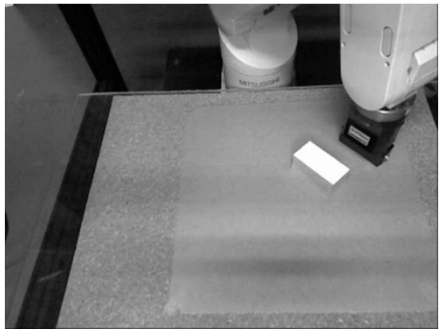
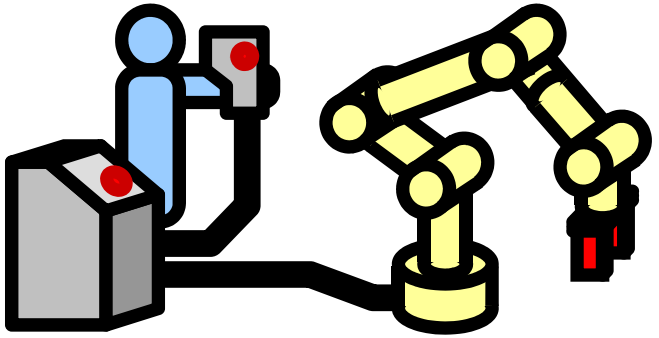
Lighting- and Occlusion-robust View-based Teaching/Playback for Model-free Robot Programming

*Yusuke MAEDA
(Yokohama National University)
Yoshito SAITO
(Ricoh Corp.)



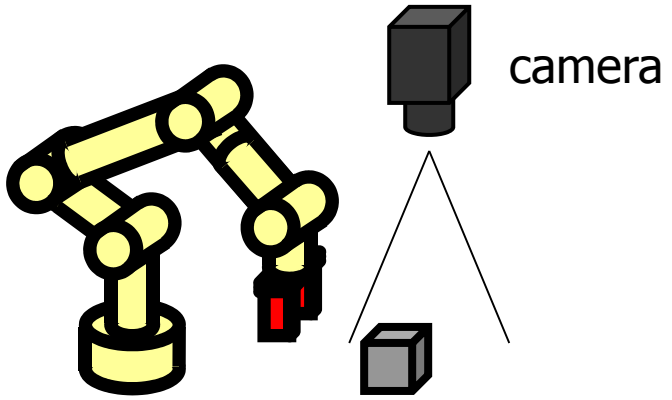
Background

- Conventional Teaching/Playback
  - still widely used
  - **model-free**: neither task-specific models nor object-specific models are necessary
  - for constant task conditions
    - e.g.) initial pose of object does not change



When the initial object pose is not constant…

- Object localization with cameras
  - Model-based image processing
    - Geometric feature extraction: edge, vertex, …
    - Pattern matching
  - Object-specific: model-freeness is lost



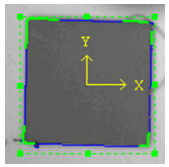
Motivation

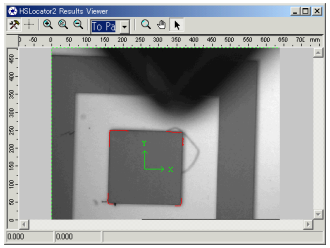
- To develop a **model-free** robot programming method that can cope with change of task conditions

"View-based teaching/playback": robot programming with **view-based** image processing [Maeda 2011 ICRA]

Model-based vs. View-based

- Model-based approach
  - with object-specific models
  - accurate
  - cumbersome
- View-based (Appearance-based) approach
  - without object-specific models
  - versatile





View-based Teaching/Playback [Maeda 2011 ICRA]

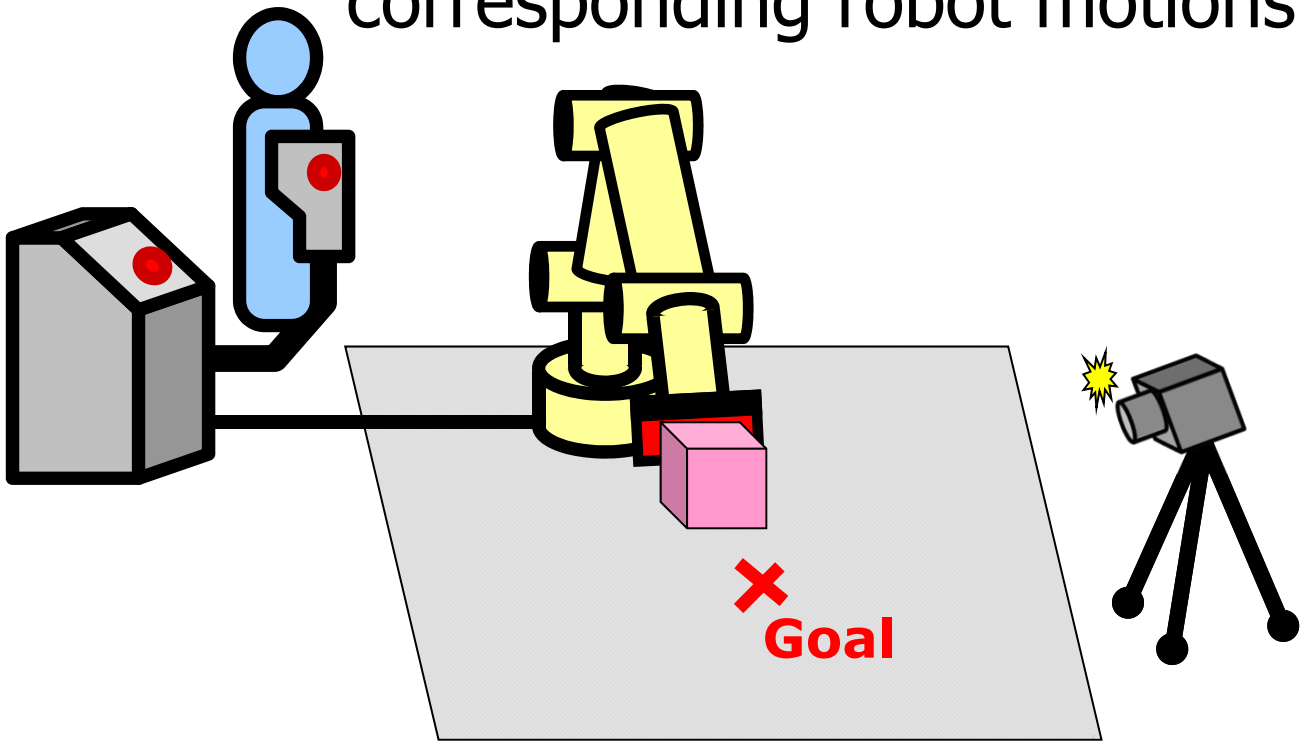
- View-based image processing using PCA
  - not object-specific
  - no need for camera calibration
- Adaptability to change of initial object pose using the generalization ability of neural networks
  - generalization from multiple demonstrations

Related Works

- [Zhang et al. 2000]: View-based fine positioning for grippers
- [Zhao et al. 2008]: View-based visual servoing for relative positioning
- [Levine et al. 2016]: View-based grasping of novel objects through massive learning

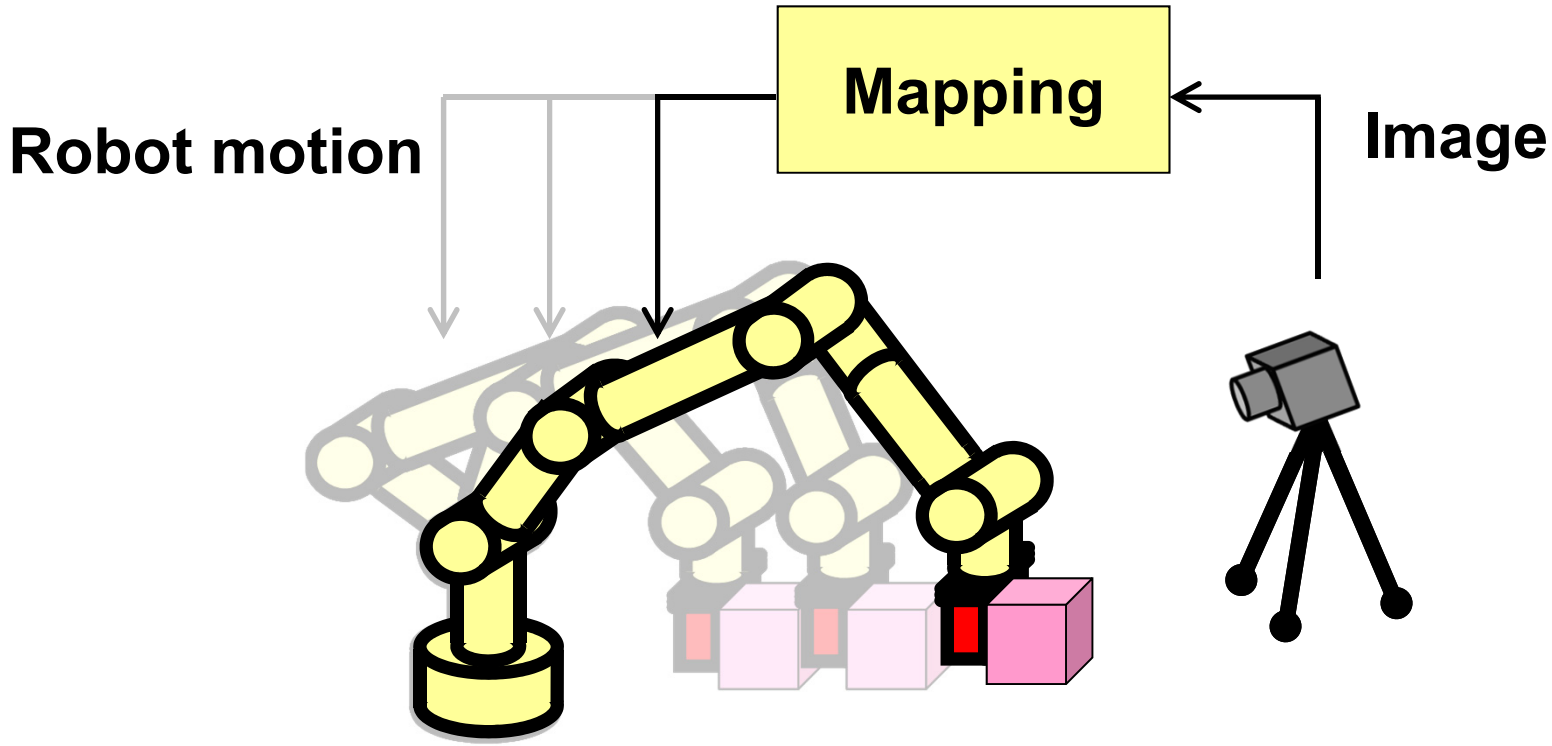
Overview of View-based Teaching/Playback (1/3)

- Human Demonstration
  - Record all the scene images and corresponding robot motions



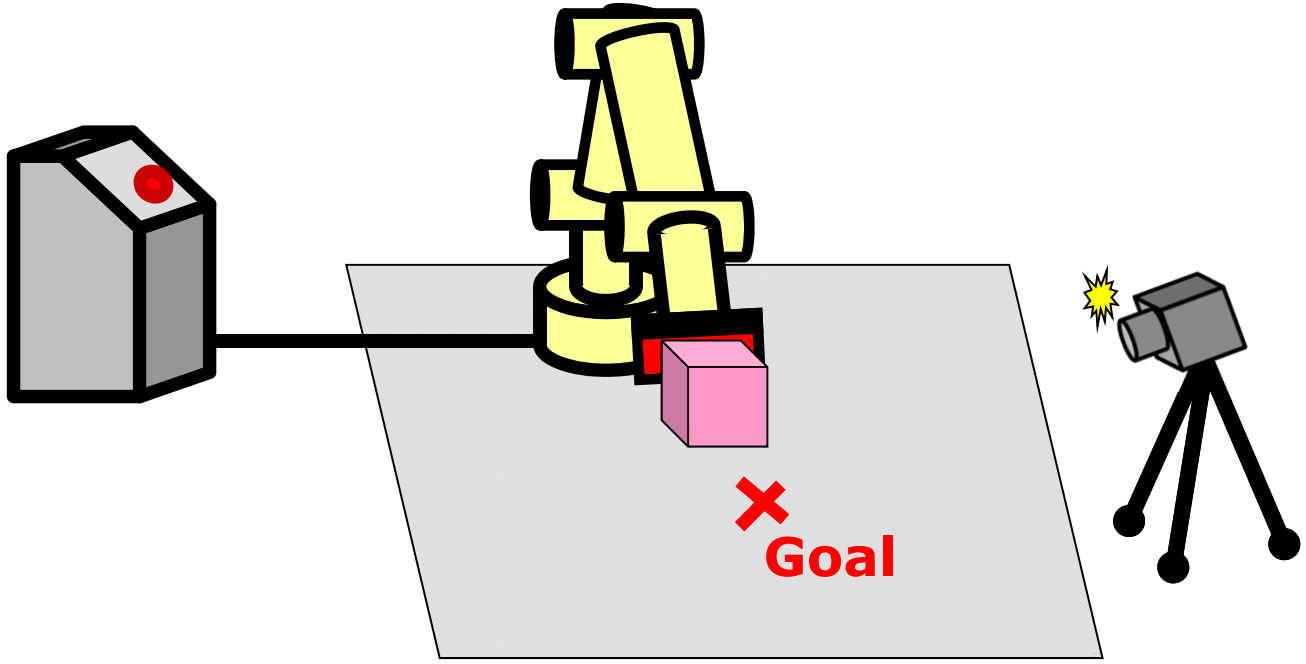
Overview of View-based Teaching/Playback (2/3)

- Mapping Acquisition
  - From scene image to robot motion

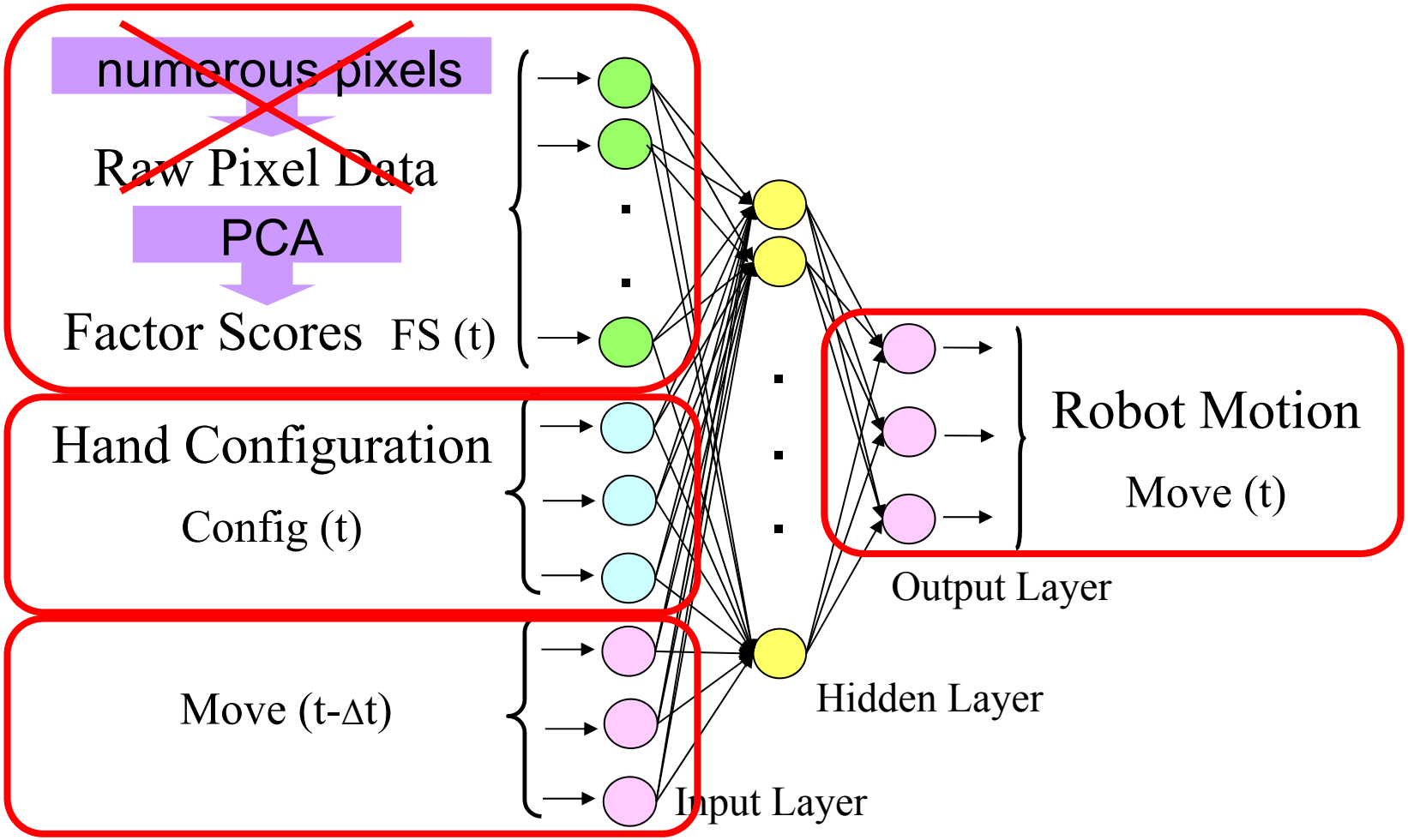


Overview of View-based Teaching/Playback (3/3)

- View-based Playback
  - Autonomous task execution using the acquired image-to-motion mapping

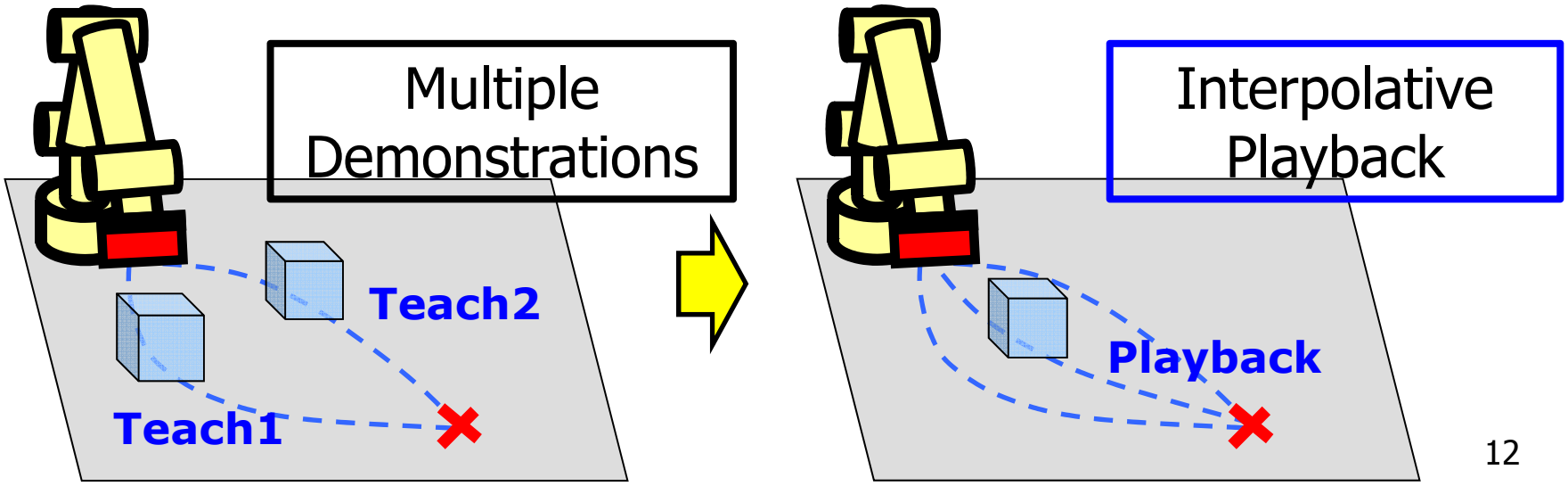


Neural Network for Mapping



Advantage of View-based Teaching/Playback

- Possible to cope with changes of task conditions (to some extent)
  - Use of generalization ability of NN
  - Model-free



Objective

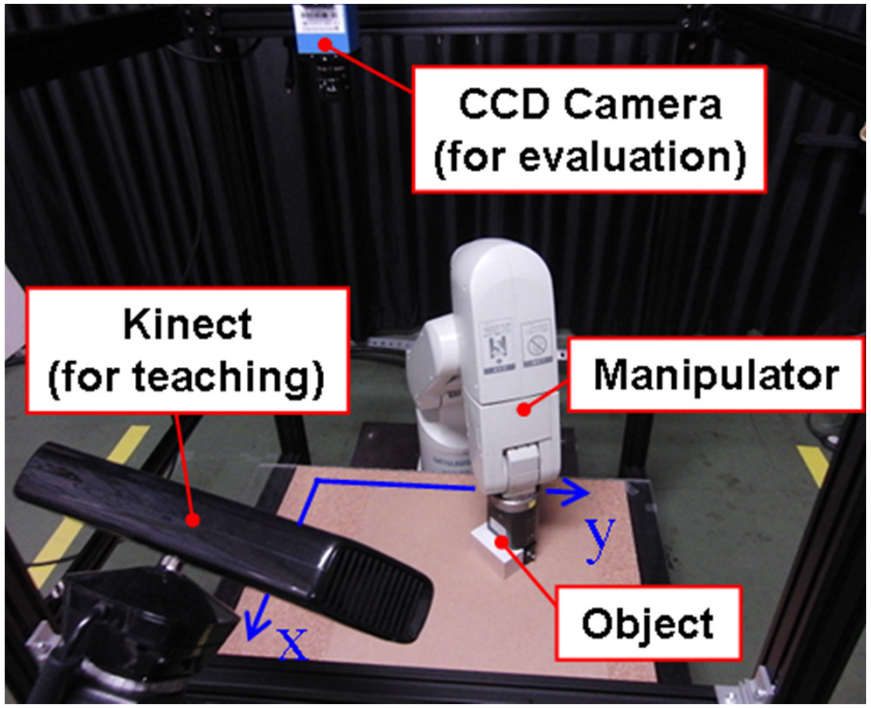
- Original view-based teaching/playback [Maeda 2011 ICRA]
  - sensitive to lighting conditions
  - sensitive to occlusions

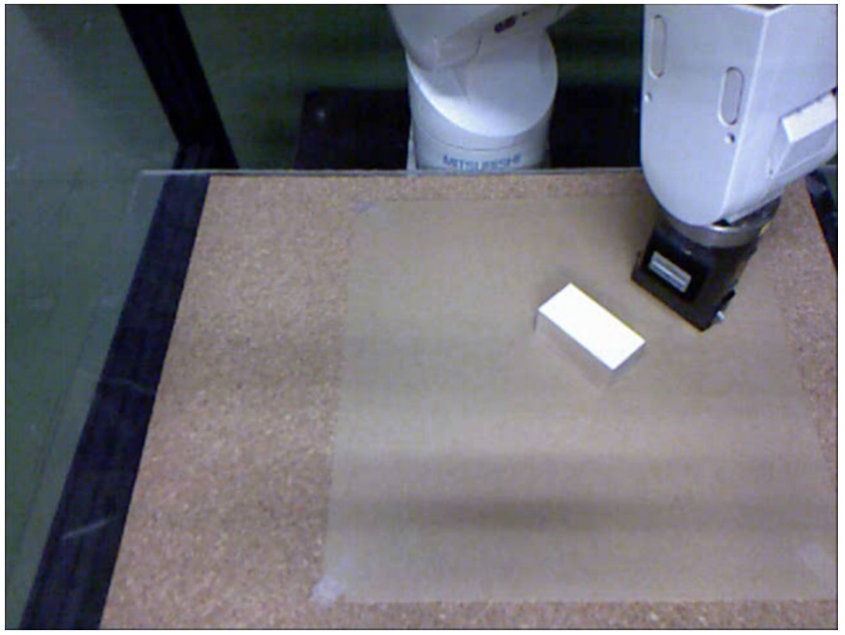
To make our view-based teaching/playback lighting- and occulusion-robust

- Use of range images
- Use of subimages

Experimental Setup

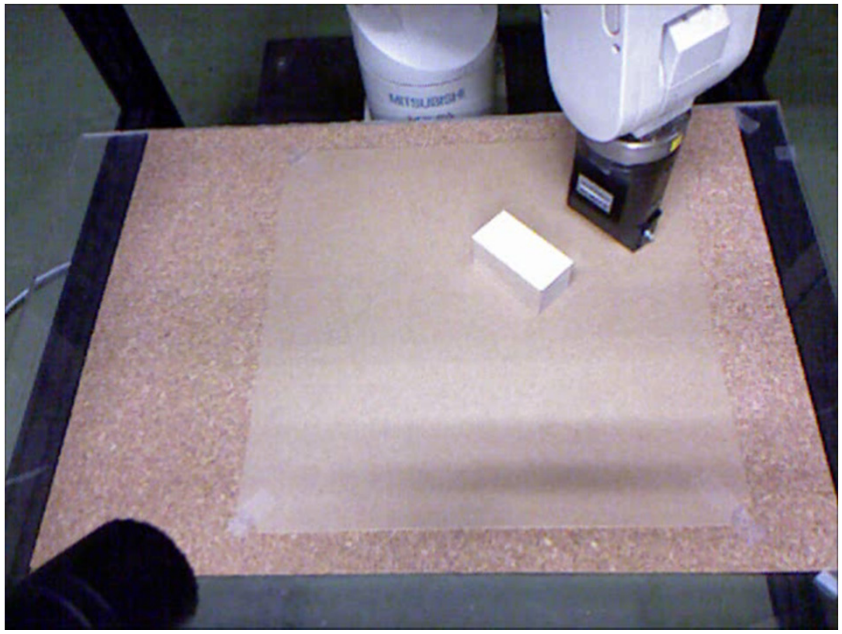
- Target task: Pushing on a plane
- Kinect for grayscale and range images



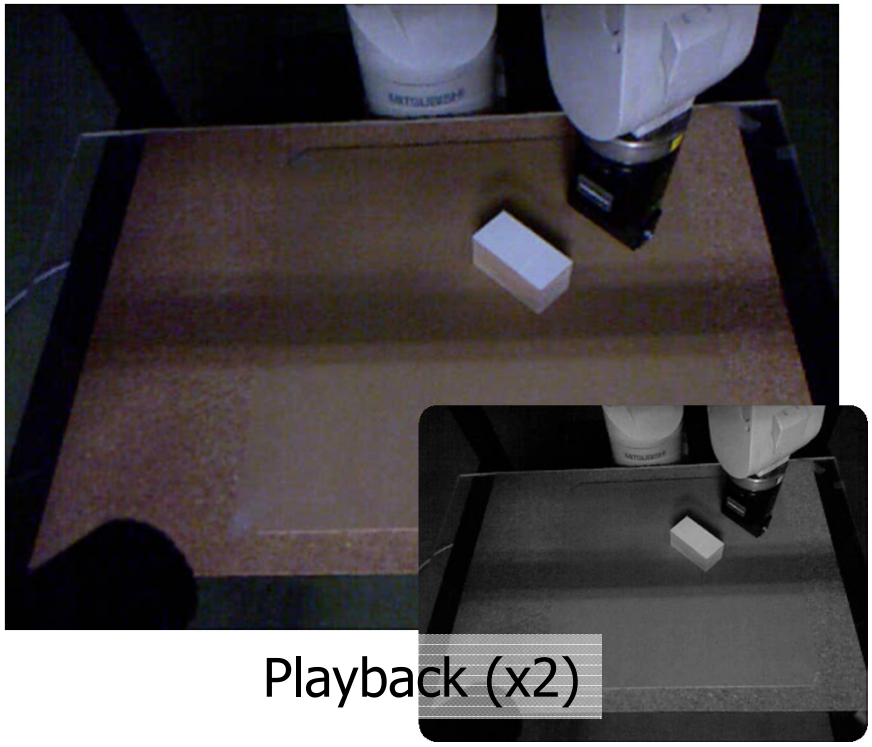


Teaching (x2)

Change of Lighting Condition (playback with grayscale images)

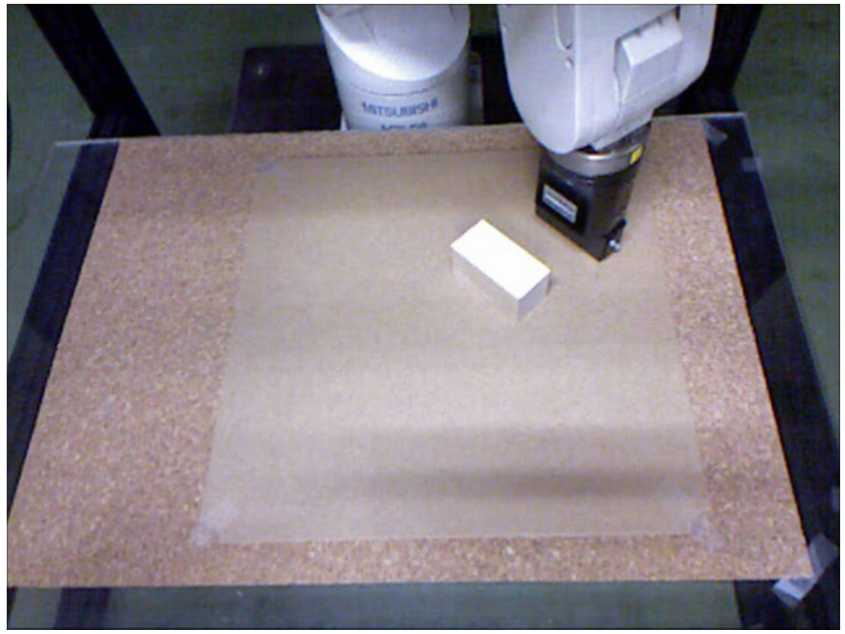


Teaching (x2)

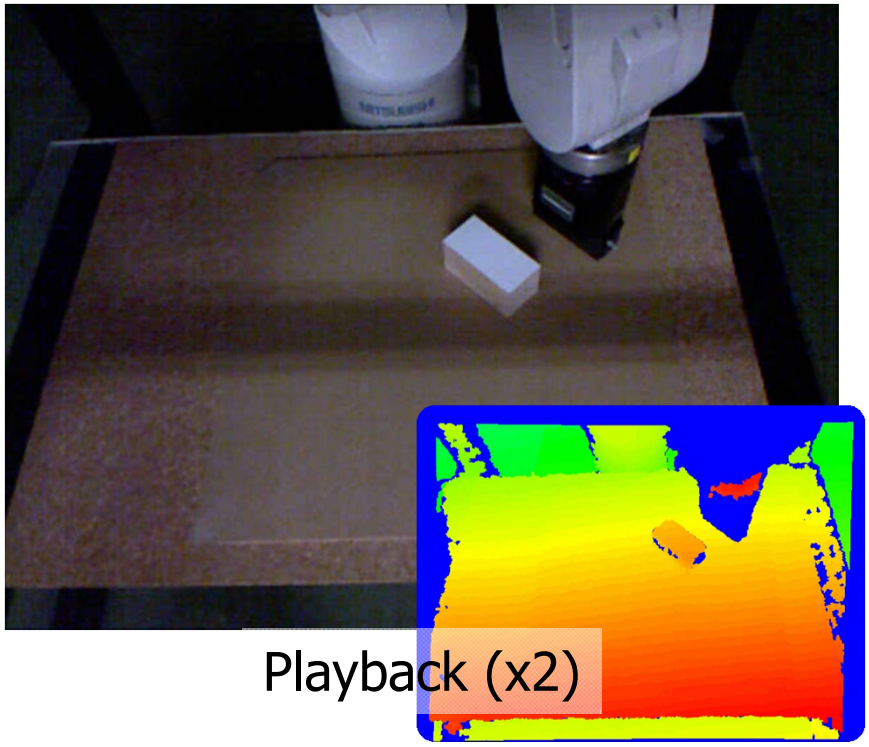


Playback (x2)

Change of Lighting Condition (playback with range images)

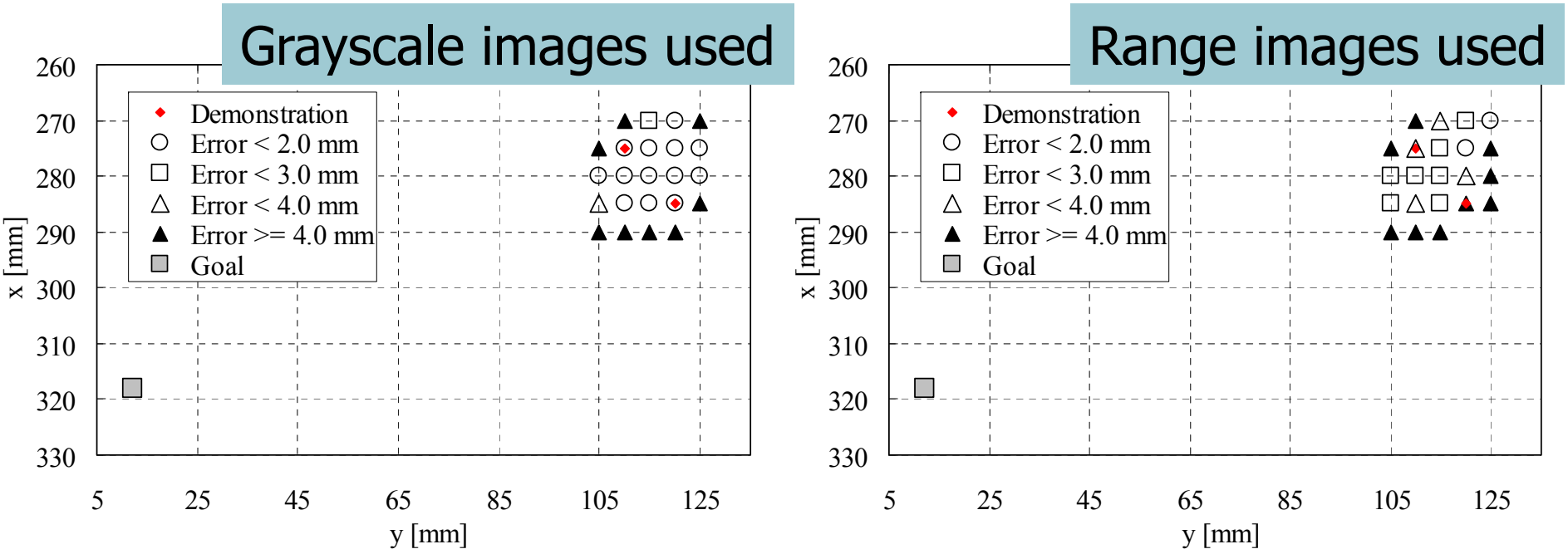


Teaching (x2)



Playback (x2)

Position Errors at Goal for Each Initial Positions



- Larger errors found for range images due to their noise

Grayscale Images vs. Range Images

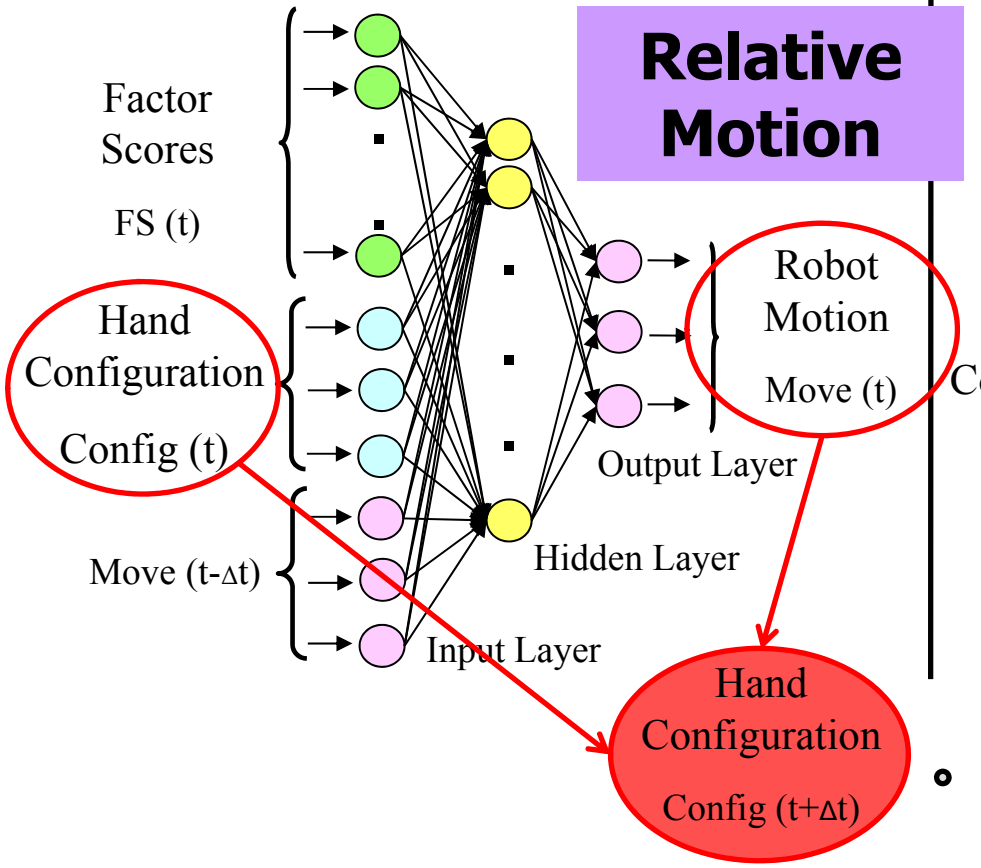
- Grayscale Images
  - Less noisy, but less robust
- Range Images
  - Noisy, but robust

Adaptive online switching of used images

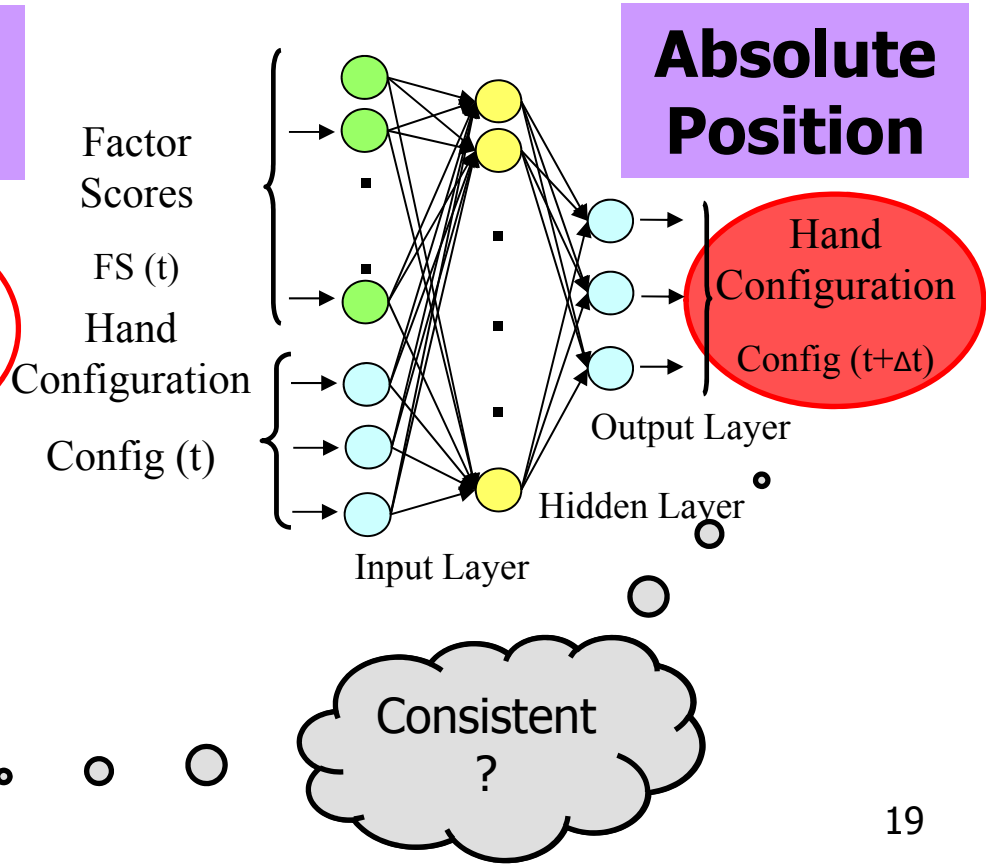
Achieve both accuracy and robustness in view-based teaching/playback

Switching Grayscale/Range Images (1/2)

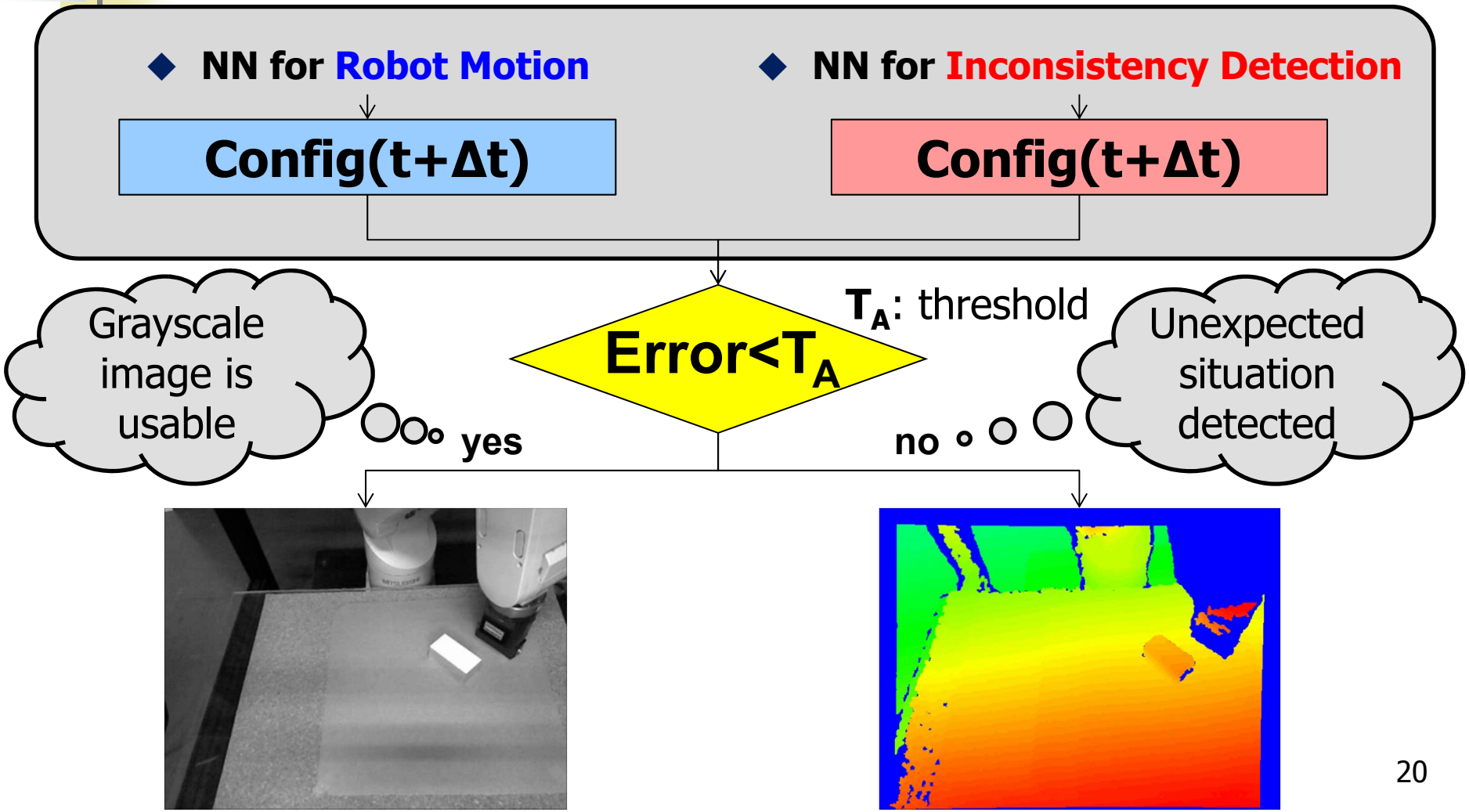
- **Neural Network for Robot Motion**



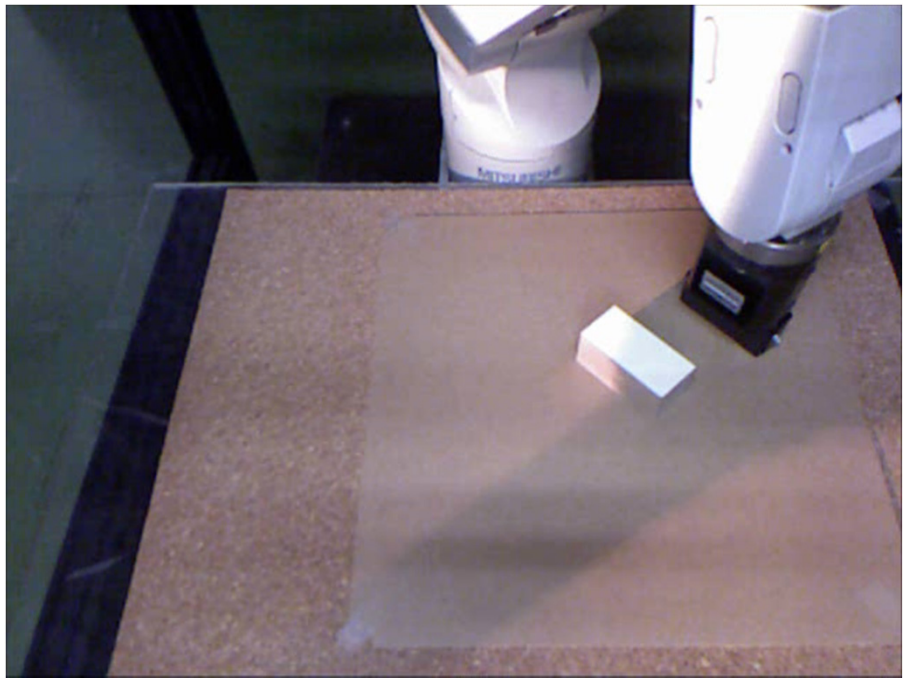
- **Additional Neural Network for Inconsistency Detection**



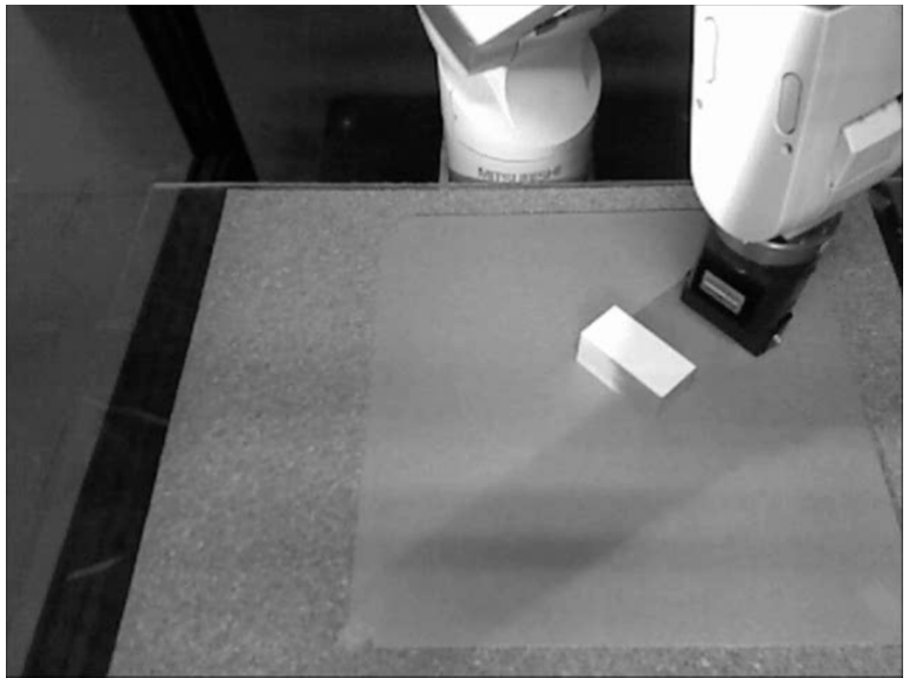
Switching Grayscale/Range Images (2/2)



Lighting-robust View-based Playback (light added)



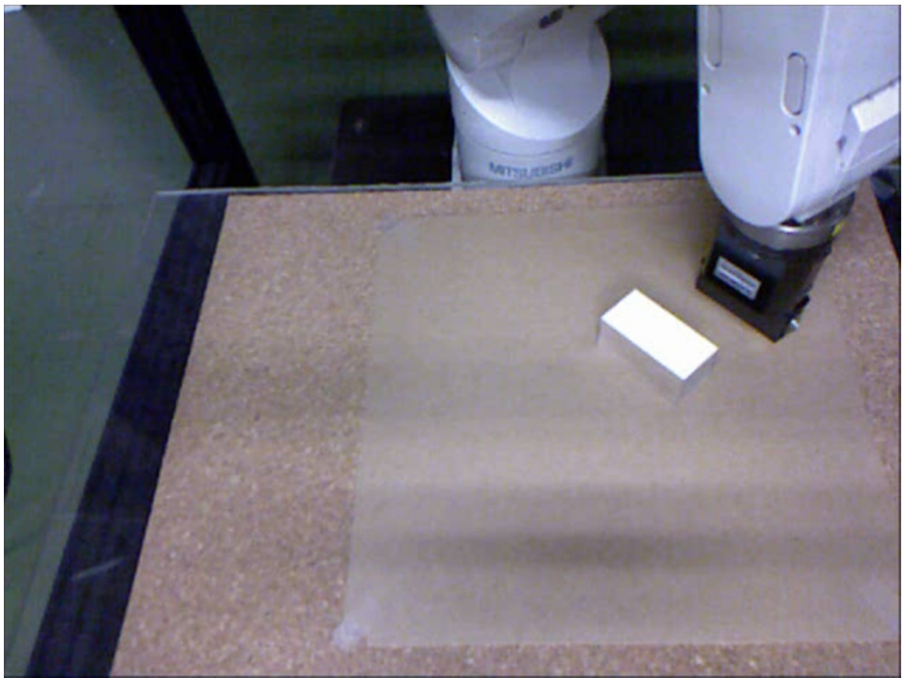
Playback (x2)



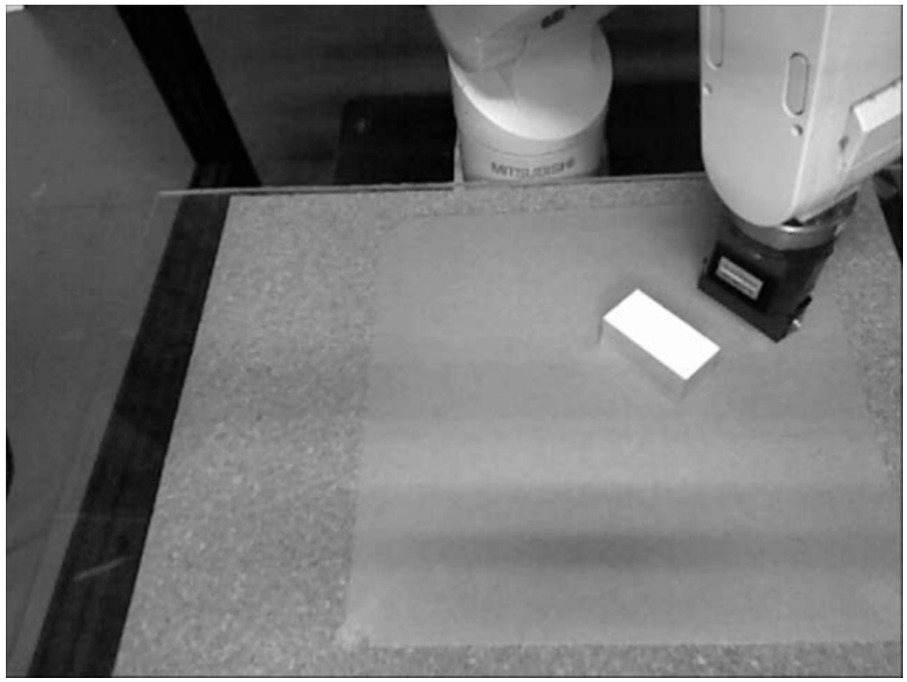
Used Images (x2)

Range images were used automatically due to change of lighting condition

Lighting-robust View-based Playback (light on/off)



Playback (x2)

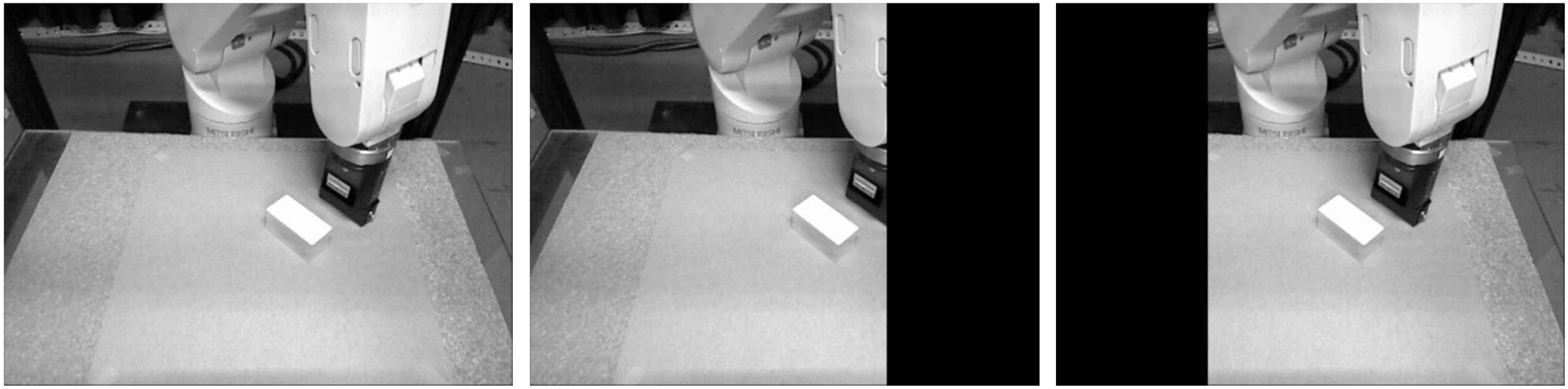


Used Images (x2)

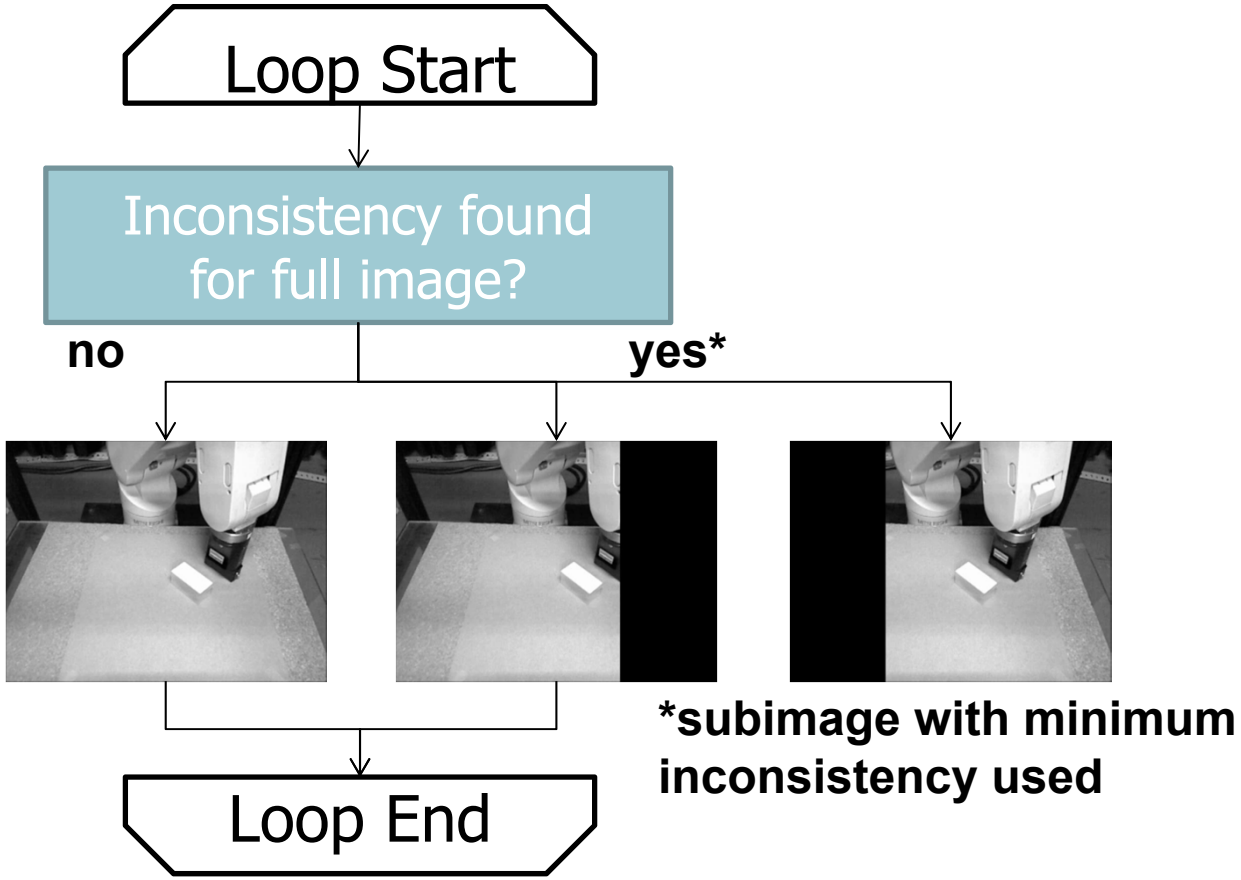
Successfully adapted to light on/off by switching neural networks

Making View-based Teaching/Playback Occlusion-robust

- Neural networks are trained for not only full images but also subimages
  - to overcome partial occlusions



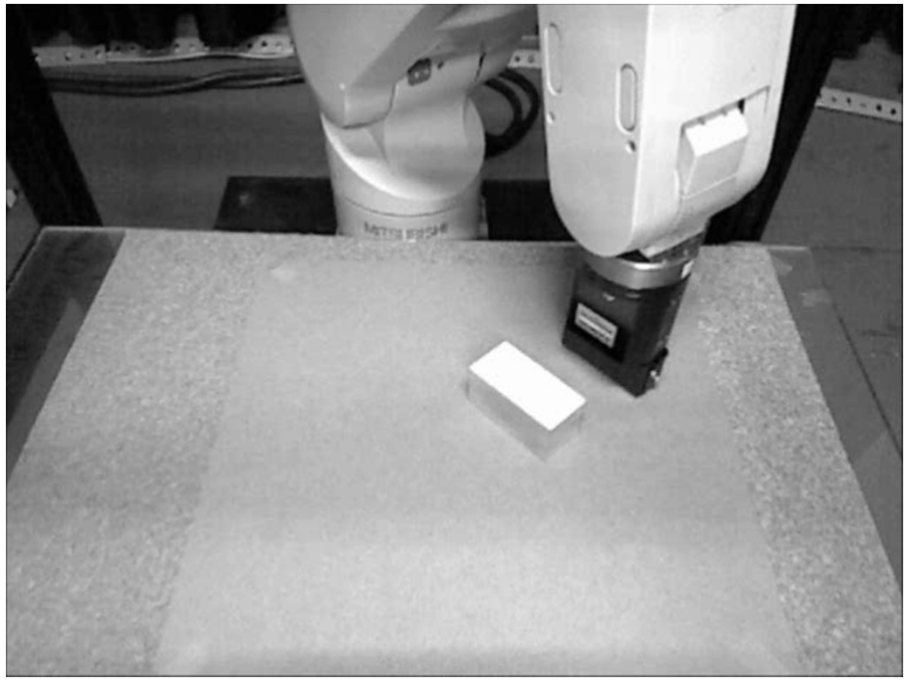
Switching Full/Sub Images



Occlusion-robust View-based Playback

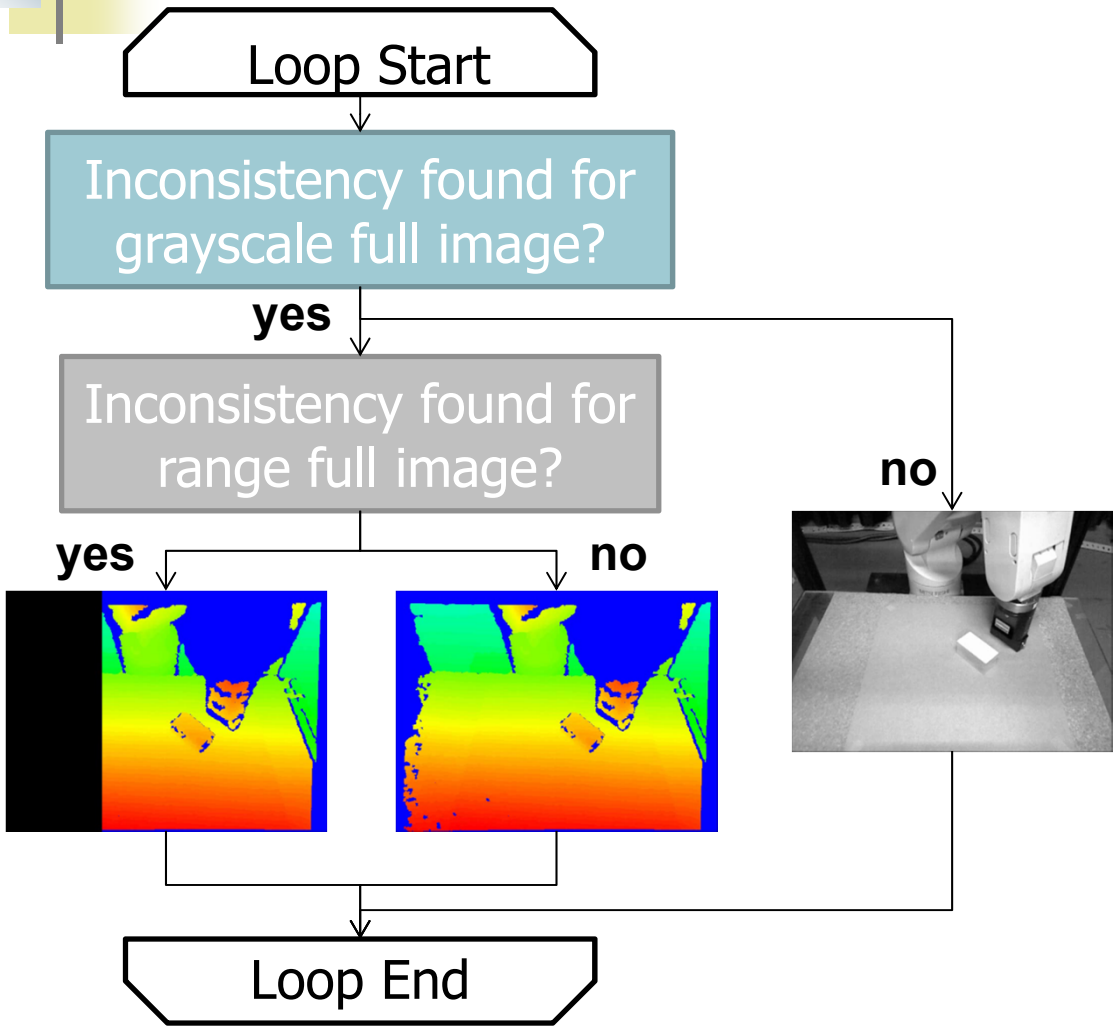
- Use of subimage-based NNs

Playback (x2)



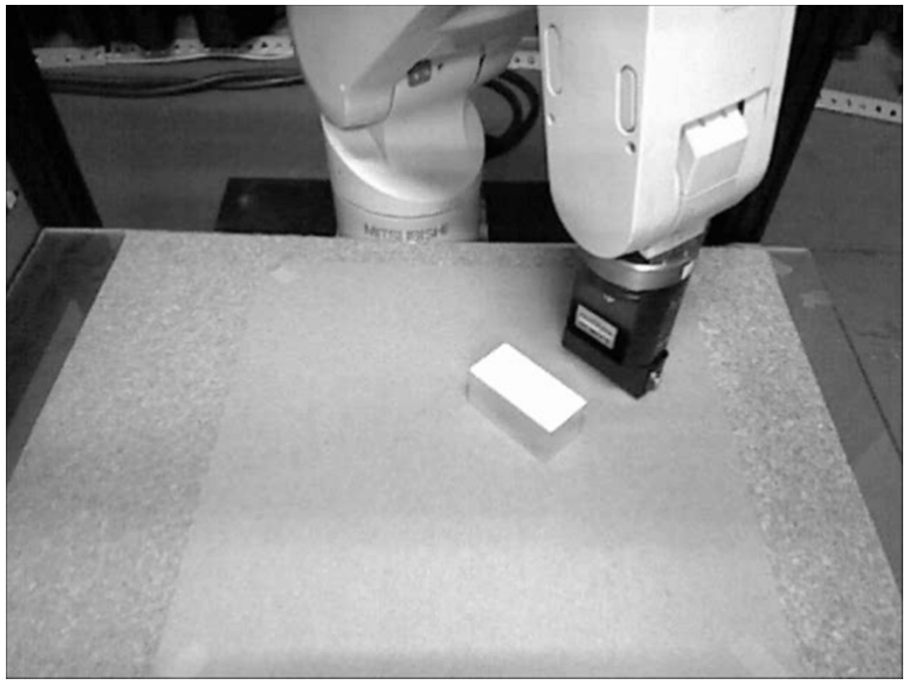
Used Images (x2)

Switching Grayscale Image, Range Full image and Range Subimages



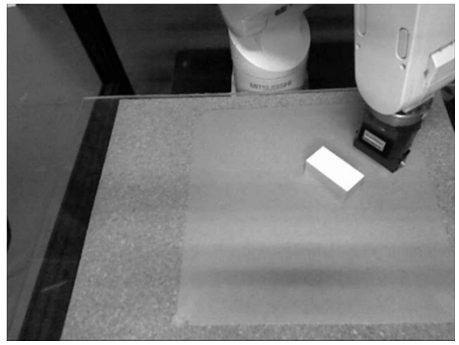
Lighting- and Occlusion-robust View-based Playback

Playback (x2)

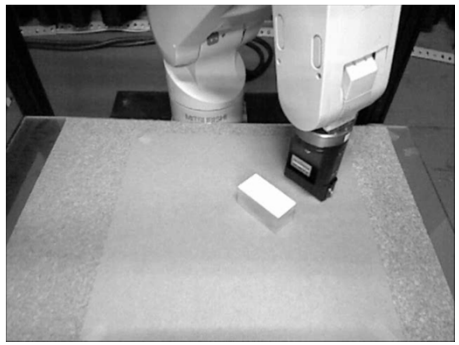


Used Images (x2)

Conclusion



- Lighting-robust view-based teaching/playback
  - Use of range images
  - Switching between range and grayscale images for better accuracy
- Occlusion-robust view-based teaching/playback
  - Use of subimages



Future Work

- Application to various robotic tasks that require higher DOF
- Incorporation of various sensor modalities
  - Force information [Nakagawa et al. 2016 IAS-14]