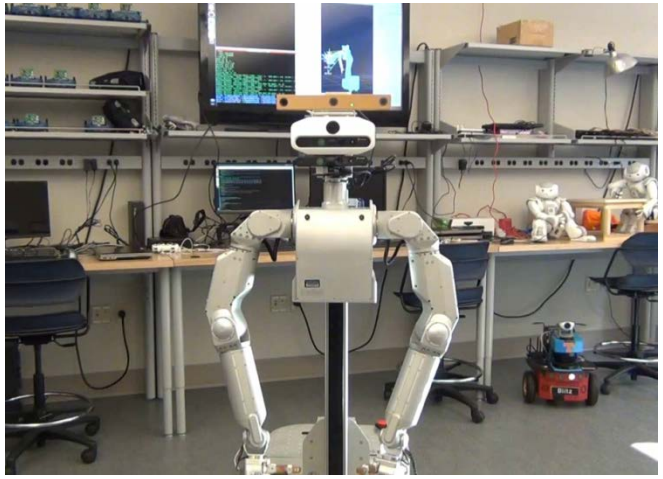
Overview of Kinect for Robotics

Nov. 4, 2014



(a) Lifting (b) Removing (c) Waving (d) Pushing (e) Walking (f) Signaling

Kinect – inexpensive media interface

- Made by Microsoft
- Available as product since late 2010
- Kinect is a part of Microsoft Xbox game platform but it can be bought separately
- Costs ~ $100

Kinect applications

Games and interactive playing (sports, dancing)

More applications: exercising, rehbilitation, child development

Control of devices by voice, gestures

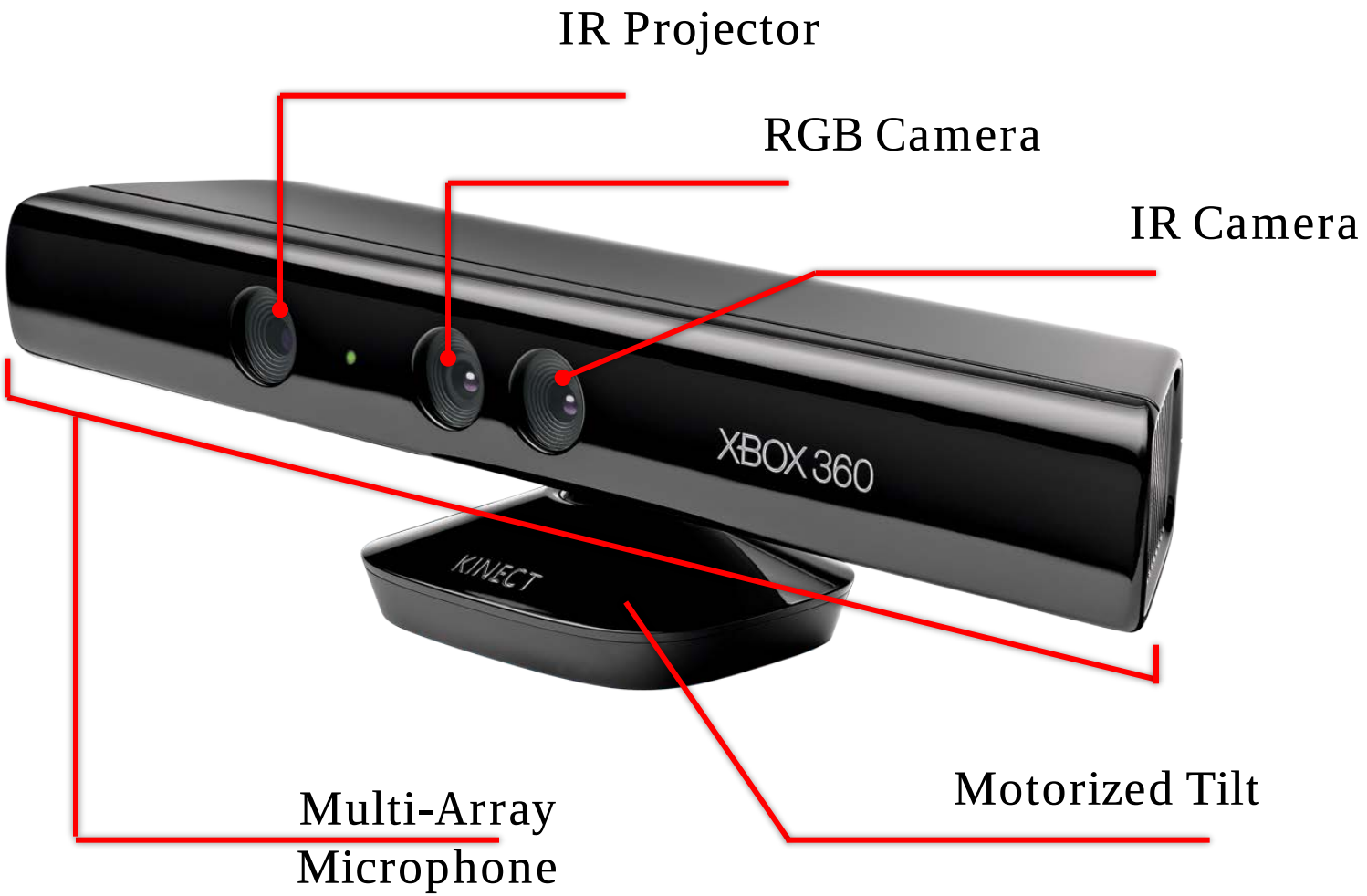
Automation, robotics

More….

Kinect – Why use it?

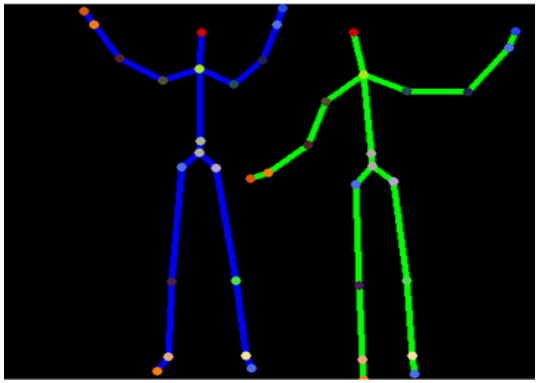
- Powerful
  - Capable of acquiring color, depth, and audio
- Inexpensive
- Accessible
  - Easily available at game stores, computer stores, on Amazon.com
- Easy to set up and use

Kinect Hardware



Kinect Sensor Data

RGB Camera



Skeleton Data

Depth

Demos

- Humanoid Robot Control and Interaction

  https://www.youtube.com/watch?v=GdepIXZTJsw

- Human Tracking and Following

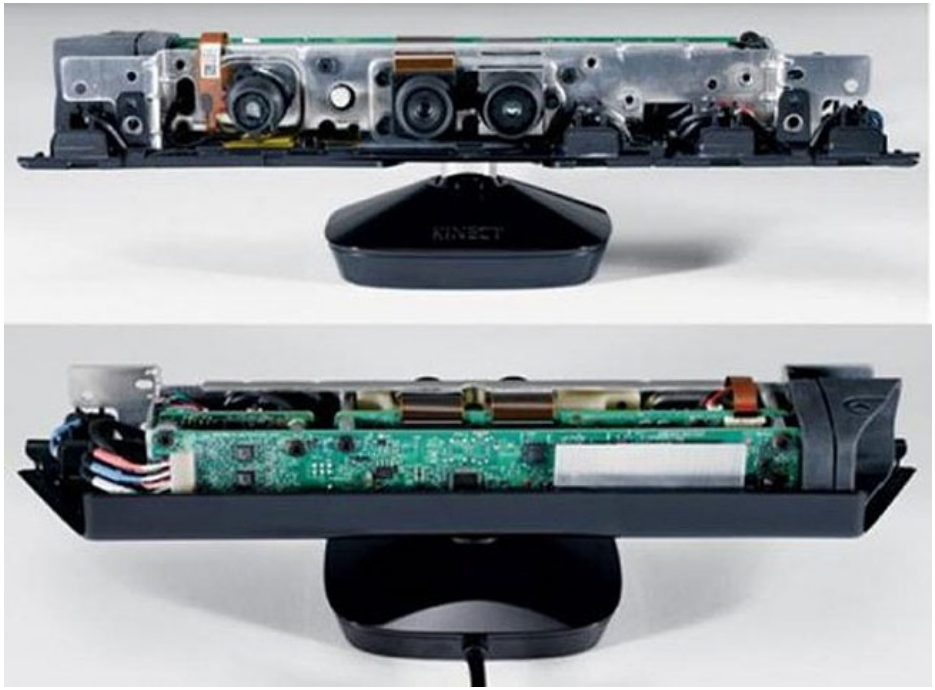
  https://www.youtube.com/watch?v=3Z56JV9g6y4

- Simultaneous Localization and Mapping

  https://www.youtube.com/watch?v=XejNctt2Fcs

How does Kinect sense depth?

- The IR emitter projects an irregular pattern of IR dots of varying intensities
- The Depth Camera reconstructs a depth image by recognizing the distortion in this pattern.



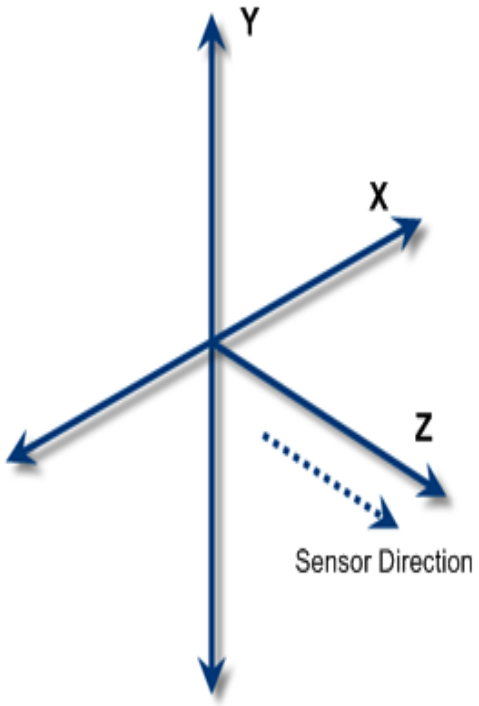
What is the accuracy of a Kinect sensor?

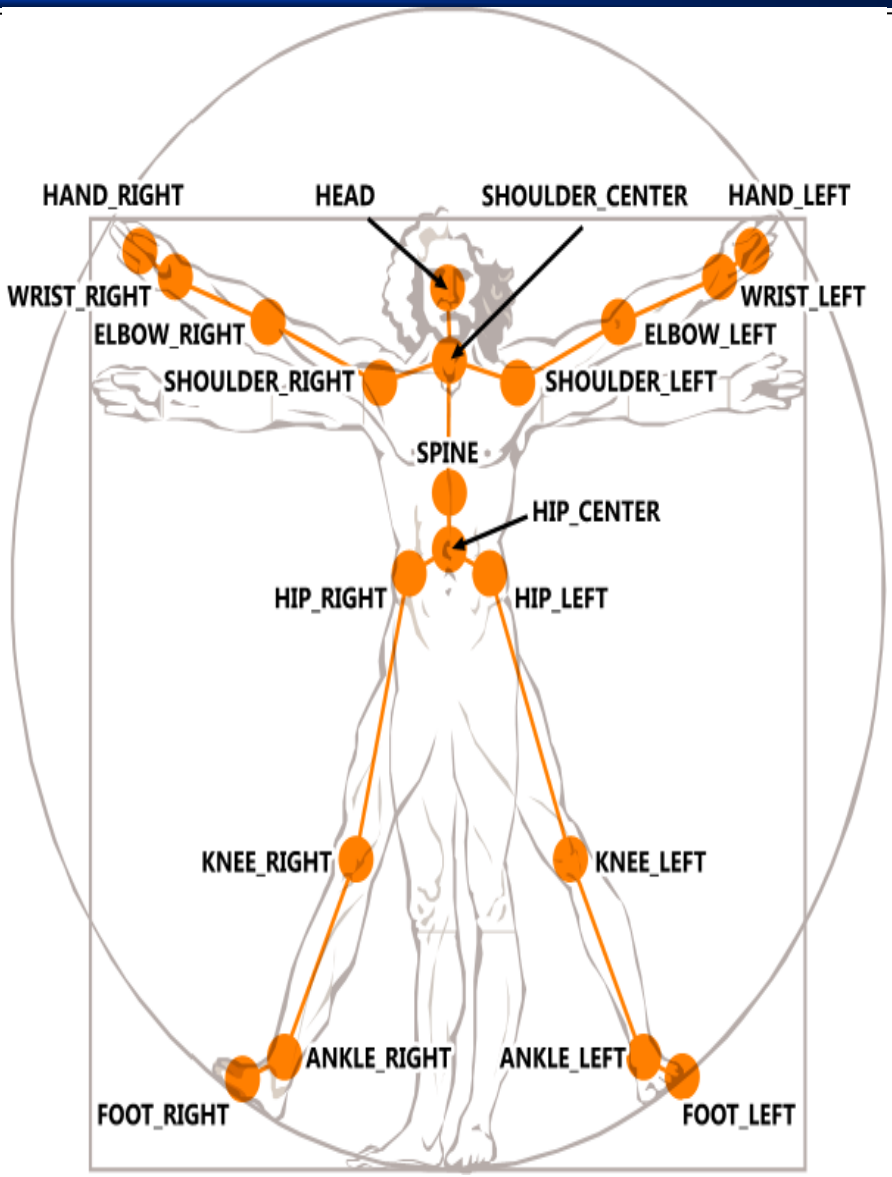
- Data Stream:
  - 640x480, 320x320 in Linux and Mac
  - 1024 x 768, 640x480, 320x240 in Windows 7
  - 30 frames/second

- Depth Camera:
  - Field of View:
    - Horizontal: 58^{o}, Vertical: 45^{o}, Diagonal: 70^{o}
    - Spatial x/y resolution: 3mm
    - Depth z resolution: 1 cm
    - Operational range: 0.8m – 3.5 m

- Physical Tilt Range: $\pm$ 27 degrees

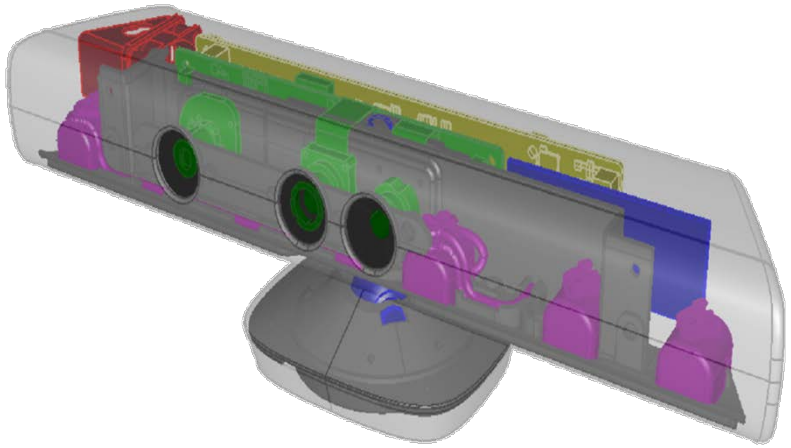
Skeleton Data





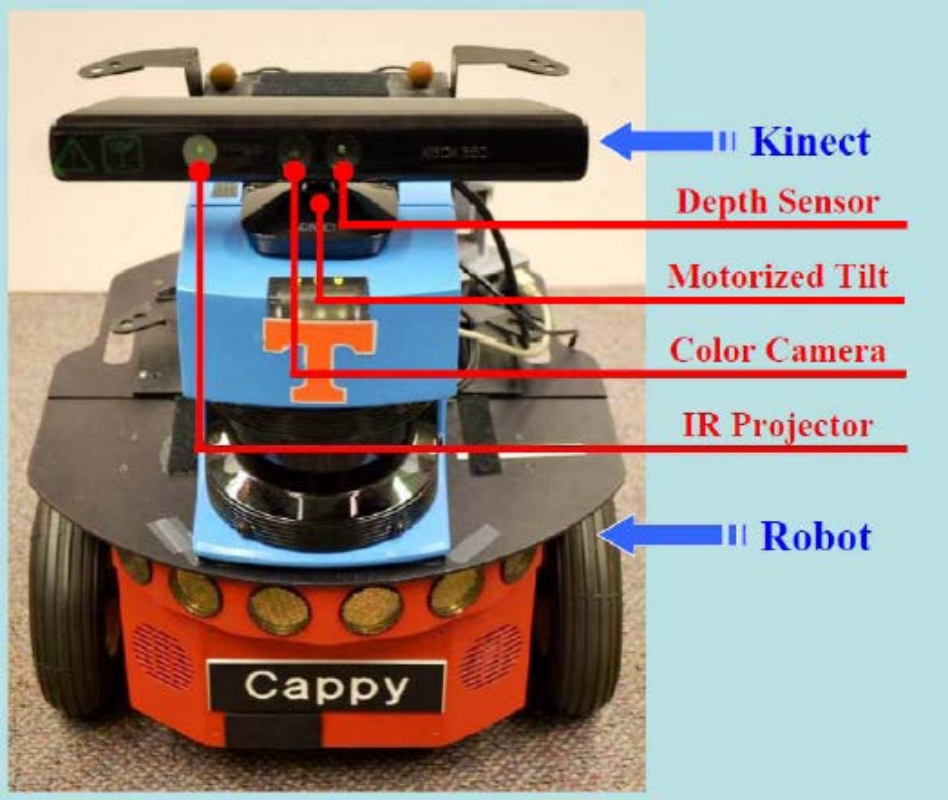
Kinect Audio

- Four-microphone array with hardware-based audio processing
  - Multichannel echo cancellation (MEC)
  - Sound position tracking
  - Other digital signal processing (noise suppression and reduction)

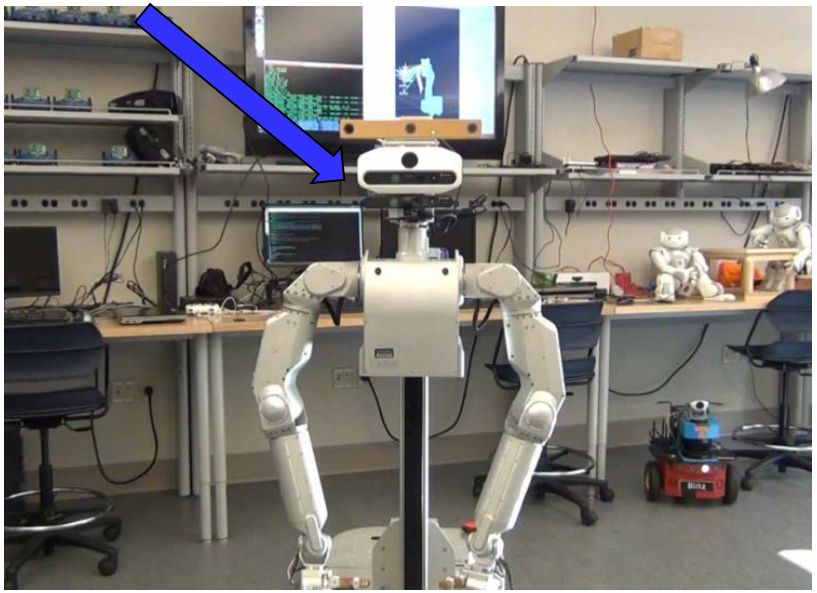


Our UTK (DILab) Research: Human Activity Recognition

- First: Install on robots:

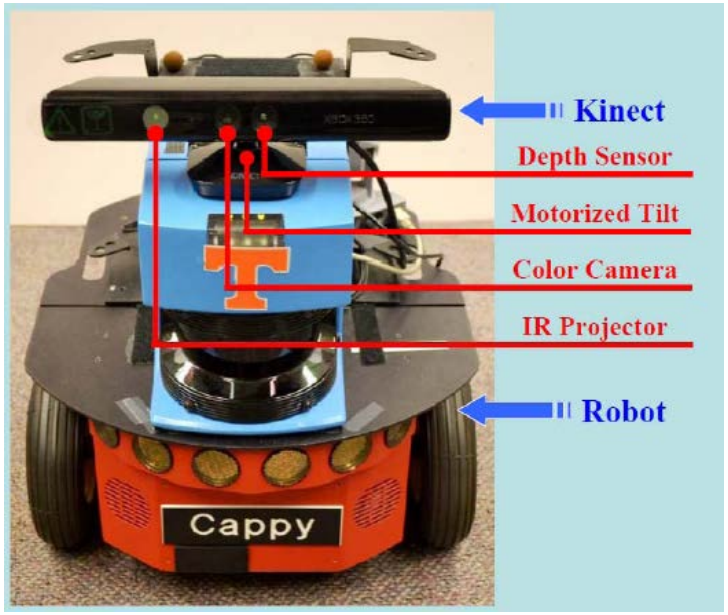


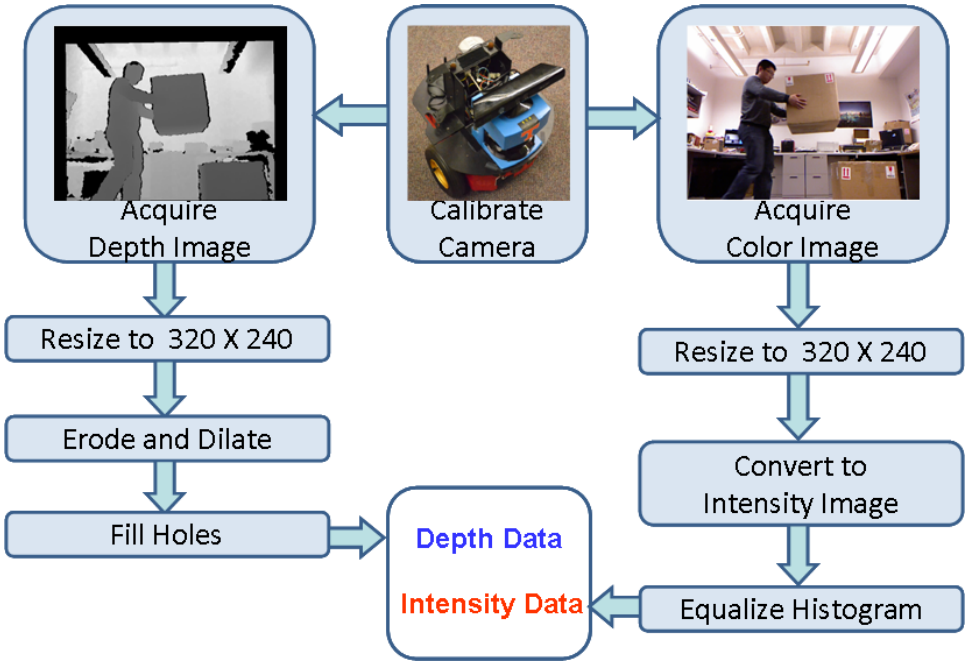
Kinect



Feature Extraction: Data Acquisition and Processing

- Data Acquisition:
  - Use Microsoft Kinect installed on a Pioneer robot
  - Collect both color and depth information
- Image processing:
  - Compute intensity image from raw color image
  - Make the data cleaner and ready for feature extraction





Feature Extraction: Feature Detection

1. Apply Gaussian filter along spatial dimension

$$D_s(\boldsymbol{x}_o, t) = \big(D(\boldsymbol{x}, t) \circ f(\boldsymbol{x}, t|\delta)\big) * p(\boldsymbol{x}|\sigma)\big|_{\boldsymbol{x}=\boldsymbol{x}_o}$$

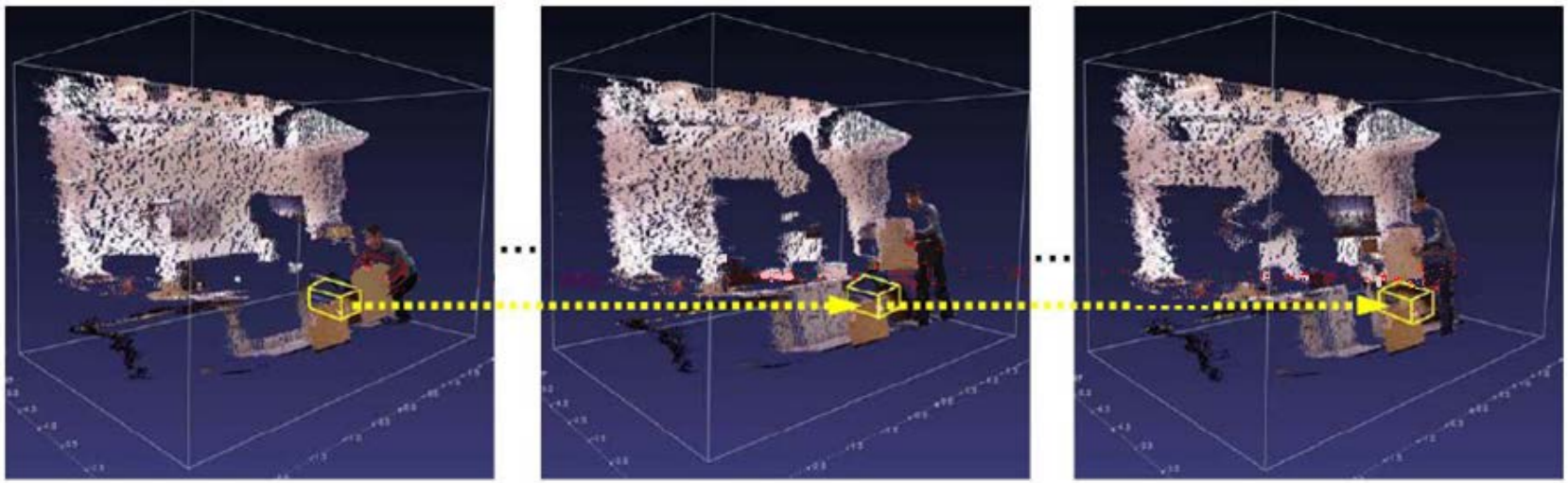
$$\text{where} \quad f(\boldsymbol{x}, t) = \mathbf{1}(\big|D(\boldsymbol{x}, t) - D(\boldsymbol{x}_o, t)\big| \leqslant \delta)$$

$$\text{and} \quad p(\boldsymbol{x}|\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{\|\boldsymbol{x}\|^2}{2\sigma^2}}$$

2. Apply Gabor filter along temporal dimension

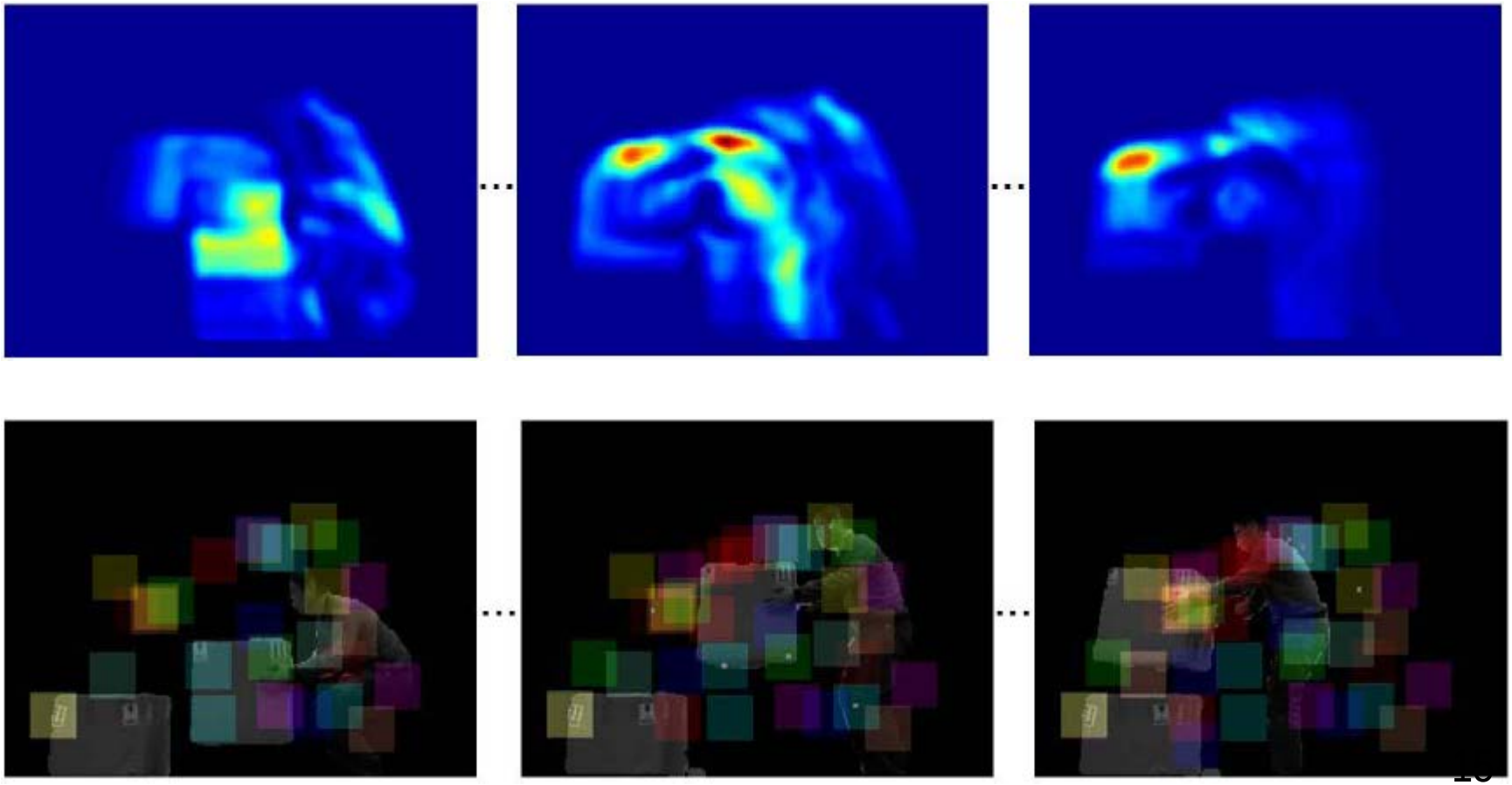
$$D_{st}(\boldsymbol{x}_o, t) = D_s(\boldsymbol{x}_o, t) * g(t|\tau, \omega)\big|_{t=t_o}$$

$$\text{where} \quad g(t|\tau, \omega) = \frac{1}{\sqrt{2\pi}\tau} \cdot e^{-\frac{t^2}{2\tau^2}} \cdot e^{i(2\pi\omega t)}$$



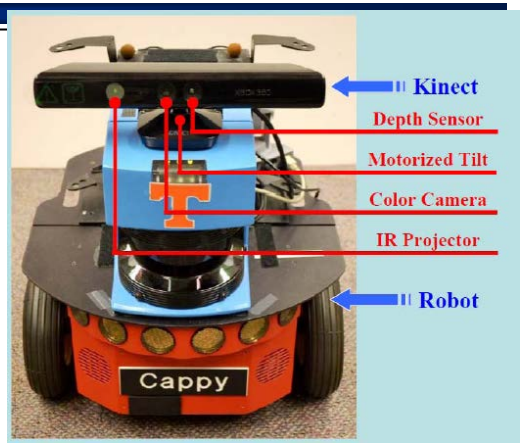
Feature Detection

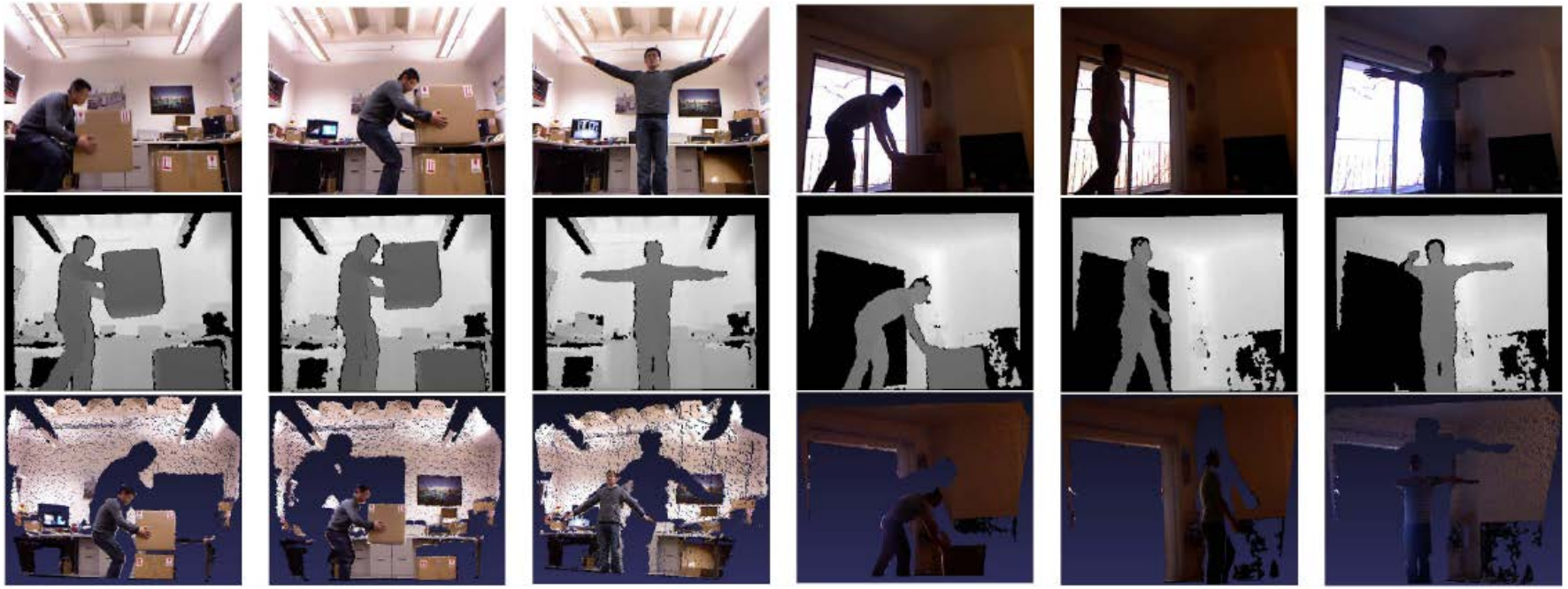
$$R(\boldsymbol{x}_o) = \alpha \cdot \|I_{st}(\boldsymbol{x}_o)\|^2 + (1-\alpha) \cdot \|D_{st}(\boldsymbol{x}_o)\|^2$$



Evaluation: Activity Dataset

- Activity Dataset
  - 6 types of human activities
  - 33 samples for each activity
  - 2 ~ 4 seconds of each sample
  - Office and home environments





(a) Lifting (b) Removing (c) Waving (d) Pushing (e) Walking (f) Signaling

Video Demonstrating Results

- (video)

For more information on our Kinect-based research:

- H. Zhang, W. Zhou, and L. E. Parker, “Fuzzy segmentation and recognition of continuous human activities,” in *ICRA*, 2014.
- H. Zhang, C. Reardon, and L. E. Parker, “Real-Time Multiple Human Perception with Color-Depth Cameras on a Mobile Robot,” *IEEE Trans. Cybernetics*, vol. 43, no. 5, pp. 1429–1441, Oct. 2013.
- H. Zhang and L. E. Parker, “4-dimensional local spatio-temporal features for human activity recognition,” in *IROS*, 2011.

How to Get Started Using Kinect?

- Install a driver for Kinect sensor and related dependencies
  - Kinect for Windows SDK

  - OpenKinect
    - OpenNI Kinect
    - Libfreenect
      - Supports Windows, Mac, and Linux
      - Also combined with ROS

OPENKINECT