Slide 1

Real-time Vision Task with a High Degree of Freedom Robot Arm

Haochen Shi, Rui Pan, Chenhao Lu

{hshi74, rpan33, clu92}@wisc.edu
CS 639: Computer Vision Fall 2020 Course Project

Slide 2

Project Overview

TASK

Taking a stable Panorama

Real-time Object Detection

Object Tracking

X

Robot Control Method

Manual Control with a Joystick

Fully Autonomous

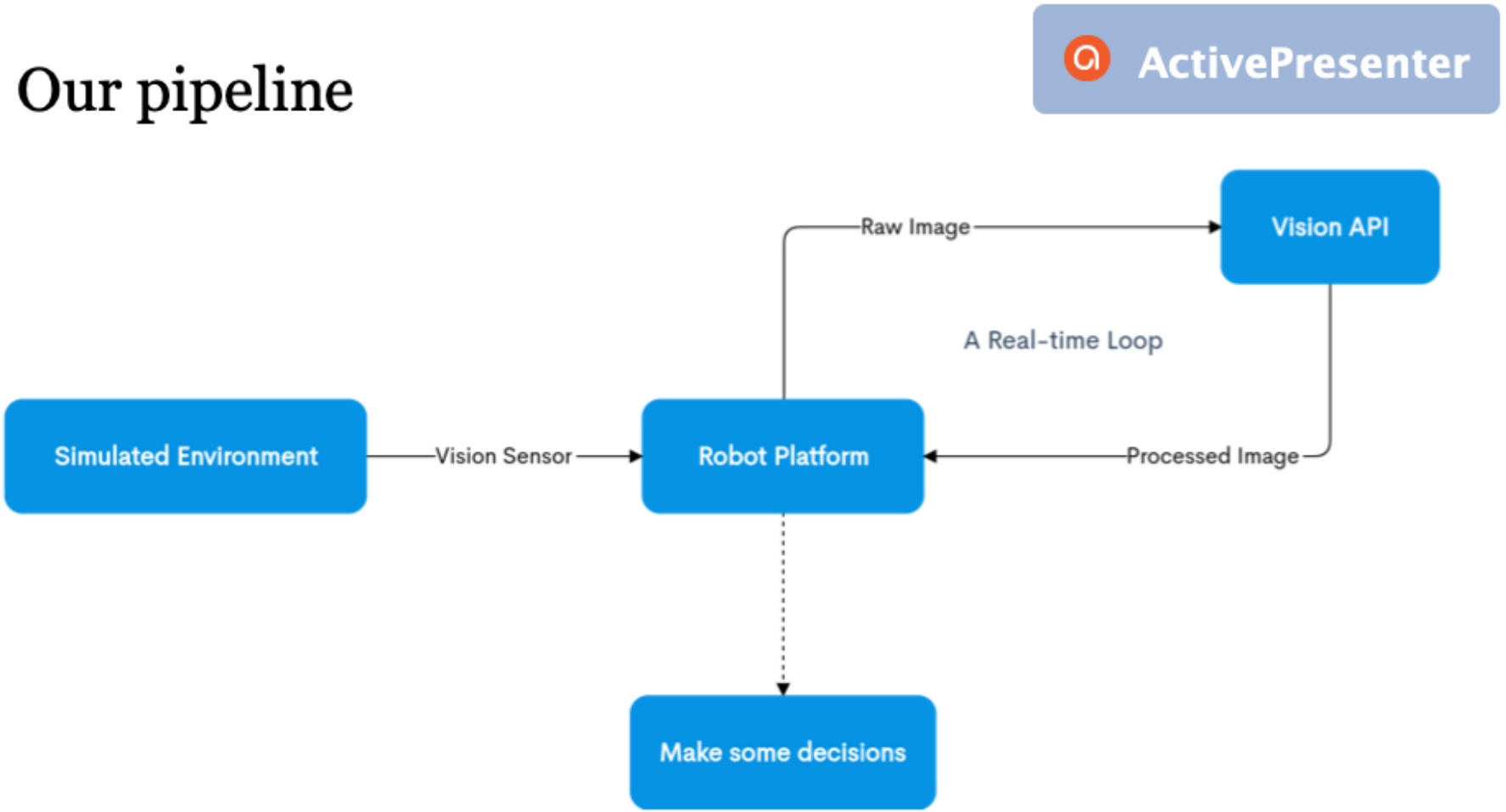
Slide 3

Specifications

- All tasks are tested on the Rethink Robotics Sawyer (7-DOF) in a simulated scene in CoppeliaSim
- We used a motion planning framework called RelaxedIK (Relaxed Inverse Kinematics) to calculate the motion of the robot arm

Slide 4

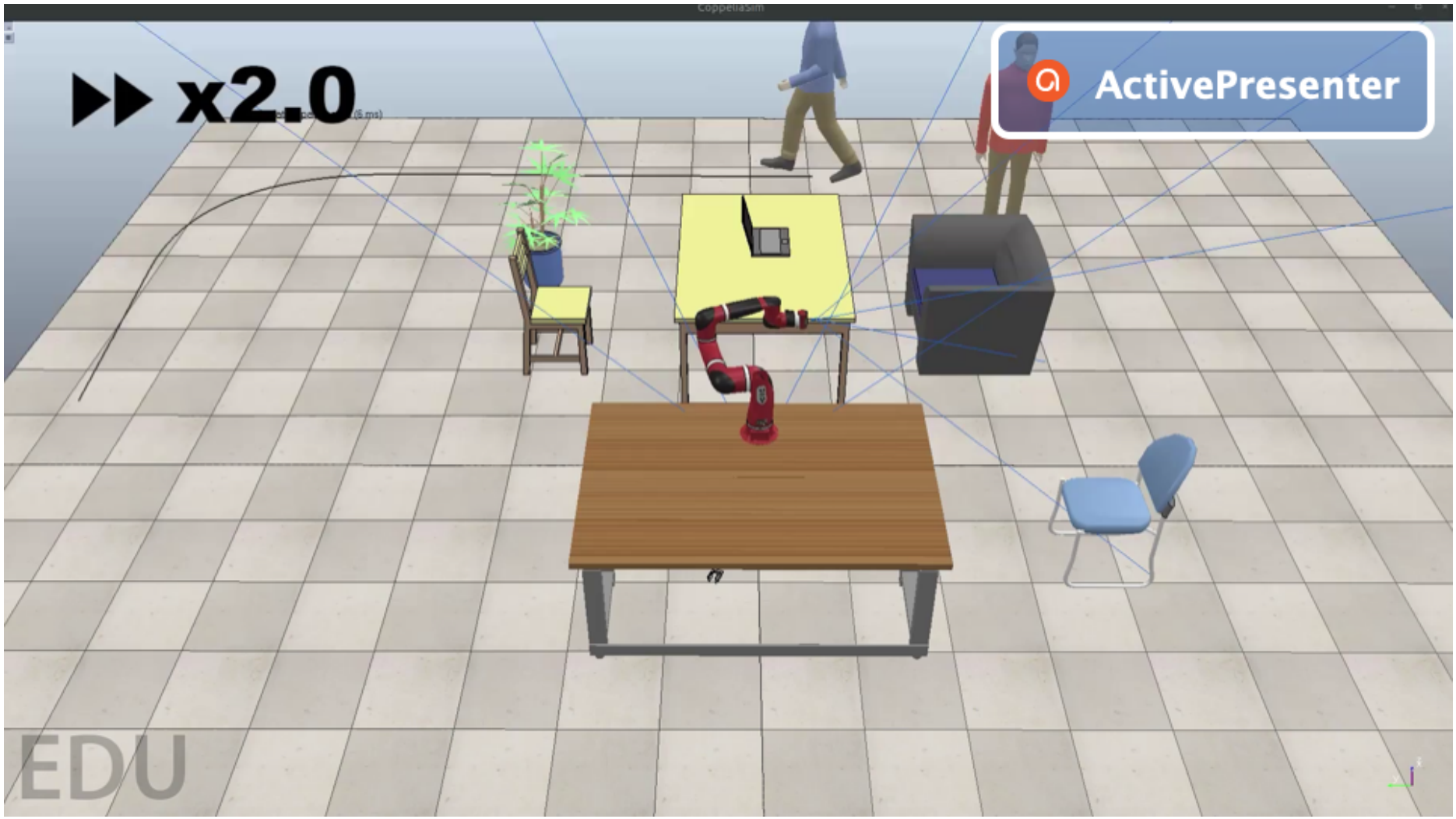
Our pipeline



Slide 5

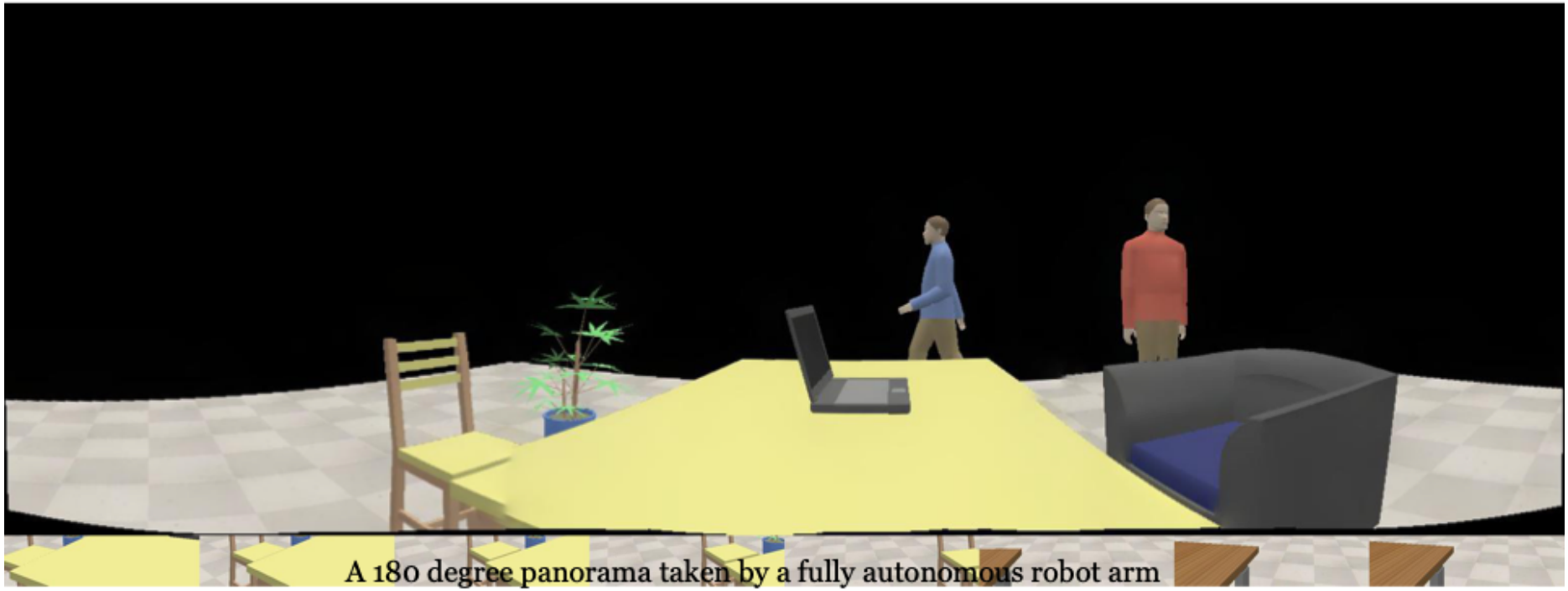
1/3: Taking a Stable Panorama

Slide 6



Slide 7

Task 1 Result



A 180 degree panorama taken by a fully autonomous robot arm

A montage of all the frames captured

Slide 8

2/3: Detecting Objects in Real Time

Slide 9

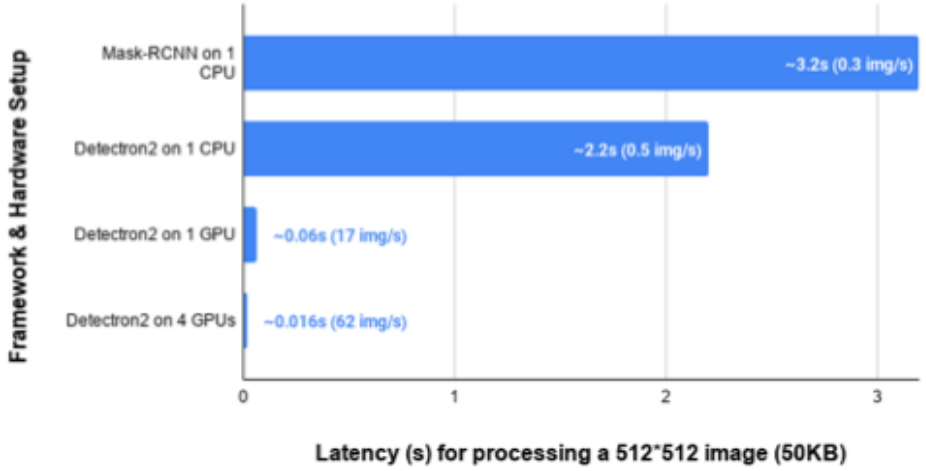
First Attempt: Mask-RCNN + COCO Dataset

- PyTorch + Mask-RCNN
- Downloaded pre-trained model on the COCO dataset
- Latency: ~4 seconds per image

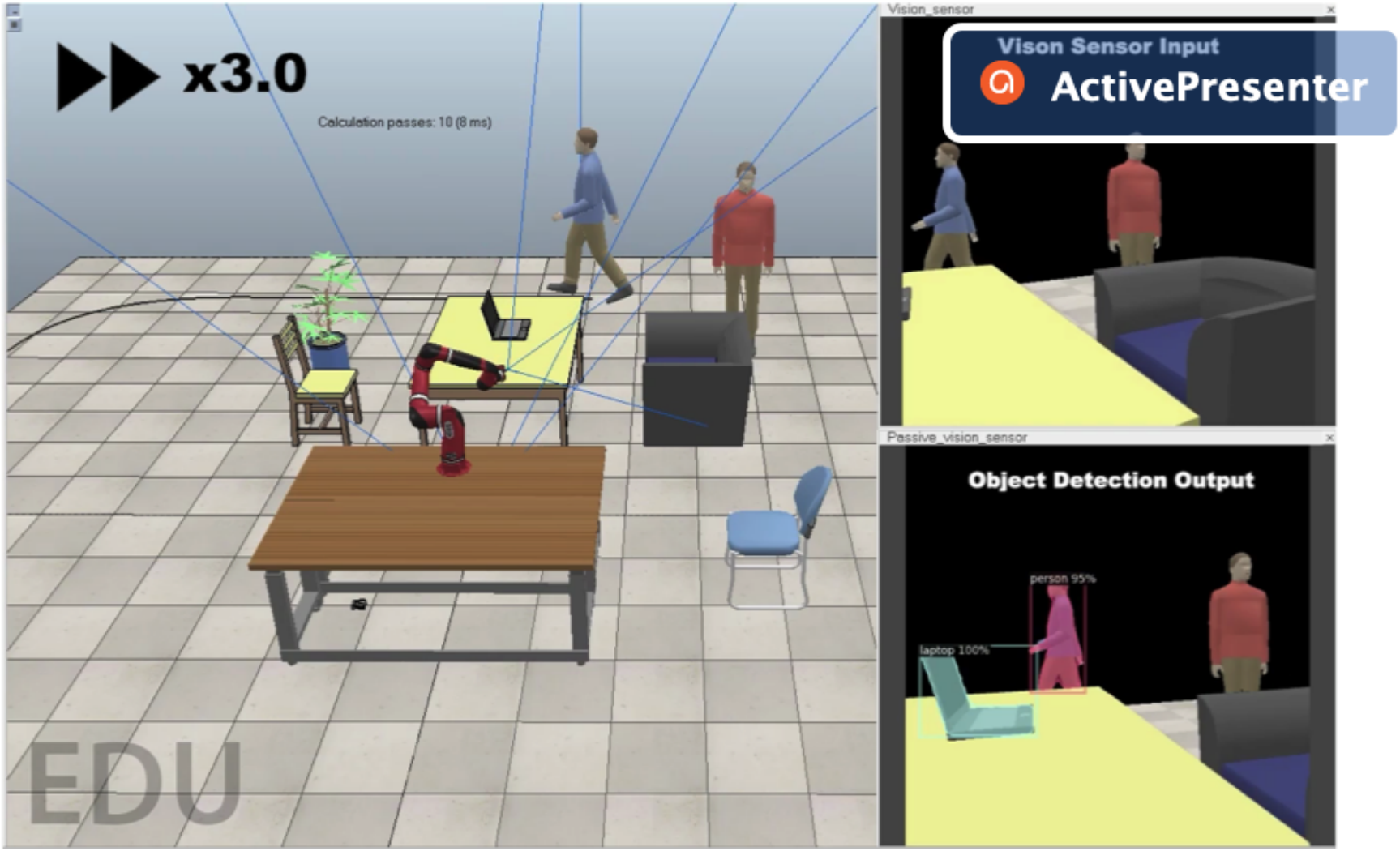
Slide 10

Second attempt: Detectron2 & GPU on Azure

- Detectron2 by Facebook: faster speed
- Theoretical throughput: ~60 img/s on 4 NVIDIA Tesla P100 GPUs
- State-of-the-art detection methods has a ~0.1s turnaround time with 1 GPU (GTX 1080 Ti)



Slide 11


x3.0
Calculation passes: 10 (8 ms)
EDU
Vision_sensor
Vison Sensor Input
ActivePresenter
Passive_vision_sensor
Object Detection Output
person 95%
laptop 100%

Slide 12

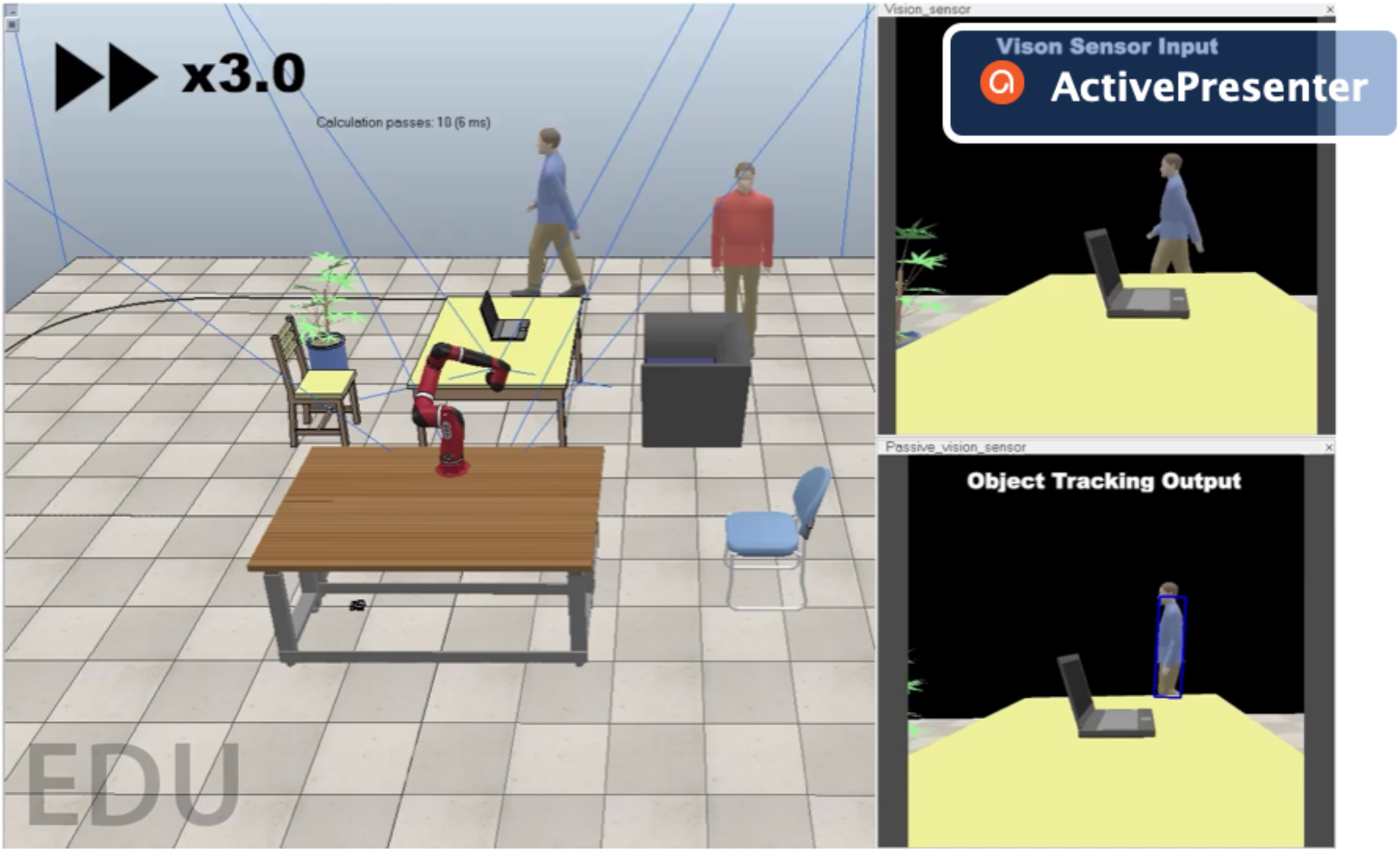
3/3: Object Tracking

Slide 13

Algorithms: MeanShift and CAMShift

- MeanShift:
  - Histogram-based template tracking
  - Disadvantages: Not robust to changes in window size and rotation
- CAMShift (our final choice):
  - Continuously Adaptive Mean Shift
  - Applies MeanShift first
  - Once MeanShift converges, CAMShift updates the size and the orientation of the window

Slide 14


x3.0
Calculation passes: 10 (6 ms)
EDU
Vision_sensor
Vison Sensor Input
ActivePresenter
Passive_vision_sensor
Object Tracking Output

Slide 15

Conclusions

- Computer vision tasks work very well on high degree of freedom robot arms
- Different control methods works well for each vision task

Slide 16

Future Directions

- Task 2 (Object detection):
  - Use GPUs to speed up our real-time detection
  - Test in real-world settings
- Task 3 (Object tracking):
  - Make the camera follow the object being tracked automatically
  - Improve the robustness to changes in scale and orientation
- Thank you!