28

rbonatti

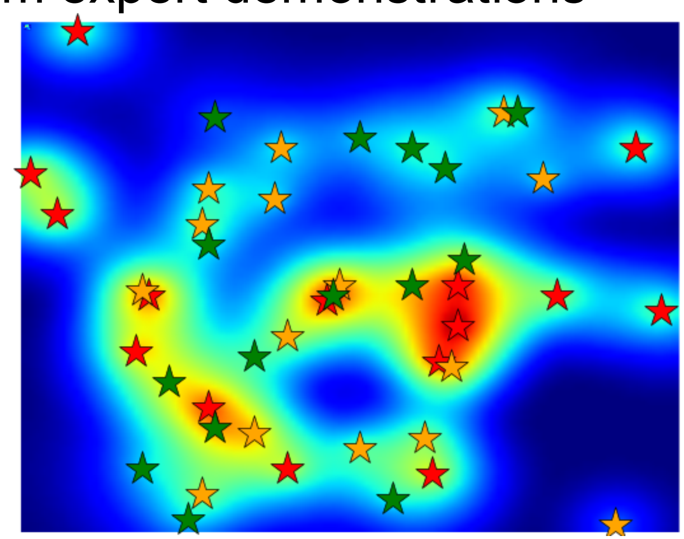
ratneshm

arjuns2

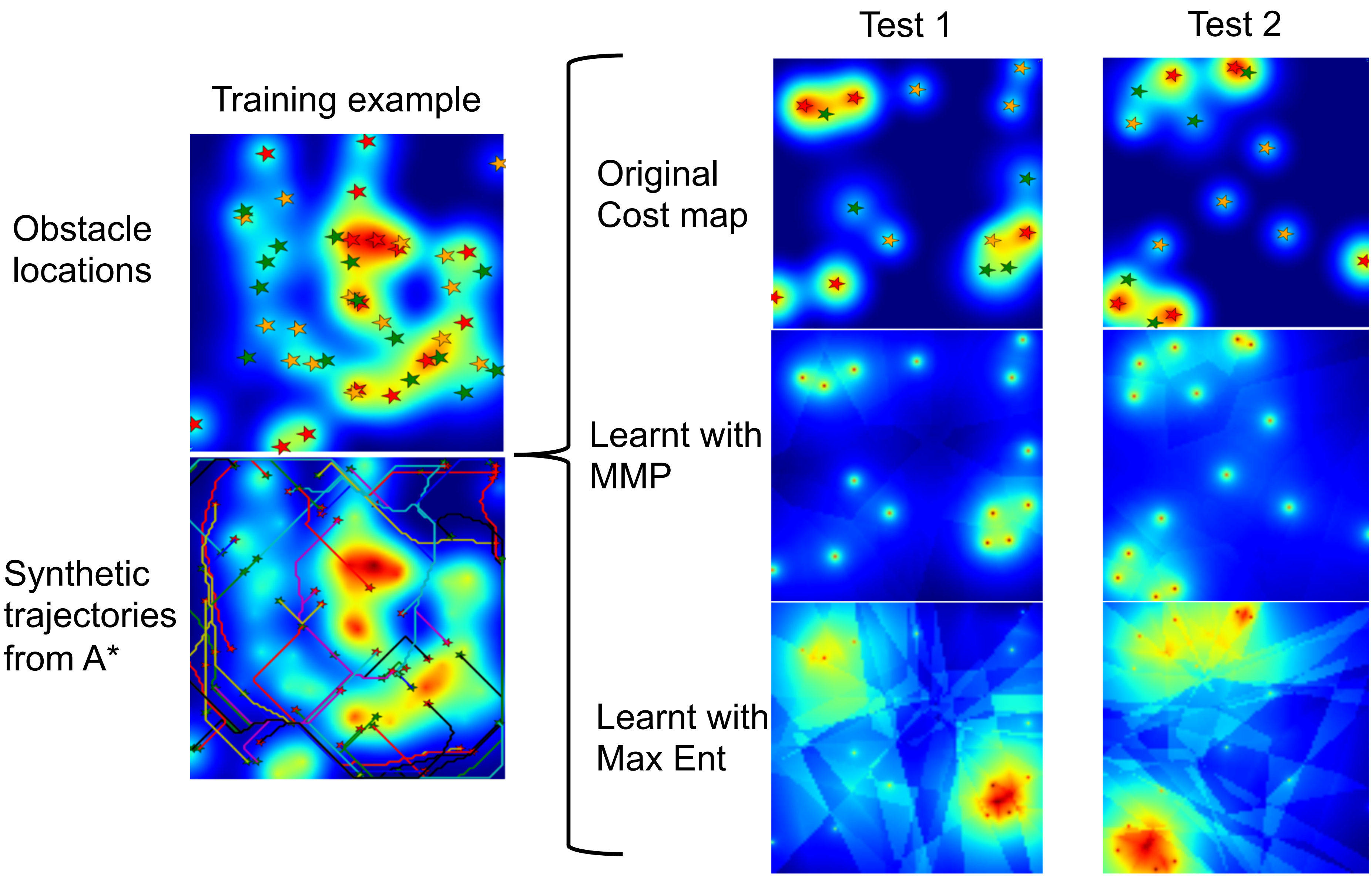
Semantic Obstacle Avoidance for UAVs by Learning Cost Functions via Inverse Reinforcement Learning

- Problem: how to encode high-dimensional cost functions for different obstacles? How to mimic an intuitive behaviour?

- Proposal: apprenticeship learning from expert demonstrations



MMP learned cost maps with greater accuracy than Max Ent



Theory

τ - sequence of (s,a) pairs $\quad \bar{f}(\tau_i) = \sum_{t=1}^{T} f(s_{i,t}, a_{i,t}) \quad r(\tau_i) = w^T \bar{f}(\tau_i)$

- Maximum Entropy Inverse Reinforcement Learning

p_{τ_i} - Probability of τ_i

$$\min_{p} \quad \sum_{\tau} p_\tau \log p_\tau$$

$$\text{subject to} \quad Fp = \frac{1}{n}\sum_{i=1}^{n} \bar{f}(\tau_i)$$

$$1^T p = 1$$

Solution, $p(\tau) = \dfrac{e^{w^T \bar{f}(\tau)}}{Z(w)}$

Dual :

$$\min_{w} \quad \log Z(w) - \left\langle w, \frac{1}{n}\sum_{i=1}^{n} \bar{f}(\tau_i) \right\rangle$$

- Maximum Margin Planning

$$\min_{w,\eta_i} \quad \|w\|_2^2 + C\sum_i \eta_i$$

$$\text{subject to} \quad w^T \bar{f}(\tau_i) + \eta_i > \max_{\tau} w^T \bar{f}(\tau) + \mathcal{L}(\tau)$$