CAtNIPP: Context-Aware Attention-based Network for Informative Path Planning - Supplemental Material

Yuhong Cao
National University of Singapore
caoyuhong@u.nus.edu

Yizhuo Wang
National University of Singapore
wy98@u.nus.edu

Apoorva Vashisth
Indian Institute of Technology, Kharagpur
av1999@iitkgp.ac.in

Haolin Fan
National University of Singapore
e0816265@u.nus.edu

Guillaume Sartoretti
National University of Singapore
mpegas@nus.edu.sg

Through a series of variants and generalization studies, our additional results presented here highlight that a) CAtNIPP has natural scalability to various graph sizes even beyond the scale of the training discretization, b) increasing the graph fineness and number of sampled trajectory generally leads to a higher-quality solution at the cost of more computing time and memory usage, and c) CAtNIPP can robustly generalize to other drastically changed Gaussian and even non-Gaussian mixture distributions from training environments, while maintaining its improved performance over other IPP planners.

A Variants Analysis

Table 1: **Comparison between different CAtNIPP variants (10 trials on 30 instances for each budget).** Tr(P) is the average covariance matrix trace after running out of budget (standard deviation in parentheses). T(s) is the average total planning time in seconds.

Method	Budget 6		Budget 8		Budget 10		Budget 12	
	Tr(P)	T(s)	Tr(P)	T(s)	Tr(P)	T(s)	Tr(P)	T(s)
g.(200)	**20.35**(±6.50)	0.30	8.46(±2.90)	0.35	4.49(±1.38)	0.42	8.46(±30.60)	0.41
g.(400)	22.67(±6.60)	0.45	7.70(±3.37)	0.53	**3.81**(±1.37)	0.78	2.53(±1.18)	0.88
g.(600)	21.43(±5.40)	0.76	**7.62**(±2.99)	1.07	3.94(±1.30)	1.34	**2.49**(±1.00)	1.70
g.(800)	22.86(±6.42)	1.23	7.72(±2.77)	1.68	3.97(±1.46)	2.20	2.70(±1.18)	2.52
ts.(4)	20.19(±3.88)	90.31	7.04(±1.44)	123.56	3.82(±0.61)	158.44	2.52(±0.41)	194.97
ts.(8)	19.29(±3.64)	158.93	6.84(±1.44)	215.76	3.72(±0.64)	275.59	2.52(±0.41)	339.72
ts.(16)	**18.36**(±**3.18**)	295.48	**6.36**(±1.18)	402.50	**3.69**(±0.68)	512.90	**2.41**(±0.38)	627.27

Table 1 shows the performance of different greedy variants of CAtNIPP on different-sized graphs, as well as the effect of different number of samples for our sampling-based variant on fixed-sized (400-nodes) graphs. We first note that our trained model naturally scales to larger graphs. Intuitively, using a more complex graph will improve solution quality, since it makes the planned trajectory more fine-grained. However, at the same time the trajectory space also grows exponentially with this fineness. Despite this, our greedy variants keep solving the IPP in less than a few seconds due to their learning based nature, while the computing time of CAtNIPP only increases linearly with the budget constraint and graph size. Note that we do not include the time for computing the minimal cost to the destination from each node, since these costs can be pre-computed offline and do not require re-computation if the route graph is unchanged for different IPP instances (for 200,

6th Conference on Robot Learning (CoRL 2022), Auckland, New Zealand.

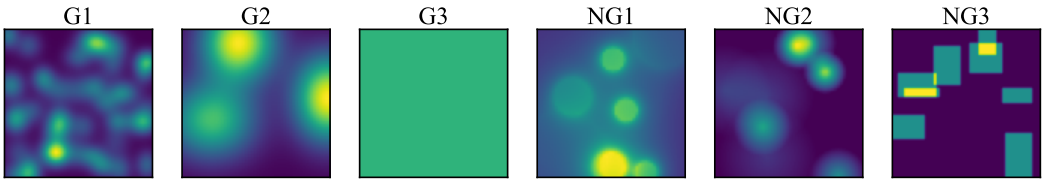

Figure 1: **True interest maps in our different generalization evaluation cases.** From left to right: G1/G2 - Gaussian mixture with more/fewer peaks but smaller/larger variance compared to training environments, G3 - uniform distribution, NG1 - mixture of inverse square decay functions when $r^i > r_0^i$, NG2 - mixture of linear decay functions, NG3 - mixture of random rectangles.

Table 2: **Generalization capability to different ground truth distributions (10 trials on 30 instances for budget 10).** Numbers are the average covariance matrix trace after running out of budget (standard deviation in parentheses). Note that despite some lacks of rigor for non-Gaussian distributions, here we still adopt covariance matrix trace as our metric as it reveals the confidence level of our belief.

Method	Case G1	Case G2	Case G3	Case NG1	Case NG2	Case NG3
RAOr	6.40(±3.07)	5.57(±2.37)	15.32(±8.96)	11.21(±3.48)	4.68(±2.55)	4.83(±2.74)
CMA-ES	6.64(±4.22)	5.95(±3.32)	11.63(±4.55)	12.61(±4.44)	4.34(±2.93)	4.95(±3.69)
g.(400)	4.08(±1.40)	3.77(±1.44)	9.19(±1.55)	9.30(±2.51)	2.60(±1.06)	2.93(±1.79)
ts.(4)	**3.76**(±1.22)	**3.38**(±1.26)	**8.91**(±1.38)	**8.70**(±2.43)	**2.40**(±1.09)	**2.62**(±1.33)

400, 600 and 800 nodes graphs, computing the minimal cost of all nodes requires 0.8s, 4.5s, 14.0s, and 31.4s respectively). Trajectory sampling variants exhibit improved solution quality over greedy variants (10% better in average). However, the computing time of trajectory sampling variants is more than one order of magnitude longer than greedy variants and close to non-learning IPP solvers. Nevertheless, we note that trajectory sampling could be implemented as an anytime algorithm, which makes it very appealing for many real-world applications.

B Generalization Analysis

Since the ground truth is commonly represented as a mixture of Gaussian for its nice mathematical properties, the majority of IPP solvers rely on Gaussian Processes as an approach of vicinity interpolation. However, some spatial distributions that consist of point source (e.g., electric field strength, radiation intensity, sound level, etc.) decay as the inverse square of the distance. Therefore, we evaluate CAtNIPP on a number of handcrafted scenarios where the parameters of Gaussian mixture significantly differs from the environments seen at training, and also test its compatibility with non-Gaussian distributions.

Figure 1 shows the ground truth maps considered, featuring either Gaussian mixtures with different parameter set, different decay functions, or even random rectangles. Nonetheless, we find that the isotropic fixed kernel that we adopted during both training and evaluation is still able to yield a good approximation of the true interest map, and highlights CAtNIPP's generalization capability to various kind of ground truth distributions including Gaussian (case G1-G3) and non-Gaussian (case NG1-NG3) mixture models, as shown in Table 3. Interestingly, we find that none of these cases lead to CAtNIPP's failure, while still showing that CAtNIPP outperforms other baselines by a large margin, without any retraining.

Table 3: **Ablation study of CAtNIPP WITHOUT encoder.** Tr(P) is the average covariance matrix trace after running out of budget (standard deviation in parentheses).

Method	Budget 6	Budget 8	Budget 10
g.(800)	44.01(±28.37)	18.40(±13.48)	11.80(±11.37)
ts.(4)	29.88(±14.80)	9.97(±5.73)	5.86(±2.84)

C Ablation Study

Our ablation results (see Table 3) confirm that the presence of the encoder is critical: without it, performance is drastically degraded, as the agent mostly sequences locally greedy decisions into suboptimal search paths. Nevertheless, we note that by relying on receding-horizon optimization, our trajectory sampling variant drastically improves the solution quality of the model without encoder.

D Numerical Simulations

Due to size limitations of the arena used for our experimental results, and to the limited number of ground truth instances, we performed additional simulations in Gazebo to further test the use of CAtNIPP on robot. Keeping the robot (TurtleBot3-Burger) and sensor (Rasberry Pi Camera V2) identical to our experimental results, we implemented our algorithm in a larger planning domain of $8 \times 8\mathrm{m}^2$. The test results are shown in Figure 2 and Table 4.

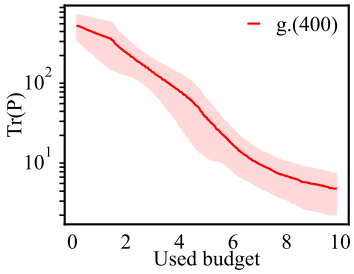

Figure 2: **Uncertainty reduction in our Gazebo simulations (9 instances).** The performance is aligned with our evaluation results.

Table 4: **Uncertainty reduction in our Gazebo simulations (9 instances for each budget).** Numbers are the average covariance matrix trace after running out of budget (standard deviation in parentheses).

Method	Budget 6	Budget 8	Budget 10	Budget 12
g.(400)	17.13($\pm$5.99)	7.39($\pm$4.54)	3.49($\pm$1.69)	2.06($\pm$1.13)