SCIENCE CHINA
Information Sciences

May 2019, Vol. 62 050208:1–050208:3
https://doi.org/10.1007/s11432-018-9653-7

• LETTER •
Special Focus on Human-Robot Hybrid Intelligence

Long-term adaptive informative path planning for scalar field monitoring using cross-entropy optimization

Yang LI[1,2], Rongxin CUI[2*], Weisheng YAN[1,2] & Demin XU[2]

[1]*Research & Development Institute of Northwestern Polytechnical University in Shenzhen, Shenzhen* 518057*, China;*
[2]*School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an* 710072*, China*

Received 14 September 2018/Accepted 7 October 2018/Published online 27 February 2019

Citation Li Y, Cui R X, Yan W S, et al. Long-term adaptive informative path planning for scalar monitoring using cross-entropy optimization. Sci China Inf Sci, 2019, 62(5): 050208, https://doi.org/10.1007/s11432-018-9653-7

Dear editor,

Robots are extensively used in various civil and military applications [1, 2]. In the last decade, informative path planning (IPP) has been one of the most important research areas in modern robot engineering [3], especially in long-term monitoring of a particular area of interest. In such scenarios, it is expected that an optimal path will be chosen along with the collection of maximal information about the scalar field with a limited fuel budget. To the best of our knowledge, state-of-the-art IPP methods do not consider adaptive long-term scalar field monitoring tasks where the reliability on historical data is gradually decreasing. The Gaussian process (GP) is an attractive tool because its mean and covariance functions can describe a given scalar field. In this study, we design a time-varying likelihood of measurements to handle the concept of descending belief. Moreover, we utilize a clamped B-spline curve to analytically parameterize the continuous path. Here, the path is determined by a sequence of control points. When new measurements are received, an adaptive replanning scheme that makes the trade-off between information gain and fuel budget is required [4]. Further, the results of simulations are depicted to display the effectiveness of our algorithms.

Modified Gaussian process. The Gaussian process is a powerful non-parameter tool used for describing an underlying scalar field that can be formulated as $f : \mathbb{R}^d \to \mathbb{R}$ on d-dimensional space $\mathcal{X} \subseteq \mathbb{R}^d$ [5]. Here, we focus on a scalar workspace where $d = 2$. A typical Gaussian process is defined as

$$f(x) \sim \mathcal{GP}(m(x), k(x, x')), \tag{1}$$

where $m(x)$ is the mean function and generally set to zero. $k(x, x')$ denotes the corresponding covariance function, which describes the relation between x and x'.

In practice, the scalar filed $f(x)$ is defined only through noisy measurements $y = f(x) + \varepsilon$, where ε is generally white Gaussian noise, i.e., $\varepsilon \sim \mathcal{N}(0, \sigma_n^2)$. This implies that the likelihood of measurements $y(x)$ is also Gaussian. Unfortunately, in our long-term monitoring application, the previously collected data would be inadmissible. This certainly results in a non-Gaussian likelihood of $y(x)$.

To consider the non-Gaussian likelihood, we design a new covariance function $k(x, x')$ as

$$k(x,x') = \begin{cases} \sigma_f^2 \exp(-\frac{1}{2}(x-x')^{\mathrm{T}} M (x-x')) \\ +\sigma_n^2 \min\{m_{\mathrm{cov}}, \exp(t_c^2 t(x)^2)/(2\nabla_t^2)\}, \\ \qquad \text{when} \quad x = x', \\ \sigma_f^2 \exp(-\frac{1}{2}(x-x')^{\mathrm{T}} M (x-x')), \\ \qquad \text{other}, \end{cases} \tag{2}$$

where $\theta_{\mathcal{GP}} = \{\sigma_f, M, \sigma_n\}$ is the Gaussian process hyperparameter set. Specifically, $M = \mathrm{diag}(l)^{-1}$,

* Corresponding author (email: r.cui@nwpu.edu.cn)

where l is a positive vector. $\theta_{\mathcal{GP}}$ can be obtained using simple descending algorithms.

$\theta_t = \{t_c, \nabla_t\}$ is the time hyperparameter set of the Gaussian process. $t(x) \geqslant 0$ is the time interval between now and the point of collection. Importantly, the influence of time-varying measurement reliability is decreasing and limited to $m_{\text{cov}}\sigma_n^2$. It is important to note that if maximum $m_{\text{cov}}\sigma_n^2$ is achieved, the corresponding measurements will be discarded.

Suppose we already have H history measurements $Y_H = \{y_{h1}, y_{h2}, \ldots, y_H\}$ at locations $X_H = \{x_{h1}, x_{h2}, \ldots, x_H\}$, and we want to predict values at locations of interest $X_T = \{x_{t1}, x_{t2}, \ldots, x_T\}$. Therefore, using properties of Gaussian process, we can obtain the posterior distribution of $f(X_T)$ using

$$\begin{aligned}\hat{f}_{X_T} &= K(X_T, X_H) \\ &\quad \times [K(X_H, X_H) + \sigma_n^2 I]^{-1}(Y_H - m(X_H)), \\ \text{cov}(f_{X_T}) &= K(X_T, X_T) - K(X_T, X_H) \\ &\quad \times [K(X_H, X_H) + \sigma_n^2 I]^{-1} K(X_H, X_T).\end{aligned} \tag{3}$$

Given the model of the scalar field, we utilize a mutual information method to quantize the value of collected measurements. Suppose we have F forthcoming sampling locations on the candidate path $X_F = \{x_{f1}, x_{f2}, \ldots, x_F\}$. The information gain by X_F can be formulated using

$$\begin{aligned}&I(f(X_T), Y(X_F)) \\ &= H(f(X_T)|X_H, Y_H) - H(f(X_T)|X_H, Y_H, X_F, Y_F) \\ &= \frac{1}{2}\log\frac{|\text{cov}(f(X_T)|X_H, Y_H)|}{|\text{cov}(f(X_T)|X_H, Y_H, X_F, Y_F)|}.\end{aligned} \tag{4}$$

Adaptive informative path planning. Clamped B-spline curve's continuous availability of any order derivatives makes it a powerful tool that can be used to represent a continuous path [6]. It is parameterized by a sequence of n control points $X_C = \{x_{c1}, x_{c2}, \ldots, x_C\}$, which is formulated using

$$\begin{aligned}\tau(w) &= \sum_{i=c1}^{C} x_i B_{i,P}(w), \quad \text{where} \\ B_{i,0}(w) &= \begin{cases} 1, \text{ if } v_i < w < v_{i+1}, \\ 0, \text{ otherwise}, \end{cases} \\ B_{i,p}(w) &= \frac{w - v_i}{v_{i+1} - v_i} B_{i,p-1}(w) \\ &\quad + \frac{v_{i+p+1} - w}{v_{i+p+1} - v_{i+1}} B_{i+1,p-1},\end{aligned} \tag{5}$$

where $B_{i,P}$ is the basis function, and $\{v_i\} \in [0, 1]$ is the knot vector, which is equally divided in $[0, 1]$. Briefly, we also define the curve as τ_{X_C}, and the first derivative can be formulated as a B-spline curve using

$$\frac{\text{d}\tau(w)}{\text{d}w} = \sum_{i=c1}^{C-1} B_{i+1,P-1} Q_i, \tag{6}$$

where $Q_i = \frac{P}{v_{i+P+1} - u_{i+1}}$. The measurements are collected with a fixed interval l. Thus, the sampling locations are obtained using

$$X_S = \left\{ \tau(w_i) \middle| \int_{w_{i-1}}^{w_i} \left| \frac{\text{d}\tau(w)}{\text{d}w} \right| \text{d}w = l \right\}, \tag{7}$$

where $w_0 = 0$.

Note that the B-spline curve is hardly divided into exact segments of length l. If the curve is partitioned into several sub-curves with length l, and the last one is less than l then the last location is abandoned. Now, we introduce our informative path planning scheme. First, the local sequence of control points X_C is optimized according to our designed objective function $O(X_C)$ in (8) using the cross-entropy method.

$$O(X_C) = I(X_C) u_{\text{obs}}(X_C) u_c(X_C), \tag{8}$$

where $I(X_C)$ is the mutual information carried by X_C. u_{obs} is the penalty of obstacles. $u_{\text{obs}}(X_C) = \frac{1}{1+\exp\{-\lambda_{\text{obs}} \min(|\tau_{X_C} - X_{\text{obs}}| - \epsilon_{\text{obs}})\}}$. $u_c(X_C)$ is the penalty of the limited resource. $u_c = \frac{1}{1+\exp\{-\lambda_c(c(\tau_{X_C}) + \tilde{c}(\tau_{X_C}) - c_{\text{left}} - \epsilon_c)\}}$. X_{obs} denotes the obstacle region. ϵ_{obs} is the obstacle clearance parameter. $c(\tau_{X_C})$ denotes the length of τ_{X_C}, and $\tilde{c}(\tau_{X_C})$ approximates the length for returning to start location. c_{left} is the remaining source. Second, the measurements are collected along the path, which is determined by X_C. Finally, the posterior estimation of the scalar field is obtained based on (3).

Because we have quantified the information about the scalar field in terms of $I(X_C)$, it is critical to dynamically choose X_T according to the collected data. Instinctively, measurements are taken where the variance is high. Simultaneously, the total variance of the scalar field is kept at a low level. We divide X_T into two sets: X_{T_s} and X_{T_d}. X_{T_s} is fixed and averagely distributed in $\mathcal{X}$. Contrastingly, X_{T_d} is adaptively chosen in the high posterior variance region. We generate X_{T_d} with distribution expressed as

$$p(x_{T_d}) = \frac{1}{\text{cov}(f_{x_{T_d}})} \Big/ \int_{\mathcal{X}} \frac{1}{\text{cov}(f_{x_{T_d}})}. \tag{9}$$

Here we formally propose the long-term adaptive IPP algorithm in Algorithm 1.

Algorithm 1 Long-term adaptive IPP algorithm

Input: starting point x_{start}, planning horizon ϱ, historical sequence of control points X_H, obstacle region $\mathcal{X}_{\text{obs}}$.
Output: posterior estimation of scalar field $\tilde{f}$, local optimal path τ_{X_C}.

```
[μ, s_2] ← init_para(X_H, ϱ);
local optimal sequence of control points distribution
[μ*, s_2^*] ← CEoptimize(μ, s_2, X_obs, c_left, X_T);
X_C ← Sample(μ*, s_2^*); X_F ← τ_{X_C};
f̃ ← solveGP(X_H, X_F);
adaptive re-plan X_{T_d} ← Sample(p(x_{T_d})).
```

The number of historical measurements X_H is increasing with in-situ data accumulation, which results in infinite dimensional X_H. Thus, X_H is truncated to a fixed scale according to the time label of the measurements. The scale is empirically chosen, where a trade-off of estimation error and computation efficiency must be managed.

Cross-entropy optimization was first proposed in [7]. It is widely used in multi-extremal optimization problems, which do not require a convex target function. Thus, it is appropriate in our path optimization problem. Suppose we need to maximize $O(X_C)$ in (8) and write the maximum as $\gamma^* = O^*(X_C) = \max_{X\subseteq\mathcal{X}} O(x)$. We define a family of pdfs $\{\varsigma(X, v), v \in \mathcal{V}$ is pdfs parameter$\}$ on $\mathcal{X}$. The probability of the rare event $\{O(X_C) \geqslant \gamma\}$ can be formulated as $l(\gamma) = \mathbb{P}_v(O(X) \geqslant \gamma) = \mathbb{E}_v I(O(X) \geqslant \gamma)$. Thus, it is possible to optimize the pdfs family parameter v, until the rare event probability is small enough. We formally present the optimization algorithm in Algorithm 2.

Algorithm 2 Cross-entropy optimization algorithm

Input: v_1, quantile η, size N, max iterations M.
Output: pdf of x^*, i.e., $g(v^*)$.

```
for t = 2, ..., M do
    generate samples from pdf g(X, v_{t−1}), and sort as
    O_1 ⩽ O_2 ⩽ ··· ⩽ O_N;
    γ_t ← O_{⌈(1−η)N⌉};
    v_t ← arg max_v (1/N) I{O(X) ⩾ γ_t} ln g(X; v);
    if γ_t < γ_{t−1} then
        break;
    end if
end for
```

Simulations. Here, the proposed algorithms are demonstrated using simulations. The scalar field is randomly generated in an area of 3 m× 3 m. Suppose the start location is $[0, 0]$. Four circle obstacles with the same radius $r = 0.5$ m are separately located at $[1, 1]$, $[1, 2]$ $[2, 1]$, $[2, 2]$. All parameters include three parts, i.e., GPs, cross-entropy optimization, and path curve, where $t_c = 1$, $\nabla_t = 2$, $\eta = 0.1$, $N = 20$, $n = 4$, $l = 3$. As shown in Figure 1, the estimation variance is decreasing along with the optimized path.

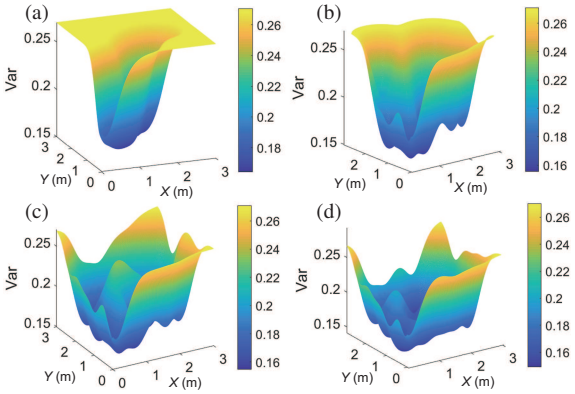

Figure 1 (Color online) Estimation variance at different iterations. (a) Iteration = 1; (b) iteration = 5; (c) iteration = 10; (d) iteration = 15.

Conclusion. Here, we present a new long-term adaptive IPP algorithm for scalar field monitoring in which time influence and energy limitations are considered in a long-term situation. Additionally, the interested target locations are adaptively chosen when new measurements are collected; and the cross-entropy method is utilized to obtain the local optimal path segment, which is more appropriate in online cases. Thus, it is more practical for real-world applications.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant Nos. U1813225, 61472325, 61733014, 51579210) and Science, Technology and Innovation Commission of Shenzhen Municipality (Grant No. JCYJ20170817145216803).

References

1 Li Z, Huang B, Ye Z, et al. Physical human-robot interaction of a robotic exoskeleton by admittance control. IEEE Trans Ind Electron, 2018, 65: 9614–9624
2 Pan Y, Yang C, Pan L, et al. Integral sliding mode control: performance, modification, and improvement. IEEE Trans Ind Inf, 2018, 14: 3087–3096
3 Ma K C, Liu L, Heidarsson H K, et al. Data-driven learning and planning for environmental sampling. J Field Robotics, 2018, 35: 643–661
4 Yang C, Jiang Y, He W, et al. Adaptive parameter estimation and control design for robot manipulators with finite-time convergence. IEEE Trans Ind Electron, 2018, 65: 8112–8123
5 Seeger M. Gaussian processes for machine learning. Int J Neur Syst, 2004, 14: 69–106
6 Yang K, Moon S, Yoo S, et al. Spline-based RRT path planner for non-holonomic robots. J Intell Robot Syst, 2014, 73: 763–782
7 Reuven R. The cross-entropy method for combinatorial and continuous optimization. Methodol Comput Appl Probabil, 1999, 1: 127–190