Allocating Limited Sensing Resources to Accurately Map Dynamic Environments

Derek Mitchell[1] and Nathan Michael[1]

***Abstract*— This work addresses the problem of learning a model of a dynamic environment using many independent Hidden Markov Models (HMMs) with a limited number of observations available per iteration. Many techniques exist to model dynamic environments, but do not consider how to deploy robots to build this model. Additionally, there are many techniques for exploring environments that do not consider how to prioritize regions when resources, in terms of robots to deploy and deployment durations, are limited. Here, we consider an environment model consisting of a series of HMMs that evolve over time independently and can be directly observed. At each iteration, we must determine which HMMs to observe in order to maximize the gain in model accuracy. We present a utility measure that balances a Pearson's χ^2 *goodness-of-fit* of the dynamics model with *Mutual Information* (MI) to ensure that observations are allocated to maximize the convergence rate of all HMMs, resulting in a faster convergence to higher steady-state model confidence and accuracy than either χ^2 or MI alone.**

I. Introduction and Related Works

Autonomous exploration is a well studied problem with many applications in domains such as search-and-rescue [1], surveillance [2], and mapping [3]. At its core, the exploration problem involves deploying one or more robots to build an accurate model of an unknown environment. This is of particular importance for disaster response, where a persistent understanding of the environment is necessary to facilitate the distribution of resources and avoid dangerous regions. Flooding is a prime example, as roads needed to transport supplies are initially blocked or destroyed, but can become passable again as the water ebbs and rescue teams effect repairs. Likewise, areas where rubble is regularly shifting or where people are congregating must be continuously monitored to account for future dangers. Efficient operation in such a domain requires a mapping strategy that is robust to changes despite the limited resources available in the field.

There are two main topics of study that are relevant to mapping explicitly or implicitly dynamic environments: methods for modeling environments and methods to plan for and prioritize observations. Approaches such as Gaussian Process representations [4] or Hierarchical Gaussian Mixture Models [5] can generate very accurate representations of static environments by leveraging the properties of a normal distribution to model the space between sparse observations, but are not easily extended to handle environment dynamics. Occupancy grids [6] simplify the problem by discretizing the space into a regular grid, modeling the likelihood of intersecting an object at each cell. Meyer-Delius et al. [7] extend the transitional occupancy grid to use Hidden Markov Models (HMMs) to represent the transition probability between the free and occupied states at each cell. The work we present here compliments this model by prioritizing observations to maximize accuracy when the rate observations can be collected is limited.

When the environment topography is represented with a probability distribution, many approaches seek to maximize reduction of model uncertainty [8], expressed as the Mutual Information (MI) [9] provided by updating the map with an additional measurement. Various approaches build on this notion by incorporating MI with a beam-sensor model [10] or factoring in human input [11]. However, these approaches address static or stationary environments and will eventually converge to a high confidence, low accuracy state if insufficient measurements are provided. The work most closely representing the described problem [12] models long-term dynamics as a Fourier series, where the key frequencies are extracted using an augmented Discrete Fourier Transform designed for sparse, irregularly spaced observations. Their proposed approach uses HMMs to model transient dynamics and presents several exploration strategies, among which are those that direct robots towards regions of high entropy in both the observation likelihoods and the transition probabilities. However, our proposed approach uses an objective that blends MI with how well the transition probabilities match observations, more directly allocating observations where they can best improve model accuracy and confidence.

The focus of this work is to investigate the implications of limited sensing resources directly with a simplified problem formulation and easily scalable experiments. Environments are modeled as an occupancy grid, where each cell is an independent HMM that is observed directly, and a limited number of observations are available at each time step. When the number of cells exceeds the number of observations available, we must determine how to properly allocate observations such that each model is viewed often enough to capture the inherent dynamics. The main contribution of this work is an observation allocation strategy that balances improving model confidence and accuracy to quickly develop an understanding of environment dynamics, resulting in a steady-state distribution of observations that reflects the distribution of environment dynamics.

II. Problem Definition

In this work, we seek to accurately model the physical structure of a dynamic environment where the number of sensing actions that can be taken at any given time is limited.

[1]Derek Mitchell and Nathan Michael are affiliated with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA {derekm, nmichael}@cmu.edu

When the environment is initially unknown, the inherent dynamics must be discovered online before observations can be allocated appropriately. Here, we choose to model the environment as a time-varying occupancy grid $m_t = \{m_{1,t}, \ldots, m_{N_c,t}\}$, where any cell i can be in the occupied $(m_{i,t} = occ)$ or free $(m_{i,t} = free)$ state at any time-step t. To model both the environment topography and how it changes, each cell stores a unique and independent HMM parameterized by the probability the cell is occupied

$$p(m_{i,t} = occ) = 1 - p(m_{i,t} = free) \in [0,1]$$

and the probabilities that the cell transitions between states

$$p(m_{i,t}|m_{i,t-1}) \in [0,1] \quad \forall m_{i,t}, m_{i,t-1} \in \{free, occ\}.$$

Note that we assume the transition probabilities are locally stationary, where the value of $p(m_{i,t}|m_{i,t-1})$ is held consistent for all t in the absence of observations, but can change when the observed behavior disagrees with the model.

At each time-step, N_r sensing actions are taken that update the parameters of each HMM in m_t, as depicted in Fig. 1. Each sensing action produces an observation $o_j \in \{hit, miss\}$, where $\gamma_{\text{hit}} = p(o_j = hit|m_{i,t} = occ)$ and $\gamma_{\text{miss}} = p(o_j = miss|m_{i,t} = occ)$ describe the likelihood of sensing the corresponding observation conditioned on the cell being occupied. For this work, we assume each observation observes one cell directly with no occlusion and that there is no additional cost incurred when transitioning between sensing actions (as would occur when robots travel between locations to collect observations). While these assumptions are restrictive, the results we generate provide clearer intuition as to the influence of limitations and parameter choice in our approach directly.

The problem we address, then, is how to allocate the $N_r < N_c$ observations available at each time step to maximize the quality of the resulting model m_t, where quality is defined relative to the following three measures: 1) *Entropy*, expressed as:

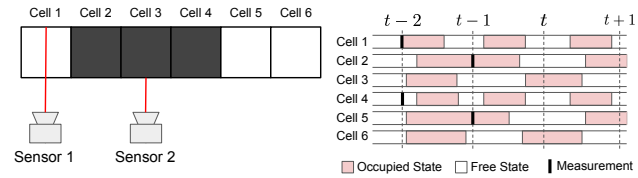
$$H(m_t) = -\sum_{m_{i,t} \in m_t} p(m_{i,t}) \log_2 p(m_{i,t}) + (1 - p(m_{i,t})) \log_2 (1 - p(m_{i,t})), \quad (1)$$

provides a measure of model confidence as it decreases while the occupancy likelihood of all cells trend towards 0 or 1. 2) Model accuracy is expressed through the *Kullback-Leibler Divergence* between the learned environment model $p(m_{i,t})$ and an oracle model where $p_o(m_{i,t} = occ) = 0$ or 1 according to the ground truth state:

$$D_{KL}(t,m) = \sum_{m_i \in m} p_o(m_{i,t}) \log_2 \frac{p_o(m_{i,t})}{p(m_{i,t})} + (1 - p_o(m_{i,t})) \log_2 \left(\frac{1 - p_o(m_{i,t})}{1 - p(m_{i,t})} \right). \quad (2)$$

3) The responsiveness of our approach is expressed through the *response time* $d^i_{r,t}$ which denotes the amount of time that passes after cell i changes state before it is observed.

$$d^i_{r,t} = (\max(t^i_{obs} - t^i_{change}, t - t^i_{change}))dt, \quad (3)$$



(a) Diagram of example scenario (b) Occupancy state time-series

Fig. 1: This work explores an example scenario where N_c cells are directly observed by N_r sensors and $N_r < N_c$. The objective is to appropriately allocate measurements, where N_r are available at each time-step t, such that the inherent environment dynamics can be accurately modeled.

where dt is duration between time-steps. In this work, we show how our approach outperforms standard techniques according to these measures despite the limited observations available.

III. Methodology

This work proposes an observation allocation strategy that scores cells based on the utility an additional observation would provide and chooses which N_r of N_c cells to observe each time-step based on that score. Section III-A describes the occupancy grid formulation in detail. Section III-B describes the utility measure which factors in the projected information gain and current *goodness-of-fit* to allocate observations. Section III-C discusses how we might extend this work to be utilized with a beam sensor that can observe multiple cells with one sensing action. In Section IV, experiments highlight the benefits of this approach and evaluate the effect of varying parameters on the method's performance.

A. HMM Occupancy Grid

The model we use in this paper is chosen to mimic the typical implementation of an occupancy grid. However, a static model regresses all measurements over time directly into the occupancy likelihood value, making it unsuitable for dynamic environments. This can result in historical bias having the dominant effect on occupancy likelihood, eventually forcing the model to stay in one state, as depicted in Fig. 2a.

Meyer-Delius et al. [7] generalize the occupancy grid approach to account for environment dynamics by defining an HMM at each cell. In this manner, the occupancy likelihood of each cell evolves over a series of discrete time steps as an independent Markov process according to the recursive update function:

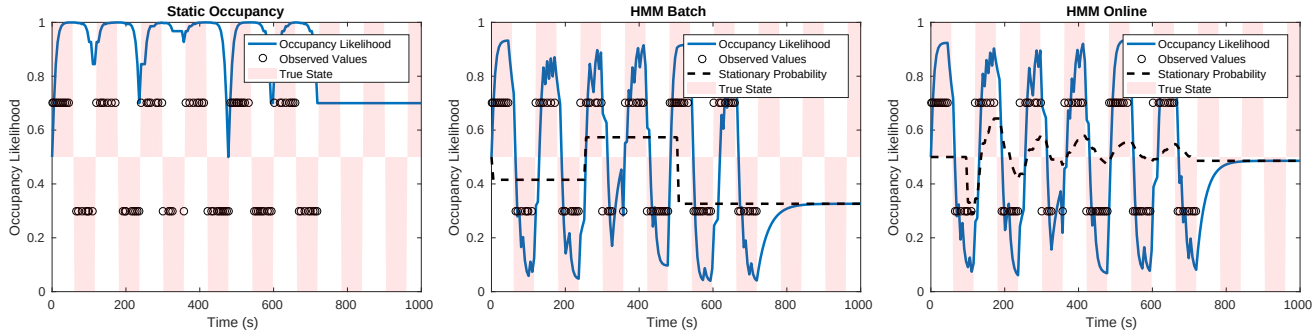
$$p(m_{i,t}|o_{1:n}) = \eta\, p(o_n|m_{i,t}) \sum_{m_{i,t-1} \in \{free,occ\}} p(m_{i,t}|m_{i,t-1}) p(m_{i,t-1}|o_{1:n-1}), \quad (4)$$

where η is a normalization constant.

Typical HMMs [13] are updated using Expectation Maximization (EM) to maximize the accuracy of the model over a series of observations:

$$p(m_{i,t}|m_{i,t-1}) = \frac{\sum_{\tau=1}^{t} p(m_{i,\tau-1} = \alpha, m_{i,\tau} = \beta|o_{1:n}, \theta)}{\sum_{\tau=1}^{t} p(m_{i,\tau-1} = \alpha|o_{1:n}, \theta)}$$

where $\alpha, \beta \in \{free, occ\}$ and θ are the HMM parameters learned via the standard forward-backward procedure. We



(a) Static Single Cell Model (b) Windowed Single Cell HMM (c) Online Single Cell HMM

Fig. 2: A time series comparison of binary-state models given sparse observations. A static model has difficulty matching the true value when it oscillates, but the HMM versions learn transition probabilities that allow for the change. The batched HMM performs sufficiently well at the cost of storing a window of observations, but we see similar performance with the online HMM approach without needing to store as much data.

can implement this method as an online approach by computing transition probabilities over a sliding window, resulting in the batched HMM update depicted in Fig. 2b. While this may be sufficient in most cases, it still requires storing a significant amount of data for each cell to accurately represent transition probabilities. Instead, we rely on the online version derived by Mongillo et al [14] to update parameters as each data point arrives, as shown in Fig. 2c.

Given that the measurements of cells arrive asynchronously, the Markov process evolves in the absence of observations according to:

$$p(m_{i,t}|o_{1:n}) = \sum_{m_{i,t-1}\in\{free,occ\}} p(m_{i,t}|m_{i,t-1})p(m_{i,t-1}|o_{1:n-1}), \quad (5)$$

which serves to drive the occupancy likelihood asymptotically towards the stationary probability

$$p_{\text{stat}} = \frac{p_{o|f}}{p_{o|f} + p_{f|o}},$$

where $p_{o|f} = p(m_{i,t} = occ|m_{i,t-1} = free)$ and $p_{f|o} = p(m_{i,t} = free|m_{i,t-1} = occ)$. We can see this process in the progression of occupancy likelihood in Fig. 2, where the likelihood exponentially decays towards the stationary probability in the absence of observations.

B. Observation Utility Measure

Mutual Information is a useful measure of utility as it indicates the amount of information gained through an additional sensing action. It can be expressed in the form:

$$I_{MI}[m;o] = -\sum_{i,j} p(m_i,o_j)\log\frac{p(m_i)p(o_j)}{p(m_i,o_j)},$$

or interpreted as the change in entropy, or uncertainty, as the result of a measurement:

$$I_{MI}[m;o] = H[m] - H[m|o],$$

where $H[m]$ is the entropy of the map and $H[m|o]$ is the map entropy conditioned on the new observation.

Given that this expression of MI is designed for static environments, it is insufficient when the occupancy state is subject to change. While MI can be formulated to sufficiently address the influence of transition probabilities, the additional parameters make the computation complex and prohibitively expensive. Alternatively, we consider using the *goodness-of-fit* of the transition probability parameters in the utility measure.

Here, we leverage the Pearson's χ^2 test to determine how well the transition probabilities match the history of observations. The χ^2 test is a used to evaluate whether or not an observed frequency distribution conforms to a theoretical distribution. Kalbfleisch and Lawless [15] show how this test can be used to determine how accurate the transition probabilities are modeled.

The χ^2 value is computed using:

$$\chi^2(t) = \sum_{ij\in\{occ,free\}} \frac{(n_{ijt} - e_{ijt})^2}{e_{ijt}},$$

where n_{ijt} is the number of observed transitions from state i to state j from time step $t-L$ to t and e_{ijt} is the expected number of transitions, where e_{ijt} can be expressed as:

$$e_{ijt} = (p_{\text{stat}}L)p(m_{i,t}|m_{i,t-1}).$$

The variable L, here, denotes the window size, or the total number of time steps we search for transitions. One can interpret this value as a scaled, square-distance measure of the transition probability parameter. As we see more or less transitions than expected, we can expect the value of χ^2 to increase. If we use this value to modulate the observation rate of a cell, then cells that see the incorrect number of observations will receive higher priority.

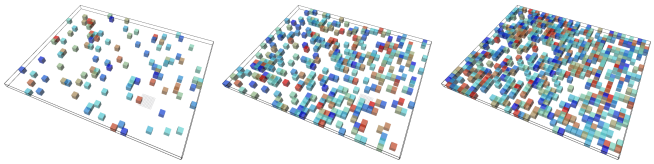
To incorporate the *goodness-of-fit* into our allocation objective, we compose the weighted objective function:

$$u_i = I_{MI}[m_i;o_j] + \alpha\frac{\chi^2}{L}, \quad (6)$$

using the value α to balance the relative influence of the component objectives. Note that we normalize χ^2 by the window size L to mitigate its influence on the choice of α. This ensures that the effect of α is consistent as the window grows over time (until $t > L$) or when L is changed between runs, as is evidenced in Sect. IV-C.

C. Beam Utility Measure

While the proposed approach works for single cell models, observations traditionally impact more than one cell at a time. For a beam sensor, there are well studied models [16][17] that incorporate the notion of visibility to measure the likely



(a) 10% dynamic (b) 30% dynamic (c) 50% dynamic

Fig. 3: The set of test environments. Each dynamic cell is attributed an oscillation period between 300s and 2000s, drawn from a uniform random distribution. The percentage of dynamic cells varies between environments as listed.

utility a sensing action could provide. The visibility of a cell c_i the set of cells $c \subset m_t$ intersected by a beam can be expressed as:

$$\text{vis}(c_i) = \prod_{j<i}(1 - p(c_j)) \tag{7}$$

where $j < i$ defines the set of cells in c that precede i along the ray's direction vector. Using this value, we can determine the expected utility of visible cells using the equation:

$$E[u] = \sum_i \text{vis}(c_i) u_i. \tag{8}$$

Thus, we will able to incorporate *goodness-of-fit* directly into the beam-based utility measure in the same manner as proposed in this paper, with the weight α applied to the per-beam utility instead of per-cell.

IV. Results

In this section, we compare the performance of the proposed allocation objective against alternatives and evaluate performance, as defined in Sect. II, when parameters and conditions vary. Section IV-A describes the general format of experiments and provides details on the evaluation criteria. Section IV-B shows how the proposed approach outperforms random selection and the component objectives used alone. Section IV-C highlights the change in performance as the relevant parameters are adjusted. Finally, Section IV-D evaluates how performance varies as a function of environment conditions.

A. Experiment Setup

In this work, we consider scenarios where the occupancy state of cells evolve over time at regular intervals, where the state of each cell is represented by a square wave with the frequency f_i. As such, each cell will begin in the free state for a duration of $T_i/2$, where $T_i = 1/f_i$ is the associated square wave period, before transitioning to the occupied state for the same duration. Static cells are a special case where the cell exists in its initial state for all time. Excepting the illustrative example in Fig. 4, environments are constructed in a grid of 1073 cells with the set of dynamic cells and their T_i values sampled randomly. For each environment we test, a ratio of dynamic to static cells is chosen and a random sampling of this fraction are assigned a value of T_i sampled from a uniform distribution between 300 and 2000 seconds. To perform a run, we allocate N_r observations to the cells chosen based on the specified allocation objective each time step, advancing $dt = 10s$ after each step. Cells that are observed are updated according to (4), while other cells evolve according to (5).

TABLE I: Comparison of approaches for runs in the 30% dynamic environment (Fig 3b) for 10,000 seconds. Each row corresponds to a separate run with the listed parameters. The μ_H and μ_D terms correspond to the average entropy and KL-divergence of the last 2000s of operation. $E[d_r]$ is evaluated over the last 2000s, where a value of greater than 1.0 suggests that a majority of state changes go unobserved. N_u indicates the number of cells that are unobserved over the 2000s window. Here, we see that the proposed approach outperforms all other allocation objectives, with significantly smaller μ_H, μ_D, and $E[d_r]$ values as well as a sufficiently small N_u value.

Objective	N_r	μ_H	μ_D	$E[d_r]$	N_u
Random	100	337.7	340.4	0.2644	0
Random	150	284.6	299.8	0.1634	0
Random	200	250.5	270.9	0.1238	0
MI only	100	572.5	438.2	0.9315	98
MI only	150	490.1	319	0.4498	38
MI only	200	392.8	264.9	0.4319	33
χ^2 only	100	276.1	320.8	0.7028	12
χ^2 only	150	249	282.2	0.4779	1
χ^2 only	200	229.3	245.9	0.347	2
Proposed	100	297.6	325.7	0.1238	1
Proposed	150	222.3	217.3	0.05703	0
Proposed	200	**177.8**	**157.8**	**0.03326**	0

B. Comparison of Objective Functions

We evaluate performance according to the entropy (1), KL-Divergence (2), and response time (3) as formulated in Sec. II with one caveat. While we seek to reduce the response time for all cells, the value which best reflects our response to changes over the whole map is the average maximum response time fraction

$$E[d_r] = \frac{1}{N_c} \sum_{i=1:N_c} \max_t \bar{d}^i_{r,t}.$$

over all cells, where we define the response time fraction as:

$$\bar{d}^i_{r,t} = \frac{d^i_{r,t}}{T_i},$$

interpreted as the response time normalized by the cell's associated period of change. As a cell changes state at each $0.5T_i$, values of $\bar{d}^i_{r,t} > 0.5$ indicate that cell i is receiving observations less frequently than changes occur. Providing the average maximum $\bar{d}^i_{r,t}$, as opposed to the maximum over all cells, permits outliers to reduce the quality of performance without being the dominant influence.

Figure 4 shows how the proposed allocation objective compares against alternatives. Here, a small test of four independent cells is run for each allocation objective with $N_r = 1$. A random allocation objective provides a solid baseline, which exhibits a reasonable approximation of the time-series occupancy likelihood, but is still capable of missing changes. We further evaluate performance in Table I, where tests are performed on the grid shown in Fig. 3b. The listed values for average, steady-state entropy μ_H, divergence μ_D, and response time fraction $E[d_r]$ for the random allocation case provide a target to exceed for our desired model confidence, accuracy, and response time. Additionally, the number of cells that go unobserved in the steady-state N_u highlights when an objective is not effectively distributing focus.

Allocation by MI, while appropriate for maximizing the confidence of individual cells, performs significantly worse than random allocation when applied to the environment model as a whole. Once an HMM reaches a stationary probability with high enough confidence, priority is given to

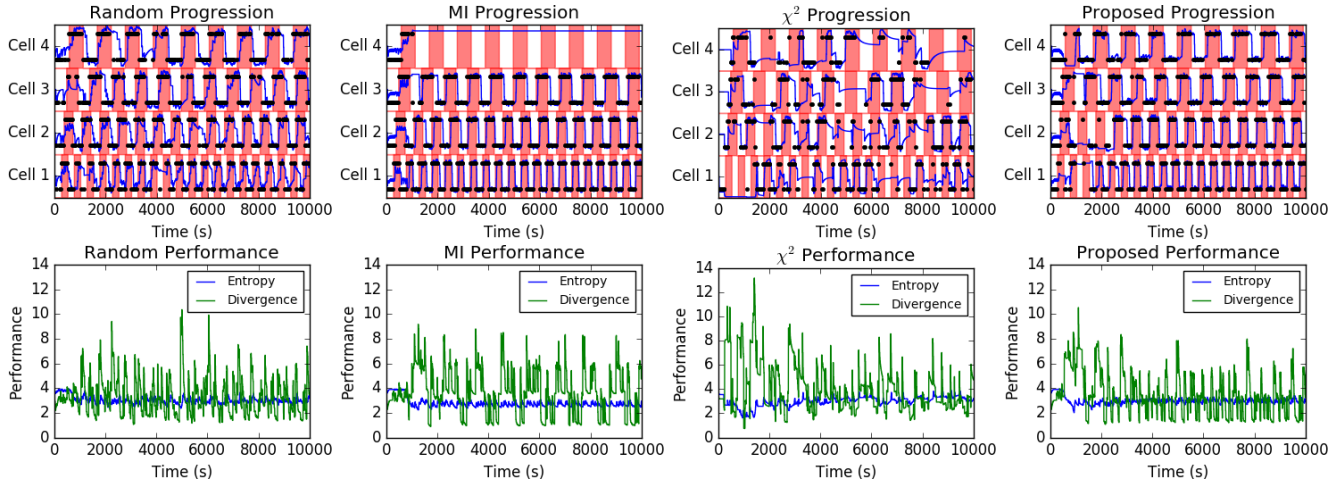


Fig. 4: Comparison of allocation objectives. The top row shows a time series in the same manner as Fig. 2 where measurements are applied to one of four cells every 10 seconds and each row within the plot corresponds to a different HMM cell. The bottom row plots entropy and KL-Divergence for the set of cells in each approach. Random allocation serves as a baseline for comparison. Allocating based on MI results in accurate models, but can abandon cells once confidence is high, resulting in low entropy but high divergence. While χ^2 alone produces poor results, the proposed balanced objective can distribute observations based on the learned dynamics resulting in low entropy and a fast decrease in divergence whenever a change induces a spike.

cells whose observations suggest dynamics. As time passes without subsequent observations for a cell, it appears to be static and is ignored. This results in high accuracy for a limited subset of cells, but low accuracy and confidence overall. We see the same trend in Table I, where N_u is significantly higher than other approaches. As many cells go unobserved in the steady-state, the environment model exhibits a higher steady-state entropy and divergence as well as poor response times.

Alternatively, using the χ^2 *goodness-of-fit* measure alone results in measurements being distributed more evenly, but with less focus on ensuring high confidence. Measurements are drawn to cells that do not observe the number of transition their HMMs predict. However, measurements that occur immediately succeeding a change will drastically reduce model confidence for the cell, making it more difficult to recognize subsequent transitions.

In using the proposed objective function, we can infer that the χ^2 component serves to draw attention to cells whose dynamics are inappropriately modeled while MI recovers confidence lost in adapting to the newly recognized dynamics. Additionally, we note that our approach will bias early measurements towards static cells, as these produce no transitions, until the transition probabilities are low enough and confidence high enough that other cells draw focus. This process is highlighted in Fig 5, where our approach is applied to an environment with cells of corresponding dynamics artificially clustered to show how attention is allocated. By expending effort to initially identify static regions, we allow ourselves significantly more observations for dynamic regions, eventually settling on a distribution of attention that reflects the true environment dynamics.

C. Varying Parameters

Table II shows how performance changes relative to parameter choice. The first thing to note is that altering the time window L does not appear to have an appreciable effect on performance. As L simply defines the range of time in which we search for transitions, it is sufficient to define L large enough to account for the expected dynamic range and small enough to quickly forget data collected before the model has converged.

While we treat N_r as a parameter here, it is more accurate to consider N_r as a problem constraint as the rate of observation collection is usually limited by the available number of robots and how quickly they can respond. As can be expected, increasing this value results in a drastic improvement in performance.

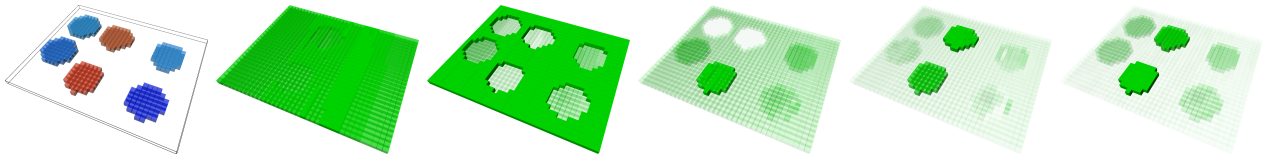
Varying the objective weight predictably scales the influence of the associated component objectives. As α becomes smaller, performance mimics a pure MI objective with poorer performance measures and a drastic increase in steady-state N_u. Larger α values mimic the pure χ^2 objective, with reasonable performance, but less than when properly balanced.

D. Varying Environments

To evaluate performance as environment conditions change, we run simulations on randomly generated environments of varying density of dynamic cells. Figure 3 shows the three environments tested in this work, with 10%, 30%, and 50% of cells oscillating between the occupied and free state with periods drawn from a uniform distribution between 300s and 2000s. As expected, the more the environment exhibits change, the harder it is to model. However, we note that the best choice of α appears to vary depending on the environment. More static environments prefer an approach closer to pure MI, while environments with dense dynamics rely more heavily on the χ^2 objective.

V. Discussion

There are a couple of limitations when using this approach that we note here. First, the approach is sensitive to an accurate determination of the number of observed transitions, n_{ijt}. If the sensor noise is significant, false positives can be registered, which promote further measurements of false



(a) Heatmap (b) $t = 400$ (c) $t = 1200$ (d) $t = 3000$ (e) $t = 6000$ (f) $t = 10000$

Fig. 5: Evolution of priority as the environment model is learned when allocating 150 observations per time-step according to the proposed objective. A heatmap of the true dynamics is shown in a, where red indicates high frequency changes and blue indicates low frequency changes. The subsequent images reflect the relative number of measurements assigned to the cell over a 1000s time window preceding time t, with cells appearing more transparent the fewer measurements they are allocated over the window. Initially, measurements are spread evenly over all cells. Then focus is directed towards finding the static cells in the environment. Finally, measurements are distributed with a bias towards dynamics cells until the distribution accurately reflects the underlying dynamics.

TABLE II: Evaluation of performance as a function of parameter choice. Again, each run operates for 10,000 seconds on the environment in Fig. 3b. The objective performs as expected, with L not having a significant impact, N_r improving performance as more observations are available, and α mimicking the component objectives as it is tuned in either direction.

L	N_r	α	μ_H	μ_D	$E[d_r]$	N_u
1000	100	10	420	416.8	1.332	117
3000	100	10	392.9	451	1.632	130
5000	100	10	392.8	452.4	1.604	132
1000	150	10	311	343	1.202	85
3000	150	10	306.8	350.2	1.298	88
5000	150	10	305.2	350.6	1.369	88
1000	200	10	261.6	268	1.011	54
3000	200	10	256	267.7	1.052	53
5000	200	10	256.4	267.8	0.9191	50
1000	100	100	290.9	253.5	0.06865	0
3000	100	100	297.6	325.7	0.1238	1
5000	100	100	297.8	347.8	0.1504	1
1000	150	100	219.5	163.9	0.0434	2
3000	150	100	222.3	217.3	0.05703	0
5000	150	100	229.4	238.8	0.06176	0
1000	200	100	185.6	**136.7**	0.04156	2
3000	200	100	**177.8**	157.8	**0.03326**	0
5000	200	100	182.6	169.1	0.03602	0
1000	100	200	334.5	287.9	0.1551	0
3000	100	200	314.3	347.9	0.2521	0
5000	100	200	318.5	345.1	0.2763	0
1000	150	200	228.5	172.2	0.0362	0
3000	150	200	242.3	259.2	0.0742	0
5000	150	200	246.5	265	0.1001	0
1000	200	200	197.5	148.4	0.04378	2
3000	200	200	192	194.5	0.04598	0
5000	200	200	198.1	206.6	0.05416	0

positives, creating a self-sustaining loop that biases priority towards cells that produce noisy measurements. Strategies for determining n_{ijt} in the presence of noise will allow this approach comparable performance that is robust to noisy measurements.

Second, cells with low transition probabilities will be revisited at very low frequencies. This trait is acceptable when the dynamics do not change, but less so when a static region becomes dynamic at some future time (e.g. a region is designated for parking). We expect that inducing decay in the confidence of our dynamics model as well as the occupancy likelihood will serve to promote revisitation of cells that have been considered static for too long a duration.

VI. Conclusion

This work presents a strategy for allocating limited sensing resources to mapping a dynamic environment. As expected, the best performance is achieved when the distribution of observations reflects the dynamics being observed, which is most easily achievable when the dynamics are learned quickly. The results show that our proposed approach outperforms both the pure MI and χ^2 objectives, which suggests that environments can only be modeled accurately if there is a balanced effort between reinforcing poorly modeled dynamics and reducing uncertainty in occupancy likelihood. When this balance exists, we are capable of ensuring an average response time significantly less than $0.5T_i$ when as little as 14% of cells are observed each time step. While this value will vary based on the dynamic properties of the environment, future works can extend these results to provide performance guarantees relative to the available number of robots and their limited energy capacities while considering the influence of travel time.

TABLE III: Comparison of performance for different concentrations of dynamics. Tests are performed on the environments depicted in Fig. 3 for 10,000s each. The weight α is varied for each run, while the remaining parameters are fixed to $L = 3000$ and $N_r = 150$. Environments with less dynamics cells prefer an allocation objective closer to pure MI while more dense dynamic cells lean towards pure χ^2.

α	Dyn %	μ_H	μ_D	$E[d_r]$	N_u
10	10	114.4	**46.56**	**0.01938**	0
50	10	96.6	49.41	0.02549	0
100	10	**94.87**	57.34	0.03464	0
150	10	99.82	67.95	0.04216	0
200	10	106.3	80.01	0.06319	0
10	30	306.8	350.2	1.298	88
50	30	232.3	235.9	0.1475	6
100	30	**222.3**	**217.3**	**0.05703**	0
150	30	230.8	235.1	0.05913	0
200	30	242.3	259.2	0.0742	0
10	50	442.5	704.2	2.286	247
50	50	**383.3**	551.3	0.5649	58
100	50	388.7	479.9	0.111	7
150	50	395	**472.5**	**0.09815**	1
200	50	401.2	482.4	0.1074	0

References

[1] Y. Liu and G. Nejat, "Robotic urban search and rescue: A survey from the control perspective," *Journal of Intelligent and Robotic Systems*, vol. 72, 11 2013.

[2] S. A. Stoeter, P. E. Rybski, M. D. Erickson, M. Gini, D. F. Hougen, D. G. Krantz, N. Papanikolopoulos, and M. Wyman, "A robot team for exploration and surveillance: Design and architecture," in *Int. Conf. on Intelligent Autonomous Systems*, 2000, pp. 767–774.

[3] A. Howard, L. E. Parker, and G. S. Sukhatme, "Experiments with a large heterogeneous mobile robot team: Exploration, mapping, deployment and detection," *The International Journal of Robotics Research*, vol. 25, no. 5-6, pp. 431–447, 2006.

[4] S. T. O'Callaghan and F. T. Ramos, "Gaussian process occupancy maps," *The International Journal of Robotics Research*, vol. 31, no. 1, pp. 42–62, 2012.

[5] S. Srivastava and N. Michael, "Approximate continuous belief distributions for precise autonomous inspection," in *IEEE Int. Sym. on Safety Security and Rescue Robotics*, Oct 2016, pp. 74–80.

[6] S. Thrun, "Learning occupancy grid maps with forward sensor models," *Autonomous Robots*, vol. 15, no. 2, pp. 111–127, Sep 2003.

[7] D. Meyer-Delius, M. Beinhofer, and W. Burgard, "Occupancy grid

models for robot mapping in changing environments," in *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, ser. AAAI'12. AAAI Press, 2012, pp. 2024–2030.

[8] A. Singh, F. Ramos, H. D. Whyte, and W. J. Kaiser, "Modeling and decision making in spatio-temporal processes for environmental surveillance," in *Proc. of the IEEE Intl. Conf. on Robot. Auto.*, May 2010, pp. 5490–5497.

[9] *Entropy, Relative Entropy and Mutual Information*. John Wiley & Sons, Ltd, 2001, ch. 2, pp. 12–49.

[10] B. J. Julian, S. Karaman, and D. Rus, "On mutual information-based control of range sensing robots for mapping applications," *The International Journal of Robotics Research*, vol. 33, no. 10, pp. 1375–1392, 2014.

[11] H. Liu and M. A. Hsieh, "Neural network aided information theoretic exploration," in *IEEE Int. Sym. on Safety Security and Rescue Robotics*, Aug 2018, pp. 1–7.

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

[13] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, Feb 1989.

[14] G. Mongillo and S. Deneve, "Online learning with hidden markov models," *Neural Comput.*, vol. 20, no. 7, pp. 1706–1716, July 2008.

[15] J. D. Kalbfleisch and J. F. Lawless, "The analysis of panel data under a markov assumption," *Journal of the American Statistical Association*, vol. 80, no. 392, pp. 863–871, 1985.

[16] D. Crispell, "A continuous probabilistic scene model for aerial imagery," *PhD thesis*, 2010.

[17] B. Charrow, S. Liu, V. Kumar, and N. Michael, "Information-theoretic mapping using cauchy-schwarz quadratic mutual information," in *Proc. of the IEEE Intl. Conf. on Robot. Auto.*, May 2015, pp. 4791–4798.