Model-based Informative Path Planning

Gabriel Margolis, Massachusetts Institute of Technology

May 5, 2019

Abstract

Surrogate model construction can enable autonomous agents to quantify the uncertainty of their beliefs about the world - to know what they do not know. This is valuable because it can allow those agents to plan to gain information about the world using informative path planning. Prior art on informative path planning models exclusively spatial correlation, resulting in highly local predictions. As algorithms for online non-myopic planning improve, it is desirable to integrate them with models that can utilize expert knowledge to make better long-range predictions. We present Model-based Informative Path Planning, an algorithm which uses non-myopic active learning to rapidly estimate general relationships between intensive properties. We demonstrate that this method outperforms passive parameter estimation and Single-feature Informative Path Planning in a benthic habitat environment where such general relationships are found.

1 Introduction

The efficient planning of surveys is critical for exploratory science missions in remote regions, such as the deep ocean or foreign planets, where the survey region is vast and resources are limited. The Woods Hole Oceanographic Institution (WHOI), in collaboration with the Computer Science and Artificial Intelligence Laboratory at MIT, embarked on one such exploratory science mission in December 2018 in Costa Rica, where multiple Autonomous Underwater Vehicles (AUVs) (Figure 1) were deployed to sense oil seeps on the ocean floor. Another such mission will take place in late 2019 in Santorini, Greece.

Figure 1 The Slocum Glider, an Autonomous Underwater Vehicle frequently used by the Woods Hole Oceanographic Institution and MIT to conduct science missions in remote regions of the deep ocean.

Surrogate models allow AUVs and other autonomous agents to survey their environment in a methodical manner, which we refer to as informative path planning (IPP). The true state of a survey environment such as the ocean floor can be modeled as a function, mapping location coordinates to the value of the feature of interest. This function is black-box, meaning no closed-form expression is known for it, and is expensive to evaluate, meaning that observing the function consumes some limited resource such as energy or time. In this context, a survey plan is expressible as a sampling strategy for this expensive black-box function. When sampling an expensive black-box function, it is worthwhile to expend some computational effort to construct a surrogate model of the function, which represents our belief about the mean and uncertainty of the true function's value, and to query that surrogate to decide where subsequent exploration should be made. Regions of high information entropy, where the value of a feature is predicted with the least certainty under the surrogate, can be expected to yield a greater increase in knowledge about the overall environment once they are explored, so an intelligent agent should recognize the value in visiting these regions. Past science missions carried out by WHOI and MIT have featured the deployment of path planning algorithms which explore using this principle of information entropy on surrogate functions.

The informative path planning surrogate function is most commonly implemented as a Gaussian process [13]. Gaussian processes are adept at modeling spatial correlations, which are often an invariant property of physical features. They also provide an explicit representation of uncertainty, in the form of a probability distribution over output values. From this uncertainty representation, an acquisition function such as information entropy or the Upper Confidence Bound (UCB) can be computed at a set of points across a field [7]. Such an acquisition function $a(x)$ quantifies the utility of collecting an observation at a location of interest x.

One shortcoming of the Gaussian process as a surrogate model is that in a field with no prior information, the model estimate returns to the prior distribution at points not physically close to an observed location. This is prone to produce relatively simple planning behavior, akin to lawnmower patterns for an explorative acquisition function such as information entropy, or to gradient-following for an exploitative acquisition function such as UCB [6]. A well-specified prior distribution on our Gaussian process surrogate can elicit more interesting and effective behavior from a path planner.

To construct an informative Gaussian process prior for large survey domains, we propose Model-based Informative Path Planning (MIPP). An agent performing traditional informative path planning actively learns a surrogate function; An agent performing MIPP actively learns both a surrogate function and a knowledge model describing the relationship between intensive properties in the survey region. We implement our agent's knowledge model as a Gaussian Graphical Model with known structure and actively learned parameters. We demonstrate that in an environment where some side information – expert information about features correlated with the feature of interest – is provided, MIPP correctly estimates a knowledge model and achieves lower predictive error than its model-free alternative.

Once a predictive surrogate model has been constructed for a survey field, an autonomous agent must plan a sequence of observation locations to visit which cumulatively maximize the acquisition function while satisfying some constraints such as range, time, or energy. Previous work in informative path planning has commonly utilized greedy observation selection, known as myopic path planning [3, 10]. Recently, advances from the planning and search community have been applied to perform long-horizon, or nonmyopic, observation selection in an informative path planning context. Binney (2012) used branch-and-bound search to speed up the selection of nonmyopic informative paths across a Gaussian process field [3]. Nguyen et al. applied Monte Carlo Tree Search (MCTS) to perform nonmyopic informative path planning for a resource constrained-aerial glider [11]. We implement Monte Carlo Tree Search [5] over a fixed time horizon to perform active, nonmyopic informative path planning, which enables us to select efficient and high-value extended observation plans in real time. We introduce a novel formulation of the agent's action space as a graph on the acquisition function maxima, which is shown to provide an inexpensive approximation to the optimally informative path in a domain with sparse expert side information.

2 Background and Related Work

2.1 Informative Path Planning

The selection of observations which maximize information gain, also known as adaptive sampling, has been an extensively studied problem in the contexts of optimal experimental design, algorithm optimization, and robotics. Lindley [7] proposed the use of entropy reduction as a utility or acquisition function for optimal experimental design in 1956. Shewry and Wynn [14] in 1987 built upon this work to define an algorithm for the optimal placement of sensors to maximize entropy reduction in a survey field. In 2006, Krause et al. [8] proposed a Gaussian process-based sensor placement approach which used mutual information as its information gain function and exploits the submodularity property of Gaussian processes for significant speed-up. This approach has been widely applied in sensor placement and remains among the state-of-the-art approaches to the sensor placement problem.

In the informative path planning (IPP) scenario, a mobile sensing agent must autonomously select its next observation site based on the observations it has made previously. As in optimal sensor placement, the agent seeks to maximize the cumulative information gain achieved by its observations. However, unlike in optimal sensor placement, IPP considers that a vehicle must expend time and energy as a function of the distance it travels between observation locations. An agent performing IPP also makes its observation decisions sequentially, so an online learning approach can be adopted which improves the agent's plan at each step. In 2010, Binney et al. published an IPP algorithm based on greedy mutual information maximization and demonstrated that it gathers information with considerably better efficiency than exhaustive search [3].

2.2 Gaussian Processes

A Gaussian process is a random process in which the correlation between two input points x_1 and x_2 is defined by a kernel function $k(x_1, x_2)$ [13]. Given a set of observations in a Gaussian process, we can predict the mean and variance for any unobserved input, conditioned on those observations (Figure 2). Gaussian processes are particularly effective surrogate models for geospatial mapping, since the value of a continuous-valued feature across physical space commonly exhibits consistent spatial correlation that are modeled well by a kernel covariance

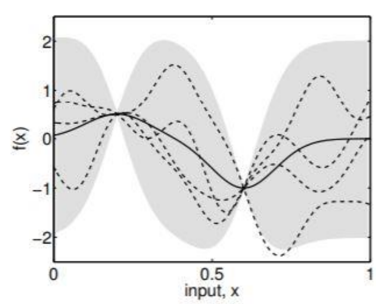

Figure 2 Gaussian Process Regression on two training points. The black line represents the function mean. The grey region represents points within two standard deviations of the mean. [7]

Typically, a Gaussian process model with multiple output variables assumes that those output variables are independent of one another. The Linear Model of Coregionalization (LMC) [17] represents linear correlative relationships between outputs of a multi-output Gaussian process. The predictive mean and variance of each output variable under LMC is conditioned on the training data collected about all output and input variables at all locations, according to a kernel covariance in the inputs and a linear-Gaussian conditional dependence in the outputs. Typically, learning a multi-output Gaussian process under the LMC is an expensive maximum likelihood optimization problem. However, given an external estimate of the feature covariance matrix, we can fix the hyperparameters of the LMC and avoid this optimization step.

2.3 Gaussian Graphical Models

A Gaussian Graphical Model (GGM) [12] is an undirected graph defined by a set of vertices and edges, G = <V, E>. Each vertex in a GGM represents a different Gaussian-distributed variable, and each edge represents a conditional dependence between the variables it connects (Figure 3). A GGM's edges may be defined by its precision matrix. The precision matrix of a fully connected GGM is the inverse of the covariance matrix of the joint Gaussian distribution over all variables. Employing a sparse approximation of the precision matrix approximates independence relations between variables in the graph to reduce the size of its edge set.

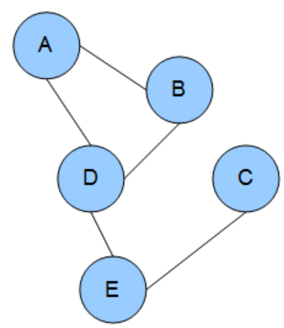

Figure 3 A Gaussian Graphical Model. Vertices represent variables, undirected edges represent conditional dependency

The estimation of Gaussian Graphical Model parameters can be achieved by maximum-likelihood estimation or Bayesian estimation techniques [12]. Maximum-likelihood techniques offer greater speed and flexibility than Bayesian estimation, formulating parameter estimation as an optimization problem. However, Bayesian estimation produces distributions over parameter values which explicitly represent uncertainty, facilitating integration of GGMs with active learning and informative path planning approaches. Tong and Koller (2000) presented a theoretical framework for the active learning of Bayesian Network parameters and structure and demonstrated that this approach improved the sample-efficiency of learning in both cases [15, 16].

2.4 Nonmyopic Planning

Early work on informative path planning [3, 10] focused on the selection of sampling actions one at a time by selecting the action which would immediately return the greatest reduction in entropy. This style of planning is known as greedy or myopic planning.

Recent developments on state space search within the planning community have brought nonmyopic path planning closer to practicality. Binney and Sukhatme (2012) applied branch and bound search, a heuristic search algorithm, to generate optimal nonmyopic plans in the informative path planning scenario [4]. Nguyen et al. applied Monte Carlo Tree Search to perform nonmyopic informative path planning for a resource constrained-aerial glider [11].

3 Method

3. 1 Method Overview

We propose a method, Model-based Informative Path Planning, for an autonomous agent to gather information about a survey field in a resource-efficient manner. We achieve this improved resource efficiency by maintaining an informative surrogate function computed from a 'knowledge model' which is informed by an expert. In contrast to previous work on IPP [2, 3, 4, 9, 10, 11], in which an agent seeks to reduce uncertainty about a single feature's value at all locations in a survey field, our algorithm directs an agent to additionally reduce uncertainty about its knowledge model, which encodes the conditional relationships between multiple intensive features in a survey field. When adequate expert knowledge is initially supplied, our algorithm enables an agent to quickly infer the distribution of a new feature in locations where it has not yet been observed, leading the agent to plan a more effective survey. Our algorithm is demonstrated to correctly estimate the conditional relationships between features in a synthetic survey environment, resulting in a surrogate function with improved predictive accuracy.

3.2 Belief Modeling

We represent our agent's knowledge model as a Gaussian Graphical Model, which represents the joint Gaussian distribution of features at any given location in the survey region. We learn a GGM from our past observations by finding a precision matrix that describes those observations with maximum likelihood. In addition, we determine the agent's confidence in its GGM representation of the world by tracking the variance of each element of the precision matrix based on the number of times the relevant features have been observed together [5, 12].

3.3 Inference

We infer a prior distribution over our feature of interest as the output of a multi-output Gaussian process with a Linear Model of Coregionalization. The feature covariance hyperparameters in the LMC are fixed as the inverse of the precision matrix estimated in our GGM belief model. Each feature in our belief model is thus modeled as an output of the multi-output Gaussian process. The multi-output Gaussian process is trained on all available expert-provided observations as well as observations made by the agent. The input covariance kernel of each output is modeled as a Gaussian kernel with expert-provided lengthscale and variance.

3.4 Planning

We demonstrate the impact of our improved prior on agent inference capability when performing myopic planning, then extend our agent to perform nonmyopic planning and examine the impact of our improved prior on agent behavior. In the simplest problem we consider, we restrict our agent to take one of four actions, which consist of moving either North, South, East, or West by a fixed distance and then making a sensor observation of the feature of interest.

In the myopic planning case, our agent selects its action to greedily minimize the total model entropy achieved after the action is taken. If the uncertainty in the agent's belief model is taken to be fixed, this is the same mutual information acquisition function used in Binney et al. [3]. With nonzero uncertainty in the belief model, the acquisition function is equal to the mutual information acquisition function, plus the knowledge model information gain scaled by some tunable constant.

In the nonmyopic planning case, our agent selects its action based on the results of a fixed-horizon Monte Carlo Tree Search [5]. We perform simulation rollouts for MCTS by sequentially selecting random actions and updating our belief model and multi-output Gaussian process at each step. The reward signal used is the final entropy of the belief model after the time horizon has been reached.

Although MCTS offers significant speedup over exhaustive search techniques for nonmyopic planning, planning over a large horizon with MCTS is still slow, particularly when rewards are sparse. In our domain with sparse expert knowledge (Section 4.3), we modify the action space of our agent to include only travel between nearby locations where expert observations are available, since they are likely to produce a much higher information gain by informing the knowledge model. This modified action space produces suboptimal paths but is more practical for real-world application of MCTS than the original action space.

3.5 Agent Behavior

After taking an action and collecting an observation, our agent first updates its belief model by estimating its parameters given the new data it has received. It also decrements the uncertainty of any correlative relationships involving exclusively the features known at the newly sampled location. Next, it updates the feature covariance parameters of the LMC to match those estimated in its belief model, augments the LMC training data with the new observation, and infers the new prior distribution over the feature of interest. Depending on the setting, greedy planning or MCTS is used to select an entropy-minimizing action and observation pair to take next, and the process repeats.

Algorithm 1. Single-feature Informative Path Planning

```
Load expert observations of the sensed feature (obs)
Initialize agent
while time_remaining > 0:
    inference_model = GP(obs)
    prior = Infer-Prior(inference_model)
    action = argmin{actions}(Entropy-After-Action(prior))
    agent.Move(action)
    obs = obs + agent.Make-Observation()
    time_remaining = time_remaining - action.Duration
end while
```

Algorithm 2. Model-based Informative Path Planning

```
Load expert observations of all features (obs) and knowledge model structure (M)
Initialize agent
while time_remaining > 0:
    knowledge_model = GMM(obs, M)
    inference_model = LCM-GP(obs, K)
    prior = Infer-Prior(inference_model)
    action = argmin{actions}(Entropy-After-Action(prior, knowledge_model))
    agent.Move(action)
    obs = obs + agent.Make-Observation()
    time_remaining = time_remaining - action.Duration
end while
```

4 Results

We evaluate the performance of our approach, Model-based Informative Path Planning, by simulating its ability to facilitate an agent's exploration of three simulated underwater environments. In our first simulated environment, only a single sensor is available to the agent for use, and no expert knowledge is available to the agent during planning. In this domain, we demonstrate that our algorithm performs equivalently to the Single-feature Informative Path Planning formulation, referred to as S-IPP, that is well-established in prior art [3, 10, 11]. This is because our algorithm is identical to S-IPP in the single-sensor case with no expert knowledge. We demonstrate that both our algorithm and S-IPP outperform a random walk policy on the metric of predictive accuracy in this domain.

We introduce expert knowledge in our second environment in the form of multiple side information features and a linear-Gaussian Bayesian Network "knowledge model" with known structure, but unknown parameters. Here, we compare M-IPP so S-IPP as well as to a best-case "expert" who knows the true parameters of the "knowledge model" from the start.

Our third environment illustrates the impact of active parameter learning in M-IPP, where information gain in the parameters of the linear-Gaussian Bayesian Network "knowledge model" is incorporated into the acquisition function in addition to the typical IPP information gain term. The effect of incentivizing the knowledge model information gain is most evident in an environment where expert measurements of 'side information' state variables are sparsely available over disjoint spatial regions of the explored domain, and an agent must decide which of these regions to visit in which order to learn an accurate knowledge model more quickly.

4.1 Single-sensor environment

The physical range of our simulated single-sensor environment is a two-dimensional rectangular region 600 meters across. In this environment, we model temperature as our only observed variable, with no expert knowledge about the domain encoded. The true temperature distribution is a random sample drawn from a Gaussian process prior, with a Gaussian kernel having a variance of 1.0 and a lengthscale of 225 meters.

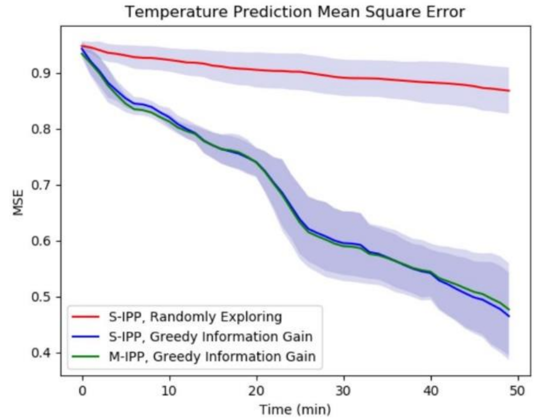

Figure 4 Mean square predictive error over the temperature field in the single-sensor environment. The mean of 10 trials is plotted with shading within 1 standard deviation of the mean.

Our agent is represented as an AUV restricted to move at 0.5 m/s, and for simplicity we assume full observability of its location and velocity as well as perfect control over its velocity. Our agent has four actions available to it: motions for one second to the North, South, East, or West. The agent employs a policy to select one of these four available actions at any given minute, take that action, update its model, and replan at each step.

We compare three types of agent in this environment: an agent employing S-IPP to greedily maximize information gain, an agent employing M-IPP to greedily maximize information gain, and an agent employing S-IPP but exploring the map with a random walk. We observe that the greedy M-IPP and greedy S-IPP models exhibit nearly identical learning, and both outperform the random-walking agent (Figure 4). Since M-IPP and S-IPP models are mathematically identical in the single-variable case, the result that they produce nearly identical learning curves is to be expected. In addition, the greedily information-maximizing agents are shown here to explore more effectively than random-walking agents in the single-sensor environment, a result commonly supported by previous work on S-IPP type algorithms [1, 2, 3].

4.2 Side-information environment

Our simulated side-information environment is the same as our single-sensor baseline, but with three additional ‘side information’ state variables which are known over the entire field before exploration begins (Figure 5). Two of these variables, sea floor depth and luminosity, are strongly linearly correlated with the temperature signal field, while the third, ocean current intensity, is uncorrelated with the temperature signal field. Temperature and ocean current are independent draws from a Gaussian process prior, with a Gaussian kernel having a variance of 1.0 and a lengthscale of 225 meters.

In M-IPP, the training inputs for our Gaussian process models are (x, y) coordinate locations. Temperature, sea floor depth, luminosity, and ocean current intensity are modeled as GP outputs. The “knowledge model”, or relationship between these outputs, is modeled as a linear-Gaussian Bayesian Network, which is the main innovation of our approach. This parametric model allows extrapolative inference on locations with partial data and new values of the output variables

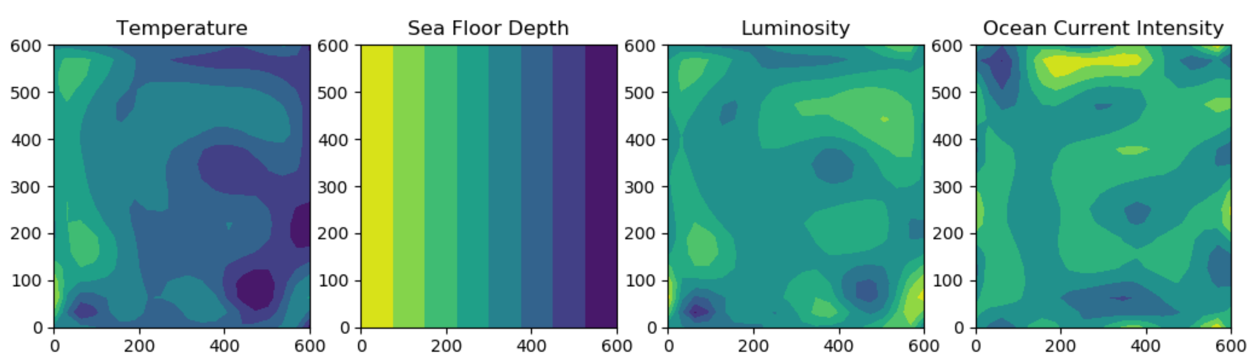

Figure 5 Contour maps of the four state variables in a sample from the posterior distribution of our side-information environment.

in a way that is appropriate for the output variables but not for the coordinate input variables, the coordinates not being physical properties of the system.

We compare this M-IPP model with Bayesian Network parameter learning to the baseline S-IPP, which assumes all output variables to be independent, as well as to an M-IPP model where the Bayesian Network parameters are provided by an expert. M-IPP initially matches the predictive accuracy of S-IPP, since its prior on the Bayesian Network parameters is independence between the outputs. As the agent explores and improves its estimate of the covariance matrix, the predictive accuracy of M-IPP with parameter learning approaches the predictive accuracy of the model with parameters fixed to their true expert-provided value. The predictive accuracy of S-IPP will also eventually approach that of the fully expert model, but is shown here to do so at a much slower rate than M-IPP with parameter learning (Figure 6).

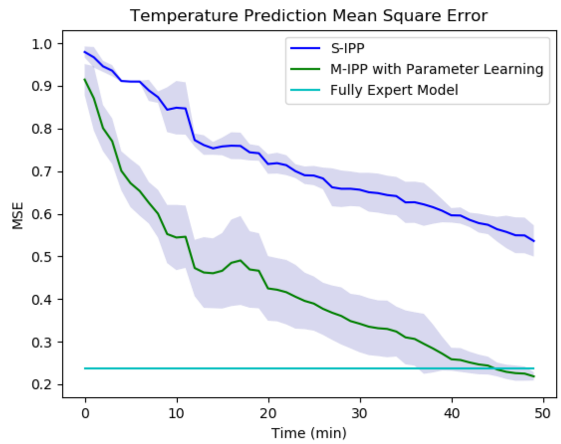

Figure 6 Mean square predictive error over the temperature field in the side information environment. The mean of 10 trials is plotted with shading within 1 standard deviation of the mean.

4.3 Sparse side-information environment

Our simulated sparse side-information environment is constructed in the same way as our side-information environment in section 4.2, but instead of being provided with measurements of the side variables (sea floor depth, luminosity, ocean current intensity) over the entire field before exploration begins, our agent is provided with the values of these variables in sparsely distributed disjoint spatial regions.

We compare two different types of agent in this environment, both of which use M-IPP as specified in section 4.2 and greedily maximize information gain. The difference between these agents lies in their information gain formulae. Agent 1 measures information gain as the reduction of entropy in a Gaussian process model of the sensed variable (temperature). This is the type of M-IPP agent demonstrated in sections 4.1 and 4.2. Agent 2 measures information gain as the sum of two terms: the first being the reduction of entropy in a Gaussian process model of the sensed variable, and the second term being the reduction of entropy in the linear-Gaussian Bayesian Network "knowledge model" of the M-IPP agent. Since this "knowledge acquisition" term of the information gain is increased by observing the sensed variable in a location where a side variable is known, the second agent is incentivized to follow a long-term plan which visits those locations where the value of a side variable has been provided, while the first agent is not incentivized to follow such a plan.

Because the "knowledge acquisition" term of agent 2's information gain is only elevated when we visit locations of sparse side information, the problem of planning to maximize the gains from this term is a problem of planning in a domain of sparse reward. In such a domain, locally greedy algorithms are generally ineffective. So, we employ Monte Carlo Tree Search to evaluate plans up to a horizon of five actions [10]. This makes planning significantly slower, but is necessary to develop plans which maximize the information gain of agent 2.

(Results in the sparse side-information environment are pending but should be available reasonably soon. I expect that Agent 2 will learn significantly faster than Agent 1, which would be a novel and interesting application result for this type of model.)

5 Conclusions and Future Work

In this work, we addressed the problem of producing a descriptive prior for informative path planning from expert-provided knowledge. Our proposed solution, MIPP, actively estimates a probabilistic map of a survey field based on expert knowledge of the GGM dependence structure between known and surveyed features. We demonstrated that MIPP learns the true parameters of the feature GGM in a survey region exhibiting strong correlation between features, and that this results in improved predictive performance for an autonomous surveillance agent under a fixed time budget.

Our result has implications for improving the resource-effectiveness of science surveys conducted with autonomous vehicles. We hope that this will enable such surveys to gather more meaningful data in remote environments where limited expert knowledge is available. The deep ocean and deep space are two exciting and large domains which humanity has yet to map in significant detail.

Several extensions in knowledge model learning may be possible beyond the algorithms presented in this paper. In our formulation of M-IPP, we assume that the structure of our GGM is provided by an expert, and actively learn the GGM parameters. The estimation of GGM structure would be a valuable extension to our method, reducing the degree of expert knowledge necessary to apply M-IPP. Tong and Koller (2000) demonstrated active learning in a structural dependence model [15]. Furthermore, a GGM is an undirected dependence graph, meaning it does not encode causality but rather correlation. An extension to M-IPP might enable an agent with interventional experiment-running capability to actively learn the direction of causal relationships between features in its environment.

Works Cited

1. Alvarez, M.A. et al. “Kernels for Vector-Valued Functions: A Review.” Foundation and Trends in Machine Learning, vol 4, no 3 (2011): 195–266.
2. Ayton, B.J. “Risk-Bounded Autonomous Information Gathering for Localization of Phenomena in Hazardous Environments.” Thesis, Massachusetts Institute of Technology (2017).
3. Binney, J. et al. “Informative path planning for an autonomous underwater vehicle.” IEEE International Conference on Robotics and Automation (2010).
4. Binney, J. and Sukhatme, G.S. “Branch and bound for informative path planning.” IEEE International Conference on Robotics and Automation (2012).
5. Browne, Cameron et al. “A Survey of Monte Carlo Tree Search Methods.” IEEE Transactions on Computational Intelligence and AI in Games, vol 4, no 1 (2011).
6. Poloczek, M. et al., “Warm Starting Bayesian Optimization.” Winter Simulation Conference (2016).
7. Krause, A. et al., “Near-Optimal Sensor Placements in Gaussian Processes: Theory, Efficient Algorithms and Empirical Studies.” JMLR (2008).
8. Lindley, D.V. “On a Measure of the Information Provided by an Experiment.”
9. Ma, Kai-Chieh, et al. “Informative Planning and Online Learning with Sparse Gaussian Processes.” IEEE Conference Publication (2017).
10. Martinez-Cantin, Ruben et al. “A Bayesian exploration-exploitation approach for optimal online sensing and planning with a visually guided mobile robot.” Auton. Robots 27 (2009): 93-103.

11. Nguyen, J.L. et al., "Nonmyopic planning for long-term information gathering with an aerial glider." IEEE International Conference on Robotics and Automation (2014).
12. Picka, Jeffrey D. "Gaussian Markov Random Fields: Theory and Applications." Technometrics 48 (2006): 146-147.
13. Rasmussen, Carl E. et al. "Gaussian Processes for Machine Learning." Advanced Lectures on Machine Learning (2009).
14. Shewry, M.C. and Wynn, H.P., "Maximum entropy sampling." Journal of Applied Statistics (1987).
15. Tong, S. and Koller, D., "Active Learning for Structure in Bayesian Networks." NIPS (2001).
16. Tong, S. and Koller, D., "Active Learning for Parameter Estimation in Bayesian Networks." NIPS (2001).
17. Williams, C. et al., "Multi-task Gaussian process prediction." NIPS (2008).