Automating Endoscopic Camera Motion for Teleoperated Minimally Invasive Surgery using Inverse Reinforcement Learning

-Ankur Agrawal

Outline

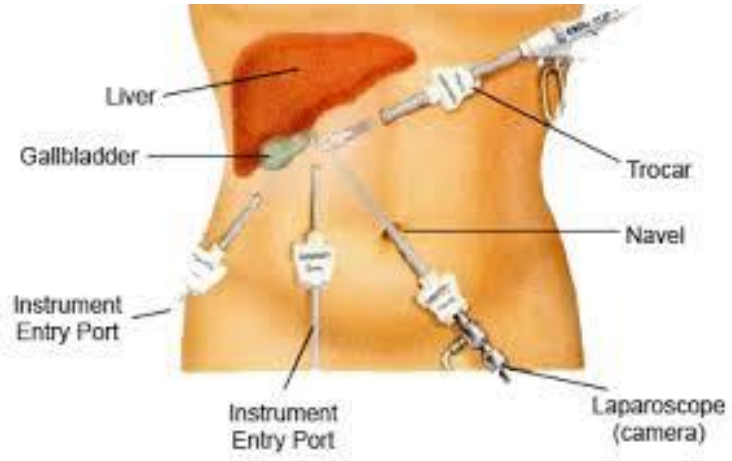
- Introduction
- Related Work
- Algorithms used
- Simulation
- User Study
- Implementation
- Results
- Conclusion and Future Work

Introduction

- Laparoscopic Surgeries
- Robotic Surgeries
- Goals
- CONTRIBUTIONS

Laparoscopic Surgeries

- Surgeries through small incisions using laparoscope/endoscope to view on monitor
- Have two surgeons – operating and assisting
- Assisting surgeon operates the camera



Schematic showing how laparoscopic surgeries are performed.
Source: [https://medium.com/@rahulsinghh/how-is-laparoscopic-surgery-done-90967cfe2b32]

Robotic Surgeries

- Surgeons teleoperate on the patient
- Uses master-slave system of robots
- Better ergonomics for surgeons, wristed instruments, smaller incisions
- Camera controlled by operating surgeon

Laparoscopic surgeries performed using the daVinci surgical system. Source: [Intuitive Surgical, Inc.]

Goals

- Detect intent of the operating surgeon by classifying task into subtasks.
- Learn optimal camera viewpoints based on this intent for a pick and place task.

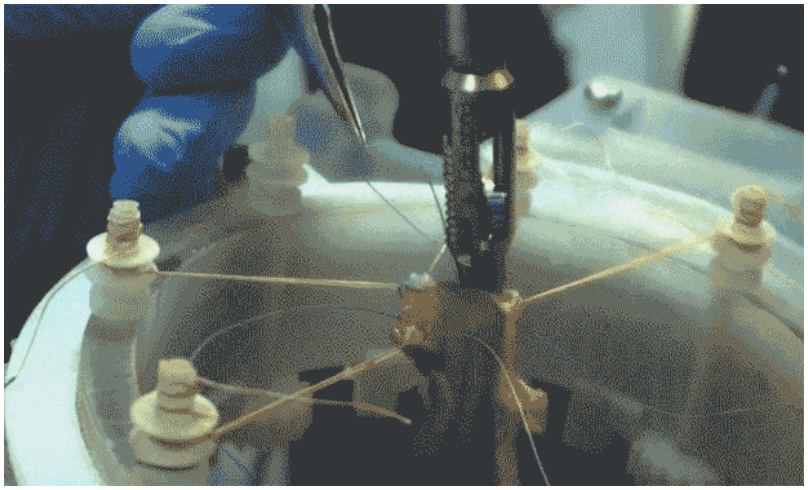
CONTRIBUTIONS

- Developed a 3D simulated virtual reality environment to pick and place a ring on a peg.
- Interfaced the Oculus Rift CV1 with the simulated environment to give an immersive experience.
- Performed a user study to collect an open source dataset recording desired camera movements (using Oculus Rift) with the movements of the master tool manipulator.
- Developing machine learning models for classifying subtask of the task to know the intent of the user.
- Using inverse reinforcement learning to learn a desired behavior to automate endoscope based on the demonstrations obtained from the user study.
- Validating the learned behaviors by comparing the obtained automated camera trajectory with a human trajectory from the user study.

Related Work

- Automation for surgical robots
- Control of endoscopic camera
- Inverse Reinforcement Learning (IRL)

Automation for surgical robots



Smart Tissue Autonomous Robot(STAR) [1],[2]

Learning by observation for surgical subtasks [3], [4]

- Supervised autonomous suturing algorithm
- Has near-infrared fluorescent (NIRF) imaging system

- Uses deep reinforcement learning to learn instrument positions for desired tensions

Control of endoscopic camera

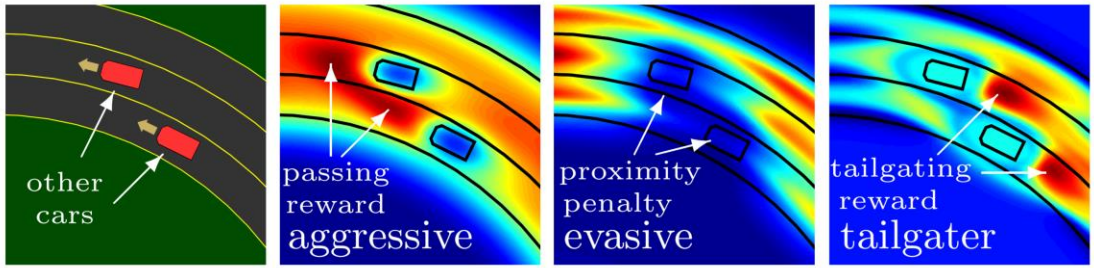
- Most common methods of camera automation include instrument tracking in the video feed [5], [6], [7]
- In [8], [9], authors use the information of entry ports and shapes to detect instrument and track them
- Camera moves according to "rules" by keeping the centroid of the instruments as its viewpoints [10], [11]. This work was performed using the daVinci Research Kit

Control of endoscopic camera

- Eye gaze tracking was used to infer surgeon's attention and used to move the camera view point to where the surgeon is gazing [12] for the AESOP endoscope system which can also be controlled using joysticks or voice commands [13].
- [14] describes a novel bending endoscope that takes the state of surgery, tools being used into account and decides whether to follow tools or stay still.
- For camera to be truly intelligent, it needs to be combination of proactive and reactive elements and needs to know the intent of the surgeon and react differently according to each intent [15].

Inverse Reinforcement Learning (IRL)

- Introduced by Ng. et. al. learns the rewards using expert policies or expert trajectories [16].
- Most of the papers in this field validate their algorithms in two environments: gridworld and autonomous driving (can be seen from survey in [17]).
- Parallels can be drawn from autonomous driving to automating endoscope.



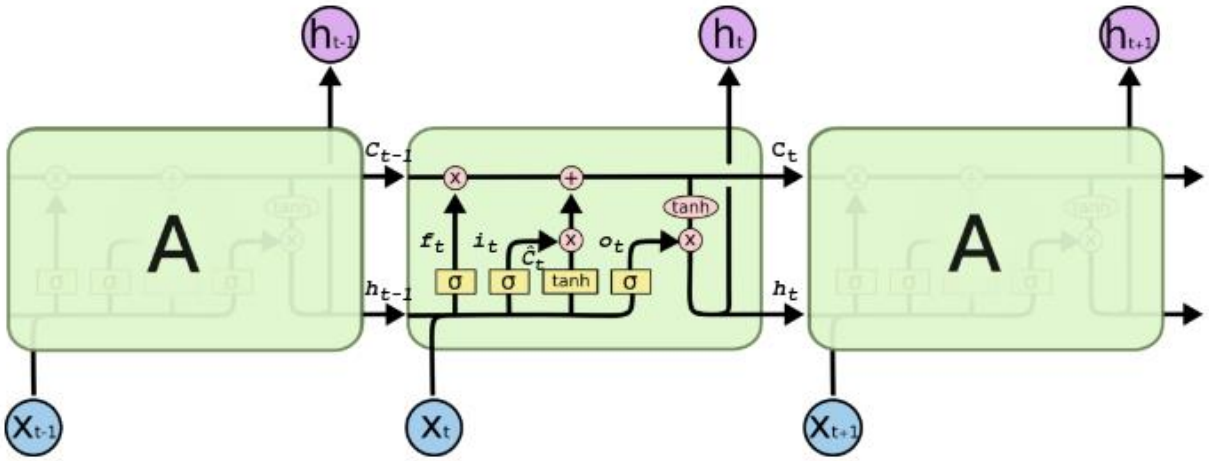
Different rewards for different driving behaviors [18]

Algorithms used

- Recurrent Neural Networks – LSTM
- Reinforcement Learning
- Markov Decision Process
- Reinforcement Leaning Algorithm
- Inverse Reinforcement Learning
- IRL- Algorithm
- Integrated Algorithm

Recurrent Neural Networks - LSTM

- Long short term memory (LSTM) [19] networks are used in sequential data for classification/prediction.
- Generally used for speech recognition [20].



Reinforcement Learning

- Reinforcement learning – Learning by experience
- Autonomous agents learn actions for different situations based on rewards/penalties received
- Based on Markov Decision Process
- Markov Property – Memoryless property

Markov Decision Process

S : A set of all possible states, s, of the environment, called as State Space.

A : A set of all possible actions, a, that an agent can take, known as Action Space.

$T(s'|s,a)$: Transition Probability Matrix, mapping probability of reaching state s' given the environment was in state s and an action a was taken by the agent.

$R(s,a,s')$: Reward as a function for a particular action a taken in state s and reaching state s'.

$V(s)$: Expected total rewards received by visiting this state.

$Q(s,a)$: Expected total rewards received by visiting this state and performing

$\pi(a|s)$: Probability of taking an action a, given a state s.

Reinforcement Learning - Algorithm

Algorithm 1: Reinforcement Learning Algorithm

Input: MDP(S, A, T, R, γ)

$V(s) \leftarrow R(s)$;

for *1 to* N **do**

  $Q(s,a) \leftarrow R(s) + \gamma \sum_{s'} T(s,a,s') V_{\pi}(s')$;

  $\pi(a|s) \leftarrow \frac{e^{\beta Q(s,a)}}{\sum_a e^{\beta Q(s,a)}}$;

  $V(s) \leftarrow \sum_a \pi(a|s) Q(s,a)$;

end

Output: Policy, $\pi(a|s)$

Inverse Reinforcement Learning

- Sometimes rewards are not easy to define for desired behaviors
- Trajectories defining behaviors are easily obtainable
- IRL can calculate reward function for given set of expert trajectories
- Reward function for the given trajectories may not be unique
- Reward function is written as a linear combination of feature functions

$$R_{\mathbf{w}}(s) = \sum_i w_i \phi_i(s)$$

Inverse Reinforcement Learning

- Maximizes the likelihood of the given expert trajectories

$$P(\tau|\mathbf{w}) \propto e^{R_{\mathbf{w}}(\tau)}$$

$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmax}} \frac{1}{|D|} \sum_{\tau \in D} R_{\mathbf{w}}(\tau) - \ln \sum_{\forall \tau} e^{R_{\mathbf{w}}(\tau)}$$

$$\nabla_w L = \frac{1}{|D|} \sum_{\tau \in D} \phi_\tau - \sum_{s \in S} P(s|w)\phi(s)$$

IRL - Algorithm

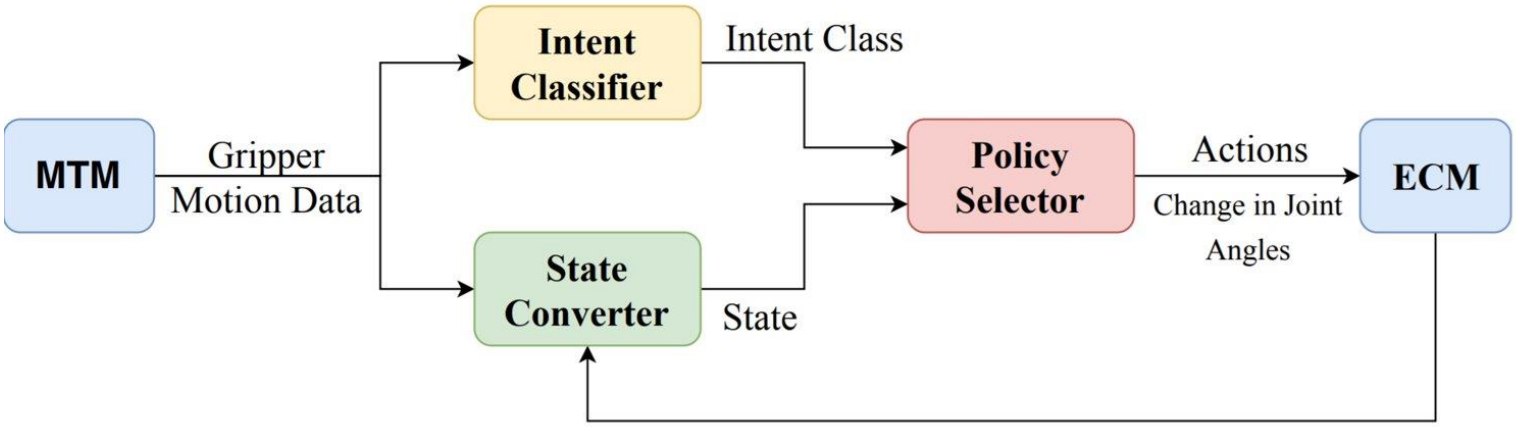
Algorithm 2: Inverse Reinforcement Learning Algorithm

Input: MDP(S, A, T, γ), expert trajectories $D = [\tau_0, \tau_1, \tau_2, ...]$, feature vectors: $\phi = [\phi_0, \phi_1, \phi_2, ...]^T$

Randomly Initialize $\mathbf{w} = [w_0, w_1, w_2, ...]^T$;
for *1 to N* **do**
  $R_{\mathbf{w}}(s) \leftarrow \sum_i w_i \phi_i(s)$;
  Calculate $\pi_{\mathbf{w}}(a|s)$ from Algorithm 1 using $R_{\mathbf{w}}(s)$;
  for $t \leftarrow$ *0 to* T **do**
    $\mu_{t+1}(s) \leftarrow \sum_a \sum_{s'} \pi_{\mathbf{w}}(a|s) T(s'|s, a) \mu_t(s')$;
  end
  $P(s|\mathbf{w}) \leftarrow \sum_t \mu_t(s)$;
  $\nabla_w L \leftarrow \frac{1}{|D|} \sum_{\tau \in D} \phi_\tau - \sum_{s \in S} P(s|w)\phi(s)$;
  $\mathbf{w} \leftarrow \mathbf{w} + \eta \nabla_{\mathbf{w}} L$
end

Output: Weights: $\mathbf{w}$, Policy: $\pi_{\mathbf{w}}(a|s)$

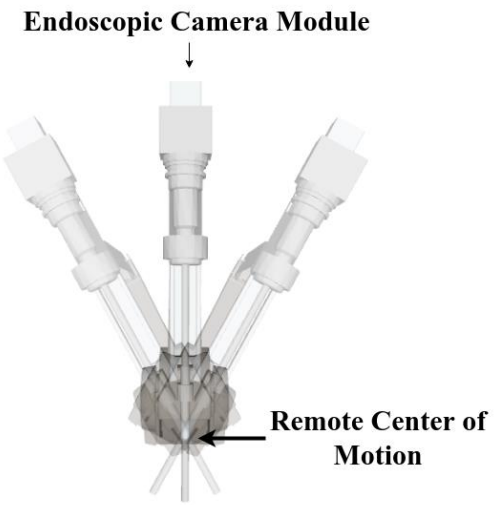
Integrated Algorithm

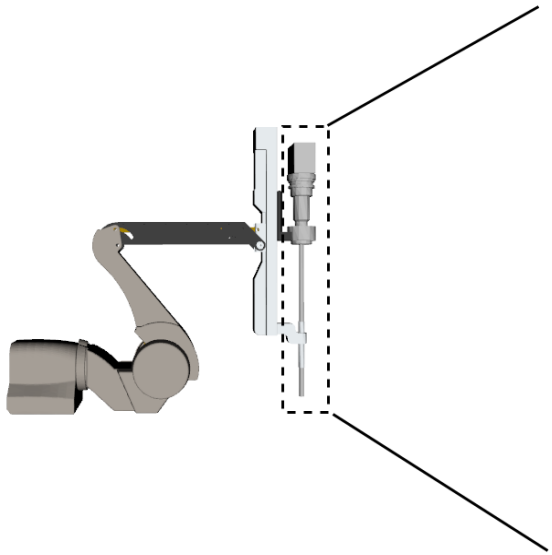


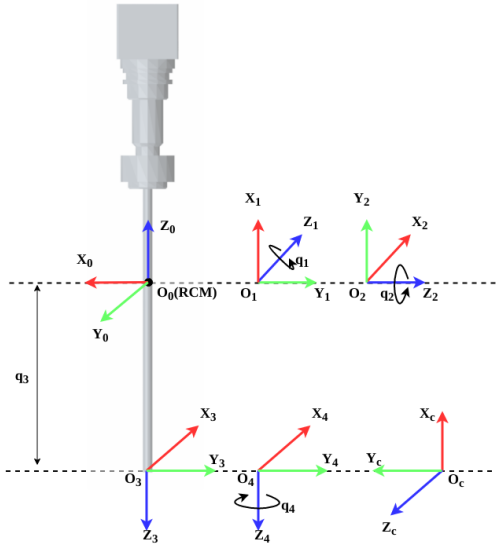
Simulation

- Endoscopic Kinematics
- Gazebo Simulation
- CHAI-3D
- Simulation Environment
- Oculus Interface
- Distortion due to Oculus Lenses

Endoscope Kinematics



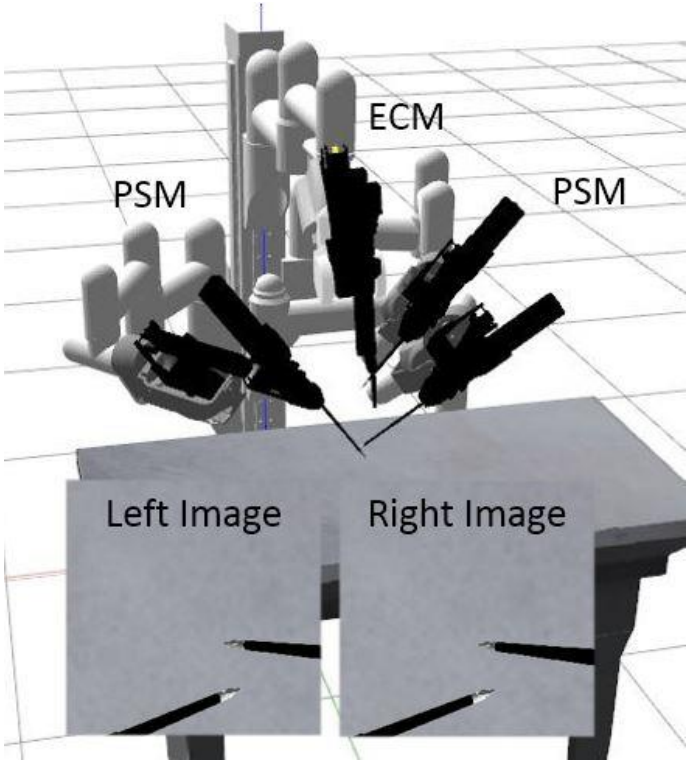




- Has 4 degrees of freedom – yaw, pitch, insert and roll.
- Has a remote center of motion mechanism.
- Modified D-H convention used to define forward and inverse kinematics.

Gazebo Simulation

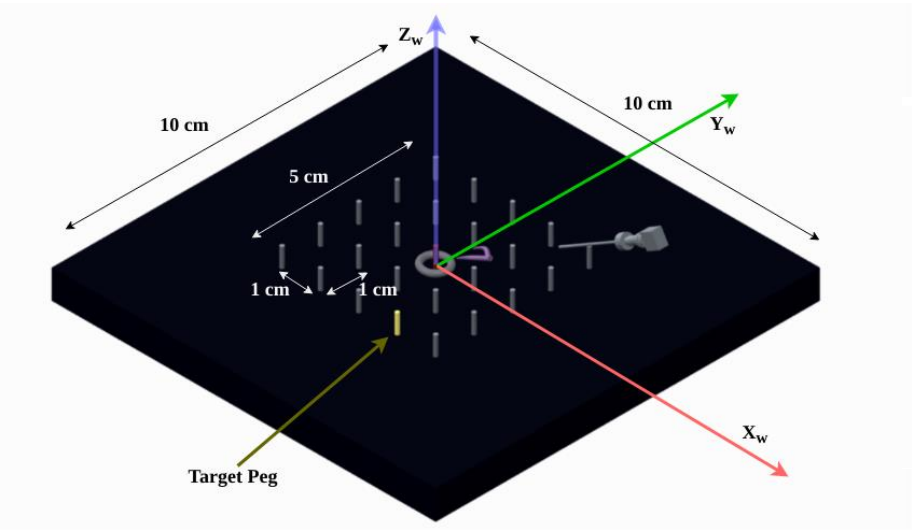
- Developed closed loop kinematic models for Gazebo simulation
- Accurate representation of the daVinci system
- Used ROS for communications
- Camera plug-ins for the stereo endoscope feed
- Models were not stable enough during teleoperation

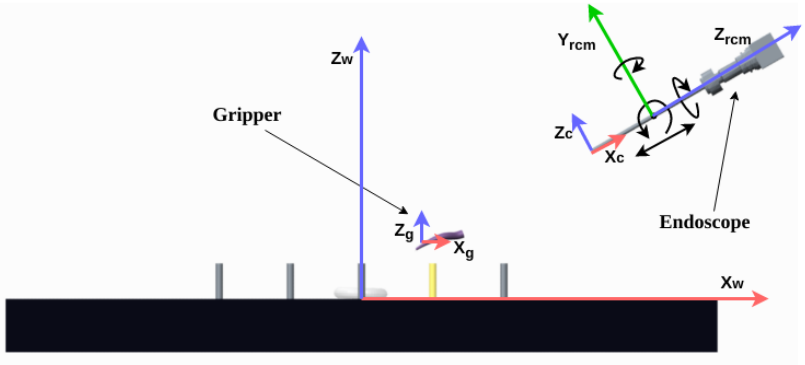


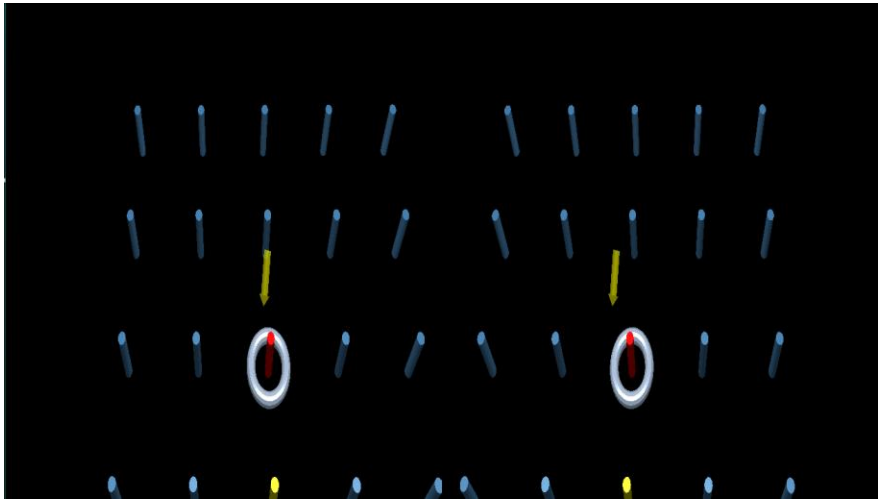
CHAI-3D

- Open source C++ libraries using OpenGL for graphics [21].
- Can interface with a variety of haptic devices.
- daVinci Master Tool Manipulator (MTM) interface for CHAI-3D was written by Adnan Munawar [22].
- Uses BULLET as the physics engine.

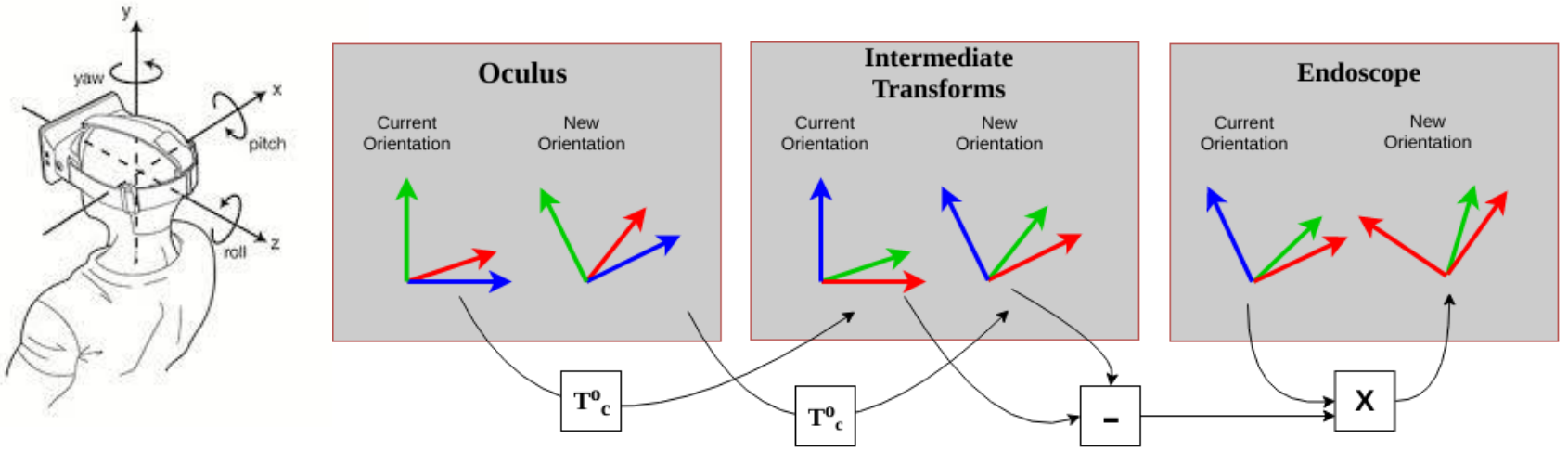
Simulation Environment





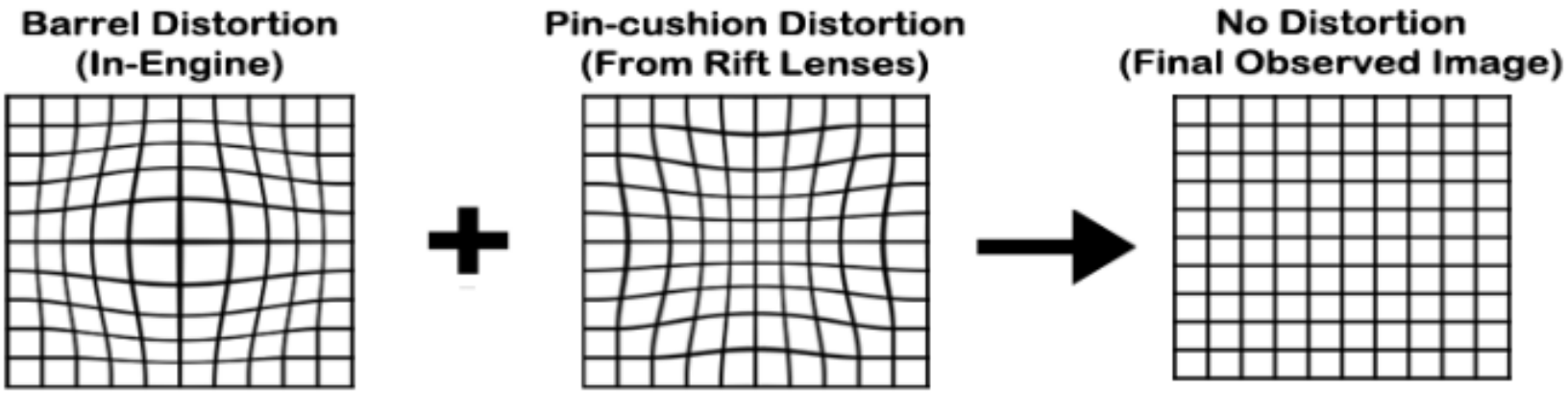


Oculus Interface



- Head tracking data is obtained from IDU and accelerometer data on the Oculus.
- These give the current orientation of the head orientation of the user.
- Mapped to endoscope motion as shown above.

Distortion due to Oculus lenses



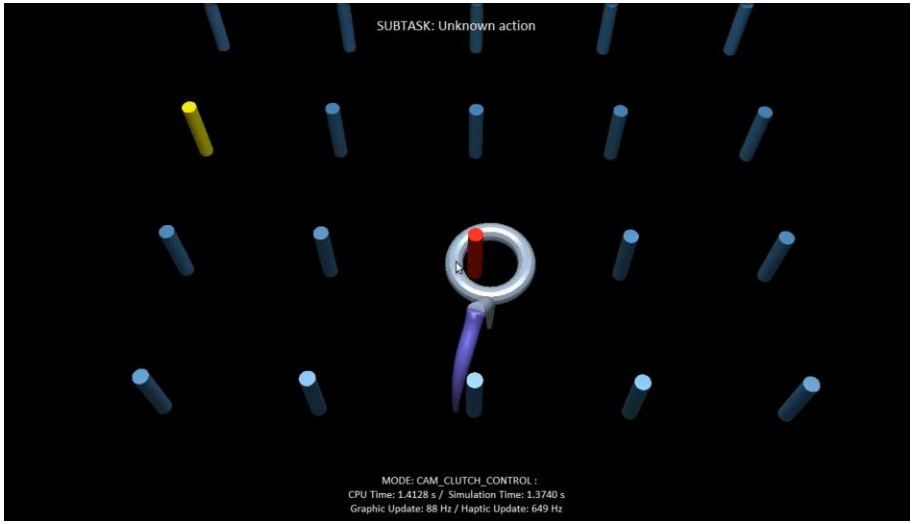
- Lenses in Oculus causes distortion in the video feed.
- An inverse distorted video feed should be given to counter this.

User Study

- User study
- Data Recorded
- Subtasks

User Study

- 9 people participated in the study – 6 male , 3 female
- Most participants had little exposure to robotic surgery
- They were asked to wear Oculus Rift and perform a pick and place task
- The task was to pick up the ring and place it on the flashing yellow peg
- During one exercise, the task was to be completed 5 times
- Data for 5 exercises was recorded for each participant



Data Recorded

- Data recorded is shown here.
- Objective was to record at 1 KHz, but was recorded at a mean frequency of 723 Hz.
- Video for each were recorded in .mkv format at 30 fps
- ROS bags were recorded to enable replaying the data

Table 5.1: Table describing the variables that were recording during the user study.

Column Indices	Number of Variables	Description of Variables
1	1	Real world time calculated using the CPU time.
2	1	Simulated time calculated using timers from CHAI library.
3-6	4	Joint angles of ECM when moved using the head tracking for oculus (j_1, j_2, j_3, j_4).
7-9	3	Position of the camera frame with respect to the world frame in simulation (xyz).
10-18	9	Orientation of the camera frame with respect to the world frame in simulation (R).
19-21	3	Position of the RCM frame of endoscope with respect to the world frame in simulation (xyz).
22-30	9	Orientation of the RCM frame of endoscope with respect to the world frame in simulation (R).
31-33	3	Position of the ring with respect to the world frame in simulation (xyz).
34-36	3	Position of the target peg with respect to the world frame in simulation (xyz).
37	1	Subtask label assigned live when the exercise was done (0 or 1 or 2 or 3 or 4).
38-40	3	Position of the gripper tip with respect to the world frame in the simulation (xyz).
41-49	9	Orientation of the gripper tip with respect to the world frame in the simulation (R).
50	1	Jaw angle of the gripper.
51	1	Subtask relabeled later using frame by frame analysis of the video (0 or 1 or 2 or 3 or 4).

Subtasks

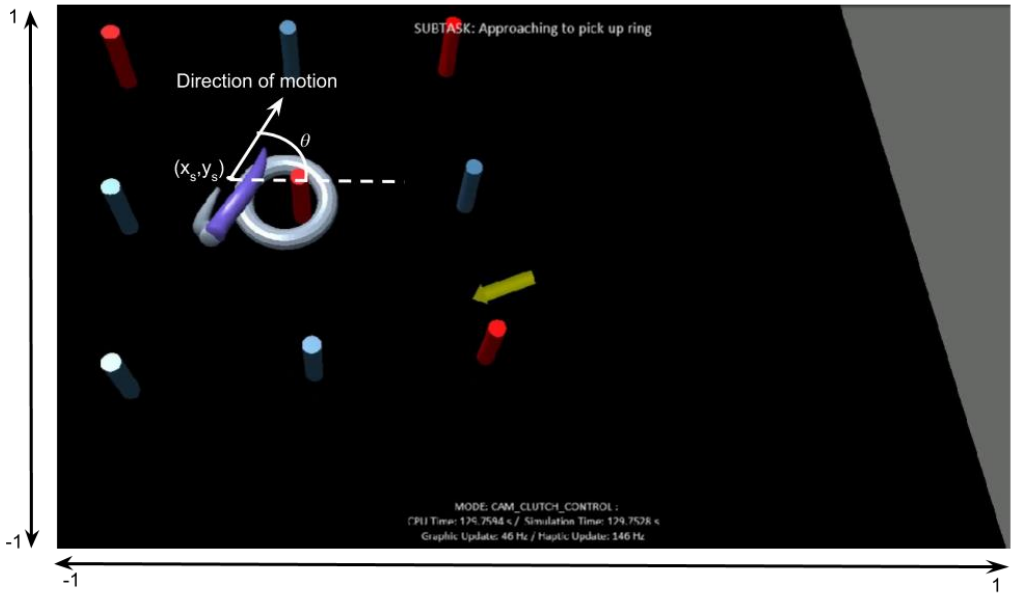
- Approaching to pick up ring
- Picking up ring
- Approaching to place ring
- Placing the ring on the peg

Implementation

- Learning camera view points
- Learning camera view points
- Features
- Validation of Learned Rewards

Learning camera view points - State

- State was the gripper position in the image frame
- Image frame was scaled to [-1,1] in x and y axes
- Other state variables were the velocity represented in polar coordinates
- State were discretized in 200 grid blocks for x and y axis
- Speed was discretized in steps of 0.01
- Direction was discretized for every 10 degrees



Learning Camera View Points - Actions

- Actions were the change in joint angles of the endoscope
- 27 total actions :
  - J1 : {-0.01, 0, 0.01}
  - J2 : {-0.01, 0, 0.01}
  - J3 : {-0.001, 0, 0.01}
- Roll was kept constant due to distortion.
- These were decided based on the data obtained from the user study.

Learning Camera View Points - Trajectories

- Trajectories were obtained by down sampling the data at 100 Hz
- 274 trajectories for subtask 1 and 2, and 246 for subtask 3 and 4
- Approximately 80% for training and 20% for validation were used.

Features

- Total of 12 feature basis functions were used
- 9 were based on the position of the gripper – Gaussians centered at {-0.5,0,0.5}
- 3 included position, speed and direction of the gripper

Validation of Learned Rewards

- A metric to compare expert trajectories with automated trajectories was needed.
- A similarity function was defined as:

$$\sigma : Q \times Q \rightarrow [0, 1]$$

$$\tau_H = [\mathbf{q}_1^H, \mathbf{q}_2^H, \mathbf{q}_3^H, ...] \quad \tau_A = [\mathbf{q}_1^A, \mathbf{q}_2^A, \mathbf{q}_3^A, ...]$$

$$\sigma(\tau_A, \tau_H) = e^{-\sqrt{\frac{1}{T}\sum_{t=1}^{T} ||\mathbf{q}_t^H - \mathbf{q}_t^A||_2^2}}$$

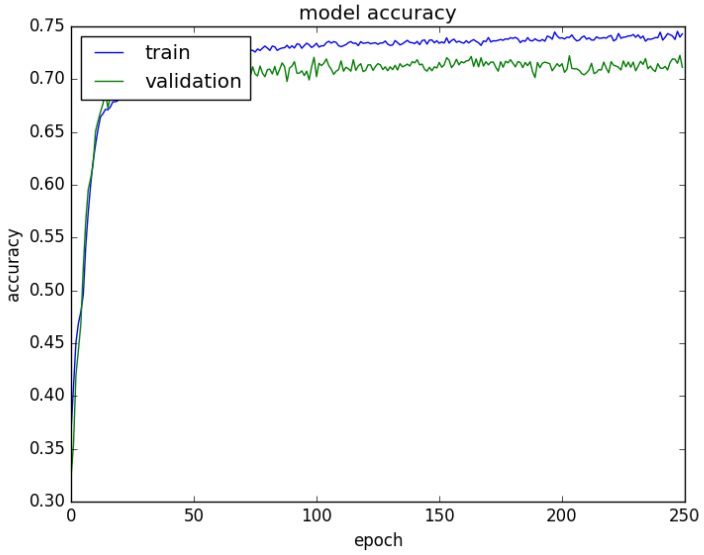
Results

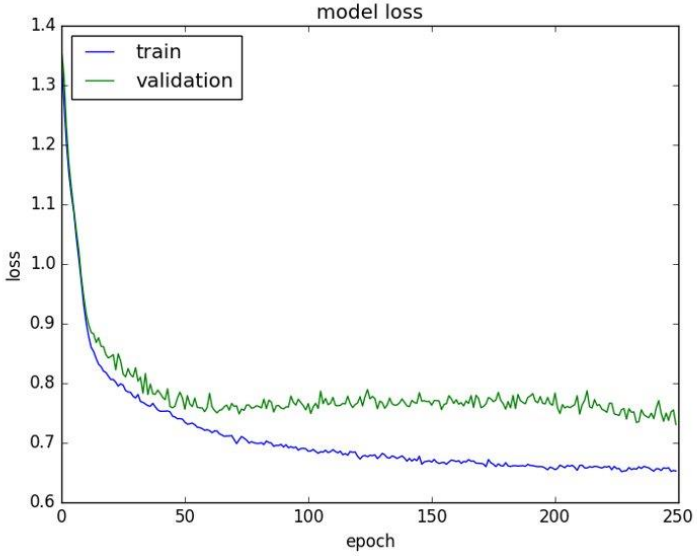
- Classification – LSTM
- Confusion Matrix
- IRL – Convergence
- Reward Function
- Value Function
- Similarity
- Automated camera control

Classification - LSTM

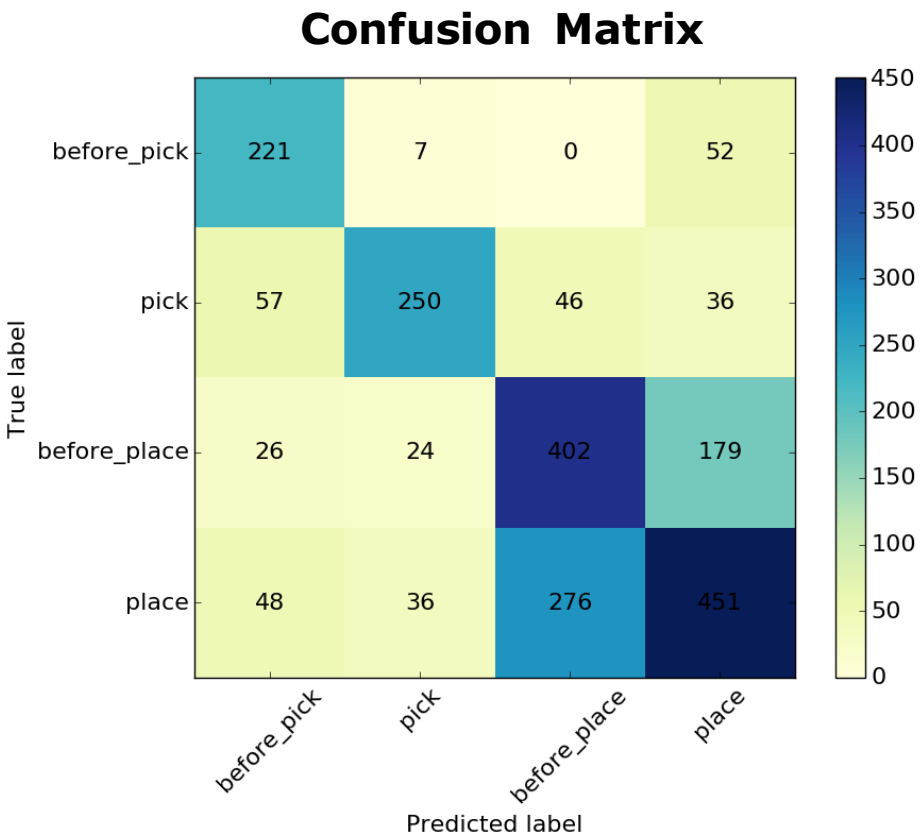
- LSTM layer with 10 neurons were used
- Inputs were: gripper position, rotation and jaw angle
- Training accuracy of 74.5% for classification was obtained.
- Test accuracy 71.3%
- Results show no overfitting and good generalization

Classification - LSTM

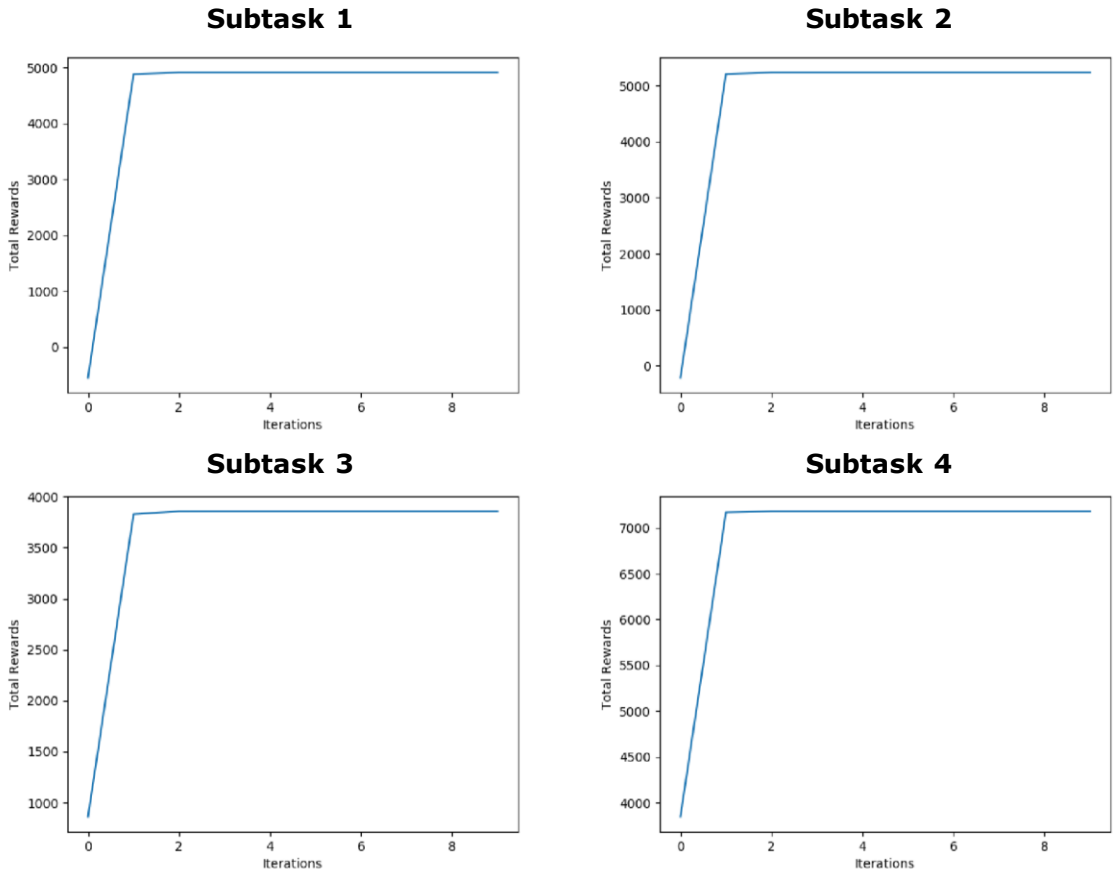




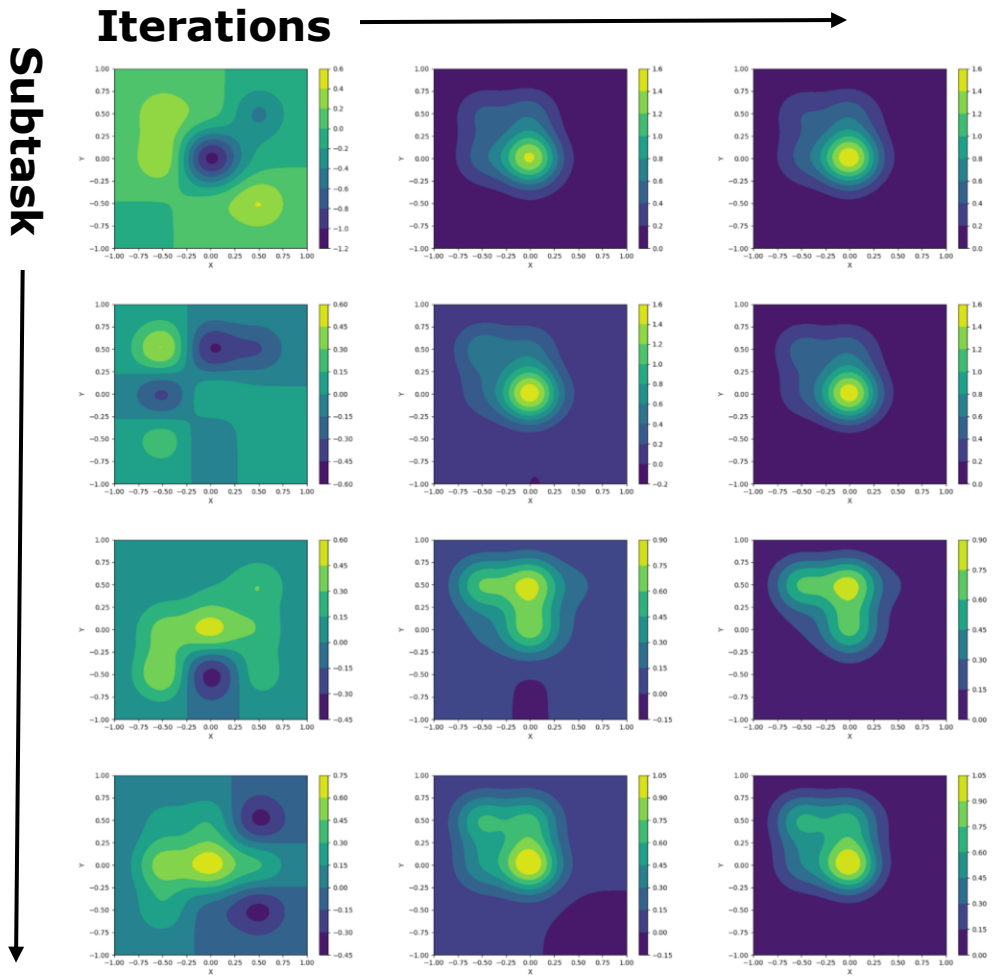
Confusion Matrix



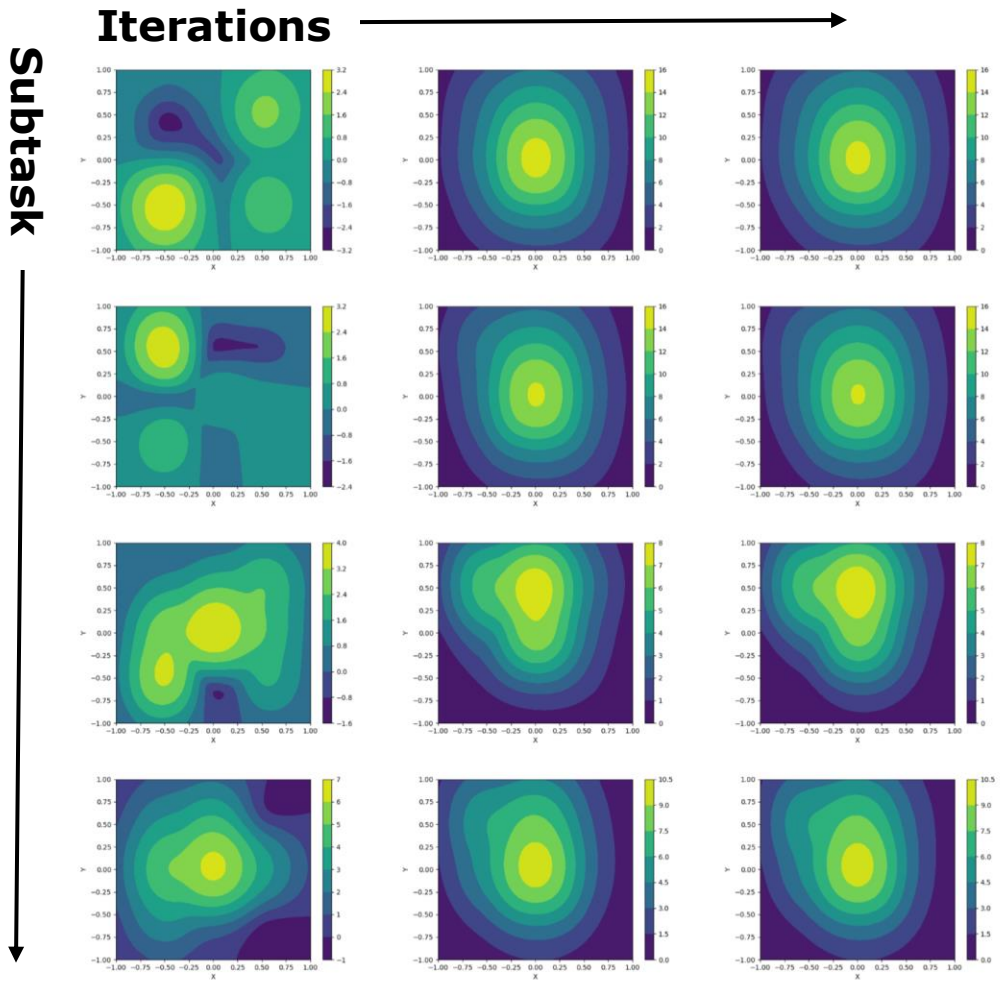
IRL - Convergence



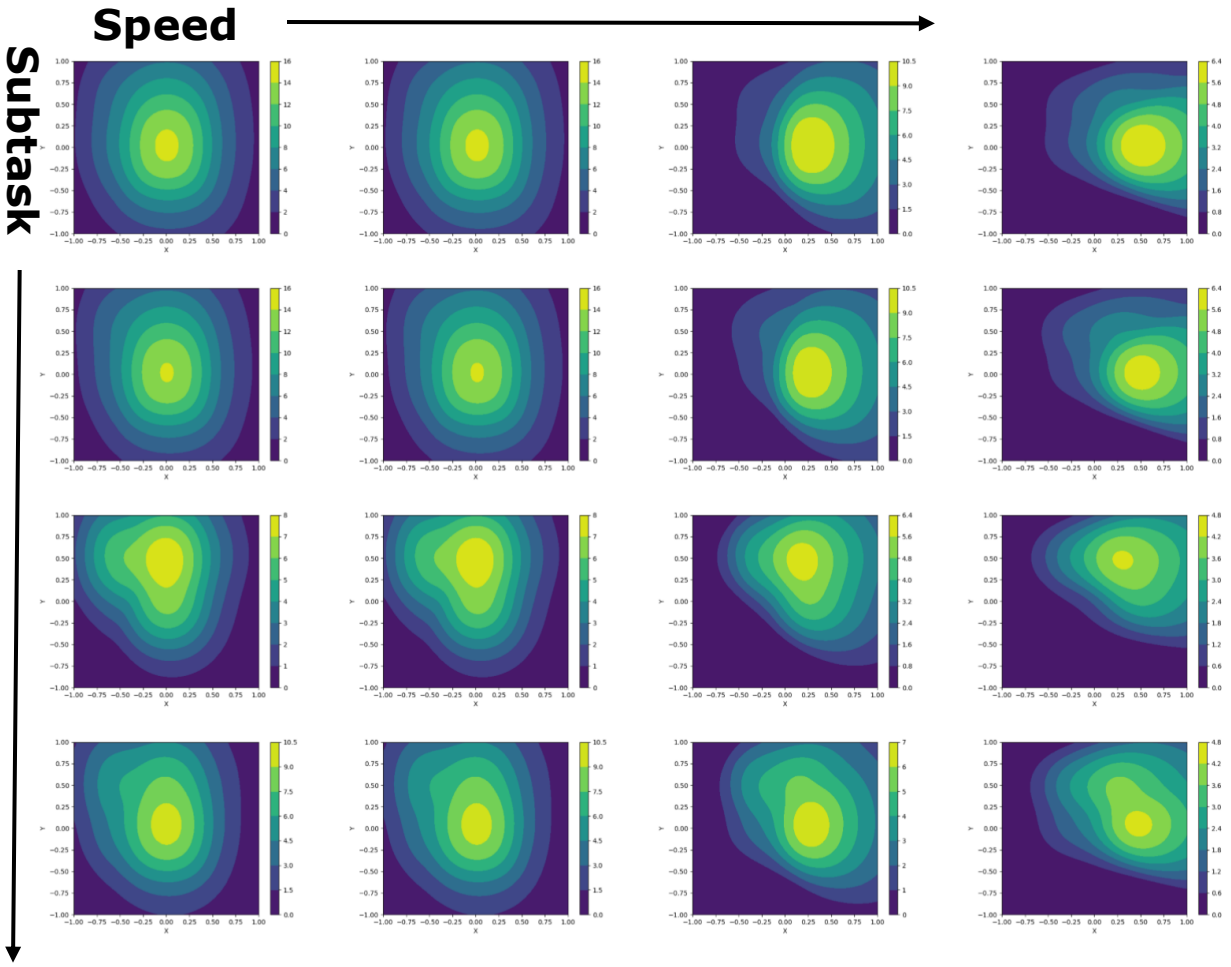
Reward Function



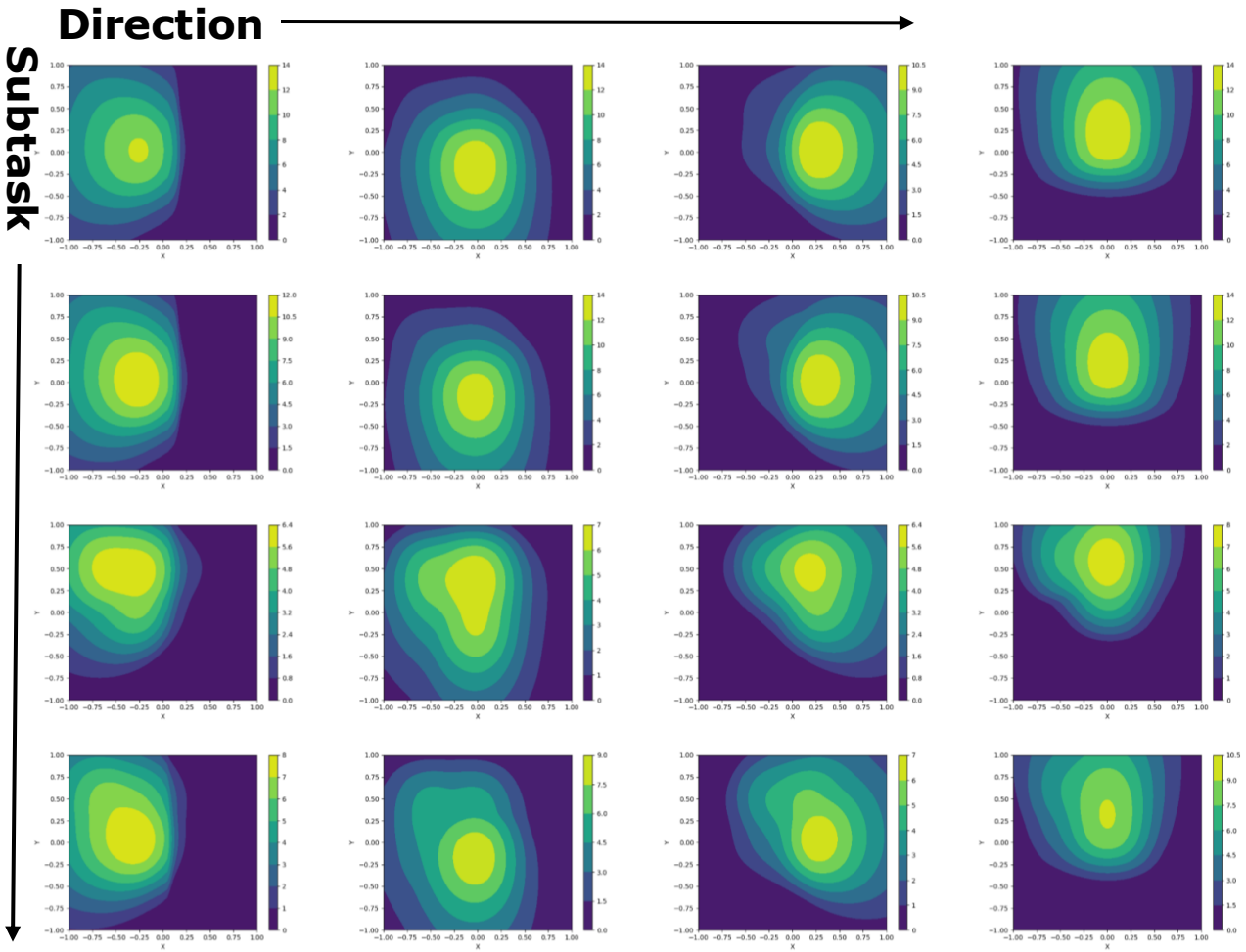
Value Function



Value Function – Varying Speed

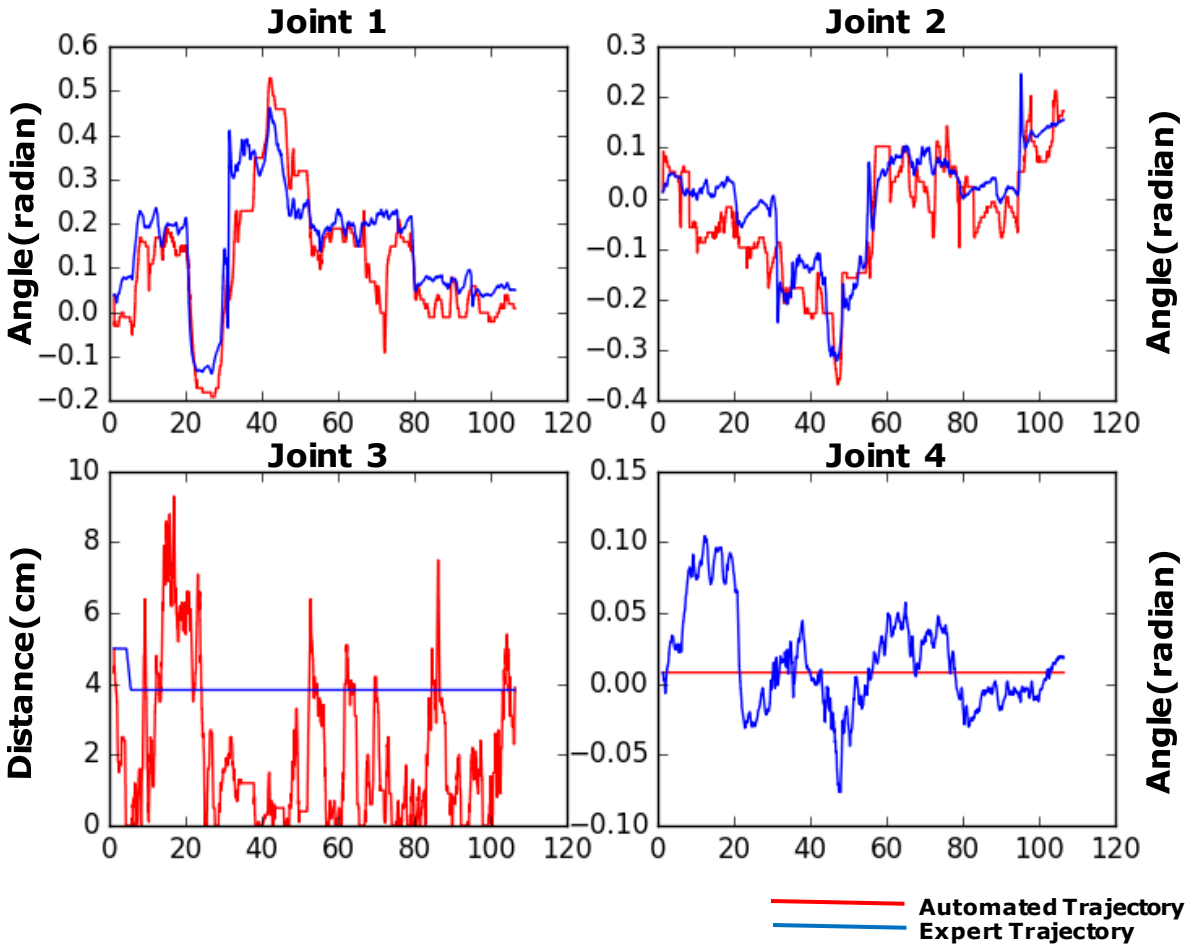


Value Function – Varying Direction

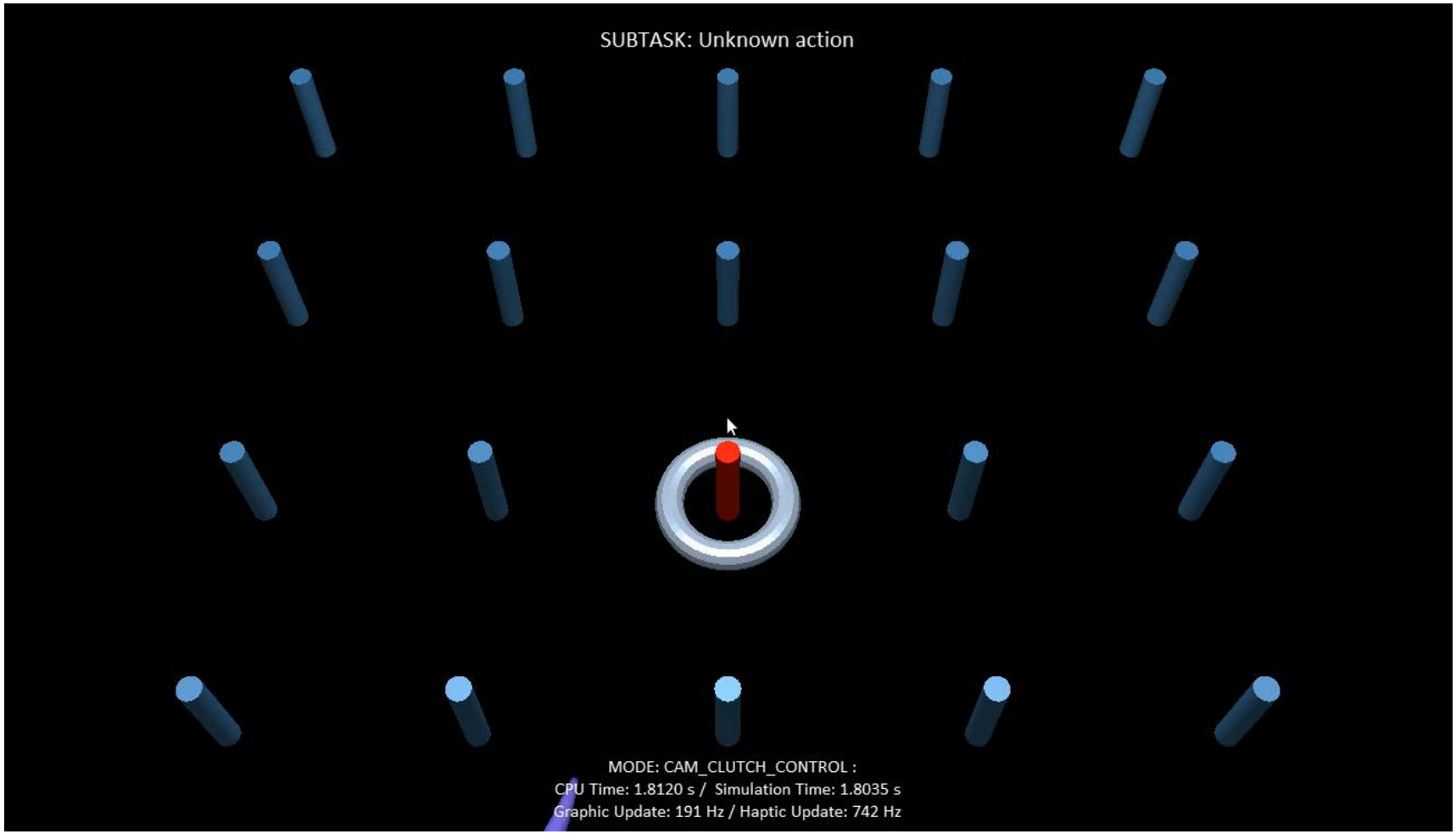


Similarity

- Similarity between automated and expert trajectories
- Unseen files by the algorithm
- Similarity metric defined earlier.
- Mean Similarity : 94.68%



Automated Camera Control



Conclusion

- Conclusion
- Future Work

Conclusion

- The approach presented in this thesis successfully automated endoscopic camera motion.
- Surgeon's intent was obtained from identifying the subtask.
- Automated camera motion was dependent on this subtask and the motion of the gripper.
- The camera did not only track the gripper motion but also was able to look ahead and show objects of importance in the environment.
- In conclusion, the thesis presented an approach in achieving an intelligent camera control.

Future Work

- Increase the accuracy of classification.
- Implement the algorithm on the hardware.
- Extend the approach presented to other tasks such as suturing, needle passing which are more dependent on camera views.
- Collect data from expert surgeons in a more realistic environment of surgery.
- If same approach can be extended to other tasks, these tasks then can be combined to achieve automation for the whole surgical procedure.

References

[1] Leonard, S., Wu, K.L., Kim, Y., Krieger, A. and Kim, P.C., 2014. Smart tissue anastomosis robot (STAR): A vision-guided robotics system for laparoscopic suturing. *IEEE Transactions on Biomedical Engineering*, *61*(4), pp.1305-1317.

[2] Shademan, A., Decker, R.S., Opfermann, J.D., Leonard, S., Krieger, A. and Kim, P.C., 2016. Supervised autonomous robotic soft tissue surgery. *Science translational medicine*, *8*(337), pp.337ra64-337ra64.

[3] Murali, A., Sen, S., Kehoe, B., Garg, A., McFarland, S., Patil, S., Boyd, W.D., Lim, S., Abbeel, P. and Goldberg, K., 2015, May. Learning by observation for surgical subtasks: Multilateral cutting of 3d viscoelastic and 2d orthotropic tissue phantoms. In Robotics and Automation (ICRA), 2015 IEEE International Conference on (pp. 1202-1209). IEEE.

[4] Thananjeyan, B., Garg, A., Krishnan, S., Chen, C., Miller, L. and Goldberg, K., 2017, May. Multilateral surgical pattern cutting in 2d orthotropic gauze with deep reinforcement learning policies for tensioning. In Robotics and Automation (ICRA), 2017 IEEE International Conference on (pp. 2371-2378). IEEE.

[5] Lee, C., Wang, Y.F., Uecker, D.R. and Wang, Y., 1994, September. Image analysis for automated tracking in robot-assisted endoscopic surgery. In Proceedings of 12th International Conference on Pattern Recognition (pp. 88-92). IEEE.

[6] Wei, G.Q., Arbter, K. and Hirzinger, G., 1997. Real-time visual servoing for laparoscopic surgery. Controlling robot motion with color image segmentation. IEEE Engineering in Medicine and Biology Magazine, 16(1), pp.40-45.

[7] Omote, K., Feussner, H., Ungeheuer, A., Arbter, K., Wei, G.Q., Siewert, J.R. and Hirzinger, G., 1999. Self-guided robotic camera control for laparoscopic surgery compared with human camera control. The American journal of surgery, 177(4), pp.321-324.

[8] Voros, S., Haber, G.P., Menudet, J.F., Long, J.A. and Cinquin, P., 2010. ViKY robotic scope holder: Initial clinical experience and preliminary results using instrument tracking. IEEE/ASME transactions on mechatronics, 15(6), pp.879-886.

[9] Voros, S., Long, J.A. and Cinquin, P., 2007. Automatic detection of instruments in laparoscopic images: A first step towards high-level command of robotic endoscopic holders. The International Journal of Robotics Research, 26(11-12), pp.1173-1190.

[10] Eslamian, S., Reisner, L.A., King, B.W. and Pandya, A.K., An Autonomous Camera System using the da Vinci Research Kit.

References

[11] Eslamian, S., Reisner, L.A., King, B.W. and Pandya, A.K., 2016, April. Towards the Implementation of an Autonomous Camera Algorithm on the da Vinci Platform. In MMVR (pp. 118-123).

[12] Ali, S.M., Reisner, L.A., King, B., Cao, A., Auner, G., Klein, M. and Pandya, A.K., 2008. Eye gaze tracking for endoscopic camera positioning: an application of a hardware/software interface developed to automate Aesop. Studies in health technology and informatics, 132, pp.4-7.

[13] Nathan, C.A.O., Chakradeo, V., Malhotra, K., D'Agostino, H. and Patwardhan, R., 2006. The voice-controlled robotic assist scope holder AESOP for the endoscopic approach to the sella. Skull Base, 16(3), p.123.

[14] Kwon, D.S., Ko, S.Y. and Kim, J., 2008. Intelligent laparoscopic assistant robot through surgery task model: how to give intelligence to medical robots. In Medical Robotics. InTech.

[15] Pandya, A., Reisner, L., King, B., Lucas, N., Composto, A., Klein, M. and Ellis, R., 2014. A review of camera viewpoint automation in robotic and laparoscopic surgery. Robotics, 3(3), pp.310-329.

[16] Ng, A.Y. and Russell, S.J., 2000, June. Algorithms for inverse reinforcement learning. In Icml (pp. 663-670).

[17] Arora, S. and Doshi, P., 2018. A Survey of Inverse Reinforcement Learning: Challenges, Methods and Progress. arXiv preprint arXiv:1806.06877.

[18] Levine, S. and Koltun, V., 2012. Continuous inverse optimal control with locally optimal examples. arXiv preprint arXiv:1206.4617.

[19] Gers, F.A., Schmidhuber, J. and Cummins, F., 1999. Learning to forget: Continual prediction with LSTM.

[20] Sak, H., Senior, A. and Beaufays, F., 2014. Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In Fifteenth annual conference of the international speech communication association.

[21] Conti, Francois, D. Morris, F. Barbagli, and C. Sewell. "CHAI 3D." Online: http://www. chai3d. org (2006).

[22] Munawar A.,2018. GitHub repository, https://github.com/WPI-AIM/chai_env


Thank You!
Questions please!


WORCESTER POLYTECHNIC INSTITUTE
LEHR UND KUNST
1865