Related Paper Introduction

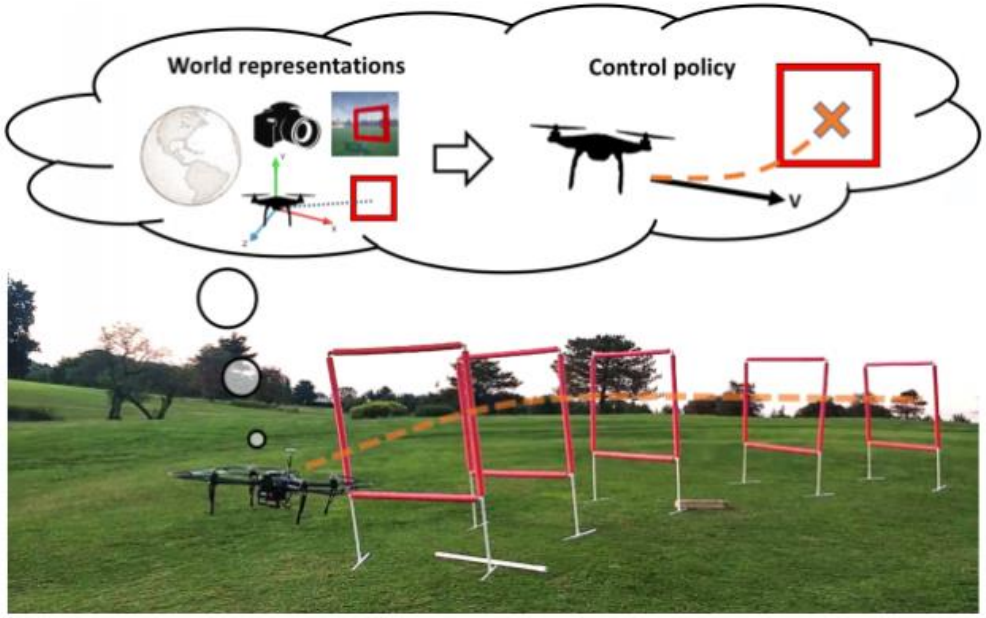
陈刚 Gang Chen

2019年10月

上海交通大學
SHANGHAI JIAO TONG UNIVERSITY

1. Learning Controls Using Cross-Modal Representations: Bridging Simulation and Reality for Drone Racing

Rogerio Bonatti, Ratnesh Madaan, Vibhav Vineet, Sebastian Scherer1, and Ashish Kapoor
The Robotics Institute, CMU & Microsoft Corporation, Redmond, WA . 2019



Target:
FPV Drone racing

Key points:
Cross-Modal Representations
Sim-to-real

Abstract

We fuse both data modalities into a novel factored architecture that learns a joint low-dimensional representation via Variational Auto Encoders. Such joint representations allow us to leverage rich labeled information from simulations together with the diversity of possible experiences via the unsupervised real-world data. We present experiments in simulation that provide insights into the rich latent spaces learned with our proposed representations, and also show that the use of our cross-modal architecture improves control policy performance in over 5X in comparison with end-to-end learning or purely unsupervised feature extractors. Finally, we present real-life results for drone navigation, showing that the learned representations and policies can generalize across simulation and reality.

Very academic!

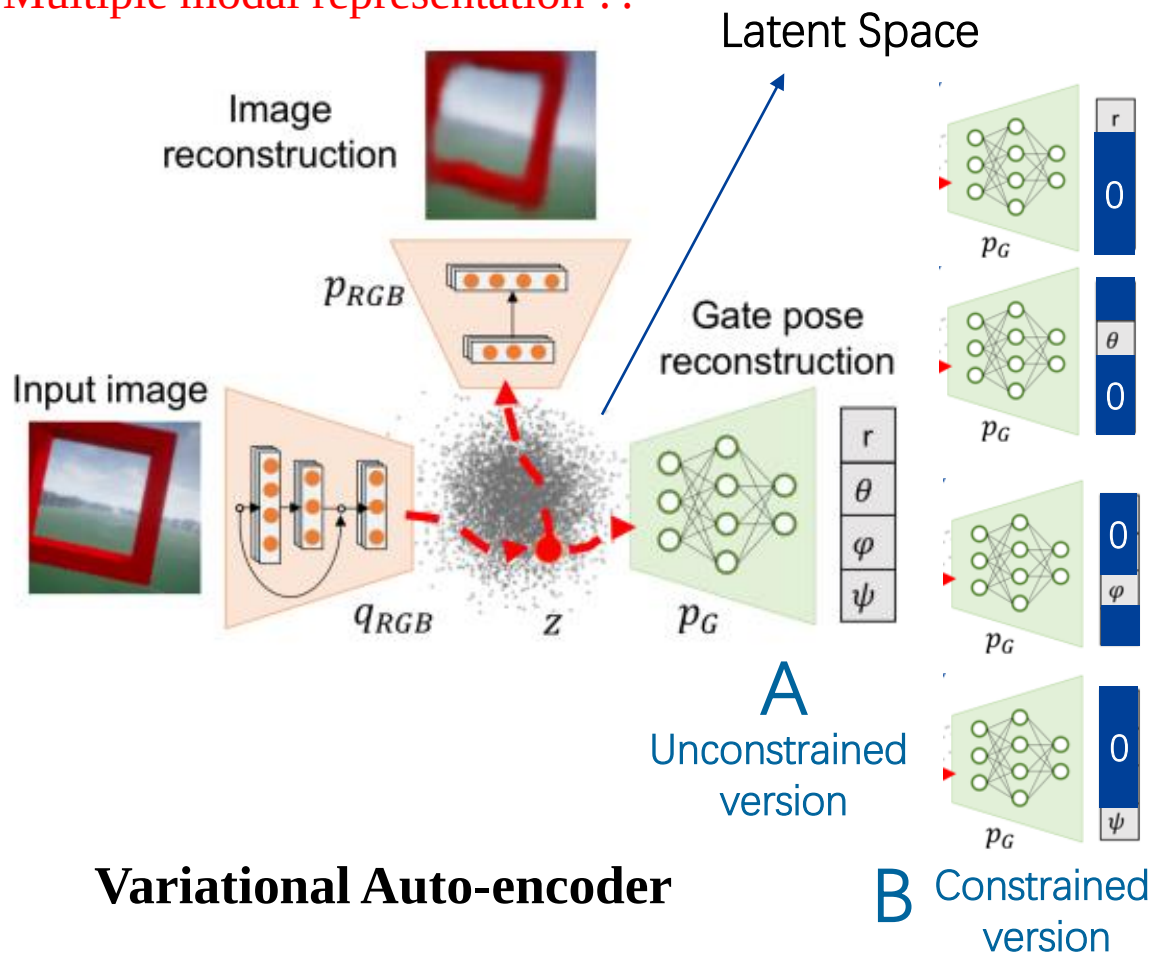
Novel!

High performance!

Method

Learning Methods: VAE (2 modes) -> Frozen Weights -> BC

Cross-modal architecture ??
Multiple modal representation ??

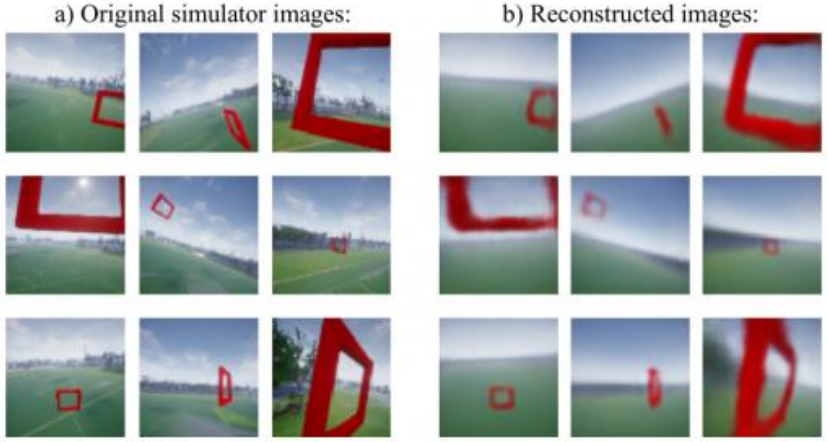


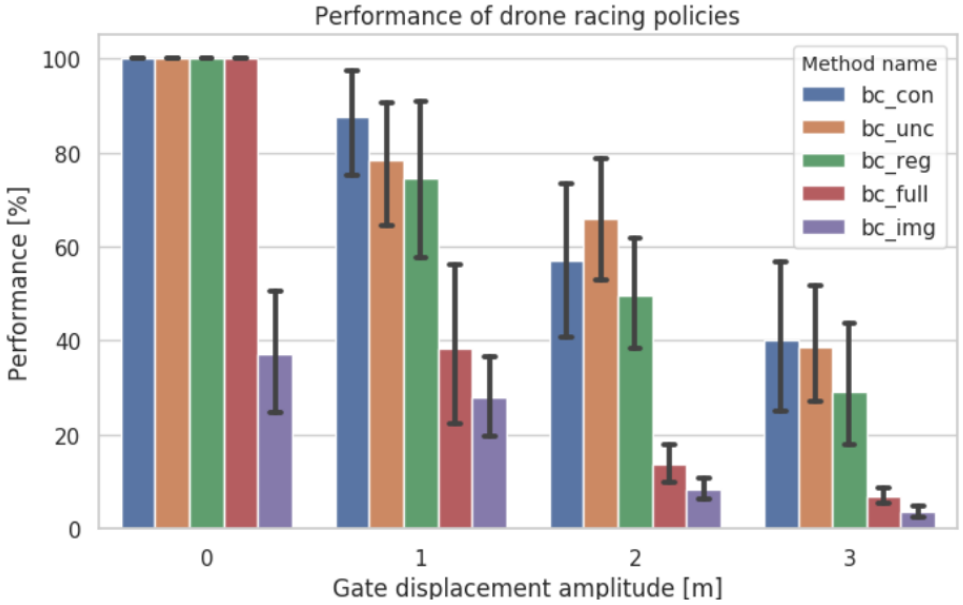
Five input paradigms

1. Image + Pose (B) (Latent Space)
2. Image + Pose (A) (Latent Space)
3. Image (Latent Space)
4. Pose (Latent Space)
5. Original image

Behavior Clone

Simulation Results





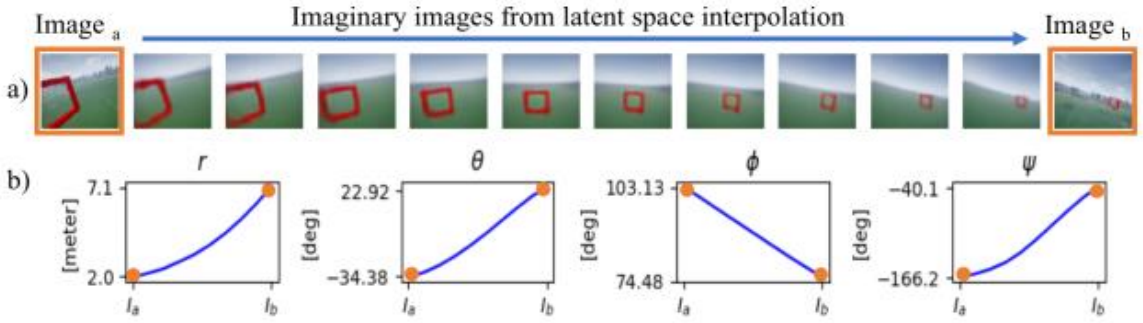
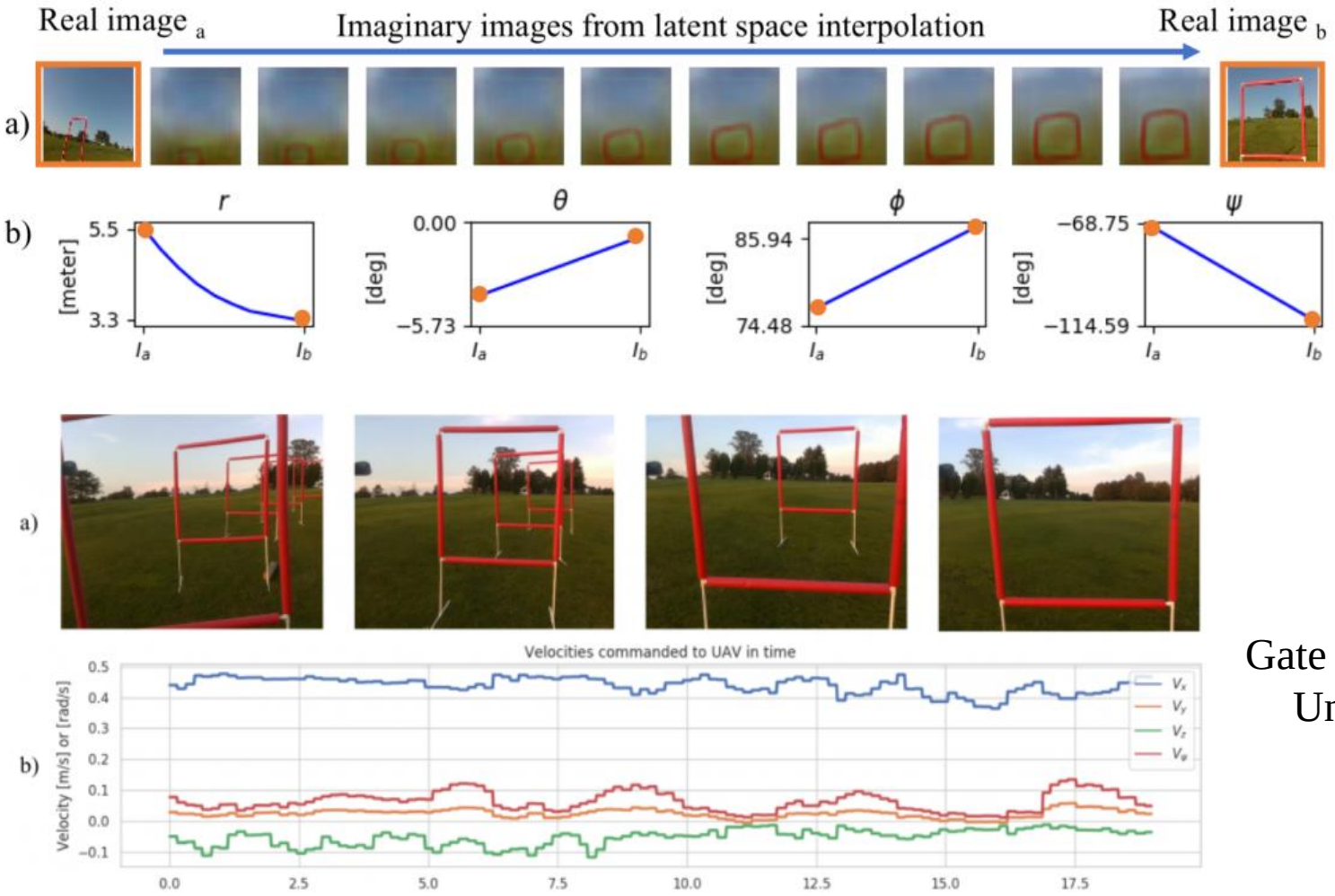


Fig. 5. Visualization of latent space interpolation between two simulated images. Smooth interpolation can be perceived in both data modalities

Real-world results



Gate amplitude Unknown

2. Weight Agnostic Neural Networks

https://weightagnostic.github.io/

谢谢！