

Spatio-Temporal Exploration Strategies for Long-Term Autonomy of Mobile Robots

João Machado Santos^a, Tomáš Krajník^a, Tom Duckett^a

^a*Lincoln Centre for Autonomous Systems, University of Lincoln, Brayford Pool, Lincoln,
Lincolnshire, LN6 7TS, United Kingdom*

Abstract

We present a study of spatio-temporal environment representations and exploration strategies for long-term deployment of mobile robots in real-world, dynamic environments. We propose a new concept for life-long mobile robot spatio-temporal exploration that aims at building, updating and maintaining the environment model during the long-term deployment. The addition of the temporal dimension to the explored space makes the exploration task a never-ending data-gathering process, which we address by application of information-theoretic exploration techniques to world representations that model the uncertainty of environment states as probabilistic functions of time. We evaluate the performance of different exploration strategies and temporal models on real-world data gathered over the course of several months. The combination of dynamic environment representations with information-gain exploration principles allows to create and maintain up-to-date models of continuously changing environments, enabling efficient and self-improving long-term operation of mobile robots.

Keywords: mobile robotics, spatio-temporal exploration, long-term autonomy

1. Introduction

As robots gradually leave the well-structured worlds of factory assembly lines and enter natural, human-populated environments, new challenges ap-

Email addresses: jsantos@lincoln.ac.uk (João Machado Santos),
tkrajnik@lincoln.ac.uk (Tomáš Krajník), tduckett@lincoln.ac.uk (Tom Duckett)

pear. One of the first problems is to operate in less structured and more uncertain environments. This challenge gave birth to the field of probabilistic mapping, which enables the representation of incomplete world knowledge obtained through noisy sensory measurements [1]. Initially, the environment models had to be created during a human-guided procedure [2], but later, the combination of probabilistic mapping and planning methods allowed robots to create the environment models themselves by means of autonomous exploration [3]. However, as robots became able to operate autonomously for longer periods of time, a new challenge appeared – their typical operating environments are subject to change.

These changes manifest themselves through sensory measurements – every perceived change causes the sensory data to disagree with the original model obtained during the exploration phase. Although probabilistic mapping methods can deal with conflicting measurements, their approach is rooted in the idea that these variations are caused by inherent sensor noise rather than by structural environment change. Thus, these conflicting measurements are generally treated as outliers caused by unwanted noise. Recently, some authors exploited these conflicting measurements in order to obtain information about the world dynamics and proposed representations that model the environment dynamics explicitly. These dynamic representations have shown their potential by improving mobile robot localization in changing environments [4, 5, 6, 7].

Similarly to traditional robotic mapping, introduction of spatio-temporal mapping naturally requires novel exploration strategies that allow to build and maintain spatio-temporal maps during the robot’s deployment. Classic exploration strategies aim at building a spatial-only model that covers the robot’s entire operational environment, ignoring the fact the environment might change after its completion. Unlike the classic exploration approaches, spatio-temporal exploration is a never-ending task, for several reasons. First, some areas of the operational environment are not exactly predictable during certain times, which requires the robot to re-observe those locations at the times when their state is uncertain. For example, even if we know the general habits of a certain person, her presence at her workplace is uncertain around the start and end of office hours, and thus, it makes sense to observe the workplace during these times. Second, the patterns in the environment dynamics might change and identification of the new patterns requires re-observation of the particular area at the right times. For example, the workplace might be occupied by a new employee with a different working pattern. Additionally,

the general structure of the environment can change due to reconstruction or displacement of furniture.

Thus, the robot needs to take repeated observations of locations in its operational environment over time in order to successfully build and maintain a spatio-temporal model. This requires the robot to continuously explore the environment in addition to the other tasks it was designed for. Therefore, spatio-temporal exploration must become a part of the robot’s daily routine that has to be carried out along with other tasks that the robot is required to perform. The ability to build and maintain the aforementioned spatio-temporal representations allows the mobile robot to better cope with changes in the environment and to perform its daily duties efficiently. Hence, being able to build, maintain and reason over such an environment representation plays a key role in achieving long-term operation without any major human intervention, i.e., long-term autonomy.

We present an exploration method that integrates sensory data captured at different times and locations into a dynamic spatio-temporal model and uses the model to determine where and when to perform future observations, while being able to cope with the other tasks the robot needs to perform. We show that the application of information-theoretic planning principles to environment models that represent uncertainties of environment states in the frequency domain results in an intelligent exploratory behaviour, which evolves as the environment knowledge becomes more refined over time. Moreover we evaluate all possible combinations of four different spatio-temporal models and five planning strategies by their long-term performance, according to their ability to provide an accurate environment model over time. To complete this study we also evaluate the impact of different exploration versus exploitation ratios on the overall accuracy of the model. The *exploration versus exploitation dilemma* means that the robot has to find a balance between the time spent exploring and the quality of its internal model [38]. The work presented in this article extends the study presented in [8] by providing a more detailed description of the spatio-temporal models and exploration strategies and by introducing a new recency-based short-term model, as well as a novelty-driven exploration strategy that takes into account the predictions of both the recency- and periodicity-based models.

2. Related work

In order to explore the environment in an efficient way, the robot not only has to be able to create a map from its sensory inputs, but also to use the map to plan its path so that it can reach previously unknown areas of the environment. Therefore, mobile robot exploration is an iterative process in which the robot integrates its observations into its world model, interprets the world model to determine which parts of the environment are unknown, and plans a path to visit and observe these unknown areas. Therefore, an efficient exploration system consists of three essential components: mapping, goal generation and path planning. For the purpose of spatio-temporal exploration, we have to use mapping methods that allow to represent dynamic environments and goal generation methods that can determine not only the position, but also the times of observations – i.e. we have to schedule the observations in such a way that the robot can perform its other tasks as well.

2.1. Exploration methods

One of the earliest and well-known methods is frontier-based exploration [2, 9, 10]. This approach represents the environment as an occupancy grid, which is processed to obtain boundaries (frontiers) between the known and unknown parts of the environment. The robot movement is then planned so that these frontiers are visited and removed. The advantage of this approach is its scalability – the frontiers can be distributed among a number of robots that can explore the environment in a cooperative manner [11]. Even though these strategies ensure the completeness of the environment model, i.e. they aim at removing all the frontiers, they do not take into account the model quality.

Another class of exploration methods is based on the notion of entropy. These methods generate a set of candidate observations and estimate the amount of information these are expected to provide [12]. The information gain is calculated as the reduction in entropy of the world model, which requires a probabilistic representation of the environment states. The lower the entropy of the environment model, the more it reflects the actual environment state.

An information-gain-based approach that integrates localization, mapping and exploration is presented in [13]. The method uses a particle filter to build the map of the environment and an entropy estimation method to plan the next location to be visited by the robot. However, the candidate

observations are not evaluated simply by their information gain. Rather, the evaluation takes into account other criteria, such as the time to reach the respective location [14]. An advantage of these methods is that they not only attempt to cover the entire environment as quickly as possible, but also plan re-observations of previously visited locations to increase the quality of the resulting map [15].

Some exploration strategies aim at building maps of the environment taking into account some *a priori* knowledge instead of building it from scratch. For instance, Oßwald et al. [16] propose a novel exploration strategy that aims at decreasing the exploration time by assuming that the layout of the environment is known, such as graphs automatically obtained from floor plans. In this method a Travelling Salesman Planner generates a global plan for the exploration run, while a frontier-based strategy is used to explore the environment at each node of the graph. In [17] an exploration strategy capable of predicting how the unexplored areas may look based on previously mapped areas is proposed. This strategy combines the knowledge obtained through previous exploration tasks (in different environments) to predict which observation points might close the loop with information-driven exploration to more map the environment more efficiently.

The aforementioned exploration strategies aim at building a map of the environment in the initial stage of the robot deployment, but fail at maintaining it over time, ignoring the changes in the environment. Thus the model accuracy will decrease as the environment changes, which would eventually lead to major localization and navigation failures.

Other strategies could include intrinsic motivation systems, which drive the robot towards situations that maximize the performance of the learning process [18, 19]. These strategies are able to actively identify anomalous or novel situations that might lead to decisions that provide more information and allows to deal with situations where the information-gain never decreases due to physical constraints. For example, novelty detection strategies, which involve the recognition of environmental stimuli that differ from those usually seen, allow the robot to gradually redirect its attention according to the evolution of its internal models [20].

2.2. *Dynamic environment representations*

Once robots have attained the ability to operate for longer periods of time, the effects of the environment changes have to be taken into account. The first approaches were aimed at short-term dynamics. These methods identify

dynamic objects and remove them from the environment representations [21, 22] or use them as moving landmarks [23] for self-localization. However, some dynamic objects do not move at the time of mapping and, consequently, the robot needs further observations to identify them. Ambrus et al. [24] propose to process several 3d point clouds of the same environment obtained over a period of several weeks to separate movable objects and refine the model of static environment structure at the same time.

Other approaches do not explicitly segment movable objects, but use representations that are able to model large-scale, substantial environment changes over long time periods. Some authors [25, 26] represent the environment dynamics by multiple temporal models with different timescales, and Dayoub and Duckett [27] use a ranking scheme that allows to identify environmental features that are more likely to be stable in long-term. Churchill and Newman [4] propose to cluster similar observations at the same spatial locations to form ‘experiences’ which are then associated with a given place and show that this approach improves autonomous vehicle localization. Tipaldi et al. [6] represent the states of the environment components (cells of an occupancy grid) with a hidden Markov model and show that their representation also improves localization. In [28], each cell in the occupancy grid stores not only the probability of it being occupied, but also the likelihood of the cell to change after a given time. Kucner et al. [29] propose a method that learns conditional probabilities of neighbouring cells of an occupancy grid to model typical motion patterns in dynamic environments. Neubert et al. [7] proposed a method that can learn appearance changes based on a long-term dataset collected across multiple seasons and use the learned model to predict the environment appearance for a given time. Another approach that possesses the ability to predict environment changes is proposed by Rosen et al. [30], which uses Bayesian-based survivability analysis to predict which environment features will still be visible after some time and which features will disappear.

Another family of algorithms aims at creating models of the environment that allow them to predict where and when to make observations of specific phenomena within the environment. Typically, these algorithms rely on Gaussian Processes [31, 32, 33], which allow the robot to learn patterns in the environment. Even though these approaches are able to build models of given phenomena, these models are not used by the robot itself to improve essential competences such as localization.

Finally, Krajník et al. [34] propose to represent the environment dynamics

in the spectral domain and apply this approach to image features to improve localization [5], to occupancy grids to reduce memory requirements [35], and to topological maps to improve both path planning [36] and robotic search [37]. While being applicable to most environment models used in mobile robotics, the aforementioned method suffers from a major drawback due to its reliance on the traditional Fast Fourier Transform (FFT) method, which requires the environment observations to be taken on a regular and frequent basis. This means that the robot’s activity has to be divided into a learning phase, when it would frequently visit individual locations to build its dynamic environment model, and a deployment phase when it would use its model to perform useful tasks. This division means that while the robot can create dynamic models, which are more suitable for long-term autonomy, it cannot maintain them during subsequent operation. Thus, the robot does not adapt to dynamics that were not present during the learning phase. This fundamental limitation is addressed by the incremental update scheme introduced in this paper.

2.3. Exploration vs Exploitation

The long-term deployment of mobile robots in human-populated environments must take into account the need to balance exploitation of what the robot already knows and exploration to select better actions in the future [38]. This issue is addressed in [39], which describes the deployment of a mobile robot in a care centre. Several tasks need to be performed by the robot but there is one that directly addresses the exploration/exploitation dilemma. Here, the mobile robot has to act as an information terminal providing information services to visitors. This task is scheduled at different locations in order to increase the number of interactions. However, the scheduler must address two different objectives: exploration and exploitation. The first one creates and maintains a spatio-temporal model of the interactions, providing interaction likelihoods for the different locations and times. The second one aims at visiting the different locations at times where the likelihood of observing interactions is uncertain. Based on the above work, Kulich et al. [40] developed several policies to schedule actions that allow to increase exploitation, or more specifically to increase the number of interactions with humans. In order to increase the interactions, the robot needs to learn human behaviours, more specifically where and when it is more likely for a human to ask for assistance. However, this needs to be achieved in parallel with the human interactions as well as the other daily tasks.

3. Spatio-temporal exploration

The primary purpose of robotic exploration is to autonomously acquire a complete and precise model of the robot’s operational environment. To explore efficiently, the robot has to direct its attention to environment areas that are currently unknown. If the world was static, these areas would simply correspond to previously unvisited locations. In the case of dynamic environments, visiting all locations only once is not enough, because they may change over time. Thus, dynamic exploration requires that the environment locations are revisited and their re-observations are used to update a dynamic environment model. However, revisiting the individual locations with the same frequency and on a regular basis is not efficient because the environment dynamics will, in general, not be homogeneous (i.e. certain areas change more often and the changes occur only at certain times). Similarly to the static environment exploration problem, the robot should revisit only the areas whose states are unknown at the time of the planned visits. Thus, the robot has to use its environment model to predict the uncertainty of the individual locations over time and use these predictions to plan observations that improve its knowledge about the world’s dynamics.

To tackle the problem of predicting environment uncertainty over time, we propose to model the probabilities and entropies of the environment states as functions of time. While the main idea is still that some of the environment’s mid- to long-term dynamics are periodic [34], the underlying mathematical representation had to be reformulated. Unlike the method in [34] that requires frequent and regular environment observations, the method proposed in this paper allows to incrementally and continuously update the spatio-temporal model from sparse observations taken at different locations and times. This eliminates the need for a separate training and deployment phase, and allows integration of spatio-temporal exploration into the robot’s daily routine. Thus, the robot can continuously refine its internal environment model and improve its efficiency from the experience gathered over long periods of time.

3.1. Problem definition

Let us represent the environment as a set \mathcal{S} of n discrete non-stationary independent binary states $s_i(t)$ that are observable by a mobile robot through its sensors. The states $s_i(t)$ might represent the occupancy of individual cells in an occupancy grid, the traversability of edges in a topological map, the

visibility of environmental features, etc. Since these states are dynamic and the robot cannot observe all the states all the time, it maintains an internal environment model that we denote as a set \mathcal{S}' , where each element $s'_i(t)$ corresponds to the real-world state $s_i(t)$. To represent the fact that the currently unobserved states are uncertain, we associate each state with a probability value $p_i(t)$ such that $p_i(t) = P(s_i(t) = 1)$. We refer to the probability function $p_i(t)$ and the way it is calculated from the past observations of $s_i(t)$ as a **temporal model**.

Let us define a **location** as a set of environment states that can be observed simultaneously, i.e. a location \mathcal{L}_j is a subset of \mathcal{S} such that by visiting location \mathcal{L}_j at time t , observations of the states that belong to \mathcal{L}_j are obtained. Given that the robot location at time t is $l(t)$, the robot can directly observe only the states s_i of location $\mathcal{L}_{l(t)}$ and states observable at other locations have to be estimated. Thus, the states of the robot’s internal environment model are

$$s'_i(t) = \begin{cases} s_i(t) & \text{if } s_i \in \mathcal{L}_{l(t)} \\ p_i(t) \geq 0.5 & \text{otherwise.} \end{cases} \quad (1)$$

The purpose of the exploration process is to obtain and maintain as faithful an environment model as possible, i.e. to minimize the difference between the states of the real environment \mathcal{S} and its model \mathcal{S}' . Technically, this corresponds to minimization of the model error $\epsilon(T)$ calculated as the difference between the real and estimated states over the time period $[0, T)$ as

$$\epsilon(T) = \frac{1}{T} \sum_{t=0}^{T-1} \sum_{i=1}^n |s'_i(t) - s_i(t)|. \quad (2)$$

Although the reduction of the error $\epsilon(T)$ can be partially achieved by visiting the relevant locations as often as possible, the robot has to perform other tasks and the number of observations is typically limited. Thus, the robot has to carefully plan where and when to perform observations so that it obtains the relevant data to create, maintain and refine its spatio-temporal models of the environment. From a technical point of view, the robot has to use its internal temporal models $p_i(t)$ to determine a sequence of locations $l(t)$. We refer to the way the robot plans the sequence of $l(t)$ from the $p_i(t)$ as its **exploration strategy**.

4. Spatio-Temporal models

The underlying spatial environment representations that we will use to test our approach are occupancy grids, topological and feature-based maps. The elementary states of these models represent the occupancy of individual cells, the presence of people at the particular areas and the visibility of image features. Unlike classic environment models that represent the probabilities of the elementary states $s(t)$ by constant values p , we represent the probability of each elementary state as a function of time $p(t)$. In particular, we model each $p(t)$ as a combination of harmonic functions that correspond to hidden periodic processes in the environment.

4.1. Spectral maps

The idea of identifying periodic patterns in the measured states and using them for future predictions was originally presented in [34]. These methods process the sequence of the measured state $s(t)$ by the Fast Fourier Transform (FFT) to obtain the corresponding frequency spectrum $s(\omega)$ and extract its most prominent spectral components $s'(\omega)$. Then, they employ the Inverse Fast Fourier Transform (IFFT) to recover the sequence of state probabilities $p(t)$, which can be used for anomaly detection [34] or state prediction [5]. However, the reliance of these methods on the Fast Fourier Transform (FFT) algorithm makes their real-world application impractical. First, the FFT can transform only the complete sequence of a state $s(t)$ or its full spectral representation $s(\omega)$. Thus, updating the spectral representation with new measurements or prediction of a single probability requires to recalculate the entire sequence of observations, which becomes computationally expensive as the observations accumulate. Most importantly, the FFT algorithm requires that the environment observations are sampled at regular intervals, which imposes an inefficient exploration scheme and goes against the concept of spatio-temporal exploration that aims at deciding when and where to observe the environment in a non-regular way.

4.1.1. Frequency map enhancement (*FreME*n)

Similarly to the aforementioned spectral representation [34], our method still aims to identify the periodic patterns of the environment states and use them for predictions. Unlike the previous representation in [34], the method proposed here allows to update the underlying dynamic models incrementally from sparse, irregular observations. The proposed method represents each

state by the number of performed measurements n , its mean probability μ , and two sets \mathcal{A} , \mathcal{B} of complex numbers α_k and β_k that correspond to the set Ω of periodicities ω_k that might be present in the modelled environment. The set Ω was chosen to cover periodicities ranging from 4 weeks to 2 hours with distribution similar to the traditional FFT, i.e. $\omega_k = \frac{2\pi k}{4 \times 7 \times 24 \times 3600}$, where $k \in \{1, 2, \dots, 4 \times 7 \times 12\}$. Initially, the mean value μ is set to 0.5 and all α_k, β_k are set to 0, which corresponds to a completely unknown state.

4.1.2. Addition of a new measurement

Each time a state $s(t)$ is observed at time t , we update its representation, i.e. the number of measurements n , the mean μ and values of \mathcal{A} , \mathcal{B} , which are actually a sparse spectral representation of $s(t)$, as follows:

$$\begin{aligned} \mu &\leftarrow \frac{1}{n+1} (n\mu + s(t)), \\ \alpha_k &\leftarrow \frac{1}{n+1} (n\alpha_k + s(t)e^{-jt\omega_k}) \quad \forall \omega_k \in \Omega, \\ \beta_k &\leftarrow \frac{1}{n+1} (n\beta_k + \mu e^{-jt\omega_k}) \quad \forall \omega_k \in \Omega, \\ n &\leftarrow n + 1. \end{aligned} \tag{3}$$

The proposed update step is analogous to incremental averaging – the absolute values of $|\alpha_k - \beta_k|$ correspond to the average influence of a periodic process (with a frequency of ω_k) on the values of $s(t)$. Note that the size of the representation of the state (i.e. the number of elements in \mathcal{A} , \mathcal{B}) is independent of the number of observations, which means that the memory requirements of the proposed representation do not grow with time. Note also that if the times of observations t are equally spaced and the frequencies ω_k are selected as described in Section 4.1.1, then (3) corresponds closely to the traditional Discrete Fourier Transform.

4.1.3. Performing predictions

To predict the value of state $s(t)$ for a future time t , we first create a set \mathcal{C} consisting of $\gamma_k = \alpha_k - \beta_k$ and then sort it descendingly according to the absolute values $|\gamma_k|$. Then, we extract the first m elements γ_l along with their corresponding frequencies ω_l and calculate the state's probability over time as

$$p(t) = \varsigma\left(\mu + \sum_{l=1}^m 2|\gamma_l| \cos(\omega_l t + \arg(\gamma_l))\right), \tag{4}$$

where $\varsigma(\cdot)$ ensures that $p(t) \in [0, 1]$. The choice of m determines how many periodic processes are considered for prediction. Setting m too low would

mean that we might omit some environment processes that actually influence the state, while setting m too high might include components of \mathcal{C} that are caused by sensor noise. To estimate the optimal value of m , we compare the predictions performed by (4) to the measured values by means of (2), and select the value of m that minimizes the prediction error ϵ . This choice of m is performed automatically during the robot operation – initially, m equals 0 and is increased only after the robot obtains enough data to verify the prediction accuracy of its spatio-temporal models. Since the most prominent periodicities in human-populated environments are related to the day/night cycle, the value of m typically equals to zero for the first two days of exploration, because inferring a day-long periodicity requires two days of data gathering – one day to build the model and one day to verify it.

One of the main advantages of the proposed representation is that the state is modelled probabilistically. This allows to calculate the time intervals when the particular states are uncertain, which is crucial to direct the robot’s attention during exploration.

4.2. Short-term memory

We propose to model the short-term dynamics using a similar model to [28]. This model is based on a Markov chain and aims not only at representing the environment states but also how likely they are to change given the last observed state and the time it was observed. Assuming that each measured state s can be occupied or free, the goal of this method is to estimate the conditional probabilities that represent the transition from one state to another, which are $p(s = 0|s = 1)$ and $p(s = 1|s = 0)$. These probabilities are estimated by means of a Poisson process, i.e., these probabilities can be approximated by the ratio between the number of state changes observed and the total number of observations. However, as described in Section 3, due to the nature of spatio-temporal exploration the observations of states are not performed uniformly in time, and consequently the discrete Markov chain described in [28] as well as the estimation of the aforementioned probabilities do not apply in our case. Thus, we propose a continuous Markov chain to model the recency of the environment states, as shown in Figure 1. In this case, the transition rates between the states 0 and 1, α and β , are inversely proportional to the average time that an observed state remains at 0 or 1. From the Markov chain shown in Figure 1, we infer the equations

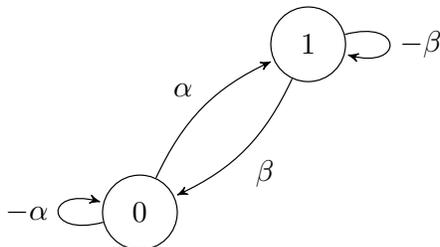


Figure 1: The underlying Markov chain in the short-term memory model.

$$\begin{aligned} \dot{p}_0(t) &= -\alpha p_0(t) + \beta p_1(t), \\ \dot{p}_1(t) &= -\beta p_1(t) + \alpha p_0(t). \end{aligned} \tag{5}$$

Since we only have two states for any time t , we have $p_0(t) + p_1(t) = 1$. Thus, by differentiating and substituting the previous set of equations we obtain Equation 6, which allows us to predict the probability of the state $s(t)$ for a given future time t , where T is the time of the most recent observation.

$$p(t) = \frac{\alpha}{\alpha + \beta} + \left(p(T) - \frac{\alpha}{\alpha + \beta}\right)e^{-(\alpha + \beta)(t - T)}. \tag{6}$$

4.3. Alternative temporal models

The most popular way to deal with the uncertainty of the environment is based on Bayesian filtering, which updates the state estimates based on the sensor noise characteristics. The typical measurement rate of the robot sensors exceeds the mid- to long-term environment dynamics, therefore the Bayesian update scheme causes the probabilities of the observed states to quickly converge towards the latest observed values. Typically, the traditional environment representations tend to reflect the latest state measurements, discarding older measurements. However, for long-term deployment it is sensible to use representations that somehow reflect the prior environment states since the initial deployment stage. To strengthen our study, we describe in this section two additional environment representations that take into account all the previous observations, a **long-term memory** model and **Gaussian Mixture Models** (GMM).

4.3.1. Long-term memory

A way to reflect the uncertainty of the observed states in the long-term is to implement a **long-term memory** (LM). The model that we propose works as a memory that takes into account all the observations and calculates the probability of a given state simply as the arithmetic mean of all its past observations.

4.3.2. Gaussian Mixture Models

Gaussian Mixture Models that can approximate multi-dimensional functions as a weighted sum of Gaussian component densities are a well-established method of function approximation. A Gaussian Mixture Model of a function $f(t)$ is a weighted sum of m Gaussian functions:

$$f(t) = \frac{1}{\sqrt{2\pi}} \sum_{j=1}^m \frac{w_j}{\sigma_j} e^{-\frac{(t-\mu_j)^2}{2\sigma_j^2}}. \quad (7)$$

GMMs find their applications in numerous fields ranging from botany to psychology [41]. The parameters of individual components of GMMs, i.e. the weights w_k , means μ_j and variances σ_j are typically estimated from training data using the iterative Expectation Maximization (EM) or Maximum A-Posteriori (MAP) algorithms. While GMMs can model arbitrarily-shaped functions, their limitation rests in the fact that they cannot naturally represent functions that are periodic.

To deal with this issue, we simply assume that people perform most of their activities on a daily basis and thus we consider the object presence in individual areas as being the same for every day. While this assumption is not entirely correct (as working days will typically be different from weekends), such a temporal model might still be better than a ‘static’ model where the probability of object presence is a constant.

Prior knowledge of the periodicity allows to transform the measured sequence of states $s(t)$ into a sequence $p'(t)$ by

$$p'(t) = \frac{k}{\tau} \sum_{i=1}^{k/\tau} s(t + i\tau), \quad (8)$$

where τ is the assumed period and k is the $s(t)$ sequence length. After calculating $p'(t)$, we employ the Expectation Maximization algorithm to find

the means μ_j , variances σ_j and weights w_j of its Gaussian Mixture approximation. Thus, the probability of occupancy of a room at time t is given by

$$p(t) = \frac{1}{\sqrt{2\pi}} \sum_{j=1}^m \frac{w_j}{\sigma_j} e^{-\frac{(\text{mod}(t,\tau)-\mu_j)^2}{2\sigma_j^2}}, \quad (9)$$

where τ is the a priori known period of the function $p(t)$ and mod is a modulo operator. The periodic-GMM-based (PerGaM) model is complementary to the FFT-based one. It can approximate even short, multiple events, but it can represent only one period (τ) that has to be known a priori. Since the dominating periodicity of human populated environments is 1 day, we chose $\tau = 86400\text{s}$.

5. Exploration strategies

As noted in Section 3.1, an exploration strategy is defined as a process that determines both which locations to visit and when to visit them. One has to assume that a real mobile robot has to perform other tasks as well and can spend only a fraction of the total time on actual exploration. We refer to this fraction as the exploration ratio e , e.g. $e = 0.2$ means that the robot can spend 20% of its operational time on exploration.

Thus, given an exploration ratio e and a set \mathcal{T} of time intervals $[t_s, t_{s+1})$, the exploration algorithm has to determine a sequence $l(t_s)$ of locations to visit. To represent situations where the time slot $[t_s, t_{s+1})$ is allocated to an unrelated activity, the value of $l(t_s)$ is set to zero, whereas a non-zero value of $l(t_s)$ signifies the location to be observed during $[t_s, t_{s+1})$.

5.1. Information-gain strategies

The information-gain strategies take into account the experiences the robot has gathered so far to plan when and which location to visit. These strategies attempt to reduce the uncertainty of the environment models by planning the observations that maximize the potential information gain. To estimate how much information is gained by a particular observation, we will use the notion of entropy. We assume that direct observation of particular states at a given time reduces the entropy of these states to zero. Thus, the information gained by a particular observation can be estimated as the sum

of the entropies of the states observed at a given location as

$$I(\mathcal{L}, t) = - \sum_{i \in \mathcal{L}} (p_i(t) \ln(p_i(t)) + (1 - p_i(t)) \ln(1 - p_i(t))). \quad (10)$$

The **Greedy** strategy calculates the potential information gains for all given time slots and locations, then assigns the best location to visit at each time slot. Then, it selects a subset \mathcal{T}' of time slots with the highest information gain such that $e = |\mathcal{T}'|/|\mathcal{T}|$. The remaining time slots are assigned to other tasks. Thus, this strategy maximizes the potential information gain obtained over the time slots in the set \mathcal{T} .

The **Monte Carlo** strategy chooses the locations randomly, but the probability of selecting a given location at a given time is proportional to the estimated information gain. At first, it estimates the $I(l, t_s)$ for all given time slots and locations and sums these values to I' . Then, it calculates the value of $I(0, t_s) = I'(1 - e)/(ne)$. Finally, it calculates the probabilities of each $l(t_s)$ as

$$P(l(t_s) = j) = \frac{I(j, t_s) + \iota}{\sum_{i \in \mathcal{L}} I(i, t_s) + \iota}. \quad (11)$$

Here, the value of $I(0, t_s)$ does not represent actual information gain, but is added to ensure that the exploration ratio e is satisfied by ensuring sufficient chance of assigning the time slots to exploration-unrelated tasks. The positive constant ι ensures that the locations will be occasionally visited even at times when the spatio-temporal model predicts their state with absolute certainty. This allows the robot to detect unexpected changes in the environment dynamics.

The **Novelty-driven** strategy follows the same principle as the Monte Carlo one. However, unlike the Monte-Carlo strategy, which strictly follows a schedule determined by Equation 11, the novelty-driven strategy uses a combination of temporal models to identify situations where a change in the Monte-Carlo schedule would result in a high amount of information obtained. To identify such situations, the novelty-driven strategy predicts the amount of information obtainable in the following time slot by:

$$I(t) = -p_{expc}(t) \ln(p_{info}(t)) - (1 - p_{expc}(t)) \ln(1 - p_{info}(t)), \quad (12)$$

where $p_{expc}(t)$ is calculated by the short-term memory model (see Section 4.2) and serves as a measure of expectation, whereas p_{info} is provided by another model and represents the amount of information expected. If $I'(\mathcal{L}, t) \gg$

$I(\mathcal{L}, t)$, i.e. the amount of information predicted by Equation 12 is significantly higher than the value calculated by Equation 10, then the location to visit in the following time-slot is changed accordingly. Thus, if the observed states at a recently visited location did not match their predictions, the robot re-observes the location again to obtain more information about this unexpected event.

5.2. Uninformed strategies

For comparison purposes, we include strategies which select the places to visit regardless of the environment dynamics. These strategies calculate the sequence of visits $l(t_s)$ simply from the values of the ratio e , number of locations n and number of time slots m .

The **Round-Robin** strategy visits all areas of the environment with the same frequency, interleaving the observations with other tasks so that the exploration ratio e is satisfied.

The **Random** strategy also attempts to visit all areas with the same frequency, but the sequence of $l(t_s)$ is not deterministic, but random. The probability of a given slot being assigned to a non-exploration task is equal to $1 - e$ and the probability of visiting the individual locations is uniform and equal to e/n .

6. Evaluation datasets

To evaluate the ability of the various temporal models and exploration strategies, we performed a comparison on two datasets gathered over several weeks. The first, ‘Aruba’ dataset was gathered by a team of the Center for Advanced Studies in Adaptive Systems (CASAS) to support their research concerning smart environments [42]. The second, ‘Brayford’ dataset was created at the Lincoln Centre for Autonomous System Research (LCAS) for their research on long-term mobile robot autonomy [5]. The aforementioned datasets were processed so that the dynamics of these environments are represented as visual-feature-based, topological and metric maps.

6.1. The Aruba dataset

The ‘Aruba’ dataset consists of maps capturing 16 week long dynamics of a large apartment that was occupied by a single, house-bound person who occasionally received visitors. An occupancy grid and a topological map were created for every minute of a 16 week long period – the resulting

dataset contains over 160 000 metric and topological maps. Since the original dataset [42] is simply a year-long collection of measurements from 50 different sensors spread over an eight-room apartment, these maps had to be created by means of simulation.

First, we processed the events from the original dataset’s motion detectors to establish the location of people in the flat for every minute of the 16 weeks. Then, we partitioned the flat into ten different areas, where eight areas represent the rooms and two correspond to corridors. This allowed us to create a topological map that indicates the presence of people in these locations. To obtain the metric representation, we created a simulated envi-



Figure 2: Aruba environment simulation.

ronment with the same structure as the ‘CASAS’ apartment, see Figure 2. Then the simulation was provided with a sequence of person locations recovered in the previous step. As a result, the simulated environment contains physical models of people at locations provided by the real-world dataset, and thus it reflects the dynamics of the real apartment. A virtual, RGB-D camera equipped robot was also introduced into the virtual environment. Every time the configuration of the simulated environment (i.e. locations of the people) changed, the robot used its 3D sensors to create occupancy grids of the flat’s individual rooms. Thus, we obtained occupancy grids that reflect the real environment dynamics minute-by-minute for 16 weeks.

6.2. The Brayford dataset

The Brayford dataset was originally collected for the purpose of benchmarking long-term mobile robot localization algorithms in dynamic environments [5]. The data collection was performed by a human-sized robot equipped with an RGB-D camera in a large, open-space office of the Lincoln Centre for Autonomous Systems. The robot was set up to obtain RGB-D images of eight designated areas every 10 minutes for a period of one week. Representative examples of the captured images are shown in Figure 3. While

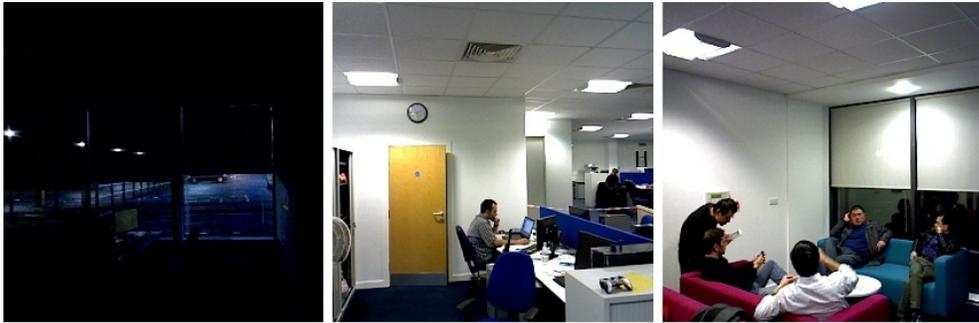


Figure 3: Examples of Brayford dataset images.

the high-level environment model of this dataset contains information about people presence at the individual locations, the states of the low-level model represent the visibilities of image features [43] established by the method presented in our earlier work on visual localisation in changing environments [5]. The resulting dataset contains more than 8000 feature-based and 8000 semantic maps collected over a period of one week.

6.3. Dataset summary

Both datasets contain a high-level model representing people presence at different locations and a low-level model based on RGB-D sensing. In the following sections, we will refer to these models as ‘symbolic’ and ‘metric’, respectively. The aforementioned datasets are available as a part of the long-term dataset collection [44].

7. Experimental results

We assume that the aforementioned datasets reflect the real state of the environments they have been captured in and thus we use the sequence of the

observations in the datasets as ground truth. To evaluate how the various temporal models and exploration strategies affect the robot’s ability to create and update its internal environment models, we emulate the exploration process using the datasets gathered. We assume that exploration can be performed during only half of the robot’s operational time (i.e. $e = 0.5$) and that a single observation takes 10 minutes. While 10 minutes might seem like a long time, creation of a 3D occupancy grid of a given location means that the robot has to position itself precisely, and capture and process approximately 50 RGB-D images from different viewpoints. This time also includes navigation to the given spot, leaving the charging station, etc.

This exploration procedure corresponds to the situation when the robot updates its spatio-temporal model and generates a new observation schedule every 24 hours at midnight. The robot starts with an empty environment model that has all probabilities constant and equal to 0.5.

First, the entropy functions of the individual locations are calculated and 72 observations for the following day are scheduled. Then, these 72 observations are retrieved from the given dataset and the temporal models of the environment states are updated. The updated temporal models are used to recalculate the spatio-temporal entropy and the next day’s observation schedule is then generated. These steps are repeated for every day of the given dataset.

7.1. Evaluating environment model error

To compare the performance of the temporal models and exploration strategies described in Sections 4 and 5, the resulting world model is compared to the actual dataset using Equation 2, which estimates the error in the environment model. Since there are 4 temporal models and 5 exploration strategies, each comparison considers 20 values that characterize the ratio of incorrectly estimated states to the total number of environment states. One dataset evaluation consist of two comparisons, each corresponding to the given environment representation.

The results of the ‘Aruba’ dataset summarized in Table 1 show that the combination of FreMEN with the novelty-driven or Monte-Carlo strategies reduces the model error by more than 40%. Nevertheless, the combination of FreMEN and the novelty-driven strategy performs slightly better than the combination of the same model with the Monte-Carlo one. One may think that the greedy strategy would be the best performer since it always chooses the room with higher entropy, but in most situations this strategy

Table 1: Aruba dataset results: Model errors for different exploration strategies and spatio-temporal models [%]

<i>Strategy</i>	<i>Spatio-Temporal model</i>							
	<i>Symbolic</i>				<i>Metric</i>			
	<i>SM</i>	<i>LM</i>	<i>FT</i>	<i>GM</i>	<i>SM</i>	<i>LM</i>	<i>FT</i>	<i>GM</i>
Round Robin	09.3	09.7	06.5	07.5	08.9	09.3	05.6	05.8
Random	08.9	09.5	09.2	07.5	08.7	09.0	08.3	07.2
Greedy	08.5	08.7	07.0	09.4	07.7	10.9	06.2	07.1
Monte-Carlo	08.5	08.9	05.8	06.4	08.0	08.3	05.0	05.7
Novelty-driven	08.5	08.9	05.7	06.1	08.0	08.4	04.9	05.4

fails to maintain an up-to-date model. For example, in the case of noisy and unpredictable signals in a given room, the robot will attempt to focus its attention mainly in that room. While this is a logical behaviour – not being able to model the location, the robot will gather the data about it through direct observation, it might not be really desirable, because the robot might not be getting valuable data at all. Also, this behaviour would mean that the robot would not observe the remaining rooms, since the entropy of the current room is higher due to the higher uncertainties.

Figure 4 shows that the FreMEn model error is lower during the first day, showing that this strategy allows quicker identification of the environment patterns. Since more than 99% of the cells in the ‘Aruba’ occupancy grids represent empty space or static objects, the model error (Equation 2) is calculated for the cells that change their occupancy at least once.

Table 2: Brayford dataset results: Model errors for different exploration strategies and spatio-temporal models [%]

<i>Strategy</i>	<i>Spatio-Temporal model</i>							
	<i>People Presence</i>				<i>Visual Features</i>			
	<i>SM</i>	<i>LM</i>	<i>FT</i>	<i>GM</i>	<i>SM</i>	<i>LM</i>	<i>FT</i>	<i>GM</i>
Round Robin	23.7	23.7	16.3	20.2	25.7	27.0	12.7	17.9
Random	23.7	23.8	23.0	23.8	25.9	27.0	25.2	20.3
Greedy	20.2	22.3	19.2	20.1	29.9	29.3	24.4	18.6
Monte-Carlo	23.5	23.5	16.4	19.3	25.6	27.0	12.3	16.9
Novelty-driven	23.4	23.5	15.2	19.4	25.6	27.0	12.1	17.1

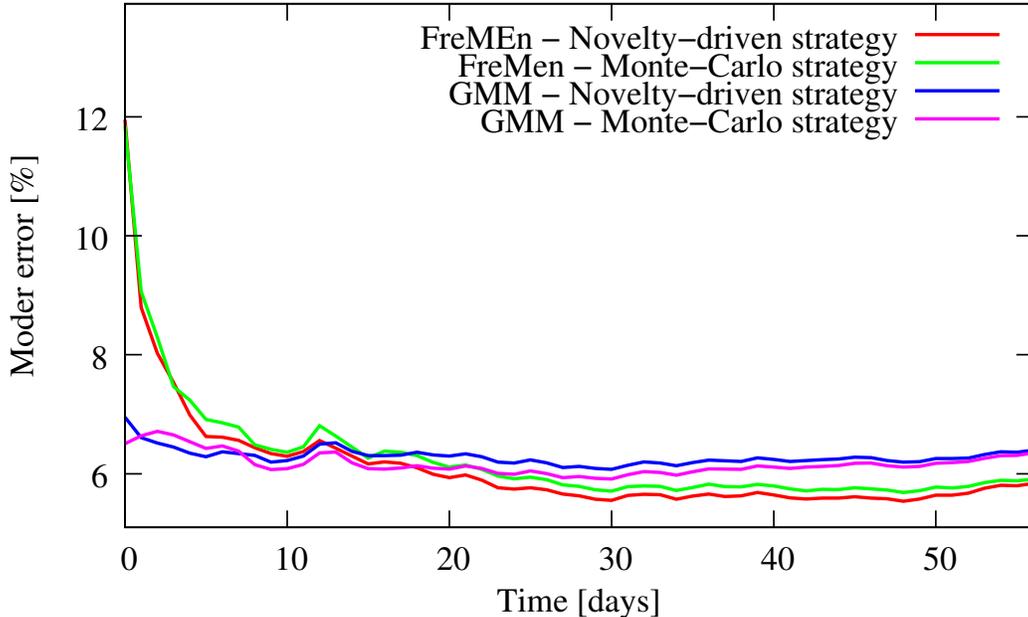


Figure 4: Comparison of the average error of the novelty-driven and Monte Carlo exploration strategies. The remaining models are not displayed due to their similar performance.

The model errors of the ‘Brayford’ dataset as shown in Table 2 again indicate that the most faithful environment representation is based on frequency-enhanced temporal models (see Section 4.1.1) in combination with the novelty-driven strategy. The improvement is more prominent in the case of people presence models. The reason for this might be that the visibility of image features tends to follow regular patterns given by the daily illumination cycle, whereas the presence of people can be influenced by unexpected events. Note that the model errors of the feature-based maps are higher than the ones reported in [8] because we used a higher number of visual features in our model.

Figure 4 shows that initially, the GMM model achieves the lowest error, but in the long-term, it is outperformed by FreMEEn. This is caused by the fact that the GMM model is tailored to represent daily periodicities, while the FreMEEn model has to identify the patterns of changes from the data by itself. After several days, FreMEEn identifies several important periodicities (not only the daily one) and its prediction capability improves, allowing it to better schedule observations and decrease the model error. Figure 4 also

shows that the novelty-driven strategy performs slightly, but consistently better than the Monte-Carlo one. In the experiments performed, we observe that the novelty-driven strategy identifies one or two unexpected observations per day.

7.2. Exploration vs. Exploitation

In the above experiments, the robot’s exploration ratio e was set to 0.5. Thus, the robot could spend 50% of its time gathering data about its operational environment. However, such a ratio is unrealistic – the robot has to spend some time replenishing its batteries and we have to assume that it should perform other tasks as well depending on the application. Moreover, we have to assume that the purpose of the robot is not in creating precise environment models, but to perform useful tasks. Thus, exploration is just an instrument to obtain and maintain knowledge to improve the robot’s performance. If the robot spends too much time on exploration, it would not be able to exploit the obtained knowledge in its everyday activities.

We evaluate the efficiency of the individual exploration strategies with different exploration ratios for predicting person presence on the Aruba dataset. We combine the Frequency Map Enhancement models with four different exploration strategies, fix the exploration ratio to a value between 0 and 1, and let the robot explore the Aruba environment for two consecutive weeks. The resulting error of the model obtained is shown in Figure 5. The results indicate that if the fraction of the time that the robot can spend on actual exploration is low, the dynamic models might make wrong assumptions about the environment changes and perform worse than their static counterparts – this is especially notable with the Greedy and Round Robin strategies. However, this effect can be mitigated by a proper exploration strategy – the graph shows that both Monte Carlo and novelty-based strategies improve the model even if the robot cannot spend too much time on exploration.

Note that the initial model error is 10% – this is caused by the fact that the Aruba dataset represents the presence of people in 10 different areas and the flat has only one inhabitant. Without any observations, the robot simply assumes that the flat is empty, which results in 10% error.

7.3. Qualitative evaluation

To gain an insight into the robot’s exploratory behaviour, we interpret the data gathered during the exploration of the ‘Aruba’ topological map. Here, the robot’s task was to create a spatio-temporal model of person presence in

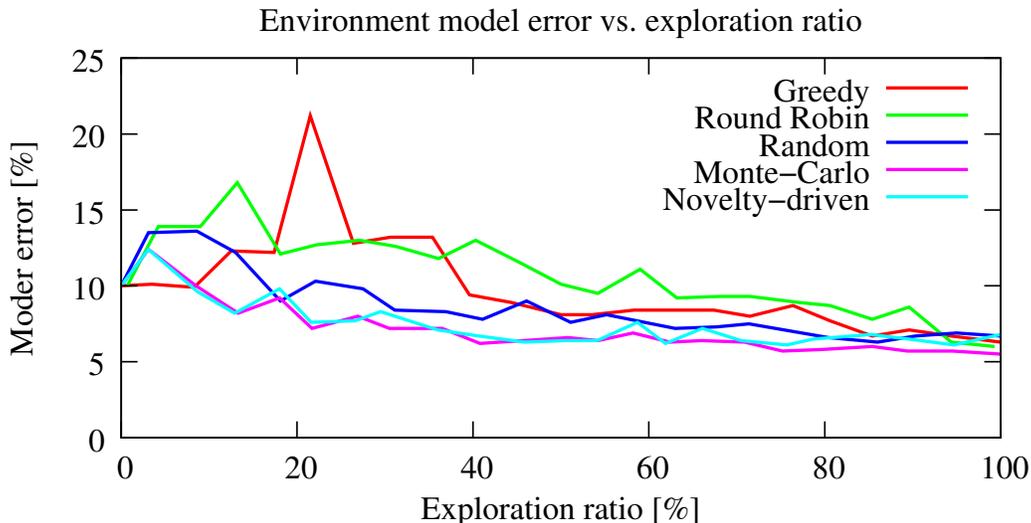


Figure 5: Exploration vs. exploitation analysis: The influence of the fraction of time spend with exploration on the performance of the exploration strategies.

the individual rooms of a small apartment. For the purpose of this explanation, let us focus on the dynamics of three rooms only – the bedroom, the kitchen and a storage room. Let the robot use the best-performing exploration method that combines the FreMEn temporal models and the Monte Carlo exploration strategy. Applying the proposed spatio-temporal exploration method to this dataset produced the behaviour in Figure 6. The top part of Figure 6 shows the real state of the environment, where the three binary functions $s_i(t)$ represent the room’s occupancies over time. The second part shows the robot’s internal model of the environment, i.e. the probabilities $p_i(t)$. The third graph displays the information that is expected to be obtained by visiting these three locations at a given time. Finally, the bottom graph shows which locations have been visited at a particular time – we assume that the exploration ratio $e = 0.5$, which reflects the situation where the robot has to spend half of its time on its charging station. Now let us explain how the robot’s understanding of the environment changes over time and how this affects its exploratory behaviour day by day.

7.3.1. Day one

Initially, the robot has no knowledge of the environment and therefore the probabilities $p_i(t)$ of the world states $s(t)$ are equal to 0.5. This means that the expected information gain from visiting any of the rooms equals 1 bit at

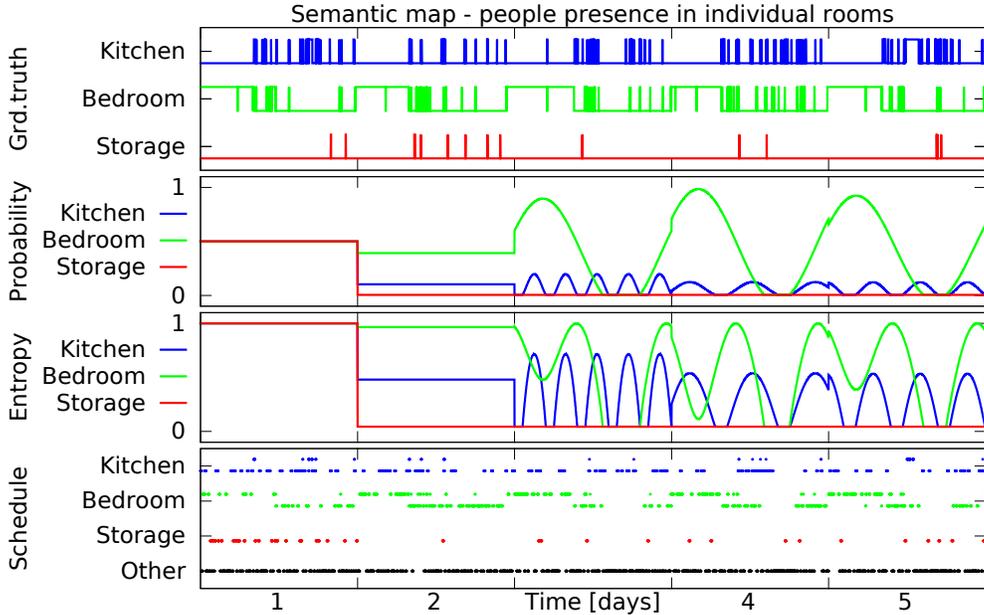


Figure 6: Spatio-temporal exploration behaviour: The robot uses its probabilistic world model (second row) and spatio-temporal entropy estimates (third row) to schedule its observations (bottom graph) and learn the environment dynamics (top). As the environment knowledge improves over time, the scheduled observations provide more information which allows for further refinement of the environment model.

any time of the first day. Thus, the robot has no room or time preference when scheduling the first day’s observations.

7.3.2. Day two

After performing the first day’s observations, the environment models provide enough evidence that the three rooms are not occupied with the same probability. This is reflected in the second day’s environment model – see the probability functions $p_i(t)$ of the second day in Figure 6. Thus the robot expects to gain more information by visiting the bedroom and kitchen than by going to the storage room. This is reflected in the second day’s observation schedule – the last row of Figure 6 shows that the first two rooms are visited more often.

7.3.3. Day three

The additional observations obtained during the second day provide information about the rooms’ dynamics: the robot assumes that the bedroom has a daily periodicity and that the kitchen is visited five times per day. This

causes the expected information gain to be time-dependent – the third day of the third row of Figure 6 shows that evening and morning observations of the bedroom provide more information than in the afternoon. This fact is rather intuitive: visiting the room at the time of its state transition allows to refine the room’s state periodicity. Thus, on the third day, the bedroom is visited mostly in the evening and morning, while the afternoon visits are scheduled to the kitchen.

7.3.4. Days four and five

Based on the data gathered during the third day, the robot modifies its hypothesis about the periodicity of activities in the kitchen and assumes that it is visited three times per day. During the following days, the robot tends to visit the kitchen and bedroom more often, and checks the storage room only occasionally. While the kitchen is visited mostly in the early afternoon, the bedroom is visited late evenings and mornings, which allows to refine the robot’s model of the person’s daily habits.

This example indicates that the combination of a probabilistic temporal model with an information-based strategy not only allows the robot to obtain knowledge about the environment dynamics, but the observations are scheduled in a seemingly logical way: at first, all the locations are visited often and with the same frequency. As the spatio-temporal environment model becomes more refined, the robot tends to visit particular locations only at times when their states are uncertain.

8. Conclusion

In this paper, we presented a method for life-long spatio-temporal exploration of dynamic environments. We assume that the robot’s operational environment is subject to perpetual change, which requires a method that can model and predict these variations. The purpose of spatio-temporal exploration is not only to obtain the environment structure and keep it up-to-date with any changes, but also to allow the robot to observe and understand the world dynamics.

We hypothesise that the problem of spatio-temporal exploration can be tackled by combining information-gain-based exploration strategies with probabilistic dynamic environment models. To verify our approach, we compare the performance of five exploration strategies and four temporal models on real-world data gathered over the course of several months. We show that the

combination of spectral-based temporal models with information-gain-based novelty-driven strategies results in an intelligent exploration behaviour that improves as the environment knowledge becomes more refined.

Analysis of the robot behaviour shows that when introduced to a new environment, the robot prefers to explore unknown locations. After it has obtained the spatial models, it starts to revisit these locations in order to learn about their dynamics. Finally, the learned dynamics allow the robot to schedule which locations to visit at which times and adapt this schedule in the case of unexpected observations.

The evaluations performed in this paper involved several assumptions to simplify the problem. The first assumption was that the time the robot spends moving to a particular location is negligible compared to the time it takes to make an observation. The second assumption was that the locations of observations were predefined and that the robot could position itself with perfect accuracy. The third assumption is that the observations are error-free, i.e. there is no noise on the sensory data. While these assumptions were needed for validation purposes in this work due to the known difficulties of ground-truthing when comparing exploration strategies, more recent work has overcome these limitations and achieved full 4D metric-based spatio-temporal exploration [45].

The analysis presented here opens several questions for further investigation, which we would like to address in the future. In particular, we will investigate not only the impact of exploration on the quality of the spatio-temporal models, but its impact on the efficiency of the robot operation over time. We will investigate how much time the robot should spend on exploration (represented by ‘exploration ratio’ e) during the initial stages of deployment, when the environment model is created, and what is the optimal e later on, when the model is just maintained or when the model needs to be re-built due to changes in the environment dynamics. We will also investigate which situations in our datasets influenced the novelty-driven strategy, so that it performed better than the Monte-Carlo strategy.

Acknowledgments

The work has been supported by the EU ICT project 600623 ‘STRANDS’.

- [1] S. Thrun, W. Burgard, D. Fox, Probabilistic Robotics (Intelligent Robotics and Autonomous Agents), The MIT Press, 2005.

- [2] B. Yamauchi, A frontier-based approach for autonomous exploration, in: Proc. of the IEEE Int. Symposium on Computational Intelligence in Robotics and Automation, 1997.
- [3] B. Kuipers, Y.-T. Byun, A robot exploration and mapping strategy based on a semantic hierarchy of spatial representations, *Robotics and autonomous systems* 8 (1) (1991) 47–63.
- [4] W. S. Churchill, P. Newman, Experience-based navigation for long-term localisation, *The Int. Journal of Robotics Research*-doi:10.1177/0278364913499193.
- [5] T. Krajník, et al., Long-term topological localization for service robots in dynamic environments using spectral maps, in: Proc. of Int. Conference on Intelligent Robots and Systems (IROS), 2014.
- [6] G. D. Tipaldi, D. Meyer-Delius, W. Burgard, Lifelong localization in changing environments, *The International Journal of Robotics Research*-doi:10.1177/0278364913502830.
- [7] P. Neubert, N. Sünderhauf, P. Protzel, Superpixel-based appearance change prediction for long-term navigation across seasons, *Robotics and Autonomous Systems* (0). doi:http://dx.doi.org/10.1016/j.robot.2014.08.005.
- [8] T. Krajník, J. Santos, T. Duckett, Life-long spatio-temporal exploration of dynamic environments, in: *Mobile Robots (ECMR), 2015 European Conference on*, 2015, pp. 1–8. doi:10.1109/ECMR.2015.7324052.
- [9] S. Koenig, C. Tovey, W. Halliburton, Greedy mapping of terrain, in: Proc. of Int. Conference on Robotics and Automation (ICRA), 2001.
- [10] D. Holz, N. Basilico, F. Amigoni, S. Behnke, Evaluating the efficiency of frontier-based exploration strategies, *ISR/ROBOTIK 2010*.
- [11] B. Yamauchi, Frontier-based exploration using multiple robots, in: Proc. of the 2nd Int. Conf. on Autonomous agents, 1998.
- [12] V. Caglioti, An entropic criterion for minimum uncertainty sensing in recognition and localization. i. theoretical and conceptual aspects, *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 31 (2) (2001) 187–196. doi:10.1109/3477.915342.

- [13] C. Stachniss, G. Grisetti, W. Burgard, Information gain-based exploration using Rao-Blackwellized particle filters, in: Proc. of Robotics: Science and Systems (RSS), Cambridge, MA, USA, 2005.
- [14] C. Stachniss, W. Burgard, Exploring unknown environments with mobile robots using coverage maps, in: Proceedings of the International Conference on Artificial Intelligence (IJCAI), 2003.
- [15] J. Fentanes, R. F. Alonso, E. Zalama, J. G. García-Bermejo, A new method for efficient three-dimensional reconstruction of outdoor environments using mobile robots, *Journal of Field Robotics*.
- [16] S. Osswald, M. Bennewitz, W. Burgard, C. Stachniss, Speeding-up robot exploration by exploiting background information, *Robotics and Automation Letters*, IEEE PP (99) (2016) 1–1. doi:10.1109/LRA.2016.2520560.
- [17] D. Perea Strom, F. Nenci, C. Stachniss, Predictive exploration considering previously mapped environments, in: *Robotics and Automation (ICRA)*, 2015 IEEE International Conference on, 2015, pp. 2761–2766. doi:10.1109/ICRA.2015.7139574.
- [18] P. Y. Oudeyer, F. Kaplan, V. V. Hafner, Intrinsic motivation systems for autonomous mental development, *IEEE Transactions on Evolutionary Computation* 11 (2) (2007) 265–286. doi:10.1109/TEVC.2006.890271.
- [19] S. Thrun, Exploration in active learning, *Handbook of Brain Science and Neural Networks* (1995) 381–384.
- [20] S. Marsland, Novelty Detection in Learning Systems, *Neural Computing Surveys* 3.
- [21] D. Hähnel, D. Schulz, W. Burgard, Mobile robot mapping in populated environments, *Advanced Robotics*.
- [22] D. Wolf, G. Sukhatme, Mobile robot simultaneous localization and mapping in dynamic environments, *Autonomous Robots*.
- [23] C. C. Wang, et al., Simultaneous localization, mapping and moving object tracking, *International Journal of Robotics Research*.

- [24] R. Ambrus, N. Bore, J. Folkesson, P. Jensfelt, Meta-rooms: Building and maintaining long term spatial models in a dynamic world, in: Proceedings of the International Conference on Intelligent Robots and Systems (IROS), 2014.
- [25] P. Biber, T. Duckett, Dynamic maps for long-term operation of mobile service robots, in: Proc. of Rob.: Science and Systems, 2005.
- [26] D. Arbuckle, A. Howard, M. Mataric, Temporal occupancy grids: a method for classifying the spatio-temporal properties of the environment, in: Proc. of Int. Conference on Intelligent Robots and Systems (IROS), Vol. 1, 2002, pp. 409–414 vol.1. doi:10.1109/IRDS.2002.1041424.
- [27] F. Dayoub, T. Duckett, An adaptive appearance-based map for long-term topological localization of mobile robots, in: Proc. of Int. Conference on Intelligent Robots and Systems (IROS), 2008.
- [28] J. Saarinen, H. Andreasson, A. Lilienthal, Independent markov chain occupancy grid maps for representation of dynamic environment, in: Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on, 2012, pp. 3489–3495. doi:10.1109/IROS.2012.6385629.
- [29] T. Kucner, et al., Conditional transition maps: Learning motion patterns in dynamic environments, in: Proc. of Int. Conf. on Intelligent Robots and Systems (IROS), 2013.
- [30] D. M. Rosen, J. Mason, J. J. Leonard, Towards lifelong feature-based mapping in semi-static environments, in: International Conference on Robotics and Automation (ICRA), IEEE, 2016, to appear.
- [31] A. Singh, F. Ramos, H. D. Whyte, W. J. Kaiser, Modeling and decision making in spatio-temporal processes for environmental surveillance, in: in Proc. IEEE Int. Conf. Robot. Autom, 2010, pp. 5490–5497.
- [32] R. Marchant, F. Ramos, Bayesian optimisation for intelligent environmental monitoring, in: Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on, 2012, pp. 2242–2249. doi:10.1109/IROS.2012.6385653.

- [33] R. Marchant, F. Ramos, Bayesian optimisation for informative continuous path planning, in: Robotics and Automation (ICRA), 2014 IEEE International Conference on, 2014, pp. 6136–6143. doi:10.1109/ICRA.2014.6907763.
- [34] T. Krajník, J. P. Fentanes, G. Cielniak, C. Dondrup, T. Duckett, Spectral analysis for long-term robotic mapping, in: Proc. of Int. Conference on Robotics and Automation (ICRA), 2014.
- [35] T. Krajník, J. Santos, B. Seemann, T. Duckett, Froctomap: An efficient spatio-temporal environment representation, in: Advances in Autonomous Robotics Systems, Springer, 2014, pp. 281–282. doi:10.1007/978-3-319-10401-0.
- [36] J. Pulido Fentanes, et al., Now or later? Predicting and maximising success of navigation actions from long-term experience, in: International Conference on Robotics and Automation (ICRA), 2015.
- [37] T. Krajník, M. Kulich, L. Mudrová, R. Ambrus, T. Duckett, Where’s Waldo at time t? Using spatio-temporal models for mobile robot search, in: Int. Conf. on Robotics and Automation (ICRA), 2015.
- [38] R. S. Sutton, A. G. Barto, Introduction to Reinforcement Learning, MIT Press, 1998.
- [39] N. Hawes, C. Burbridge, F. Jovan, L. Kunze, B. Lacerda, L. Mudrová, J. Young, J. L. Wyatt, D. Hebesberger, T. Körtner, R. Ambrus, N. Bore, J. Folkesson, P. Jensfelt, L. Beyer, A. Hermans, B. Leibe, A. Aldoma, T. Faulhammer, M. Zillich, M. Vincze, M. Al-Omari, E. Chinellato, P. Duckworth, Y. Gatsoulis, D. C. Hogg, A. G. Cohn, C. Dondrup, J. P. Fentanes, T. Krajník, J. M. Santos, T. Duckett, M. Hanheide, The STRANDS project: Long-term autonomy in everyday environments, CoRR abs/1604.04384.
URL <http://arxiv.org/abs/1604.04384>
- [40] M. Kulich, T. Krajník, L. Přeučil, T. Duckett, To explore or to exploit? learning humans’ behaviour to maximize interactions with them, in: Modelling and Simulation for Autonomous Systems Workshop (MESAS), 2016.

- [41] D. M. Titterton, A. F. Smith, U. E. Makov, et al., Statistical analysis of finite mixture distributions, Vol. 7, Wiley New York, 1985.
- [42] D. J. Cook, Learning setting-generalized activity models for smart spaces, IEEE Intelligent Systems (99) (2010) 1.
- [43] M. Calonder, V. Lepetit, C. Strecha, P. Fua, BRIEF: Binary robust independent elementary features, in: Proc. of European Conference on Computer Vision (ECCV), Springer, 2010, pp. 778–792.
- [44] T. Krajník, J. P. Fentanes, J. Santos, C. Dondrup, M. Hanheide, T. Duckett, L-CAS datasets for long-term autonomy of mobile robots, <http://lcas.lincoln.ac.uk/owncloud/shared/datasets/>.
- [45] J. Santos, T. Krajník, J. Pulido Fentanes, T. Duckett, Lifelong information-driven exploration to complete and refine 4d spatio-temporal maps, Robotics and Automation Letters, IEEE PP (99) (2016) 1–1. doi:10.1109/LRA.2016.2516594.