Active Learning with Gaussian Processes for High Throughput Phenotyping

Extended Abstract

Sumit Kumar, Wenhao Luo, George Kantor and Katia Sycara

ABSTRACT

A looming question that must be solved before robotic plant phenotyping capabilities can have significant impact to crop improvement programs is scalability. High Throughput Phenotyping (HTP) uses robotic technologies to analyze crops in order to determine species with favorable traits, however, the current practices rely on exhaustive coverage and data collection from the entire crop field being monitored under the breeding experiment. This works well in relatively small agricultural fields but can not be scaled to the larger ones, thus limiting the progress of genetics research. In this work, we propose an active learning algorithm to enable an autonomous system to collect the most informative samples in order to accurately learn the distribution of phenotypes in the field with the help of a Gaussian Process model. We demonstrate the superior performance of our proposed algorithm compared to the current practices on sorghum phenotype data collection.

KEYWORDS

Robot Learning; Gaussian Process; Active Learning; Adaptive Sampling; Crop Phenotyping

ACM Reference Format:

Sumit Kumar, Wenhao Luo, George Kantor and Katia Sycara. 2019. Active Learning with Gaussian Processes for High Throughput Phenotyping. In *Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), Montreal, Canada, May 13–17, 2019,* IFAAMAS, 3 pages.

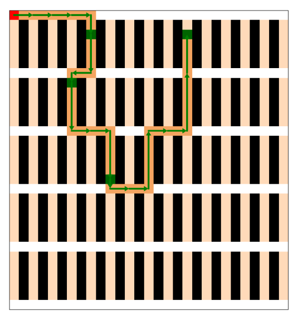
1 INTRODUCTION

In conventional adaptive sampling [6–10, 12] tasks, the robot sequentially selects a limited number of locations to collect high utility data and plans a path to reach them by optimizing a cost function like path length. The prediction of locations with high utility data relies on the utility collected from the visited locations so far during sampling. In HTP, phenotypes like plant height, stalk width, etc. are measured with the help of computer vision techniques [1, 3, 4] on images captured by on-board cameras. Note that images and hence phenotype measurements can be gathered even when the robot is in motion. This is different from conventional adaptive sampling where there is only a single source of data acquisition. In contrast, in this work, we consider two types of measurements:

- *Static* measurements: The robot stops at a sampling location to get accurate measurements (conventional sampling).
- *Mobile* measurements: As the robot is moving, it collects data from images captured from locations along its path.

Naturally, the static measurements are more accurate than the mobile ones, however, gathering them requires more time and resources than the latter. As a result, there is a trade-off between the quality of data and the required resources. In this work, we present an active learning framework where the agent first selects a set of locations with high utility data (static samples) and then plans a path to maximize joint information gain from both static and mobile samples gathered along the way.

Mueller et al. [11] presented a robotic platform for HTP. However, its efficiency can be significantly increased by collecting high utility data in a short time. Also, the amount of data needed to analyze crop genetics can be reduced by learning the distribution of phenotypes so as to estimate unobserved data as shown in this work.



(b) **Robotic platform**

(a) **Environment**

Figure 1: **(a) The robot (red cell) determines the p most informative samples (here $p = 4$) in the field which are shown in green and finds the most informative path (green arrows). (b) Robotic platform used for collecting data.**

2 GAUSSIAN PROCESS MODEL

Let $V = \{v_1, \ldots, v_n\}$ be the set of all sampling data points where $v_i = \{\text{location, vegetation index, leaf angle density}\} \in \mathbb{R}^d$ and $y = \{\text{stalk height}\}$ represents an observable feature vector. There exists a latent function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ that maps the input $v \in \mathbb{R}^d$ to the objective value $f(v)$. After sampling a set $D \subset V$ and observing the output $Y = \{y(v) \mid v \in D\}$, the robot uses GP regression to learn the underlying mapping f assuming the joint distribution of the observed readings is Gaussian. A GP is a distribution over functions fully defined by a mean function m and a covariance function k taken to be Matern Kernel with parameter $\nu = 1.5$ in this work.

The robot may have collected multiple mobile measurements for the same data point v, we combine them all into an equivalent mean measurement $\hat{y}_m(v)$. A robot may also have acquired a static measurement $y_s(v)$ for a data point v for which it has gathered mobile measurement(s) before. We fuse them together as the product

The authors are with The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Email: {skumar2, wenhao, kantor, katia}@cs.cmu.edu.

This work was supported in part by U.S. Department of Agriculture (DOA) award 2017-6700-726152.

Proc. of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2019), N. Agmon, M. E. Taylor, E. Elkind, M. Veloso (eds.), May 13–17, 2019, Montreal, Canada.

of the two probability density functions:

$$y(v) = \frac{\frac{y_s(v)}{\sigma_s^2} + \frac{\hat{y}_m(v)}{\sigma_m^2}}{\frac{1}{\sigma_s^2} + \frac{1}{\sigma_m^2}}; \quad \sigma^2(v) = \frac{\sigma_s^2 \sigma_m^2}{\sigma_s^2 + \sigma_m^2}$$

where σ_s^2 and σ_m^2 denote the static and mobile measurement variances respectively. Note that $\sigma_m^2 > \sigma_s^2$. The inter-sample covariance is modelled as:

$$\tilde{k}(v_i, v_j) = k_{\text{Matern}}(v_i, v_j) + \sigma^2(v_i)\delta_{ij} \tag{1}$$

where δ_{ij} is the Kronecker delta. The posterior target distribution of a set of samples $A \subset V$ conditioned on the sampled set D can be written as $f(A) \mid A, D, Y \sim \mathcal{N}(\mu_{A|D}, \Sigma_{A|D})$ where $\mu_{A|D} = \Sigma_{AD}\Sigma_{DD}^{-1}Y$ and $\Sigma_{A|D} = \Sigma_{AA} - \Sigma_{AD}\Sigma_{DD}^{-1}\Sigma_{DA}$. Σ_{AB} is pairwise covariance matrix. Also, the entropy of a set A conditioned on a sampled set D is $H(A|D) = \frac{1}{2}\log\left((2\pi e)^{|A|}\det\left(\Sigma_{A|D}\right)\right)$.

3 INFORMATIVE PLANNING

The robot selects a set A^* of p points (see Figure 1) with the maximum entropy conditioned on the already sampled set D. Formally,

$$A^* = \underset{A \in \mathcal{P}(V \setminus D),\ |A|=p}{\arg\max} \quad H(A|D) \tag{2}$$

where $\mathcal{P}(S)$ is the power set of S. Since, this is an NP-Hard problem [2], we use the greedy strategy proposed by Krause et al. [5] where the i^{th} sample a_i is selected as the one which results in the maximum information gain $a_i^* = \arg\max_{a_i \in V \setminus (A_{i-1}^*, D)} H\left(a_i | A_{i-1}^*, D\right)$ where $A_{i-1}^* = \{a_1^*, \ldots, a_{i-1}^*\}$. Let, $\Omega_{A,x}$ be the set of all possible paths originating from a location $x \in \mathbb{R}^2$ and passing through the plot of all samples $a \in A$. Also, let P_m be the set of mobile samples along the path P. The most informative path P^* is determined as:

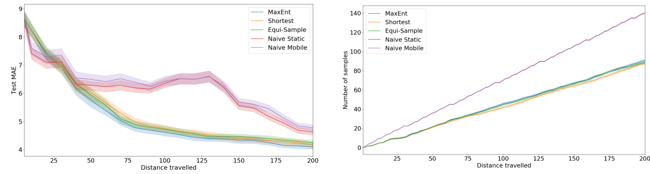
$$P^* = \underset{P \in \Omega_{A^*, x_0},\ c(P) < B}{\arg\max} \quad H\left(A^*, P_m | D\right)$$

where $c(P)$ is the path length and B is the budget set as the shortest path length plus some slack ξ, i.e., $B = \min_{P \in \Omega_{A^*, x_0}} c(P) + \xi$. The slack term ξ controls the freedom given to the robot to explore distant areas. The robot can't take a 180° turn because of narrow space between two rows. This constraint reduces the search space to a small graph G whose nodes are the junction points in the grid. We propose a heuristic to speed up the search for P^* in G by determining the bounding box formed by the robot's current location and the remaining static sampling locations. Clearly, the agent has to travel at least the distance from its current position to the nearest edge and then to the opposite edge along each of the two axes. The heuristic is the sum of the minimum distances along the two axes.

4 EXPERIMENTS

We compared our proposed Maximum Entropy algorithm [1] (*MaxEnt*) as described in Section 3 against four baseline methods:

- *Naive static*: The robot sequentially visits each column and slows down or stops in each plot to gather accurate data.
- *Naive mobile*: This strategy is same as *Naive static* except that the agent collects date while in motion.



(a) **Test MAE v/s distance** (b) **Number of samples v/s distance**

Figure 2: **The plot shows the (a) MAE of agent's prediction on the test set and (b) number of samples collected against distance travelled. For picture clarity, only** 50% **confidence interval is shown. Here,** $\sigma_s = 0.5$, $\sigma_m = 2.5$ **and** $\xi = 0$.

- *Shortest*: The agent determines selects the path with the shortest length from the set of all the feasible paths.
- *Equi-sample*: The agent selects the path with the same number of samples as the one selected by *MaxEnt* algorithm.

We used the mean stalk height in each plot as a metric for quality of produce from that plot and performed experiments on a sorghum dataset collected by our robotic platform (see Figure 1b) from a 15×25 grid field in South Carolina, USA. We reserved a set of randomly chosen 40 locations as a test set and compared the Mean Absolute Error (MAE) between their true phenotype values and the ones predicted by the model.

The graphic plot of agent's prediction against distance travelled is shown in Figure 2a. Each grid cell in the environment corresponds to 1 unit distance. All 3 informative strategies are able to quickly predict the phenotype distribution by actively visiting places with high utility. Our proposed *MaxEnt* consistently achieves the lowest prediction error on the test set indicating its ability to accurately estimate the target distribution. *Shortest* and *Equi-sample* strategies are also able to learn the target distribution and closely match the performance of *MaxEnt.* On the other hand, the two naive strategies (current practices) perform poorly and are unable to match the performance of the informative strategies.

We also compared the predictive performance of *MaxEnt* against different values of noise ratio ($k = \frac{\sigma_m}{\sigma_s}$) and slack ξ as shown in Table 1 and Table 2 respectively.

MaxEnt prediction {mean(std)} with different k				
	$k = 1$	$k = 2$	$k = 5$	$k = 10$
MAE	4.08(0.70)	4.10(0.77)	4.09(0.78)	4.40(0.68)

Table 1: **MAE on the test set for** $\sigma_s = 0.5$ **and** $\sigma_m = k\sigma_s$ **averaged over** 20 **simulations. In each simulation, the robot travels** 250 **distance units.**

MaxEnt prediction {mean(std)} with different slack ξ				
	$\xi = 0$	$\xi = 5$	$\xi = 10$	$\xi = 15$
MAE	4.22(0.58)	4.00(0.51)	4.08(0.54)	4.19(0.40)

Table 2: **MAE on the test set for** $\sigma_s = 0.5$ **and** $\sigma_m = 2.5$ **for different** ξ **averaged over** 20 **simulations. In each simulation, the robot travels** 250 **distance units.**

We observe that there is not much significant difference in the predictive accuracy of the learned model till $k = 5$. Also, increasing ξ improves the model's performance however only upto a limit.

5 CONCLUSION

We presented an active learning framework that alternates between sampling plots with high utility and learning a GP model of the target distribution for HTP. Through simulation experiments, we have demonstrated the superior performance of our proposed approach compared to the current practices.

[1] Our code is available at https://github.com/sumitsk/algp.git

REFERENCES

[1] Harjatin Singh Baweja, Tanvir Parhar, Omeed Mirbod, and Stephen Nuske. 2018. StalkNet: A Deep Learning Pipeline for High-Throughput Measurement of Plant Stalk Count and Stalk Width. In *Field and Service Robotics*. Springer, 271–284.

[2] Carlos Guestrin, Andreas Krause, and Ajit Paul Singh. 2005. Near-optimal sensor placements in gaussian processes. In *Proceedings of the 22nd international conference on Machine learning*. ACM, 265–272.

[3] Merritt Jenkins and George Kantor. 2017. Online detection of occluded plant stalks for manipulation. In *Intelligent Robots and Systems (IROS), 2017 IEEE/RSJ International Conference on*. IEEE, 5162–5167.

[4] Erkan Kayacan, Zhongzhong Zhang, and Girish Chowdhary. 2018. Embedded high precision control and corn stand counting algorithms for an ultra-compact 3d printed field robot. *Proceedings of Robotics: Science and Systems. Pittsburgh, Pennsylvania* (2018).

[5] Andreas Krause, Ajit Singh, and Carlos Guestrin. 2008. Near-optimal sensor placements in Gaussian processes: Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research* 9, Feb (2008), 235–284.

[6] Kian Hsiang Low. 2009. Multi-robot adaptive exploration and mapping for environmental sensing applications. *Ph. D. dissertation* (2009).

[7] Wenhao Luo, Changjoo Nam, and Katia Sycara. 2017. Online decision making for stream-based robotic sampling via submodular optimization. In *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 118–123.

[8] Wenhao Luo and Katia Sycara. 2018. Adaptive Sampling and Online Learning in Multi-Robot Sensor Coverage with Mixture of Gaussian Processes. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 6359–6364.

[9] Kai-Chieh Ma, Lantao Liu, and Gaurav S Sukhatme. 2016. An information-driven and disturbance-aware planning method for long-term ocean monitoring. In *Intelligent Robots and Systems (IROS), 2016 IEEE/RSJ International Conference on*. IEEE, 2102–2108.

[10] Alexandra Meliou, Andreas Krause, Carlos Guestrin, and Joseph M Hellerstein. 2007. Nonmyopic informative path planning in spatio-temporal models. In *AAAI*, Vol. 10. 16–7.

[11] Tim Mueller-Sim, Merritt Jenkins, Justin Abel, and George Kantor. 2017. The Robotanist: a ground-based agricultural robot for high-throughput crop phenotyping. In *Proceedings of IEEE International Conference on Robotics and Automation*. 3634–3639.

[12] Amarjeet Singh, Andreas Krause, Carlos Guestrin, William J Kaiser, and Maxim A Batalin. 2007. Efficient Planning of Informative Paths for Multiple Robots.. In *IJCAI*, Vol. 7. 2204–2211.