

Mohit Sharma

✉ mohits1@andrew.cmu.edu
📄 <https://mohitsharma0690.github.io>

Education

- 2016– **M.S. in Robotics**, *Carnegie Mellon University*, 4.08*.
2008–2012 **B.Tech in Computer Science**, *IIIT-Hyderabad*, 9.22, Top 5%.

Research Interests

Computer Vision: Human Activity Analysis, Scene Understanding, Active Learning, Learning with Limited Supervision

Machine Learning: Reinforcement Learning, Inverse Reinforcement Learning, Human-in-the-Loop, Learning and Planning

Publications

- Recognizing Visual Signatures of Spontaneous Head Gestures - **Mohit Sharma**, Dragan Ahmetovic, Laszlo Jeni, Kris M. Kitani – Winters Application of Computer Vision (**WACV**), **2018** (To appear)
- Inverse Reinforcement Learning with Conditional Choice Probabilities - **Mohit Sharma**, Kris M. Kitani, Joachim R. Groeger, pre-print, Available at <https://arxiv.org/abs/1709.07597> (To be submitted to **ICML**, **2018**)

Experience

- 2016– **Research Assistant**, *Carnegie Mellon University*, Pittsburgh, PA.
Advisor: **Prof. Kris M. Kitani**
Scale Inverse Reinforcement Learning to large scale computer vision problems; Deep Neural Network learning under noisy environments; Fine-grained head gesture recognition in unconstrained environments.
- 2014–2016 **Software Engineer**, *Google*, San Francisco, CA.
Google Cloud Messaging (GCM)
Worked on client (Android; iOS) and server (Data Messaging Backend; Distributed Reliable Message Queue) side components of GCM; GCM on iOS at Google I/O 2015
- 2012–2014 **Software Engineer**, *Pocket Gems*, San Francisco, CA.
Worked on game planning; path planning; player prediction; procedural content generation; multi-player networking.
- 2011–2011 **Software Engineer, Intern**, *Google*, Hyderabad.
Worked on the Infrastructure and Tools team to develop an internal Google product.

Awards & Scholastic Achievements

- **National Talent Search Scholar** (awarded to ~750 students all over India) (2006-2012)
- **IIIT-H All Round Achievement Award** for contribution in cultural, sports and academic life (2012)
- **IIIT-H Awarded Deans List** for Excellence in Academics for 7/8 semesters.
- Qualified for **Indian National Math Olympiad** (Top-1%, ~500 students all over India). Top-10 rank in Regional Math Olympiad, North-West region.

Academic Service

Reviewer for FG '18, CVPR '18

Miscellaneous Projects

Multi-Pedestrian Forecasting

Dr. Kris M. Kitani

Working on extending our Inverse Reinforcement Learning algorithm (CCP-IRL) to non-stationary models of multiple pedestrians interacting in different social environments. Considered the hypothesis of using a sequence of weak classifiers to create a strong classifier to get conditional choice probability estimates for CCP-IRL.

Inverse Reinforcement Learning with Conditional Choice Probability

Dr. Kris M. Kitani, Dr. Joachim Groeger

[\[Report\]](#)

Developed a new Inverse Reinforcement Learning algorithm (CCP-IRL) utilizing the concept of conditional choice probabilities from the economics literature. Also showed that two different formulations in the economics and robotics literature for solving a similar problem have the same algorithm although with different assumptions.

Learning under Noisy Environments

Dr. Kris M. Kitani, Dr. Laszlo Jeni

[\[Report\]](#)

Developed a novel deep neural architecture to learn under large amount of label noise. Considered the hypothesis that a neural network can report the confidence in its predictions which is used to learn the ground label confusion matrix. Compared to Sukhbaatar et al. (2015) we achieved better performance under large random noisy permutations.

Visual signatures of spontaneous Head Gestures

Dr. Kris M. Kitani, Dr. Chieko Asakawa, Dr. Laszlo Jeni

[\[Report\]](#)

Developed a novel Multi-Scale Deep Convolution-LSTM architecture, capable of recognizing short and long term motion patterns found in head gestures, from video data of natural and unconstrained conversations. Our model outperforms the state-of-the-art based models by a large margin on multiple conversation datasets.

Object Grasping using Hierarchical Reinforcement Learning

Dr. Matthew T. Mason

[\[Report\]](#)

Developed a novel hierarchical learning algorithm using a manager policy to learn to weight individual task rewards for the underlying task policy. Used Generalized Advantage Estimation (GAE) with Proximal Policy Optimization (PPO) to learn both manager and task policy. Tested our algorithm for manipulation task in Mujoco – reach and grasp.

Learning to Map from Raw Images

Dr. Abhinav Gupta

[\[Report\]](#)

Developed a deep neural network architecture for mapping environments from first person viewpoints. Used a convolutional encoder with residual connections and an LSTM layer to allow the network to make continuous free space estimates over a continuous trajectory. Used Gazebo simulation package to generate simulated trajectories for Stanford Large-Scale Indoor Spaces 3D dataset.

Coordinated Multi-Agent Reinforcement Learning

Dr. Katerina Fragkiadaki, Dr. Ruslan Salakhutdinov

[\[Report\]](#)

Extended the deep learning architecture in (Foerster et al. 2016) to learn communication policies between agents for a novel grid world environment. Used the above environmental setup to find correlation between the agents' communication and environment policies.

Ensemble Learning for Detection of Diabetic Retinopathy

Dr. Eric P. Xing, Dr. Matt Gormley

[\[Report\]](#)

Implemented an ensemble of learning classifiers to detect diabetic retinopathy from retinal images. Specifically, implemented forest of decision trees, boosting, SVM, CNN based classifiers using end to end training and finetuning pre-trained models. Compared the strengths and weakness of all these approaches together. Kaggle dataset provided by EyePACS for training and evaluation.

Game Planning and Motion Planning

Worked on predicting player moves to model certain AI behaviors to specific player moves. Used Bayesian Networks to model characteristics of the NPC behavior. Developed hierarchical and any-time path planning algorithms in different environments. Used Jump Point Search for open space environments which resulted in 10x performance improvements. Experimented with different search space representations e.g. navigation meshes, polygon maps for efficient character movement and planning.

Face Detection

Dr. C. V. Jawahar

Developed a face detection system using sliding window detection and cascade of classifiers (Viola and Jones, 2001). Compared implementation against single scale detector using HOG descriptors (Dalal and Triggs, 2005). Used filtered Caltech Faces and CMU-MIT dataset for training and testing.

Distributed and Virtualized Cloud

Dr. Saripalli Prasad

Developed a distributed, scalable virtualization platform from scratch. Worked on template provisioning and automated deployment in a distributed environment with different virtualization techniques. Additionally, developed an automated monitoring system for the environment using Simple Network Management Protocol (SNMP).

Teaching Assistant

Carnegie Mellon University

- Computer Vision, 16-720

*Spring '18**

IIT-H

- Data Structures
- Math-I
- Math-III

Spring '11, Spring '12

Fall '11

Fall '10

Relevant Graduate Coursework

Machine Learning, Computer Vision, Deep Reinforcement Learning, Visual Learning & Recognition, Language Grounding for Vision & Control, Dynamic Optimization, Mechanics of Manipulation

Skills

Languages C, C++, Python, Java, Lua, Matlab

Framework ROS, Hadoop, MapReduce, OpenAI gym, Mujoco, PyTorch, Tensorflow, CUDA, OpenCV, OpenGL