Bayesian Optimisation for Informative Continuous Path Planning

Roman Marchant
Australian Centre for Field Robotics
School of Information Technologies
University of Sydney
r.marchant@acfr.usyd.edu.au

Fabio Ramos
Australian Centre for Field Robotics
School of Information Technologies
University of Sydney
f.ramos@acfr.usyd.edu.au

1 Introduction

We present a novel approach for informative path planning applied to environmental monitoring using *Bayesian Optimisation* (BO) [1]. Autonomous robots are able to traverse over a wide range of environments and sense from several phenomena [2]. However, the problem of where and when to gather the most informative samples in an efficient manner is still an open question. This paper addresses the problem by finding informative paths in a continuous domain, solving not only the question of *where* and *when* to sample, but *how* to get there.

The decisions are chosen using a multi-layered BO algorithm. The first layer optimises over the environmental phenomenon to finds areas of interest (e.g. high pollution, high temperature). This layer uses an incremental spatial–temporal model of the phenomenon, given by a *Gaussian Process* (GP) [3–7] prior that takes into account the uncertainty and predicted values propagated in time. The second layer of BO is used to find the best set of parameters that determine a continuous path where the robot travels on.

This work builds on earlier work by the authors [8], and presents the following contributions:

1. Generalisation of the BO algorithm to optimise along continuous trajectories instead of discrete locations;
2. A layered BO for informative path planning in spatial–temporal environmental monitoring.

The proposed method has the following advantages over the previous techniques [8–15]: i) It is not a way-point greedy solution to acquiring new observations as it takes into account measurements obtained along a path with predictions propagated over time; ii) It considers a continuous action space; iii) It uses both the mean and the variance to define paths and addresses the exploration-exploitation trade off in principled Bayesian framework.

We validate our algorithm on a large-scale environment for monitoring ozone concentration in the US, and on a mobile robot that monitors the dynamics of luminosity changes.

2 Continuous Path Planning

A Bayesian Optimisation method is derived to estimate continuous paths for sampling an initially unknown environmental phenomenon. In this section we describe the general BO algorithm, to later present details on the specific algorithm for planning over continuous paths.

2.1 Bayesian Optimisation

BO is used for finding the optimal (maximum or minimum) of an unknown and costly to evaluate function f, i.e. find $\mathbf{x}^\star = \arg\max_{\mathbf{x}} f(\mathbf{x})$. To achieve this, it builds an statistical model of the unknown function using samples from it and a GP prior. An acquisition function h is evaluated over the statistical model and guides the search for the optimum. The procedure requires the maximisation of h at each iteration which is usually a much simpler optimisation problem. The generic algorithm

Algorithm 1 General BO (Discrete Locations)

1: **function** BO(f, h)
2:     **for** $t = 1, 2, 3, \ldots$ **do**
3:         Find $\mathbf{x}_t = \arg\max_{\mathbf{x}} h(\mathbf{x})$
4:         $y_t \leftarrow f(\mathbf{x}_t)$
5:         Augment training set with $(\mathbf{x}_t, y_t)$
6:         Update GP
7:     **end for**
8: **end function**

Algorithm 2 Path BO (Continuous Paths)

1: **function** PATHBO(f, h, q)
2:     **for** $t = 1, 2, 3, \ldots$ **do**
3:         $\boldsymbol{\beta}^\star \leftarrow$ BO$(r(h), q)$ $\triangleright$ Algorithm 1
4:         $\{x, y\} \leftarrow$ Sample along $\mathcal{C}(u, \boldsymbol{\beta}^\star)$
5:         Augment training set with $\{x, y\}$
6:         Update GP
7:     **end for**
8: **end function**

for BO is shown in Algorithm 1. Line 3 is the inner optimisation conducted within BO and can be solved using deterministic gradient-free optimisers [1].

The most common acquisition functions are: Probability of Improvement [16], Expected Improvement [16] and Upper Confidence Bound [17]. Dynamic allocation of these acquisition function have also been studied by [18].

For space-time problems, inputs $\mathbf{x} = (\mathbf{s}; t)$ transform an ordinary acquisition function into a space-time acquisition function.

2.2 BO for Continuous Path Planning

Let $\mathcal{C}(u, \boldsymbol{\beta})$ be a curve with u a real number $\in [0, 1]$ and parametrized by $\boldsymbol{\beta} \in \mathbb{R}^m$, such that $C : [0, 1] \rightarrow \mathbb{R}^N$, where m is the number of dimensions of the parameter space and N is the dimensionality of the input space of f. To evaluate the score, r, of a path $\mathcal{C}$, we integrate the acquisition function h over it,

$$r(\mathcal{C}(u, \boldsymbol{\beta})|h) = \int_{\mathcal{C}(u,\boldsymbol{\beta})} h(v)\mathrm{d}v. \tag{1}$$

For example, lets consider $h = \text{UCB}$, $N = 2$

$$r(\mathcal{C}(u, \boldsymbol{\beta})|h = \text{UCB}) = \int_{\mathcal{C}(u,\boldsymbol{\beta})} \text{UCB}(v)\mathrm{d}v \tag{2}$$

$$r(\mathcal{C}(\boldsymbol{\beta})|h = \text{UCB}) = \int_0^1 \text{UCB}(\mathcal{C}(u, \boldsymbol{\beta}))\|\mathcal{C}'(u, \boldsymbol{\beta})\|\mathrm{d}u \tag{3}$$

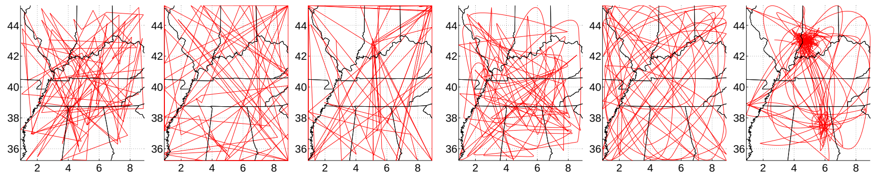
$$r(\mathcal{C}(\boldsymbol{\beta})|h = \text{UCB}) = \int_0^1 \left[\mu(\mathcal{C}(u, \boldsymbol{\beta})) + \kappa\sigma(\mathcal{C}(u, \boldsymbol{\beta}))\right] \|\mathcal{C}'(u, \boldsymbol{\beta})\|\mathrm{d}u. \tag{4}$$

The integral in equation 1 does not always have an analytical solution, depending on the definition of the acquisition and covariance functions. In this paper we use a rectangle rule quadrature-based approximation [19], which generally results in accurate approximations for the one dimensional case (since the integral is over a 1-D variable, u).

To use BO for finding continuous paths a modification to Algorithm 1 is required: Instead of finding a discrete location for taking the next sample from f (Line 3 of Algorithm 1), we find the parameters $\boldsymbol{\beta}^\star$ that define a continuous path over space and time that cumulatively delivers the best return by integrating over the acquisition function. To find the *best* set of parameters $\boldsymbol{\beta}^\star$ that defines the curve that maximises the integral over h, the following optimisation problem has to be solved:

$$\boldsymbol{\beta}^\star = \arg\max_{\boldsymbol{\beta}} r(\mathcal{C}(u, \boldsymbol{\beta})|h). \tag{5}$$

In section 2.1, the optimisation required to find the highest value of the acquisition was performed using a deterministic, derivative–free optimiser. Alternatively, equation 5 can be solved using another layer of BO as it is shown in Algorithm 2, Line 3. Since the action space $\boldsymbol{\beta}$ is five dimensional (for 2D cubic splines) and the function r is highly non-convex and expensive to evaluate, BO provides a natural solution. This second layer of BO uses a GP prior over r and a second acquisition function q to decide which path parameters to evaluate over r. This second layer of BO is over discrete locations and uses the original Algorithm 1. Note that Algorithm 2 is fast to evaluate because it does not require the robot to move and gather training samples as it uses the existing GP model of the phenomenon.



(a) RD (b) ED (c) UCBD (d) RC (e) EC (f) UCBC

Figure 1: Paths for different methods.

2.3 Path Parametrisation

In this paper we use cubic splines restricted to a two dimensional plane ($N = 2$). Therefore, the curve $\mathcal{C} = (\mathcal{C}_x(u, \boldsymbol{\beta}), \mathcal{C}_y(u, \boldsymbol{\beta}))$ is defined as:

$$\mathcal{C}_x(u, \boldsymbol{\beta}) = a_x u^3 + b_x u^2 + c_x u + d_x \tag{6}$$

$$\mathcal{C}_y(u, \boldsymbol{\beta}) = a_y u^3 + b_y u^2 + c_y u + d_y \tag{7}$$

with $\boldsymbol{\beta} = \{a_x, b_x, c_x, d_x, a_y, b_y, c_y, d_y\}$. When estimating a path to follow, the current state of the robot $\mathbf{x} = (p_x, p_y, \alpha)$ is assumed known. This defines the boundary condition reducing the action space to five free parameters, $\boldsymbol{\beta} = \{a_x, a_y, b_x, b_y, c_x\}$, only. To take into account the temporal dimension over the curve we assume that the robot travels at a constant speed. Note that our method accepts modifications to the curve parametrisation and is not strongly linked to this particular spline model.

3 Experiments

3.1 Large-scale pollution monitoring

The first experiment simulates an Unmanned Air Vehicle (UAV) monitoring ozone concentration; considered a pollutant at ground level. To simulate the environment we use a GP regression over ozone concentration measurements dating back to 1987 provided by the US Environment Protection Agency[1]. It can be noted that two peaks appear around mid-day with values that can reach up to 100ppb for the studied area. We compare six different techniques for planning the motion of the UAV while monitoring the environment: a) Random Discrete Sampling (RD); b) Entropy Discrete Sampling (ED); c) UCB Discrete Sampling (UCBD); d) Random Continuous Sampling (RC); e) Entropy Continuous Sampling (EC); f) UCB Continuous Sampling (UCBC) (Our method).

For UCB $\kappa = 0.1$ was tuned manually to balance the exploration-exploitation trade off. All strategies collect the same number of samples, therefore the differences in error will only depend on the locations where the samples were acquired. The inner optimisation for maximising among paths (Line 3 of Algorithm 2) uses $q =$ UCB as acquisition function for strategies EC and UCBC. The GP model of the inner optimisation uses a Matern3 covariance function whose hyper-parameters are optimised on each iteration using gradient decent.

Figure 1 shows the paths travelled by the robot for each case. A quick visual inspection shows that all methods were able to cover the region of interest and explore the entire environment. Entropy based techniques (ED and EC) cover the region uniformly, reducing the uncertainty of the whole area. Finally, UCBD and UCBC concentrate their samples towards the areas of higher pollution.

A very important difference is the shape of paths for discrete and continuous sampling strategies. Even though κ has the same value for the acquisition function of UCBC and UCBD, the trajectories are much more concentrated over the high pollution areas for the continuous optimisation case (UCBC). The main reason is that this method takes into account the value of the acquisition function over the entire path that is being traversed. One way of seeing this is that if a method only takes into account a discrete goal location it will not necessarily collect useful information on its way to the target location. However, if the method does take into account the information gathered while reaching the target location, then the informativeness of gathered samples will increase noticeably.

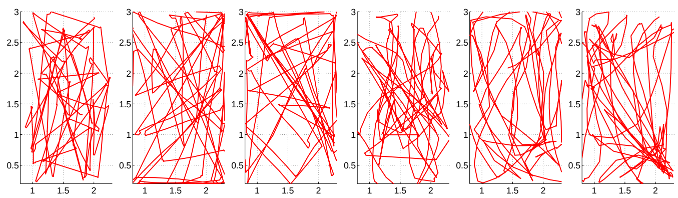
Table 1 shows the error for each method evaluated w.r.t. the ground truth on a grid over space and time for the entire duration of the experiment. It can be seen that the proposed method (UCBC)

[1] `http://java.epa.gov/castnet/reportPage.do`

Method	RMSE	WRMSE	LogLoss	WLogLoss
RD	7.3574	2.2324	3.4149	0.0956
ED	7.5225	2.2570	3.4332	0.0957
UCBD	7.1579	1.9764	3.3999	0.0937
RC	7.2238	2.0698	3.3951	0.0935
EC	7.1103	2.2958	3.3907	0.0956
UCBC	**6.7971**	**1.4981**	**3.3537**	**0.0863**

Method	RMSE	WRMSE	LogLoss	WLogLoss
RD	39.598	28.061	8.949	2.327
ED	38.389	25.106	10.013	2.779
UCBD	38.210	23.715	9.685	2.672
RC	43.390	27.045	10.197	2.754
EC	48.873	38.433	11.980	3.443
UCBC	**30.121**	**21.422**	**8.098**	**2.145**

Table 1: Results for US Ozone Monitoring (Left)



(a) RD (b) ED (c) UCBD (d) RC (e) EC (f) UCBC

Figure 2: Resulting paths for six different path planning techniques, axis in meters.

delivers the best performance for all indicators. The difference in performance between strategies is remarkable for the weighted errors WRMSE and WLogLoss. The main reason for this is the extra importance to model areas with high pollution (exploitative behaviour). It is also noticeable the difference between continuous and discrete sampling strategies. An improvement is revealed for all strategies as we are optimising over continuous paths rather than choosing discrete locations.

3.2 Luminosity monitoring

A small mobile robot was used to monitor dynamic illumination changes in an indoor environment. Two light sources with variable intensity are dimmed electronically to expose patterns with a periodic component and amplitude changes through time. The robot is gathers samples from the phenomenon every one second. Ground truth is obtained by placing static sensor boards in five locations. The same path-planning strategies from section 3.1 are compared in experimental trials that last for ten minutes.

Figure 2 shows the paths travelled by the robot using each technique. Results are similar to the experiment in the previous section. While random sampling strategies, RD and RC, derive paths mostly concentrated at the centre of the studied region, paths for entropy based strategies, ED and EC, are distributed more homogeneously over space. In constrast, UCB paths focus on areas with high luminosity while at the same time exploring the environment for unknown sources of light.

Table 1 shows numerical results. UCBC delivers the lowest error and weighted error for all the indicators. It is also shown that planning over continuous domains results in smaller error for the case of UCB acquisition function. UCB strategies have the smallest error, demonstrating a central advantage in monitoring dynamic phenomena: monitoring areas of higher pollution more intensively results on lower overall error.

The developed method takes 4.8s in each iteration, close to real-time in a standard cpu (i5 processor).

4 Discussion

This paper proposed a new technique for informative path planning over continuous paths for environmental monitoring. The main contribution is the derivation of a continuous action space strategy by integrating over an acquisition function in a principled Bayesian optimisation framework. We model space-time phenomena using Gaussian processes which enables a robot to learn periodic patterns while preserving spatial correlations between observations. A first layer of BO is used to predict regions of high concentration. Then, a second layer of BO is used to estimate the curve parameters defining the best path to collect new observations. We believe that optimising over curves for path planning can produce more informative decisions achieving longer term rewards. The method explained in this paper can significantly improve the decision making process for efficiently monitoring a wide variety of environmental phenomena.

References

[1] E. Brochu, V. M. Cora, and N. de Freitas, "A tutorial on bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning," University of British Columbia, Tech. Rep. arXiv:1012.2599, 2010.

[2] M. Dunbabin and L. Marques, "Robotics for environmental monitoring," *IEEE Robotics Automation Magazine*, vol. 19, no. 1, pp. 24–39, 2012.

[3] C. E. Rasmussen and C. Williams, *Gaussian processes for machine learning*. The MIT Press, Cambridge, Massachuset, 2006.

[4] A. Singh, F. Ramos, H. D. Whyte, and W. J. Kaiser, "Modeling and decision making in spatio-temporal processes for environmental surveillance," in *IEEE Conference on Robotics and Automation (ICRA)*, 2010.

[5] S. O'Callaghan and F. Ramos, "Gaussian Process Occupancy Maps," *International Journal of Robotics Research*, 2012.

[6] C. Stachniss, C. Plagemann, and A. J. Lilienthal, "Gas distribution modeling using sparse Gaussian process mixture models," in *Robotics Science and Systems (RSS)*, 2008.

[7] S. Vasudevan, F. Ramos, E. Nettleton, and H. Durrant-Whyte, "Gaussian Process Modelling of Large-Scale Terrain," *Journal of Field Robotics*, 2009.

[8] R. Marchant and F. Ramos, "Bayesian Optimisation for Intelligent Environmental Monitoring," in *IEEE Conference on Intelligent Robots and Systems (IROS)*, 2012.

[9] C. Guestrin, A. Krause, and A. P. Singh, "Near-optimal sensor placements in Gaussian processes," in *International Conference on Machine Learning (ICML)*, vol. 1, 2005.

[10] A. Krause and C. Guestrin, "Near-optimal Observation Selection using Submodular Functions," in *AAAI Conference on Artificial Inteligence*, 2007.

[11] J. J. Chung, N. R. Lawrence, and S. Sukkarieh, "Gaussian Process for Informative Exploration in Reinforcement Learning," in *IEEE Conference on Robotics and Automation (ICRA)*, 2013.

[12] L. A. Bush, B. Williams, and N. Roy, "Computing Exploration Policies via Closed-form Least-Squares Value Iteration," in *International Conference on Automated Planning and Scheduling (ICAPS)*, 2008.

[13] A. Singh and A. Krause, "Nonmyopic adaptive informative path planning for multiple robots," in *International Joint Conference on Artificial Inteligence (IJCAI)*, 2009.

[14] G. A. Hollinger, B. Englot, F. Hover, U. Mitra, and G. S. Sukhatme, "Uncertainty-driven view planning for underwater inspection," in *IEEE Conference on Robotics and Automation (ICRA)*, 2012.

[15] J. Suh and S. Oh, "A Cost-Aware Path Planning Algorithm for Mobile Robots," in *IEEE Conference on Intelligent Robots and Systems (IROS)*, 2012.

[16] D. R. Jones, "A taxonomy of global optimization methods based on response surfaces," *Journal of Global Optimization*, pp. 345–383, 2001.

[17] D. D. Cox and S. John, "A Statistical Method for Global Optimization," in *IEEE Conference on Systems, Man, and Cybernetics (SMC)*, 1992.

[18] M. Hoffman, E. Brochu, and N. de Freitas, "Portfolio allocation for bayesian optimization," in *Conference on Uncertainty in Artificial Intelligence (UAI)*, 2011.

[19] W. Gauschi, *Numerical Analysis*. Springer New York, 2012.