Query-Centric Inverse Reinforcement Learning for Motion Forecasting in Autonomous Driving

Muleilan Pei (mpei@ust.hk) and Shaojie Shen (eeshaojie@ust.hk)

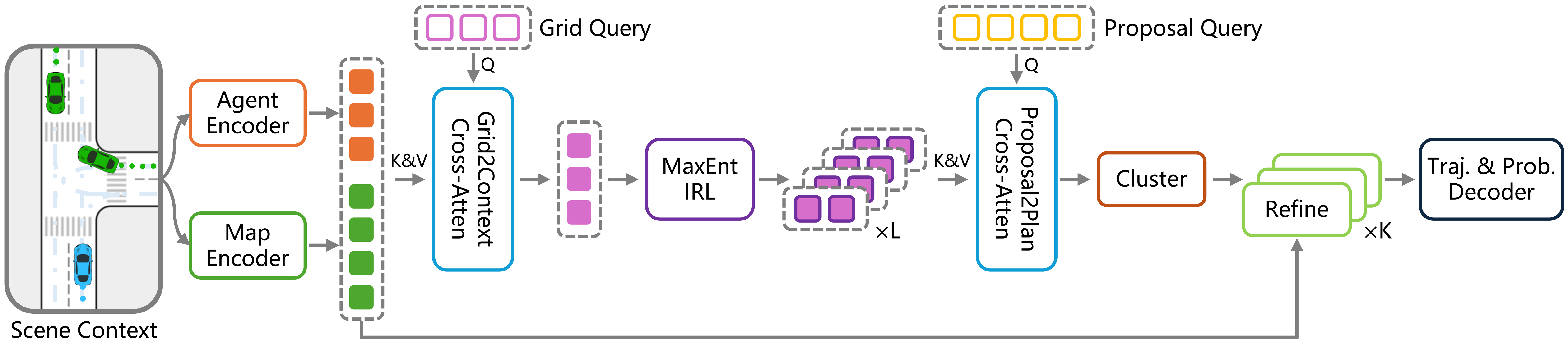
Motivation

- Motion forecasting for surrounding agents is challenging and essential in safety-critical autonomous driving, which bridges the perception and planning modules.
- Data-driven predictors try to imitate behaviors of human drivers from recorded real-world data. However, existing approaches predominantly rely on the behavior cloning (supervised) paradigm, suffering from covariate shift and modality collapse issues.
- We aim to explore the potential of inverse reinforcement learning (IRL) for motion prediction, which holds promise due to its learning from interaction mechanisms.

Contribution

- Present a Query-centric Inverse Reinforcement Learning (QIRL) framework for the motion forecasting task, which has the capability to integrate the Maximum Entropy IRL paradigm with vectorized context representation through the query-centric fashion.
- Devise a hierarchical DETR-like trajectory decoder with a refinement module to improve prediction accuracy.
- Demonstrate the highly competitive performance of IRL-based predictors on the Argoverse v1 motion forecasting benchmark compared to existing supervised models, offering a promising baseline for further investigations.

System Architecture



- The scene context features are first extracted using agent & map encoders and then aggregated to the grid-shaped queries.
- The sampled grid tokens derived from the MaxEnt IRL process are further employed to generate initial trajectory proposals.
- After clustering and refinement, multi-modal future trajectories and their corresponding confidences are finally obtained by the trajectory and probability decoder conditioned on the trajectory proposals and driving context tokens.

Methodology

Query-Centric Context Encoder

- Agent Encoder: 1-D CNN with feature pyramid network
- Map Encoder: PointNet-like architecture (Max-Pooling)

MaxEnt IRL-based Policy Generator

- Reward Model: nonlinear mapping from context tokens
- MaxEnt IRL: policy propagation & value iteration
- Plan Inference: Markov chain Monte Carlo sampling

Hierarchical DETR Trajectory Decoder

- Proposal Stage: clustered trajectories based on plans
- Refinement Stage: trajectory offsets and confidences

Training Objectives

- MaxEnt IRL optimization: negative log-likelihood loss
- Trajectory Regression: Huber loss with WTA strategy
- Probability Classification: Hinge loss

Experimental Results

Quantitative Results on the Argoverse Benchmark

Method	MR_1	$minADE_1$	$minFDE_1$	MR_6	$minADE_6$	$minFDE_6$	brier-$minFDE_6$
mmTransformer	0.6178	1.7737	4.0033	0.1540	0.8436	1.3383	2.0328
SceneTransformer	0.5921	1.8108	4.0551	0.1255	0.8026	1.2321	1.8868
HiVT	0.5473	1.5984	3.5328	0.1267	0.7735	1.1693	1.8422
MultiPath++	0.5645	1.6235	3.6141	0.1324	0.7897	1.2144	1.7932
SIMPL	0.5796	1.7501	3.9668	0.1165	0.7693	1.1545	1.7469
Wayformer	0.5716	1.6360	3.6559	0.1186	0.7676	1.1616	1.7408
QCNet	**0.5257**	**1.5234**	**3.3420**	**0.1056**	**0.7340**	**1.0666**	**1.6934**
QIRL (Ours)	0.5453	1.5824	3.4300	0.1209	0.7977	1.1652	1.7363

Qualitative Results on the Argoverse Validation Set