\text{cell}}[P(X|[t, t+\Delta t))]$ is given by (4) but the human motion counts are obtained only during a specific interval of time. For example, Fig. 3 shows the entropy values $H_{\text{cell}}[P(X|[t, t+\Delta t))]$ calculated using 1 h of nonoverlapping time intervals for a randomly picked cell.

2) Treating H_{map} as a signal over time, the next step is to compute the most prominent time correlation in the entropy values. To do so, a spectral analysis is performed. Since the map entropy input values are not equally sampled in time due to the time constraints (when $e < 1$), the nonuniform discrete Fourier transform (NUDFT) is used for this analysis. Notice that the number of input values used to calculate the NUDFT are determined by the exploration ratio defined. The higher the value of e, the more inputs are available to compute the spectra. This frequency decomposition tells us which is the most prominent time correlation over the data by checking the periodicity T with the biggest amplitude (discarding frequency 0). This can then be used to compute the averaged entropy of an interval following the correlation obtained as:

$$H_{\text{int}_j}([t_j, t_j+\Delta t)) = \\ = \frac{\sum_{z=1}^{(t_j - t_{\text{start}})/T} H_{\text{map}}([t_j - z\cdot T, t_j + \Delta t - z\cdot T))}{(t_j - t_{\text{start}})/T} \\ , j \in 1, 2, \ldots, \tau \quad (6)$$

where t_{start} is the time when the exploration activities started. For example, Fig. 4 shows the spectra of the three environments used in the experiments after gathering data for some time. For the office dataset, the most prominent time dependency is 24 h (frequency $= 7\,[1/\text{week}]$). Knowing this value, the entropy measures for a 24-h period corresponding to a day divided into 10-min intervals are computed by averaging the entropy values of the same 10-min interval across multiple days (e.g., from 10:10 to 10:20 from day 1, 2, 3,...). Note also the peak with a frequency of 1 week (frequency $= 1\,[1/\text{week}]$), which appears due to the different behavior of people between weekdays and weekends.

Algorithm 1: Choosing the Intervals to Explore.

input : I, e, τ
output: Q (array of 1s and 0s defining the intervals to explore)

begin
  Initialise Q to all 0s
  NumberOfIntervalsToExplore $\leftarrow \tau \cdot e$
  while $sum(Q) \neq$ *NumberOfIntervalsToExplore* **do**
    $int_{chosen} \leftarrow$ Choose an interval(I) `// the chances of an interval being picked are proportional to its probability`
    if $Q[int_{chosen}] == 0$ **then**
      $Q[int_{chosen}] = 1$
    end
  end
end

3) The probabilities that define the chances of each time interval to be chosen for data gathering are calculated as

$$P(\text{int}_j) = 1/\frac{H_{\text{int}_j}([t_j, t_j+\Delta t))}{\sum_{z=1}^{\tau} H_{\text{int}_z}([t_z, t_z+\Delta t))},\ j \in 1, 2, \ldots, \tau \quad (7)$$

so that an array $I = P(\text{int}_j),\ j \in 1, 2, \ldots, \tau$ containing the probabilities of all the intervals can be obtained. Finally, using I, the temporal entropy-based scheduling is determined by Algorithm 1.

D. Defining Where to Explore

This subsection explains how, given a set of ρ regions R each containing a subset of s observable cells, the entropy is used to calculate the probabilities of each region R_r being picked (in the intervals chosen by Algorithm 1) by the entropy-based spatial policy for data gathering. In contrast to the approach presented for deciding when to explore, here the temporal entropy evolution is not taken into account. Instead the entropy is calculated from all the accumulated person detections gathered in the past. In order to calculate the probabilities for each R_r the following steps are proposed.

1) First, the entropy of each region R_r is computed as the sum of the entropy of all the cells that are observable (c_r) at R_i as

$$H_{R_r} = \sum_{i=1}^{s_r} H_{cell_i}[P(X)],\ r \in 1, 2, \ldots, \rho \quad (8)$$

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

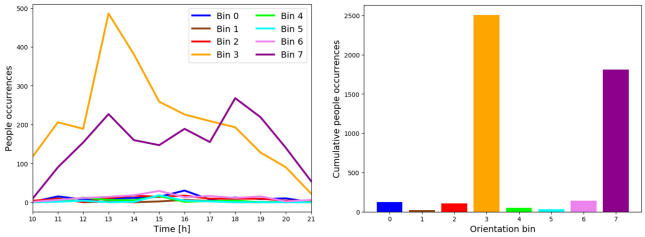


Fig. 5. Pedestrian counts in each direction over a day and the cumulative distribution for the cell a with low entropy.

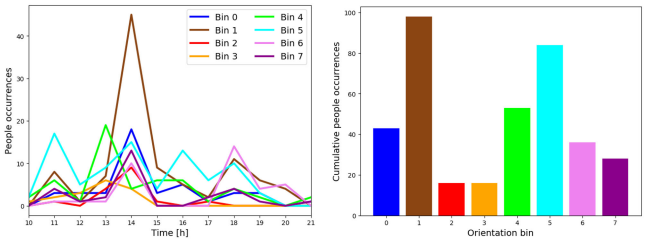


Fig. 6. Pedestrian counts in each direction over a day and the cumulative distribution for the cell b with high entropy.

where $H_{\text{cell}}[P(X)]$ is given by (4) from all the detections seen in the past distributed across the k bins. For example, Figs. 5 and 6 show the occurrences for two different cells (a and b) using $k = 8$ discrete orientations in 1-h time intervals and their corresponding cumulative distributions. The cell a in Fig. 5 shows a distribution with two clear peaks corresponding to the two dominant orientations of human motion, corresponding to a low entropy value ($H_{\text{cell}_a} = 1.58$). By contrast, Fig. 6 illustrates a cell b in another part of the environment, where each one of the $k = 8$ orientations obtains a similar number of person detections, corresponding to a more unpredictable behavior, obtaining a flatter distribution, and hence a higher entropy value ($H_{\text{cell}_b} = 2.76$).

2) Second, the probabilities that define the chances of each region R_i to be chosen to explore are calculated as

$$P(R_r) = 1/\frac{H_{R_r}}{\sum_{j=1}^{\rho} H_{R_j}},\ r \in 1, 2, \ldots, \rho. \tag{9}$$

Every time a certain interval is chosen to be explored, the region where the robot will travel to obtain information about the human motion has to be defined. Similar to the approach presented for the intervals in Algorithm 1, this decision is taken randomly, where the probability of selecting a given region R_r is proportional to $P(R_r)$.

For example, let us take an environment with just two regions where each region contains just one cell. R_1 has only the cell a from Fig. 5 and R_2 has only the cell b from Fig. 6. After gathering the information shown in Figs. 5 and 6, the associated normalized probabilities in this case would be $P(R_1) = 0.64$ and $P(R_2) = 0.36$.

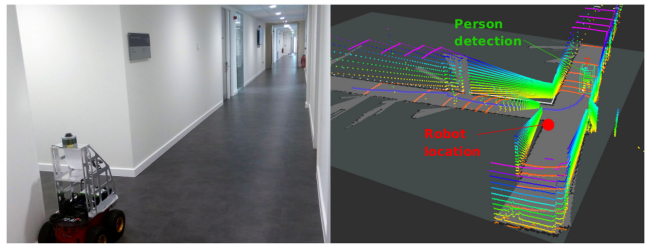


Fig. 7. Corridor dataset: Robot location in the corridor and example of a person walking seen by the Velodyne scans.

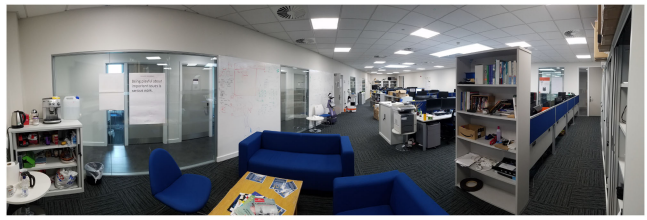


Fig. 8. Photo of the office area covered by the Velodyne VLP-16 recordings.

VI. Evaluation

A. Experimental Scenarios

In the experimental section, three exploration policies are tested (entropy, random and round robin), which define the set of rules to create the exploration sequence both in time and space.

1) *Entropy (E) Policy*: The regions/intervals are chosen following the schemes presented in the previous section. The recalculation of the entropy levels is done at the end of each day of exploration using all the data gathered on that day.
   For the two uninformed policies implemented, the environment dynamics are not taken into account. These strategies calculate the sequence of visits simply from the number of intervals τ, the number of regions ρ, and the ratio e.
2) *Random (R) Policy*: As its name indicates, the regions and intervals for exploration are chosen in a uniformly random way. Thus, all the τ intervals and all the ρ regions have the same probability of being chosen.
3) *Round Robin (B) Policy*: All the areas/time intervals of the environment are visited with the same frequency, interleaving the observations so that the exploration ratio e is satisfied.

In initial comparisons, a *Greedy (G)* algorithm was also implemented [44]. This always looks for the intervals/regions with the highest probability defined by (7) and (9), respectively. However, in all the experiments this method performed poorly compared to the other three policies aforementioned. For that reason, the results for this approach are only presented in Fig. 15 but not in the other figures containing the box-plots. This might seem a bit counter-intuitive, but the fact that the robot always goes to the area with the lowest entropy is not always the best option, as there is no chance for the other parts to be explored. This is due to the fact that in this experimental scenario, observing a cell does not imply a change in its entropy after the

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

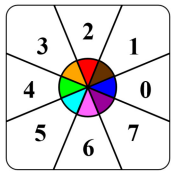


Fig. 9. Angle discretization used in every cell for building the STeF-Map models. The 8 bins color code has been kept consistent across the figures.

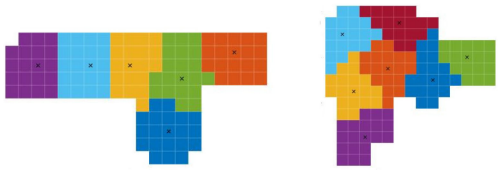


Fig. 10. Spatial division in observable regions for the corridor and office environments.

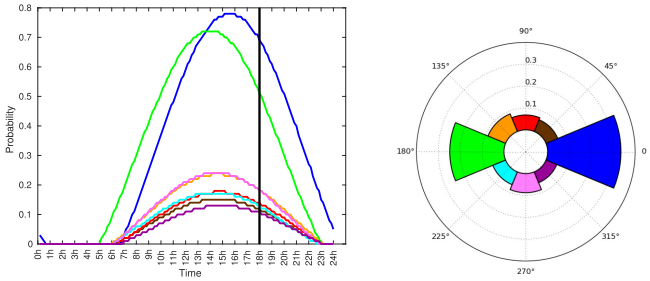


Fig. 11. Model prediction over 24 h with m=1, and probability distribution of each orientation at t = 18:00.

measurement. The entropy, random, and round robin exploration policies are compared using the nine different spatio-temporal combinations: R-R, B-R, E-R, R-B, B-B, E-B, R-E, B-E, and E-E, where the first letter indicates the policy for choosing the location and the second letter indicates the one in charge of deciding when to explore. So, for example, in the case of R-B, the region to explore is chosen in a random way, while the time to explore is deterministic. Regarding the temporal aspect, three different exploration ratios are used: $e = 0.5, 0.25, 0.125$. This percentage defines the number of time intervals that the robot will devote to gathering data from the total number of intervals available.

B. Datasets

To evaluate the approach, we ran the experiments using three real pedestrian datasets. All feature complex human movement and enough days to train the models and evaluate the different exploration strategies in the long term. The pedestrian detections in the environments are given in x, y coordinates together with the heading angle α for every timestamp t, all of them given in the same fixed world frame. Following [45], from each dataset we have taken certain days for training, some for validation, and others for testing, but none of the days for each set overlap.

1) Shopping Centre-ATC: The first dataset was recorded by tracking pedestrians at the the ATC shopping centre in Osaka, Japan [7]. The perception system consists of multiple 3D range sensors covering an area of approximately 900 m^2 (Fig. 1), which is able to detect and track all the people at every instant of time. The data were recorded on every Wednesday and Sunday between October 24th, 2012 and November 29th, 2013, resulting in a total of 92 days. From these data, we selected the first 46

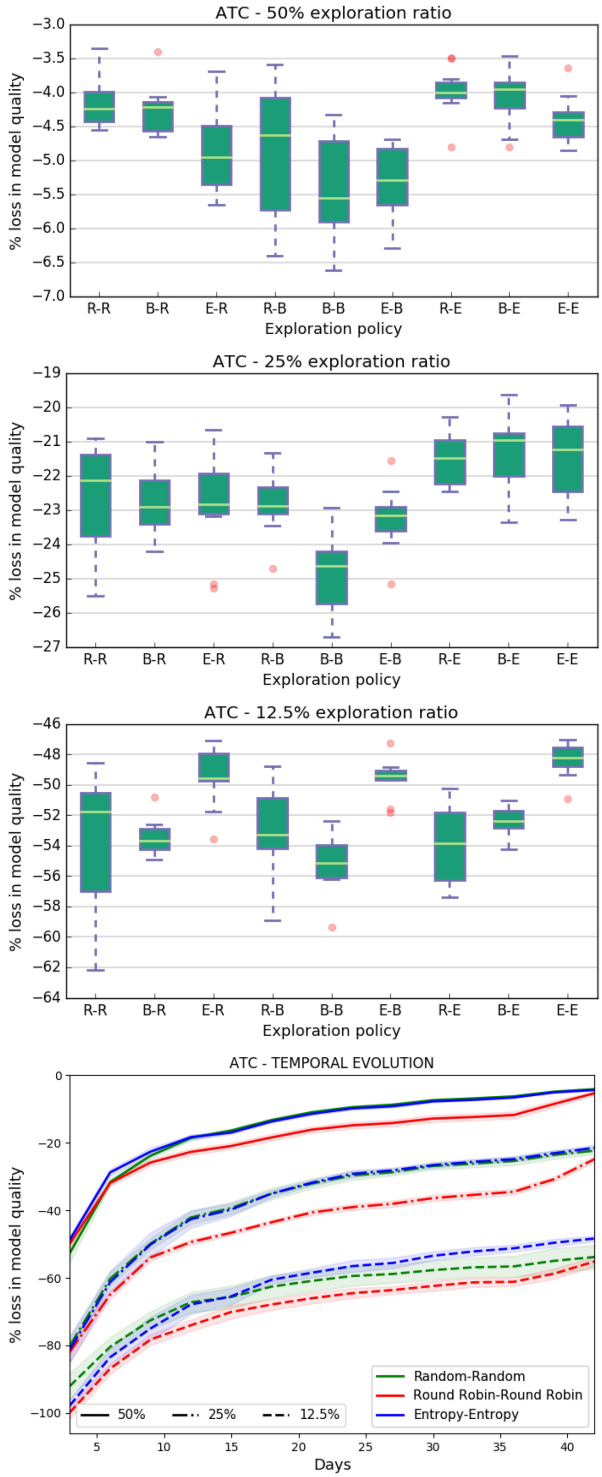


Fig. 12. Results for the ATC dataset with 50, 25, and 12.5% exploration ratios with the 9 spatio-temporal exploration strategy combinations, and temporal evolution over the exploration days for the 3 pure combinations.

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

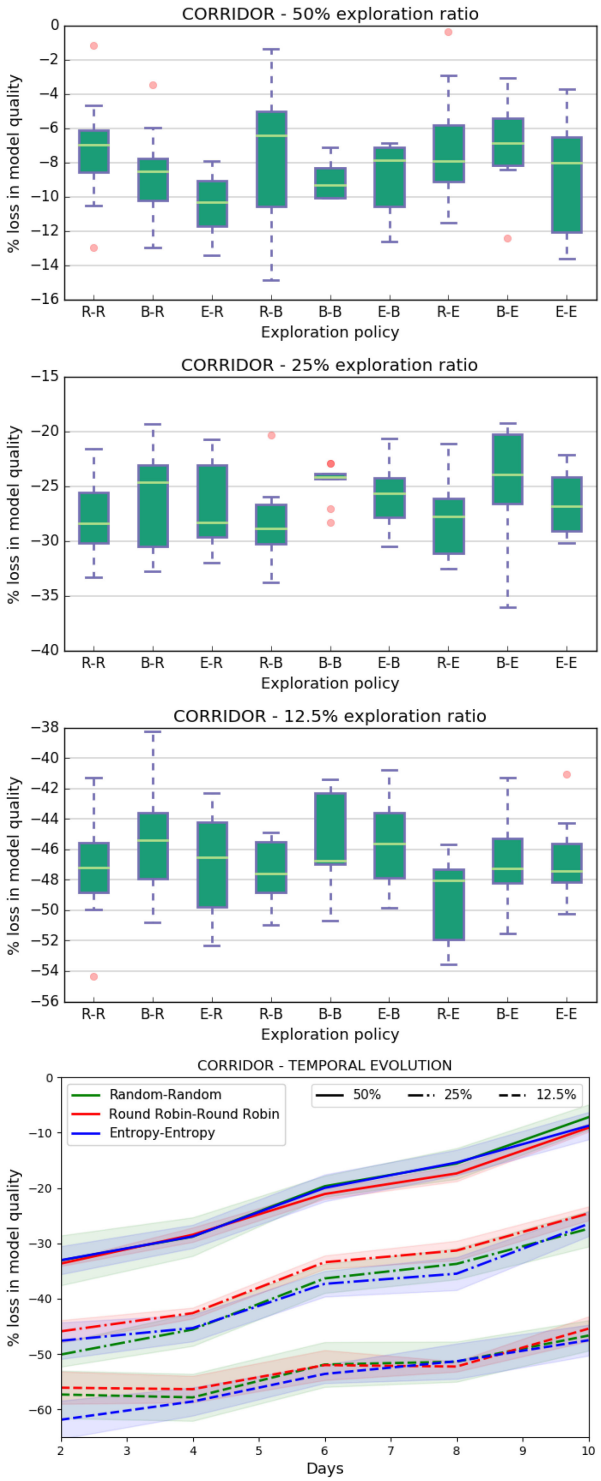


Fig. 13. Results for the Corridor dataset with 50, 25, and 12.5% exploration ratios with the 9 spatio-temporal exploration strategy combinations, and temporal evolution over the exploration days for the 3 pure combinations.

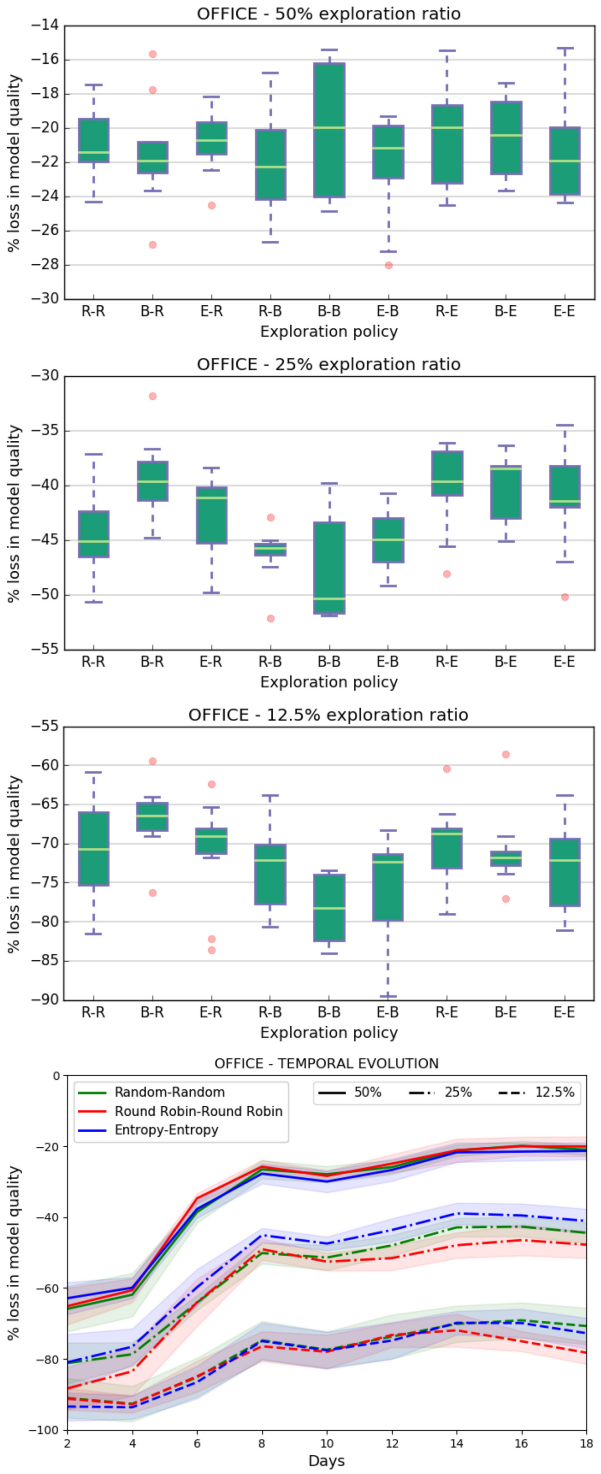


Fig. 14. Results for the Office dataset with 50, 25, and 12.5% exploration ratios with the 9 spatio-temporal exploration strategy combinations, and temporal evolution over the exploration days for the 3 pure combinations.

consecutive days (23 Wednesdays and 23 Sundays), using 42 to perform exploration, 2 days for validation and the other 2 for testing. The recording of each day provides people trajectories starting from approximately 09:00 until 21:00, so for the rest of day we assume there are no occurrences of people, simulating the shopping centre being empty. Each of the recorded days contains around 1 million detections of people.

2) Corridor: The second dataset was collected at one of corridors in the Isaac Newton building at the University of Lincoln. The data were recorded by a Pioneer 3-AT mobile robot equipped with a 3D LiDAR (Velodyne VLP-16) and a 2D

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

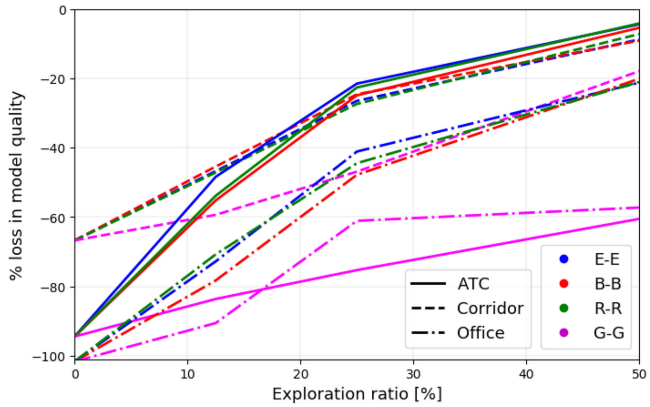


Fig. 15. Correlation between model quality loss and exploration ratio for all datasets with the 3 pure exploration strategy combinations (E-E, B-B, R-R) and the greedy-greedy (G-G) approach omitted in the previous results due to poor performance compared to the rest.

laser (Hokuyo UTM-30LX). During data collection, the robot remained stationary in a T-shaped junction, which allowed its sensors to scan the three connecting corridors simultaneously, covering a total area of around 75 m^2 (see Fig. 7). However, since the robot could not stay at the corridor overnight due to safety rules, and it was needed by other researchers occasionally, we did not collect the data on a full 24/7 basis. Instead, the data collection was performed in 10–12-h sessions starting before the usual working hours. Recharging of the batteries was performed overnight, where the building is vacant, and there are no people on the corridors. The resulting dataset is composed of 14 data-gathering sessions recorded over a span of four weeks. From these, ten days were used for training, two for evaluation, and the remaining two for testing. To detect and localize people in the 3D point cloud provided by the LiDAR, we used an efficient and reliable person detection method developed by Yan *et al.* [46]. A typical session contains approximately 30 000 detections of people walking in the monitored corridors.

3) Office: The third dataset was also collected in the Isaac Newton Building building at the University of Lincoln, but in this case, inside a large open-plan office (Fig. 8). The recordings were done with a static 3D LiDAR (Velodyne VLP-16) placed on a tripod at 1.8-m height and using the same people detector as in the corridor dataset [46]. The sensor was placed in a position covering two entrances to the office, an open area, and the coffee area, covering an area of approximately 85 m^2. The dataset contains 22 days recorded consecutively starting November 23th, 2018, and each day contains around 25 000 entries during working hours, which are usually from 08:00 to 20:00. In the experimental section, these days are divided using the first 18 for training, the following 2 for validation, and the final 2 for testing.

C. Model Parameters

In our experiments, we discretized the space into 1×1-m cells for all environments, resulting in a total of $C = 1248$, $C = 117$, and $C = 126$ active cells, respectively. The cell size can be defined arbitrarily. However, having a really fine grid (e.g., 5×5 cm like in an occupancy map) could result in many cells remaining empty during model building, depending on the number of persons (or "density" of the crowd). On the other hand, defining it too coarsely could lead to an overgeneralization of the space, merging flows that should be physically separated. Taking into account that the average walking speed of people in the datasets is around 1.1 m/s, the detection frequency of the person detections provided is 2 Hz, the size of an average person width is around 0.4 m [47] and the previous experimental results obtained when building different STeF-map representations [8], a cell size of between 0.5 and 2 m is suitable.

TABLE I
SUMMARY OF THE SPATIAL AND TEMPORAL PARAMETERS USED IN EACH DATASET

Dataset	Train	Validate	Test	# Regions	Cell size	Cells (n)
ATC	42 days	2 days	2 days	24	1×1 m	1248
Corridor	10 days	2 days	2 days	6	1×1 m	117
Office	18 days	2 days	2 days	7	1×1 m	126

The number of bins chosen to discretize the orientations in all three cases is $k = 8$, distributed at 45-degree intervals, as shown in Fig. 9. Similar to the cell size discretization, the number of bins discretizing the orientation can be chosen arbitrarily. However, from the experimental results in [8], eight bins tends to be a good number, as with higher numbers of bins the predictions results are not as good, and with lower numbers of bins such as four, some flows are lost due to overgeneralization. Also with eight bins as in Fig. 9, the probabilities for each direction can be associated with a transition cost map to the neighboring cells, as exploited in [34] for robot navigation in human populated.

As explained before, in our experiments we assume that it is not possible to observe the state of the whole environment at once, so instead we defined a set of observable regions for each dataset. During exploration, only the people passing within the boundaries of the chosen region are taken into account to update the model at a given interval in time, while the rest of the environment remains unseen. We defined 24 regions (see Fig. 1) for the ATC dataset, six for the corridor [see Fig. 10(a)], and seven for the office environment [see Fig. 10(b)], giving an average size of 52, 19, and 18 observable m^2 per region, respectively. The 3D LiDAR human detector used for creating the datasets achieves good performance up to around 10 m, which translates to an area of roughly 300 m^2. However, this is in an ideal scenario with open space and no occlusions, which is usually not the case. Therefore, for the ATC dataset we chose to have full 360-degree coverage with a downscaled radius of 4 m, corresponding to an area of $\sim$50 m^2. Using this same 50 m^2 for the smaller corridor and office datasets would mean having only 2 regions, so to make it more interesting we chose an area in the order of the typical coverage of a depth camera, i.e., 6-m range and 60 degrees of horizontal field of view, corresponding to a size of $\sim$19 m^2. In order to partition each environment into spatial regions, the k-means algorithm was used [48], optimizing such that all regions should contain a similar number of active cells. The summary of the parameters used in each dataset can be seen in Table I.

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Regarding time, we used the same parameters for the three datasets. The interval for creating the histograms used as the input for the STeF-map model creation was set to 10 min. The same interval was used to provide a single observation, i.e., every 10 min the exploration strategy can decide whether the robot should stay in the same region or move to a different one. If the robot has to move to another location, we calculate the path from the centroid (marked with crosses in Figs. 1 and 10) of the current region to the centroid of the goal region using the A* planner. During the travelling phase, we assume a constant robot speed of 1 m/s (we assume there is no interaction with the people moving around), and that the robot is only able to see what happens inside the region currently being traversed at each instant of time.

For each environment, the total available time for exploring is 12 h, corresponding to 72 time intervals (10 min each). However, for the ATC dataset the starting time is 09:00, and for the corridor and office the starting time is 08:00. The rest of the time we assume an empty environment and we set all the cells of the environment to 0. From the 72 time intervals available, data gathering only happens in a certain number of time intervals, which is proportional to the exploration ratio set in each case (e.g., if the ratio is 25%, only 18 of the 72 intervals are used for exploration, while the rest remain unused).

Every cell in the map can present different periodicities corresponding to different human activity patterns. So, in the experiments, each cell's patterns was calculated with either 1 or 2 periodicities, which is usually sufficient to represent the environment dynamics [37]. Using the validation days, the best number of spectral components for each cell is chosen with the corresponding frequencies, magnitude, and phase shifts. These are used later to compute the model predictions and loss in model quality over the testing days.

D. Evaluation Metrics

In order to compare the performance of the different models, a metric is needed to evaluate how good the predictions are when compared to the ground truth.

The output of the trained spectral models provides a function for each orientation in each cell. So, for every time t, we can obtain a normalized distribution describing how probable it is to find a person moving in each direction. For example, Fig. 11 presents the prediction for a single cell in the map (using $m = 1$ FreMEn component) over 24 h after some days of training. In this cell, there are two dominant orientations, one with higher probabilities in the morning and the other in the afternoon. At $t = 18{:}00$, the distribution of the eight normalized orientation probabilities is shown in the polar histogram on the right, which is what we compare against the ground truth. However, obtaining the same orientation distribution with real data at a single time instance t is not possible, because we cannot count sufficient detections to build a meaningful distribution. Instead, the proposed idea is to compare the distribution obtained with the predictions against the ground truth during a defined interval of time, which in our case, for all the experiments, we set to 10 min. Then, assuming that both prediction and ground truth histograms are normalized, the Chi-squared (χ^2) distance is used to compare the similarity between them.

The χ^2 histogram distance comes from the χ^2 test-statistic [49], where it is used to test the fit between a distribution and the observed frequencies. This is a common bin-to-bin distance for comparing histograms, which has been successfully applied in different domains such as texture and object categories classification [50], near-duplicate image identification [51], matching of local descriptors [52], or shape classification [53]. Despite this distance being sensitive to quantization effects, the fact that the number of bins used to discretize the orientation is low (8 in our case) makes it a robust distance measure [54]. If the number of bins was much higher, then other metrics that take into account cross-bin relationships, such as the Earth mover distance [55], would probably give a more robust approach.

In this work, the χ^2 distance indicates the level of similarity between the predicted and ground-truth discrete human motion distributions. The higher the distance, the less accurate is the model prediction compared to the ground truth. The total distance of a whole map for a single interval can be defined as

$$\chi^2_{\text{map}} = \sum_{c=1}^{C} \left(\sum_{b=1}^{k} \frac{(v_b - w_b)^2}{(v_b + w_b)} \right) \tag{10}$$

where C is the number of active cells in the map, k is the number of angular bins, v_b is the value of bin b in the predicted orientation histogram, and w_b is the value of the same bin b obtained from the ground-truth data.

Since the χ^2 distance is not a very intuitive measure as it has no units, in the results section we express the prediction accuracy as the percentage loss in model quality. The loss is based on how much worse the prediction of a model built with partial observability is when compared to the prediction obtained with the model created with full observability in time and space, i.e., with a 100% exploration ratio and always seeing all the cells in the map. Hence, the closer the value to 0, the better the performance of the spatio-temporal strategy as the model created with it gets closer to the optimal scenario where all the people detections are taken into account in the model

$$\text{loss model quality} = 100 \cdot \frac{\chi^2_{\text{full obs}} - \chi^2_{\text{partial obs}}}{\chi^2_{\text{full obs}}} \, [\%]. \tag{11}$$

VII. Results

In this section we present the result obtained for the three different datasets (Figs. 12 to 14). For each one, we tested the nine possible spatio-temporal exploration combinations: R-R, B-R, E-R, R-B, B-B, E-B, R-E, B-E, and E-E (R = Random, B = Round-robin, and E = Entropy), with three different exploration ratios e: 50, 25, and 12.5%. As mentioned in the previous subsection, all the results shown and the comparison done in this section use the *loss in model quality* indicator explained in the previous Section VI-D and defined in (11).

Since the *random* and *entropy* strategies produce stochastic policies, the values obtained are shown using a boxplot over 10 runs (median in yellow, interquartile in green, minimum and maximum with the dashed lines, and potential outliers in red).

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Even though *round-robin* produces a deterministic policy for visiting the areas/intervals, the results are also expressed with confidence intervals, since the starting region is not always the same. The values obtained are always computed at the end of the total training days corresponding to each dataset (Table I). The evaluation days are used to decide the number of periodicities for each cell, and the testing days are used to obtain the percentage loss in model quality.

Furthermore, for the three "pure" combinations possible, namely, R-R, B-B, and E-E, we computed the model quality loss at certain days during the total exploration days, always using the same evaluation/training days (at the bottom of each Figs. 12, 13, and 14). The results in this case are plotted as the average over 10 runs together with a 95% confidence interval using student's t-distribution.

A. ATC Dataset

The results obtained for the ATC dataset are summarized in Fig. 12. As a general overview, we see that the strategies with an entropy-based policy perform consistently better than their uninformed counterparts.

For low exploration ratios, as in the 12.5% case, we found that E-R, E-B, and E-E produce better results, showing that taking into account the entropy in the spatial domain is a key factor. However, as the available time to explore is increased, the strategies which take into account the entropy to decide when to explore (R-E, B-E, E-E) obtain better results, and the spatial domain becomes a secondary factor.

Observing the temporal evolution, we see that B-B exploration obtains by far the worst results, being the slowest one and only being able to catch up with the rest of the strategies in the period from days 35 to 42. For the 50% case, R-R and E-E perform very similarly, but for 25% and 12.5% after day 25, approximately, E-E obtain slightly but significantly better results.

B. Corridor and Office Datasets

For the last two datasets (Figs. 13 and 14), the results are a bit more difficult to analyze, due to fact that the deviations on the results tend to be much bigger compared with the ATC dataset. We believe this is a consequence of having a much lower number of people occurrences in each day. So, sometimes, seeing a person in a certain interval of time becomes just a matter of luck.

Nevertheless, we found that for the 25% exploration ratio in the Corridor dataset, the B-B exploration strategy is the one that obtains consistently better results, which is also confirmed by observing the temporal evolution. However, for the 50% and 12.5% exploration ratios, we cannot extract any meaningful conclusions, as all of them behave similarly and there are no patterns on either the spatial or temporal side.

For the Office dataset we found that the exploration strategies which follow the *round-robin* policy in the temporal domain perform worse for low exploration ratios, which can also be seen in the temporal evolution plot. The major difference appreciated in this dataset comes for the 25% ratio, where the combinations sharing the entropy-based exploration in the time domain (R-E, B-E, and E-E) manage to obtain consistently better results.

C. Discussion

The results suggest that the entropy-based exploration works well when we have an environment with a substantial number of people detections and somewhat regular flows, as in the ATC dataset. In scenarios with lower human encounters, like the Corridor and Office environments, it would probably be necessary to further extend the days explored to several weeks to obtain statistically significant results, or compute hundreds of runs to deal with the higher deviations obtained. Also, in the Corridor and Office environments we noticed that at any point, the exploration strategy in charge of scheduling the region to explore has a major impact on the final results. Probably the fact that there are a much lower number of regions to explore compared to the shopping centre makes no actual difference, as the simulated robot is able to visit them a lot more times, even for low exploration ratios.

However, taking the exploration ratio as clearly the factor that has the biggest impact on the total percentage loss in model quality during the exploration activities (which is clear to see in the temporal evolution plot in Figs. 12 to 14), we found that in all datasets this impact follows a similar trend. In Fig. 15 we plotted the correlation between the percentage loss in model quality and the exploration ratio for all 3 datasets, at the end of the corresponding exploration days for the 3 pure combinations (E-E, B-B, R-R). The value obtained with 0% exploration ratio corresponds to a model which has not been trained with any data, so all orientations in each cell have the same probability. The outcome shows that the loss tends to decrease exponentially as the exploration ratio is increased.

Looking at the loss in model quality also in Fig. 15, the Office dataset obtains noticeably worse results for all exploration ratios compared to the Corridor and ATC datasets. We think this is caused by the fact that the recordings in the Office dataset were also done during the weekends, days which present a very different behavior (the environment is mostly empty) when compared to the weekdays. This increased complexity makes it more difficult for the spatio-temporal model to find the overall patterns to be modeled using only sparse partial observations. In the ATC dataset, weekends are also part of the data, but in this case, there are no major differences in human motion behavior between the weekday and weekend recorded.

VIII. Conclusion

In this article, we present a comparison between multiple robot exploration strategies to build a spatio-temporal model of human motion in a given environment. Moreover, we propose the use of the data gathered by the robot to determine where and when to perform future observations based on the entropy levels in the model, which are computed from the distributions of pedestrian motion direction.

The results show that the entropy-driven policy improves performance compared to the uninformed exploration strategies in scenarios with a substantial degree of human presence and

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

rhythmic patterns of activity. On this issue, future work will aim to do a more in-depth analysis to quantify the characteristics of a given environment and analyze the differences.

Furthermore, we have demonstrated that the exploration ratio is the key factor affecting the model prediction quality and that we obtain similar trends in the correlation between model quality loss and the exploration ratio. This is interesting, considering that all three datasets tested in the experiments contained a different number of exploration days and a different number of regions to explore.

The increase in prediction accuracy means that a robot has a better understanding of how the people will behave, hence enabling the robot to plan its activities around humans in a more optimal way. The benefits of this technology should extend to a wide range of applications where socially compliant navigation is needed. Possible use cases include transport and delivery applications in crowded environments, where the robot should move with the direction of pedestrian flows in order to minimize the possibility of head-on collisions, for example, in shopping malls [7], airports [56], and museums [57].

In service robotics, the approach could be used by the robot's scheduling software to maximize and optimize encounters with selected user groups, such as providing information services to residents in care homes, or patrolling in security and surveillance applications [58], based on the expected future locations of people. For example, a robot might move to a location where a task has to be carried out some time in advance to avoid congestion at the destination point.

In safety critical applications within industrial environments, such as warehouses [59] and farms [60], the approach could be used to maintain a safe distance to the expected location of human workers, in order to minimize the probability of accidents.

In the simulation experiments we tried to mimic a real mobile robot exploring the environment using prerecorded datasets of pedestrian activities. However, we understood that the results were obtained using certain simplifications, including no noise/occlusions in the perception and perfect navigation without any physical interaction with people. These aspects would play an important role when the experiments are carried out with real mobile robots, like for example in a scenario with high people density, the robot navigation could be far from the optimal as used in this work. Nevertheless, we consider the current work necessary, as ground-truth data (having an "overhead" sensor/sensors capable of seeing the whole environment all the time) are essential to compare the exploration strategies. Using datasets that are collected by mobile robots would mean that we can only update the model from locations where the robot was actually present during the real-life recordings, which would not allow us to compare exploration strategies in the spatial domain. Also, the ability to place our simulated robot at any time in the past to start the exploration with the same human behavior over and over again, and to simulate the exploration days in a much smaller fraction of time, would be difficult if not impossible to reproduce for real-world experiments.

Future work will attempt to overcome the data ground-truthing constraint using some kind of temporal cross validation. Working with real mobile platforms certainly raises interesting problems for the multistrategy comparison, especially in the time domain. For example, in the experiments done here, the temporal strategies were free to choose any time interval during the available time as long as the total number of intervals chosen met the exploration ratio constraints. In a real environment, the available time can differ from day to day due to the other main application tasks in the robot's schedule, which adds further complexity to the system.

Another area for future work would involve exploiting the flow models obtained during exploration. Currently, the robot follows an optimal path from region A to region B, but instead the spatio-temporal model could be exploited in different ways. For example, the model could be used as a prior for path planning, so the robot can move in a more socially compliant manner, or the entropy values obtained from previous days could be used to bias the robot movements through certain regions to increase information gain even if some time is lost.

Ethical Considerations: The work presented in this paper complies with the Ethical Approval Forms EA1 and EA2 from the University of Lincoln submitted and approved at the beginning of the research activities.

REFERENCES

[1] P. T. Kucner, "Probabilistic mapping of spatial motion patterns for mobile robots," Ph.D. dissertation, Centre of Applied Autonomous Sensor Systems (AASS), Örebro Univ., Örebro, Sweden, 2018.

[2] T. Kruse, A. K. Pandey, R. Alami, and A. Kirsch, "Human-aware robot navigation: A survey," *Robot. Auton. Syst.*, vol. 61, no. 12, pp. 1726–1743, 2013.

[3] L. Sun, Z. Yan, S. Molina, M. Hanheide, and T. Duckett, "3DOF pedestrian trajectory prediction learned from long-term autonomous mobile robot deployment data," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 5942–5948.

[4] F. Jumel, J. Saraydaryan, and O. Simonin, "Mapping likelihood of encountering humans: Application to path planning in crowded environment," in *Proc. Eur. Conf. Mobile Robots*, 2017, pp. 1–7.

[5] C. S. Swaminathan, T. P. Kucner, M. Magnusson, L. Palmieri, and A. J. Lilienthal, "Down the cliff: Flow-aware trajectory planning under motion pattern uncertainty," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2018, pp. 7403–7409.

[6] L. Palmieri, T. P. Kucner, M. Magnusson, A. J. Lilienthal, and K. O. Arras, "Kinodynamic motion planning on Gaussian mixture fields," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2017, pp. 6176–6181.

[7] D. Brscic, T. Kanda, T. Ikeda, and T. T. Miyashita, "Person position and body direction tracking in large public spaces using 3D range sensors," *IEEE Trans. Human-Mach. Syst.*, vol. 43, no. 6, pp. 522–534, Nov. 2013.

[8] S. Molina, G. Cielniak, and T. Duckett, "Modelling and predicting rhythmic flow patterns in dynamic environments," in *Proc. Towards Autonomous Robotic Systems*, vol. 10965, 2018, pp. 135–146.

[9] S. Molina, G. Cielniak, T. Krajnik, and T. Duckett, "Modelling and predicting rhythmic flow patterns in dynamics environments," in *Proc. Robot. Auton. Syst.: Robots Work. Among Us*, 2017, pp. 135–146.

[10] "Software and datasets," Lincoln Centre Auton. Syst., Lincoln, U.K., 2020. [Online]. Available: https://lcas.lincoln.ac.uk/wp/research/data-sets-software/

[11] B. Yamauchi, "Frontier-based exploration using multiple robots," in *Proc. 2nd Int. Conf. Auton. Agents*, 1998, pp. 47–53.

[12] D. Holz, N. Basilico, F. Amigoni, and S. Behnke, "Evaluating the efficiency of frontier-based exploration strategies," in *Proc. 41st Int. Symp. Robot. 6th German Conf. Robot.*, 2010, pp. 1–8.

[13] F. Amigoni and V. Caglioti, "An information-based exploration strategy for environment mapping with mobile robots," *Robot. Auton. Syst.*, vol. 58, no. 5, pp. 684–699, 2010.

[14] J. Vallvé and J. Andrade-Cetto, "Dense entropy decrease estimation for mobile robot exploration," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2014, pp. 6083–6089.

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[15] J. P. Fentanes, A. Badiee, T. Duckett, J. Evans, S. Pearson, and G. Cielniak, "Kriging-based robotic exploration for soil moisture mapping using a cosmic-ray sensor," *J.Field Robot.*, vol. 37, no. 1, pp. 122–136, 2018, *arXiv:1811.05384*.

[16] C. Stachniss and W. Burgard, "Mobile robot mapping and localization in non-static environments," in *Proc. Conf. Artif. Intell.*, 2005, pp. 1324–1329.

[17] R. Marchant and F. Ramos, " Bayesian optimisation for intelligent environmental monitoring," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2012, pp. 2242–2249.

[18] A. Singh, F. Ramos, H. D. Whyte, and W. J. Kaiser, "Modeling and decision making in spatio-temporal processes for environmental surveillance," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2010, pp. 5490–5497.

[19] J. M. Santos, T. Krajník, and T. Duckett, "Spatio-temporal exploration strategies for long-term autonomy of mobile robots," *Robot. Auton. Syst.*, vol. 88, pp. 116–126, 2017.

[20] T. Krajnik, T. Vintr, S. Molina, J. P. Fentanes, G. Cielniak, and T. Duckett, "Warped hypertime representations for long-term autonomy of mobile robots," *IEEE Robot. Autom. Lett.*, vol. 4, no. 4, pp. 3310–3317, 2019, *arXiv:1810.04285*.

[21] A. Rudenko *et al.*, "Human motion trajectory prediction: A survey," *Int. J. Robot. Res.*, vol. 39, no. 8, pp. 895–935, 2020, *arXiv:1905.06113*.

[22] R. Limosani *et al.*, "Long-term human affordance maps," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2015, pp. 5748–5754.

[23] G. D. Tipaldi and O. K. Arras, "I want my coffee hot! Learning to find people under spatio-temporal constraints," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 1217–1222.

[24] R. Senanayake and F. Ramos, "Directional grid maps: Modeling multimodal angular uncertainty in dynamic environments," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2018, pp. 3241–3248.

[25] T. P. Kucner, M. Magnusson, E. Schaffernicht, V. H. Bennetts, and A. J. Lilienthal, "Enabling flow awareness for mobile robots in partially observable environments," *IEEE Robot. Automat. Lett.*, vol. 2, no. 2, pp. 1093–1100, Apr. 2017.

[26] A. Asahara, K. Maruyama, A. Sato, and K. Seto, "Pedestrian-movement prediction based on mixed Markov-chain model," in *Proc. 19th ACM SIGSPATIAL Int. Conf. Adv. Geographic Inf. Syst.*, 2011, pp. 25–33.

[27] Z. Wang, P. Jensfelt, and J. Folkesson, "Multi-scale conditional transition map: Modeling spatial-temporal dynamics of human movements with local and long-term correlations," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2015, pp. 6244–6251.

[28] Y. Han, R. Tse, and M. Campbell, "Pedestrian motion model using non-parametric trajectory clustering and discrete transition points," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2019, pp. 2614–2621.

[29] F. Jovan, J. Wyatt, N. Hawes, and T. Krajnik, "A Poisson-spectral model for modelling temporal patterns in human data observed by a robot," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2016, pp. 4013–4018.

[30] W. Zhi, R. Senanayake, L. Ott, and F. Ramos, "Spatiotemporal learning of directional uncertainty in urban environments," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2019, pp. 4306–4313.

[31] T. Wada, Z. Wang, Y. Ogawa, Y. Hirata, and K. Kosuge, "Incremental human motion map system and human walking behavior representation in indoor environment," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2012, pp. 747–752.

[32] E. Chinellato, K. V. Mardia, D. C. Hogg, and A. G. Cohn, "An incremental von Mises mixture framework for modelling human activity streaming data," in *Proc. Int. Conf. Time Ser. Forecasting*, 2017, pp. 379–389.

[33] Tomáš Vintr *et al.*, "Time-varying pedestrian flow models for service robots," in *Proc. Eur. Conf. Mobile Robot.*, 2019, pp. 1–7.

[34] T. Vintr *et al.*, "Natural criteria for comparison of pedestrian flow forecasting models," in *Proc. Int. Conf. Intell. Robots Syst.*, 2020, pp. 11197–11204.

[35] J. M. Santos, T. Krajník, J. P. Fentanes, and T. Duckett, "Lifelong information-driven exploration to complete and refine 4-D spatio-temporal maps," *IEEE Robot. Automat. Lett.*, vol. 1, no. 2, pp. 684–691, Jul. 2016.

[36] S. Molina, G. Cielniak, and T. Duckett, "Go with the flow: Exploration and mapping of pedestrian flow patterns from partial observation," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2019, pp. 9725–9731.

[37] T. Krajník, J. P. Fentanes, J. Santos, and T. Duckett, "FreMEn: Frequency map enhancement for long-term mobile robot autonomy in changing environments," *IEEE Trans. Robot.*, vol. 33, no. 4, pp. 964–977, Aug. 2017.

[38] Paul E. Smaldino, "Measures of individual uncertainty for ecological models: Variance and entropy," *Ecological Model.*, vol. 254, pp. 50–53, 2013.

[39] T. Schürmann, "Bias analysis in entropy estimation," *J. Phys. A: Math. General*, vol. 37, no. 27, 2004, Art. no. L295.

[40] George A. Miller and William G. Madow, "On the maximum likelihood estimate of the Shannon–Wiener measure of information," in *Readings in Mathematical Psychology*, Hoboken, NJ, USA: Wiley, 1963.

[41] L. Paninski, "Estimation of entropy and mutual information," *Neural Comput.*, vol. 15, no. 6, pp. 1191–1253, 2003.

[42] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," in *Classic Works of the Dempster–Shafer Theory of Belief Functions*. Berlin, Germany: Springer, 2008.

[43] G. Shafer, *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976.

[44] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.

[45] M. Oliveira, L. Torgo, and V. S. Costa, "Evaluation procedures for forecasting with spatio-temporal data," in *Proc. Joint Eur. Conf. Mach. Learn.Knowl. Discov. Databases*, Springer, 2018, pp. 703–718.

[46] Z. Yan *et al.*, " Online learning for human classification in 3D LiDAR-based tracking,"in *Proc. Int. Conf. Intell. Robots Syst.*, 2017, pp. 864–871.

[47] G. Singh, A. Singla, and G. S. Virk, "Modeling and simulation of a passive lower-body mechanism for rehabilitation," in *Proc. Conf. Mech. Eng. Technol.*, 2016, pp. 123–128.

[48] S. Lloyd, "Least squares quantization in PCM," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 129–137, Mar. 1982.

[49] WG Cochran, "The chi-square test of goodness of fit," *Ann. Math. Statist*, vol. 23, pp. 315–345, 1952.

[50] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 11, pp. 2032–2047, Nov. 2009.

[51] D. Xu, T. J. Cham, S. Yan, L. Duan, and S.-F. Chang, "Near duplicate identification with spatially aligned pyramid matching," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 8, pp. 1068–1079, Aug. 2010.

[52] P.- E. Forssén and D. G. Lowe, "Shape descriptors for maximally stable extremal regions," in *Proc. IEEE 11th Int. Conf. Comput. Vision*, 2007, pp. 1–8.

[53] H. Ling and D. W. Jacobs, " Shape classification using the inner-distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 2, pp. 286–299, Feb. 2007.

[54] O. Pele and M. Werman, "The quadratic-chi histogram distance family," in *Proc. Eur. Conf. Comput. Vision*, 2010, pp. 749–762.

[55] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vision*, vol. 40, no. 2, pp. 99–121, 2000.

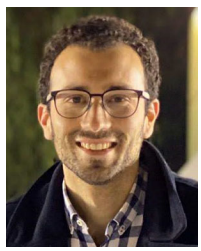
[56] R. Triebel *et al.*, " SPENCER: A socially aware service robot for passenger guidance and help in busy airports," in *Springer Tracts in Advanced Robotics*. Berlin, Germany: Springer, 2016.

[57] F. D. Duchetto, P. Baxter, and M. Hanheide, "Lindsey the tour guide robot—Usage patterns in a museum long-term deployment," in *Proc. 28th IEEE Int. Conf. Robot Hum. Interactive Commun.*, 2019, pp. 1–8.

[58] N. Hawes *et al.*, "The STRANDS project: Long-term autonomy in everyday environments," *IEEE Robot. Automat. Mag.*, vol. 24, no. 3, pp. 146–156, Sep. 2017.

[59] M.F. Carmona, T.Parekh, and M. Hanheide, "Making the case for human-aware navigation in warehouses," in *Towards Autonomous Robotic Systems*. Berlin, Germany: Springer, 2019, pp. 449–453.

[60] G. Das, G. Cielniak, P. From, and M. Hanheide, "Discrete event simulations for scalability analysis of robotic in-field logistics in agriculture—A case study," in *Proc. IEEE Int. Conf. Robot. Automat., Workshop Robot. Vision Action Agriculture*, 2018.



Sergi Molina was born in Reus, Spain, in 1991. He received the B.Sc. degree in electronic engineering from Universitat Rovira i Virgili, Tarragona, Spain, in 2014, the M.Sc. degree in robotics and automatic control from Universitat Politecnica de Catalunya, Barcelona, Spain, in 2016, and the Ph.D. degree in computer science from the University of Lincoln, Lincoln, U.K., in 2021.

Since 2020, he has been a Postdoctoral Research Associate with the University of Lincoln working within the Horizon 2020 European Project BACCHUS, with focus on robot localization and navigation in outdoor environments. His research interests include mobile robotics, artificial intelligence, spatio-temporal human motion models, and long-term mapping.

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Grzegorz Cielniak received the M.Sc. degree in robotics from the Wroclaw University of Technology, Wroclaw, Poland, in 2000, and the Ph.D. degree in computer science from Örebro University, Örebro, Sweden, in 2007.

He is currently an Associate Professor with the Lincoln Agri-Robotics centre, University of Lincoln, Lincoln, U.K. He has held a number of peer-reviewed grants from sources, including the European Commission, Innovate UK, Science and Technology Facilities Council, and Biotechnology and Biological Sciences Research Council. He is a Coordinator of the RASberry programme which brings public and industry funding to support robotic solutions for soft fruit production. His research interests include robotic perception and navigation in unstructured outdoor environments, with a focus on robotic application delivery in agriculture.

Tom Duckett received the B.Sc. (Hons.) degree in computer and management sciences from University of Warwick, Coventry, U.K., in 1991, the M.Sc. degree in knowledge based systems from Heriot-Watt University, Edinburgh, U.K., in 1995, and the Ph.D. degree in self-localisation and mapping by autonomous mobile robots from the University of Manchester, Manchester, U.K., in 1999.

He is currently a Researcher in robotics and autonomous systems based in Lincoln, U.K. From 2006 to 2021, he was with the University of Lincoln, Lincoln, U.K., where he was a Professor of Computer Science and the Director of the Lincoln Centre for Autonomous Systems. He has coauthored more than 150 scientific publications and held many peer-reviewed research grants, including the EPSRC Centre for Doctoral Training in Agri-Food Robotics (AgriFoRwArdS). He was recently listed in the world's top two percent of scientists (Web of Science, 2020). His research interests include perception systems for autonomous mobile robots, with applications in field and service robotics.

Authorized licensed use limited to: University of Lincoln -UK. Downloaded on September 15,2021 at 13:45:25 UTC from IEEE Xplore. Restrictions apply.