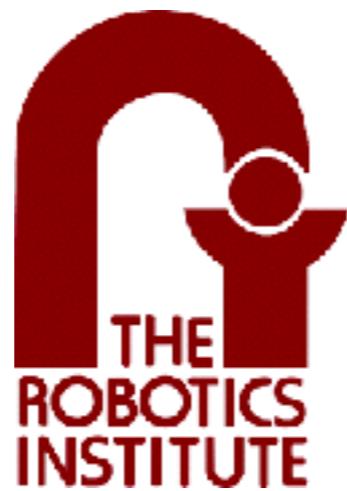


Wearable Sensing for Understanding, Forecasting and Assisting Human Activity

Kris Kitani

Assistant Research Professor
Carnegie Mellon University



Wearable sensing

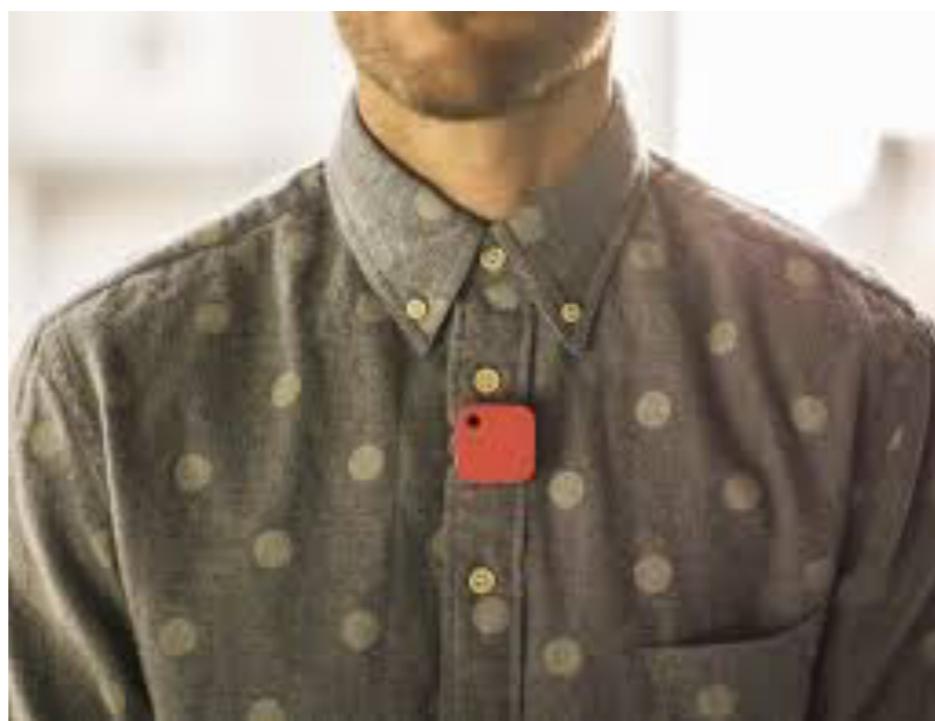


camera



smartphone

Wearable sensing



camera



smartphone

What's so unique about wearable cameras?

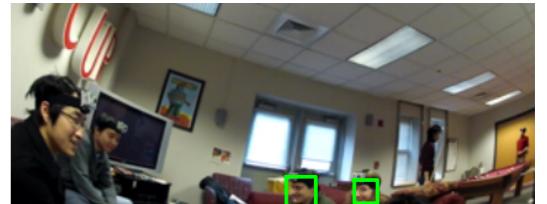
Features of FPV

Near-View Environment



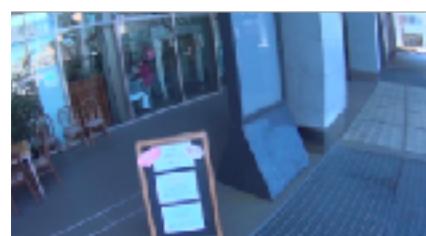
Closer view of the environment
(especially hands-object interactions)

Intimate Social Sensing



Up close observations of interactions
(faces, gestures, eye contact, touch)

True Visual Experience



Direct access to a person's visual experience
(saliency, visual motifs)

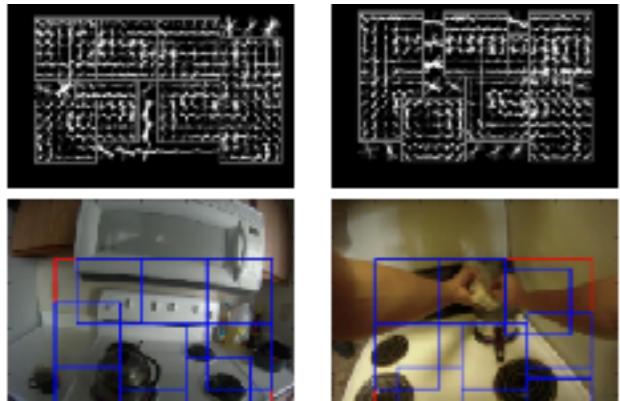
On-the-Go



Platform is mobile
(large scale visual analysis)

Research Directions

Object Recognition



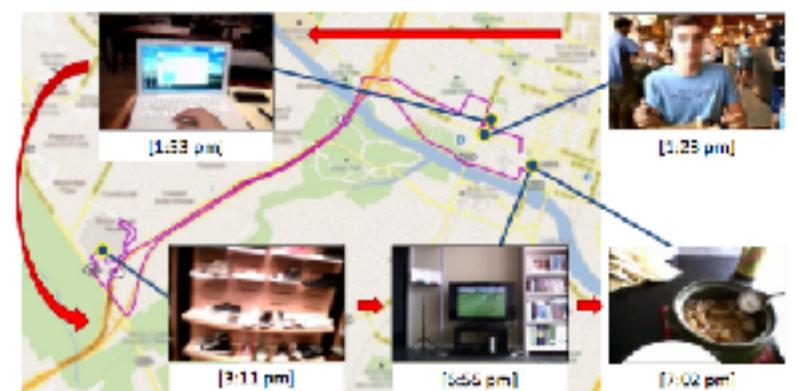
[Piriavash, Ramanan CVPR 2012]

Action Recognition



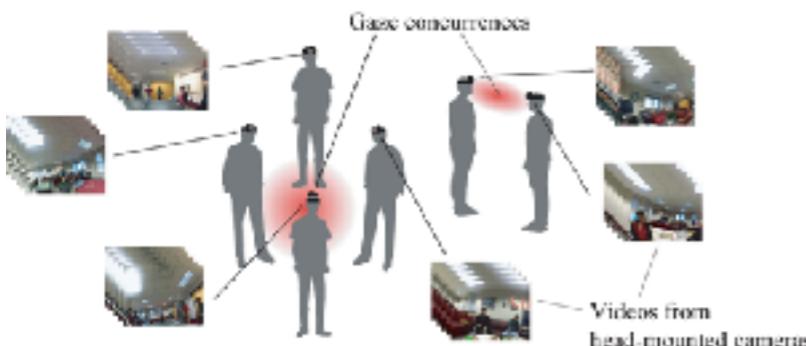
[Fathi, Faharadi, Rehg, ICCV 2011]

Video Summarization



[Lee, Gnosh, Grauman. CVPR 2012]

Social Saliency



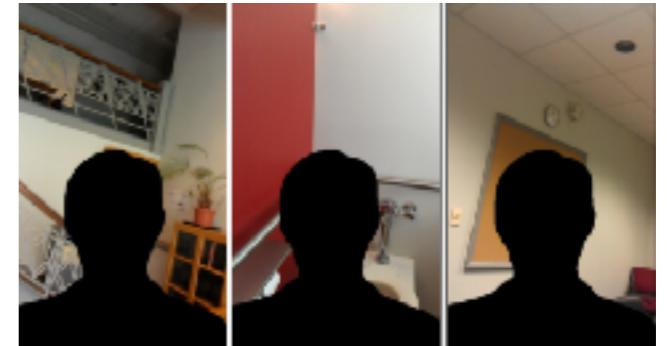
[Park, Jain, Sheikh. NIPS 2012]

Gaze Analysis



[Li, Fathi, Rehg. ICCV 2013]

Privacy Preservation



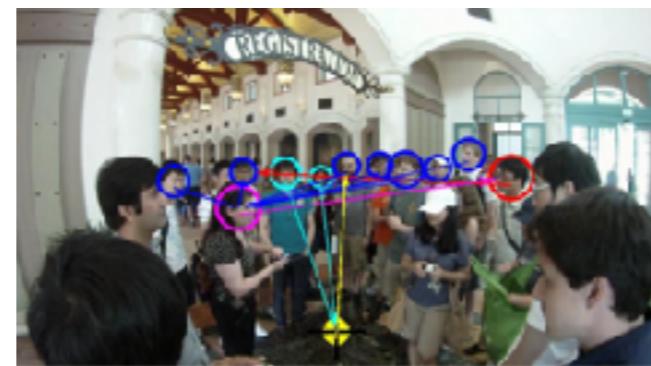
[Hoyle et al. UBICOMP 2014]

Activity Prediction



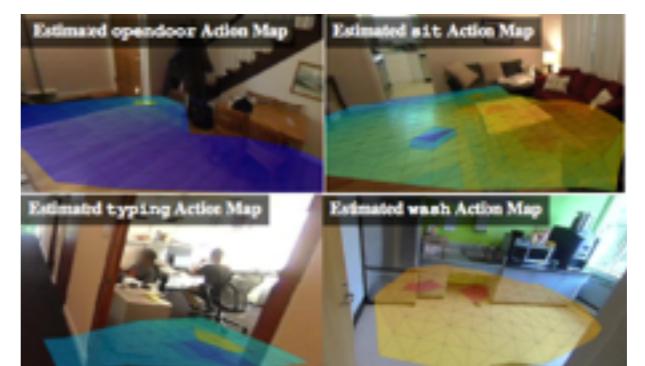
[Ryoo et al. HRI 2015]

Social Interactions



[Fathi, Hodgins, Rehg. CVPR 2012]

Affordance Estimation



[Rhinehart, Kitani. CVPR 2016]

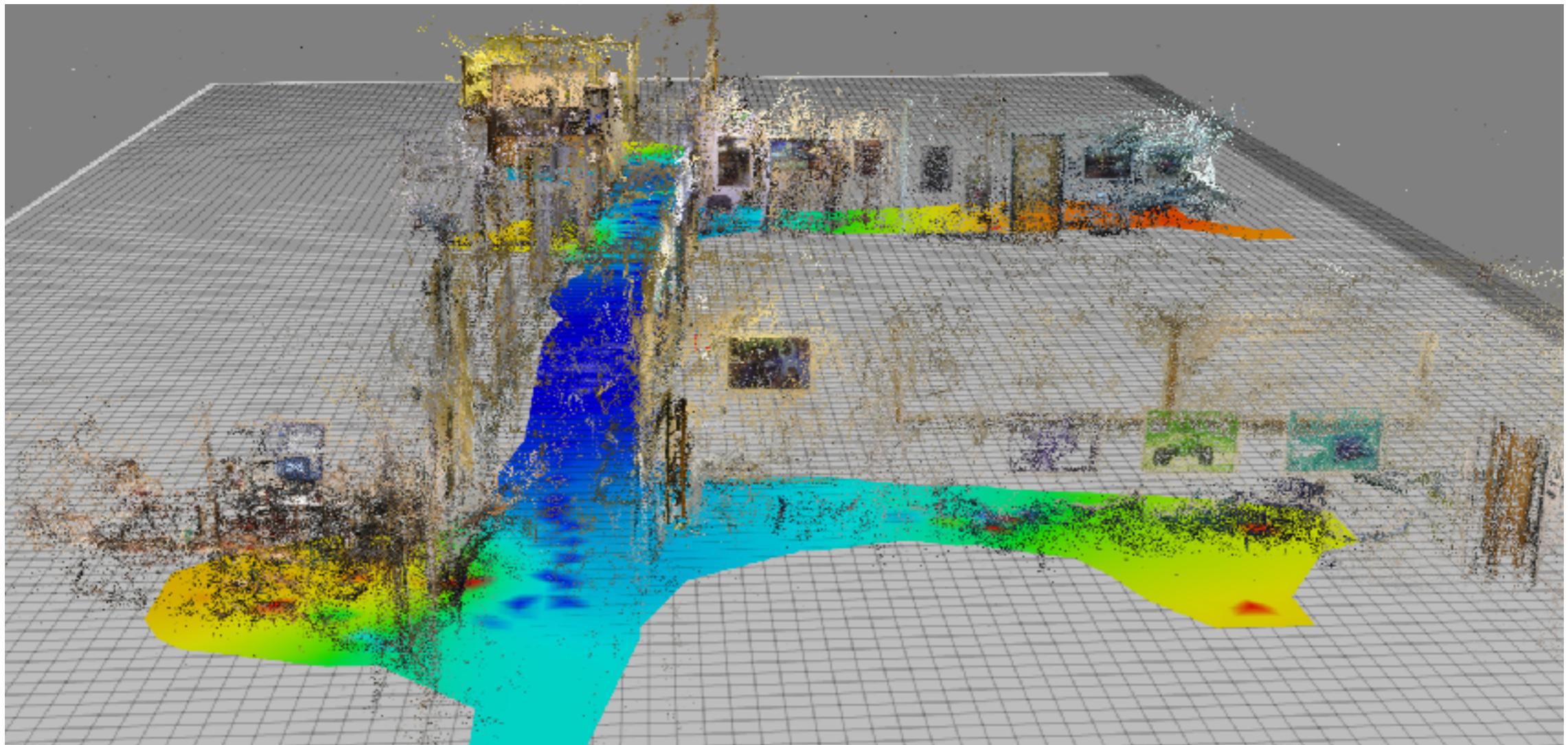
wearable sensing for
Human Activity

Understanding

Forecasting

Assisting

Learning Action Maps of Large Environments via First-Person Vision



Nicholas Rhinehart, Kris M. Kitani.
Learning Action Maps of Large Environments via First-Person Vision.
Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

When we **observe** a scene...



... we know **how we can act** in that environment

When we **observe** a scene...



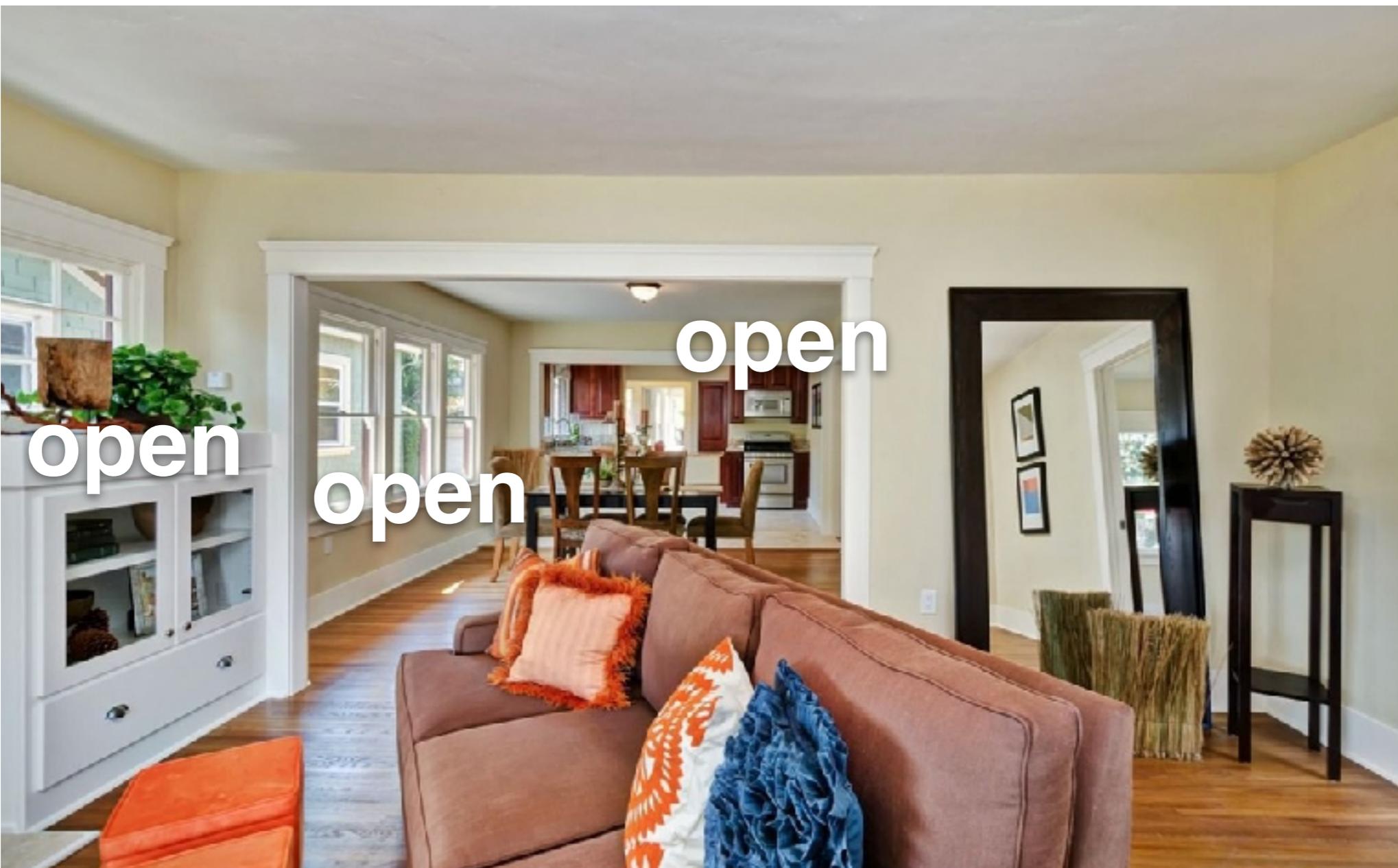
... we know **how we can act** in that environment

When we **observe** a scene...



... we know **how we can act** in that environment

When we **observe** a scene...

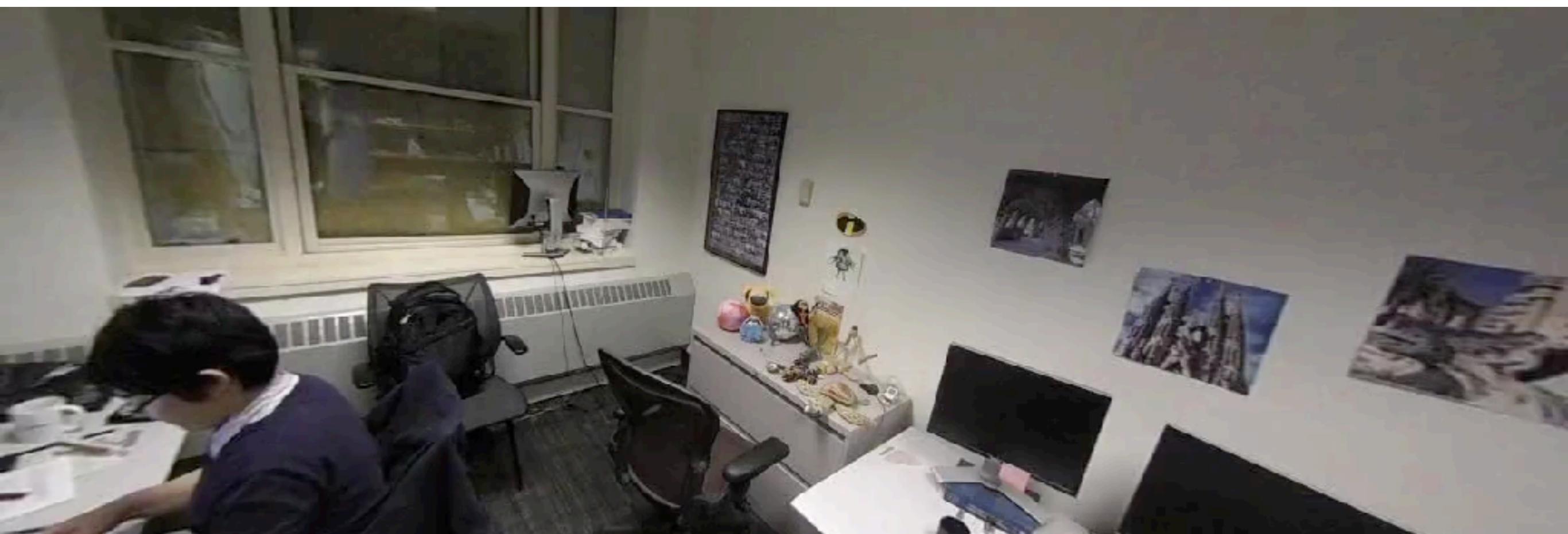


... we know **how we can act** in that environment

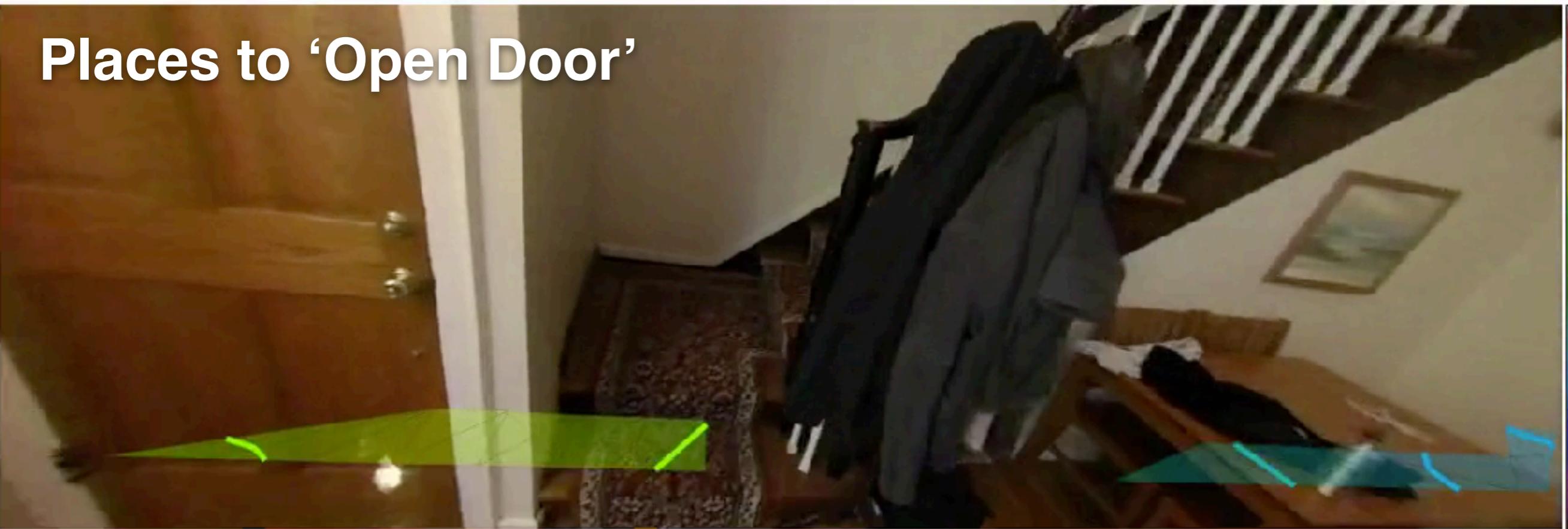
Can we teach a computer
to
understand scene functionality?



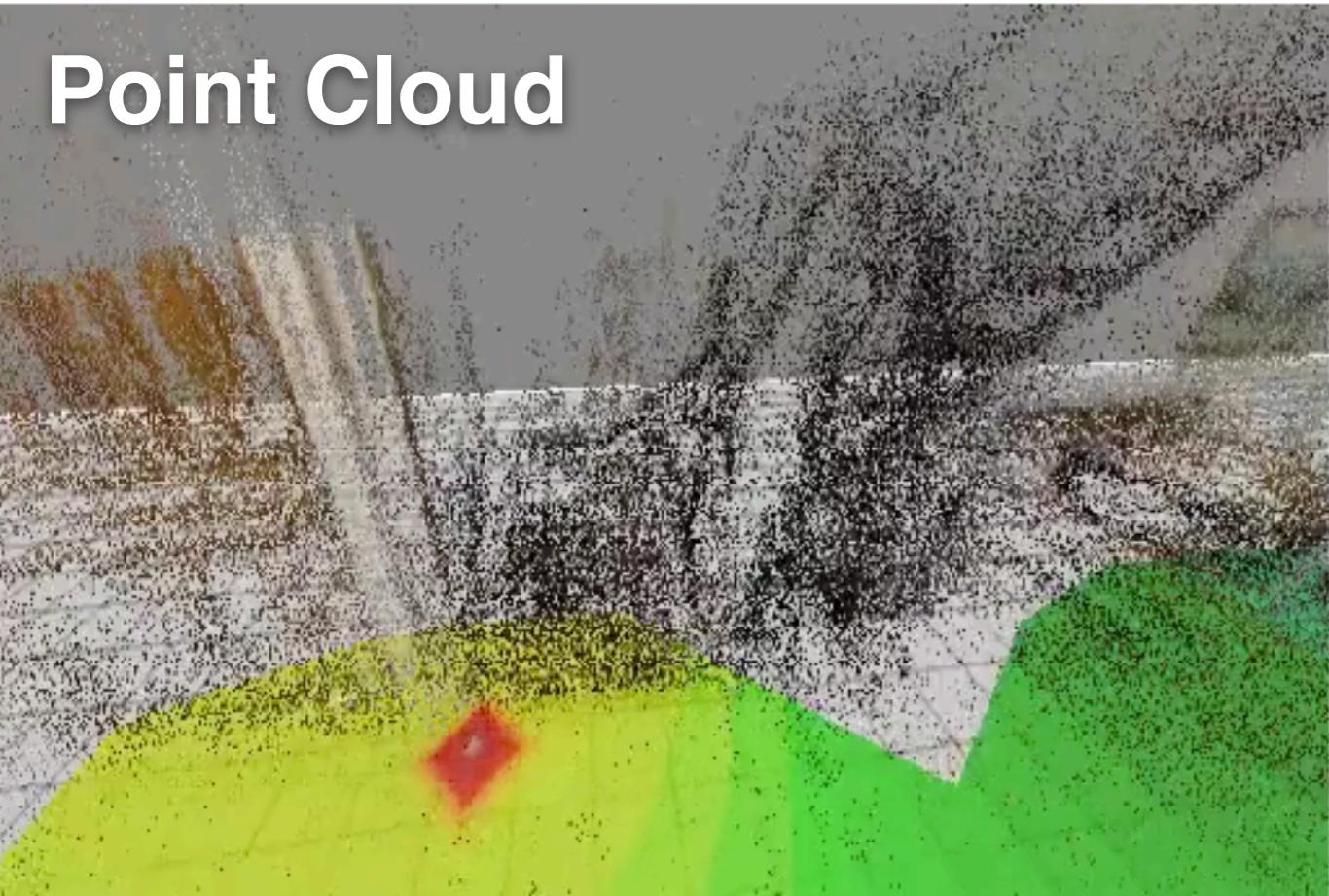
Input: Captured visual experience



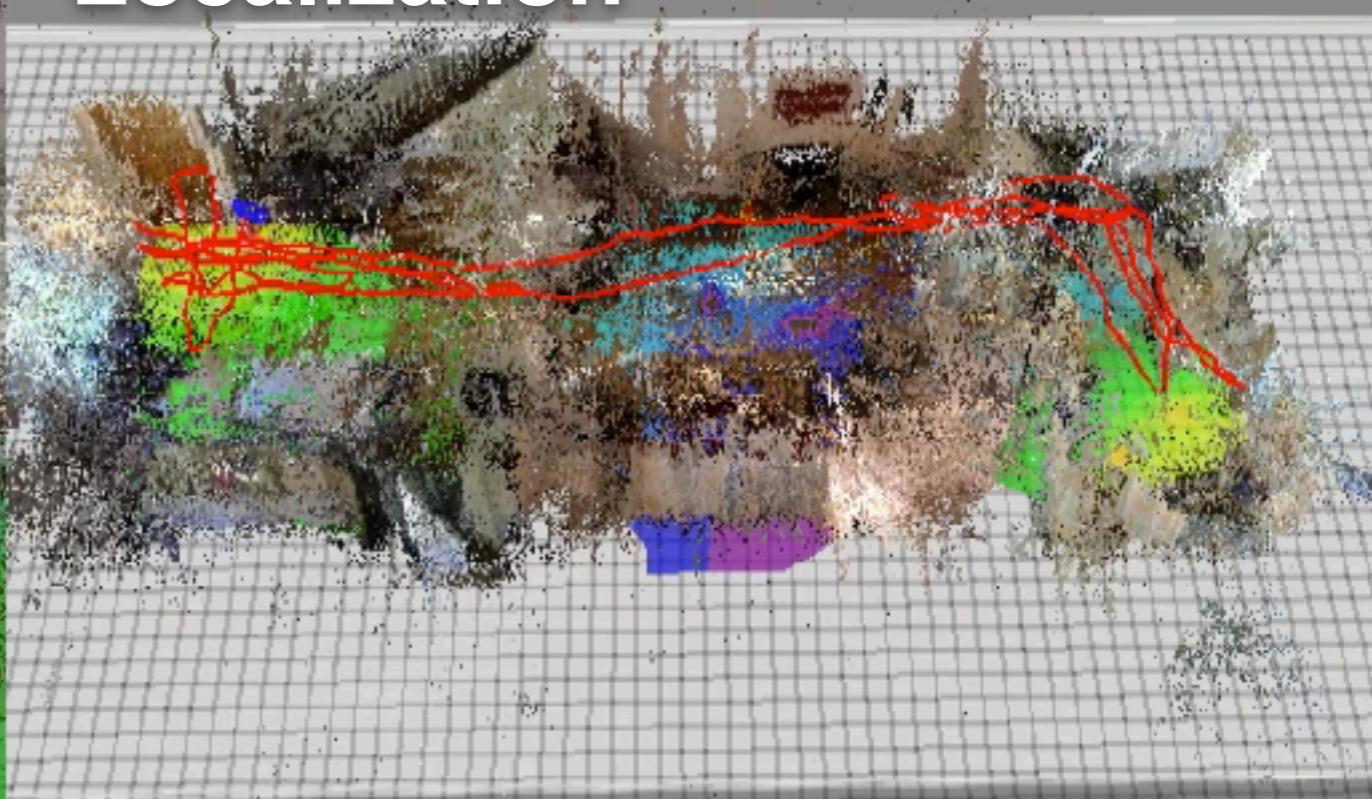
Places to ‘Open Door’



Point Cloud

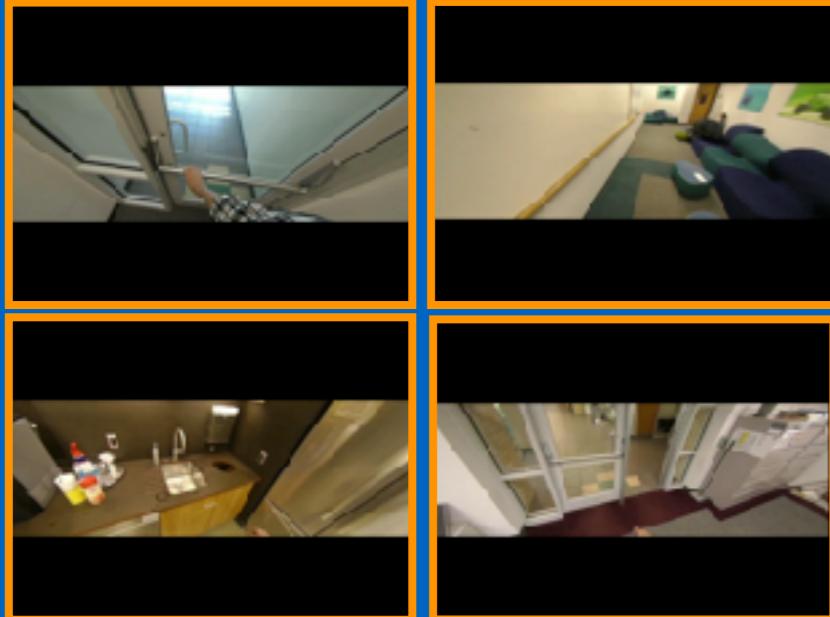


Localization

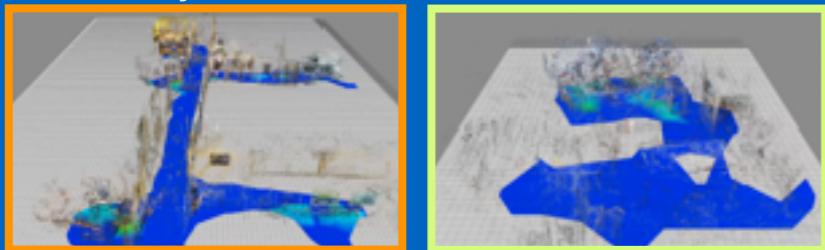


Activity Detections
In **one or more** scenes

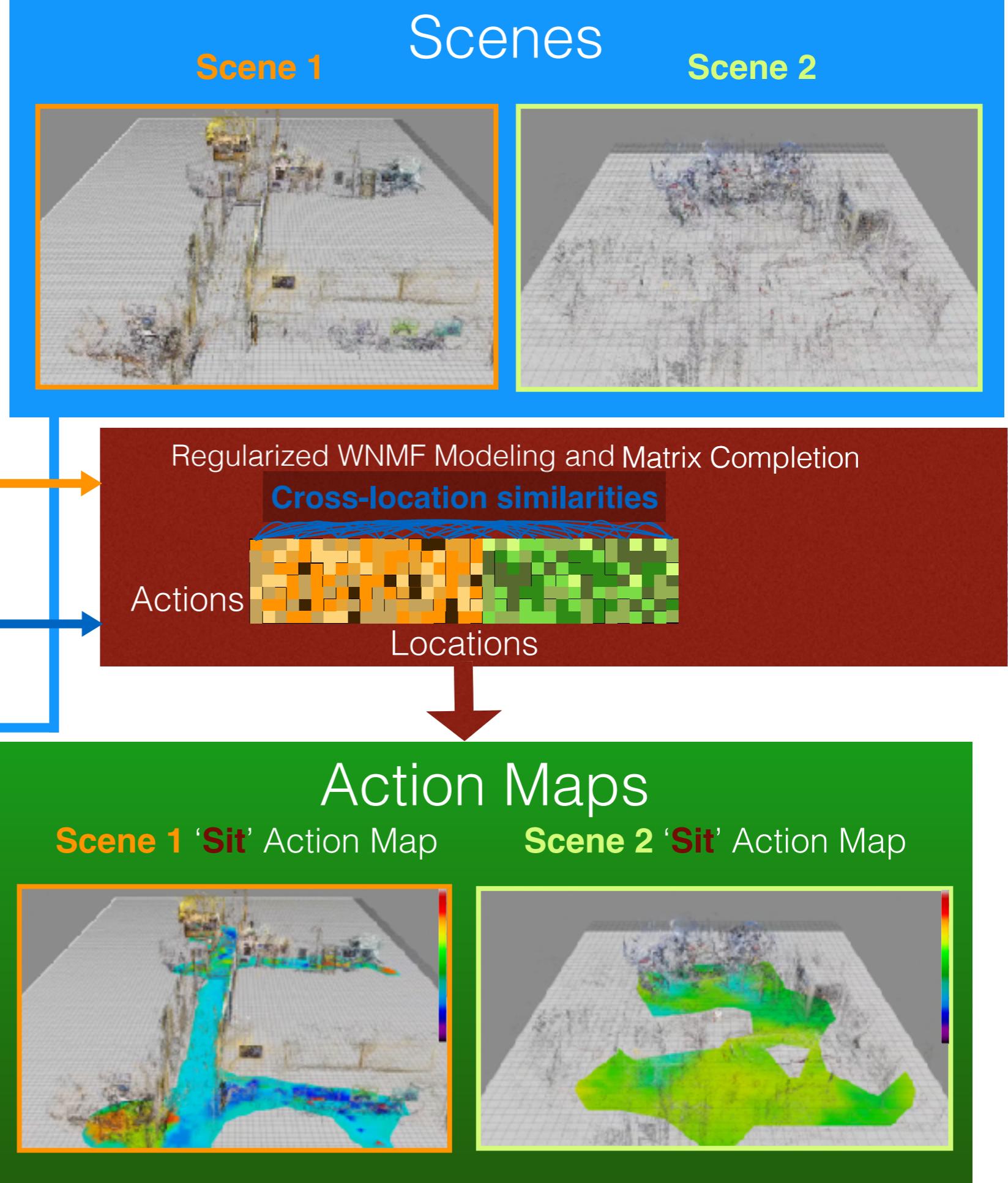
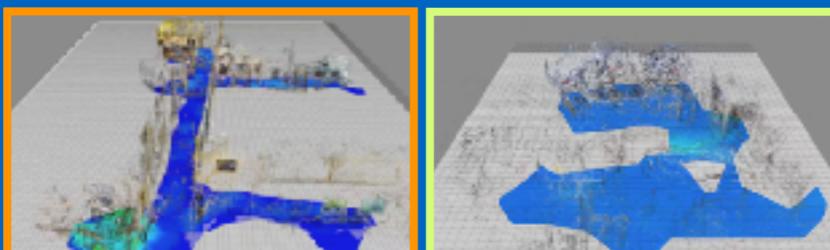
Scene 1 Detections



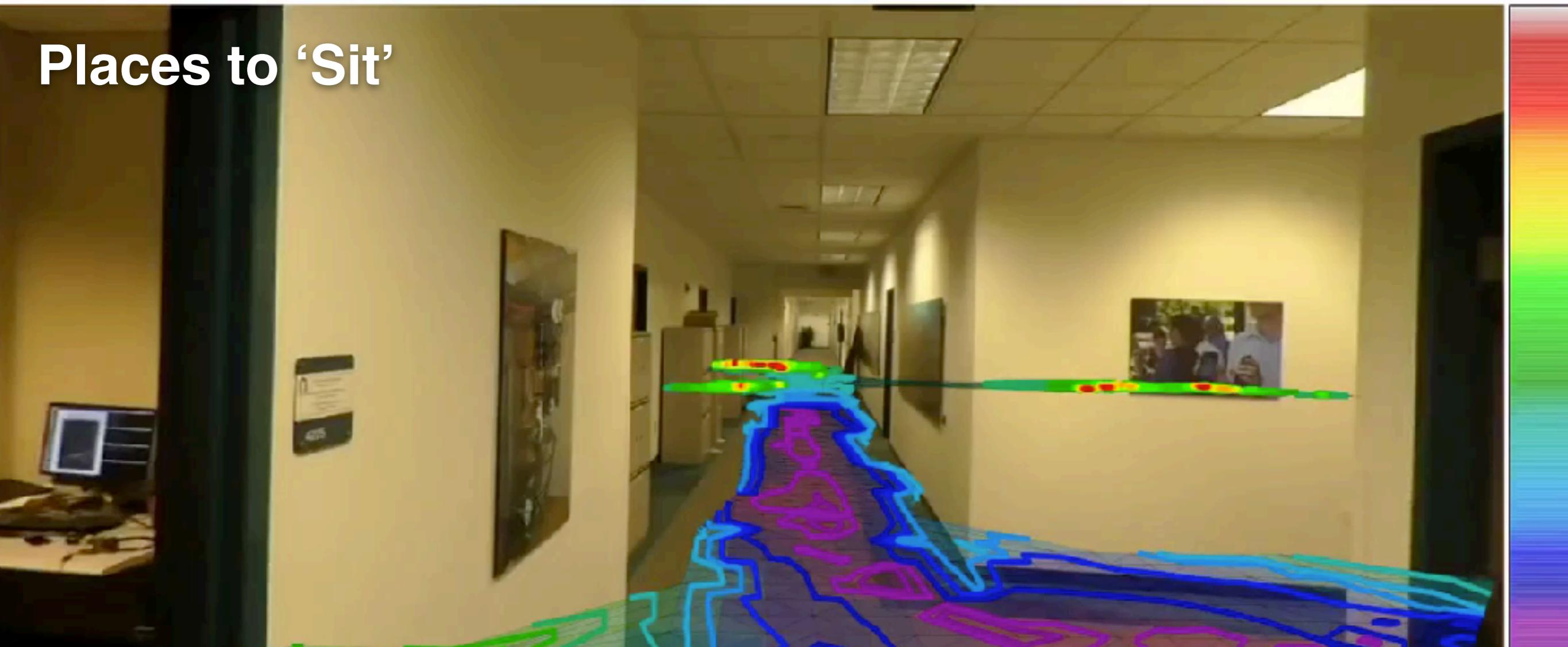
Appearance Info
Object Detection Features



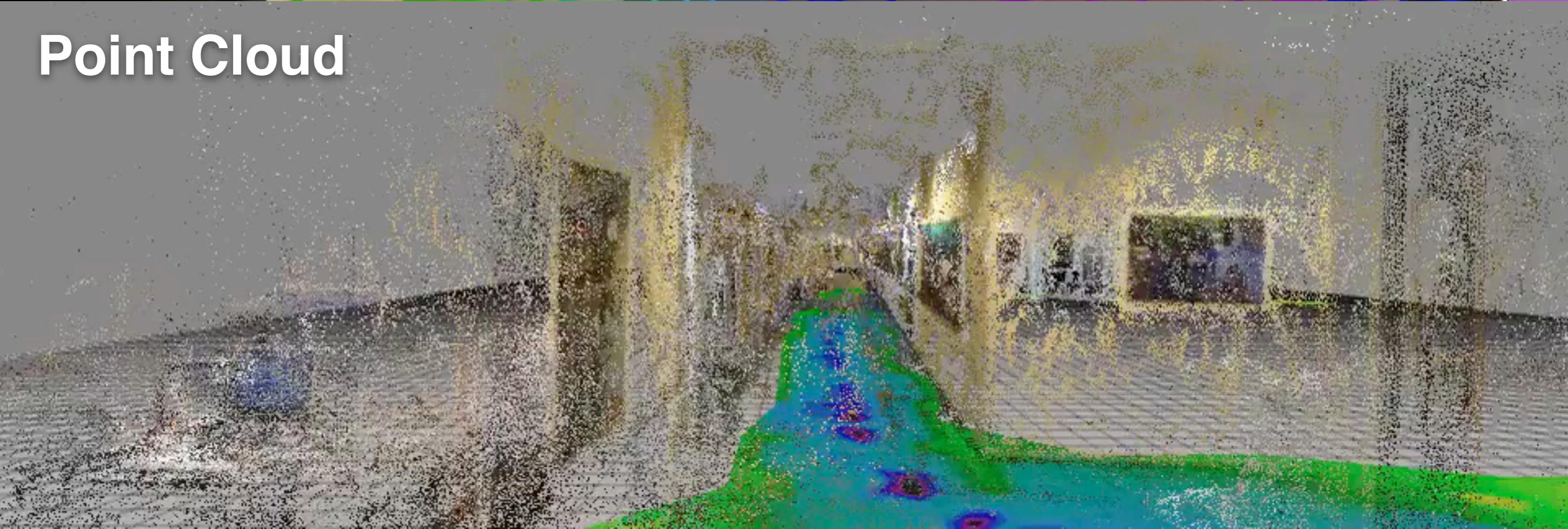
Scene Classification Features



Places to ‘Sit’



Point Cloud



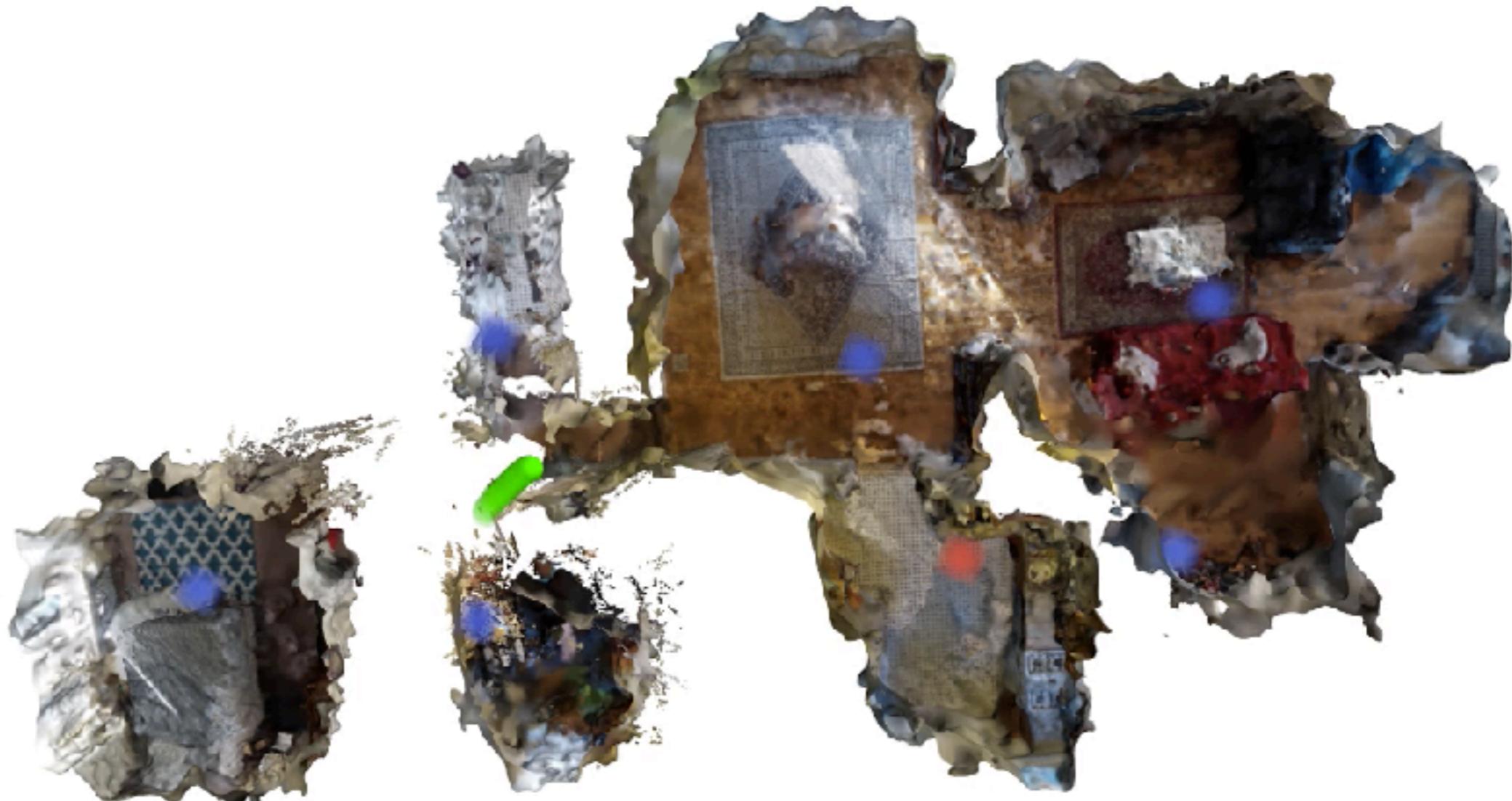
wearable sensing for
Human Activity

Understanding

Forecasting

Assisting

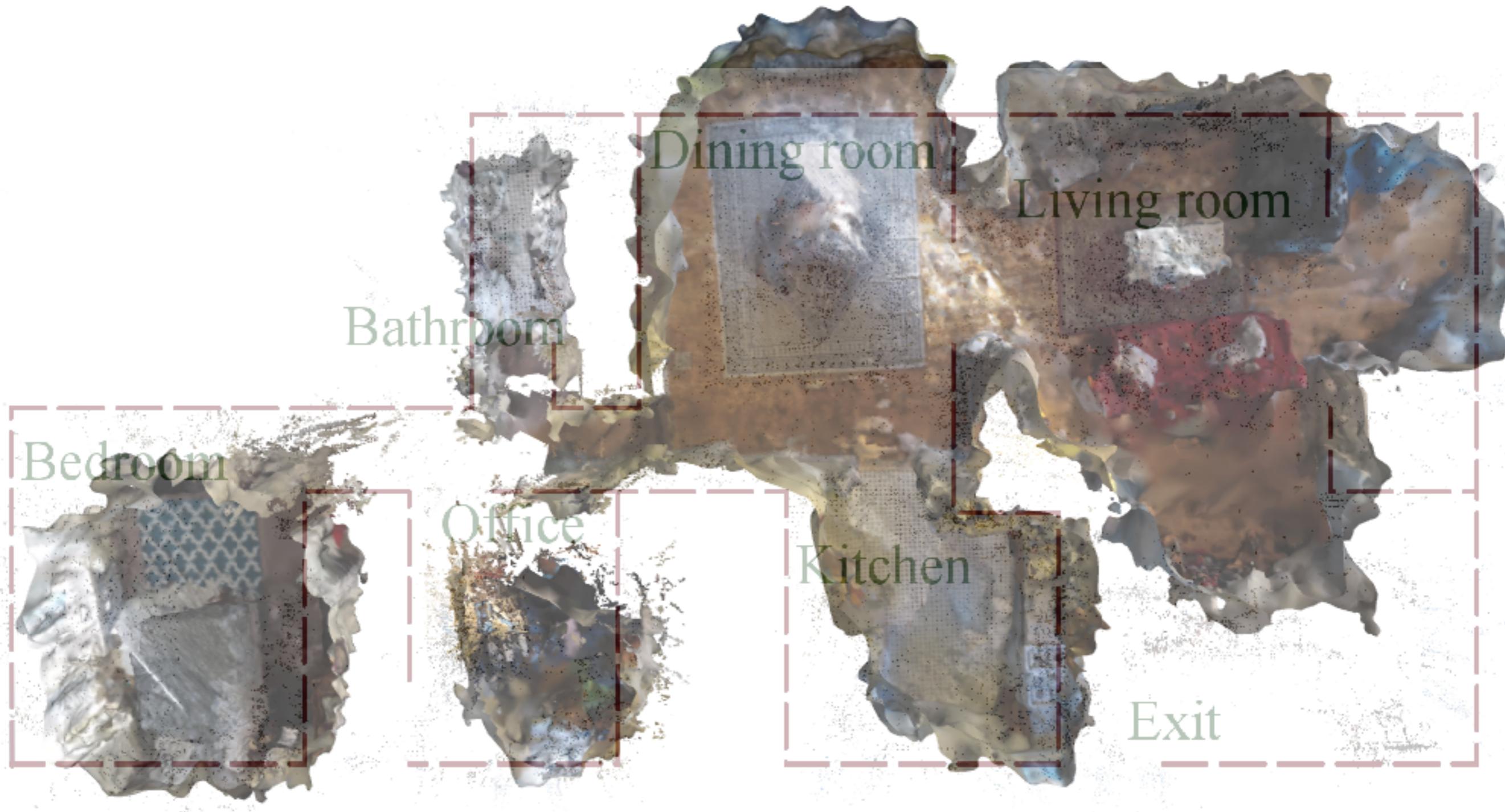
First-Person Activity Forecasting with Online Inverse Reinforcement Learning



Nicholas Rhinehart, Kris M. Kitani
The Robotics Institute, Carnegie-Mellon University
Best Paper Honorable Mention

Can we teach a computer
to
predict the future?
(using a wearable camera)

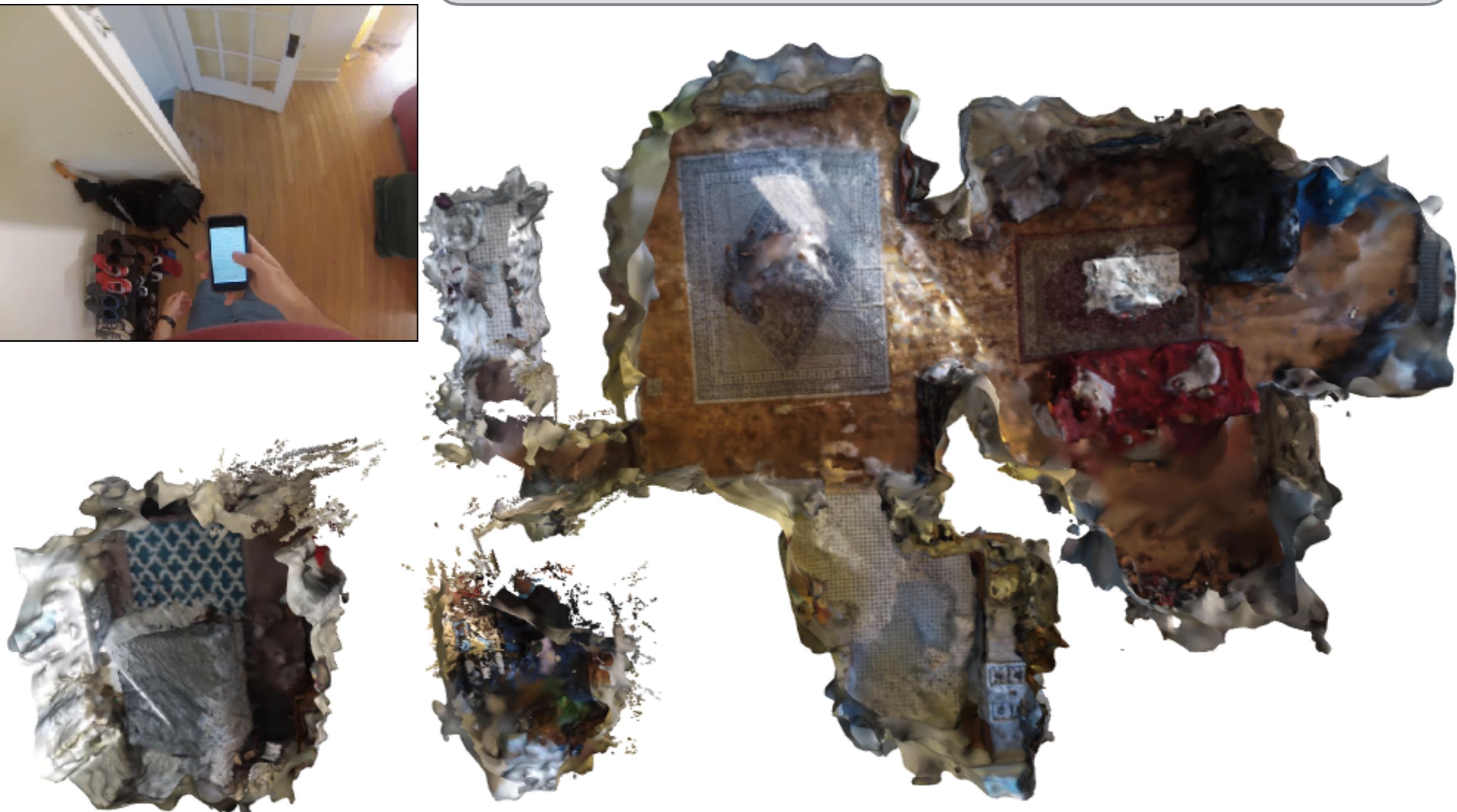






Continuously predict future goals using first-person video

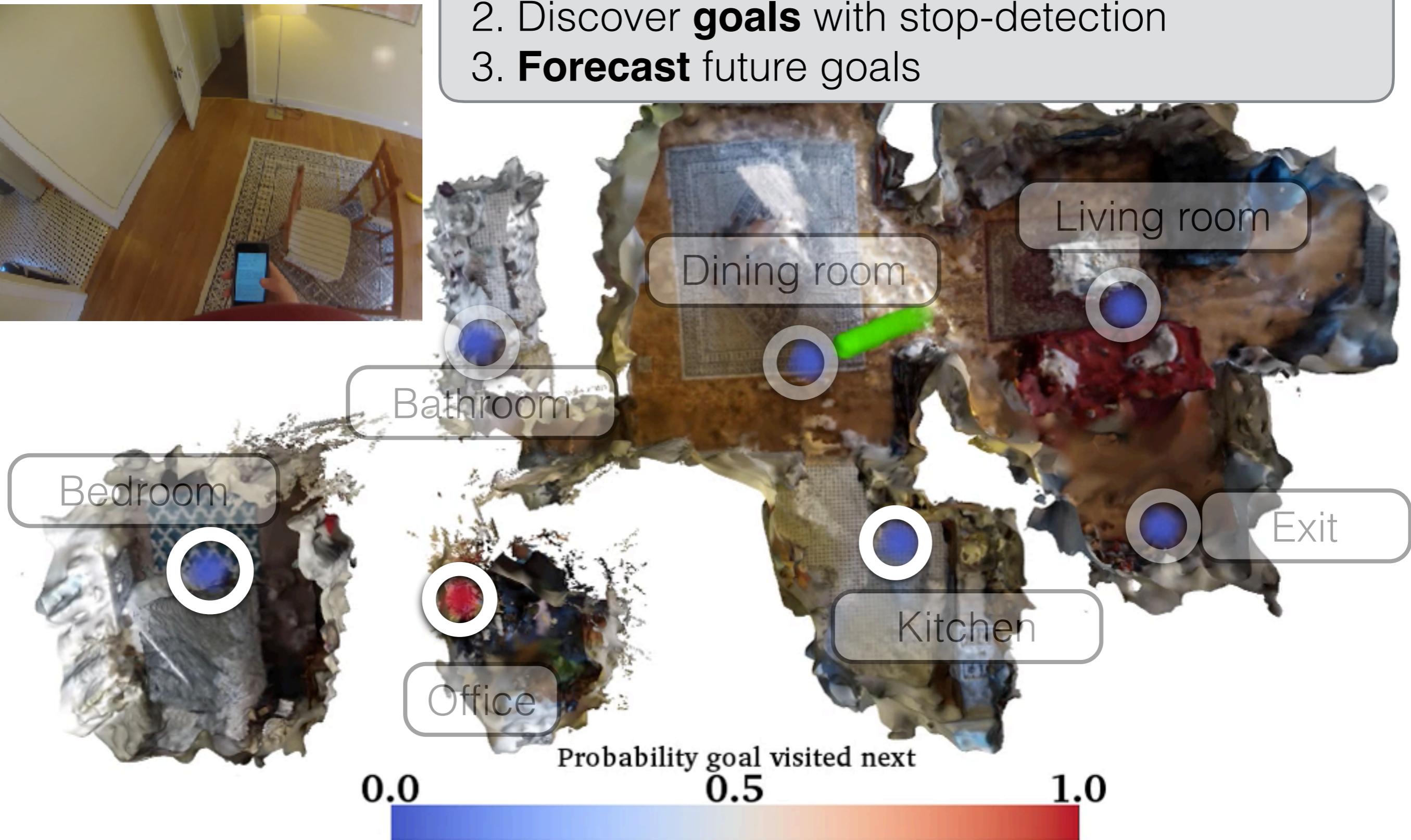
1. Track **state** with SLAM and activity detection



1. Track **state** with SLAM and activity detection
2. Discover **goals** with stop-detection



1. Track **state** with SLAM and activity detection
2. Discover **goals** with stop-detection
3. **Forecast** future goals



wearable sensing for
Human Activity

Understanding

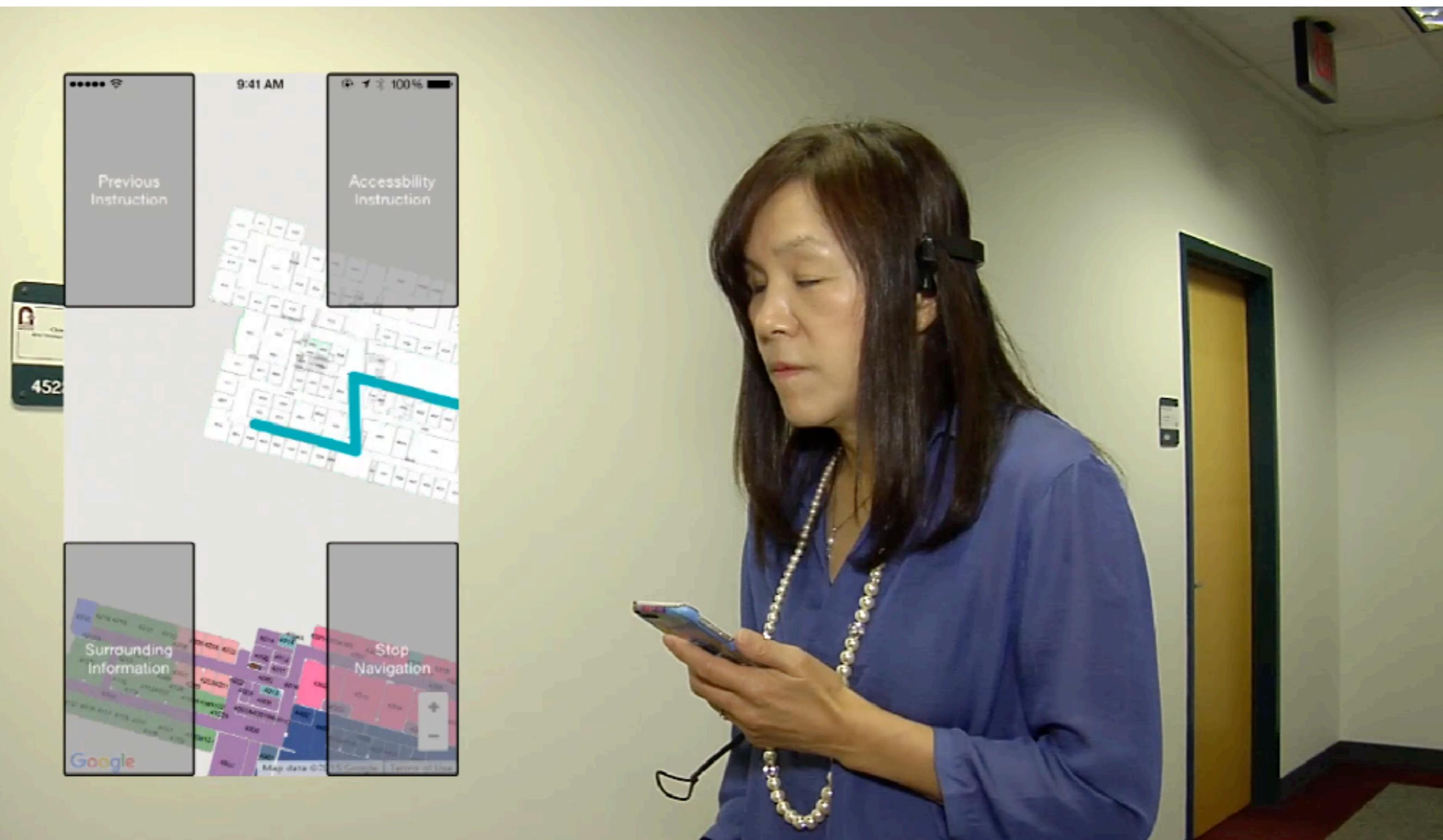
Forecasting

Assisting



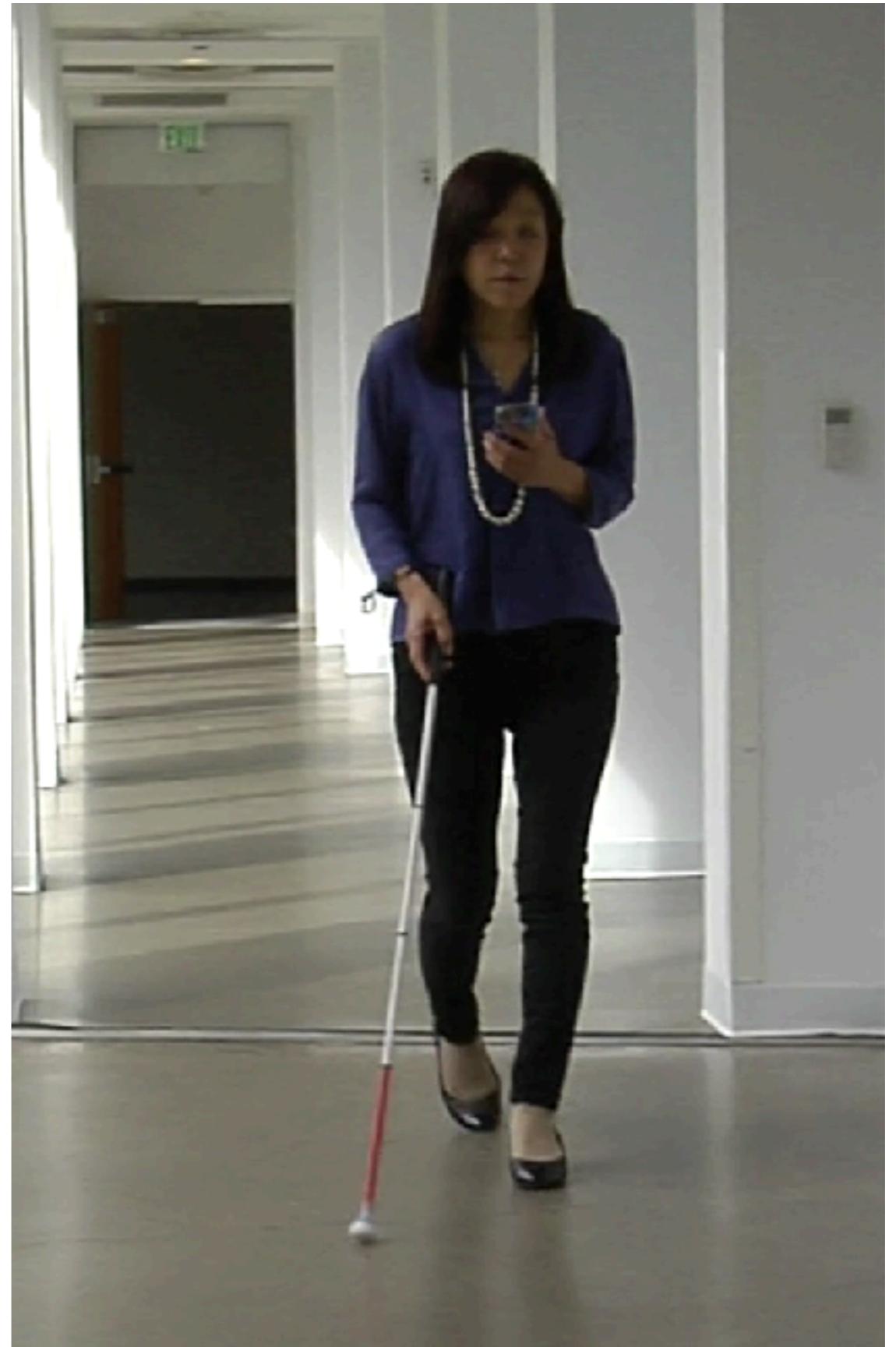
Wearable Sensing via Smartphones for Assisting Indoor Navigation for Blind People

Our Vision



NavCog:

Turn-by-turn navigational instruction

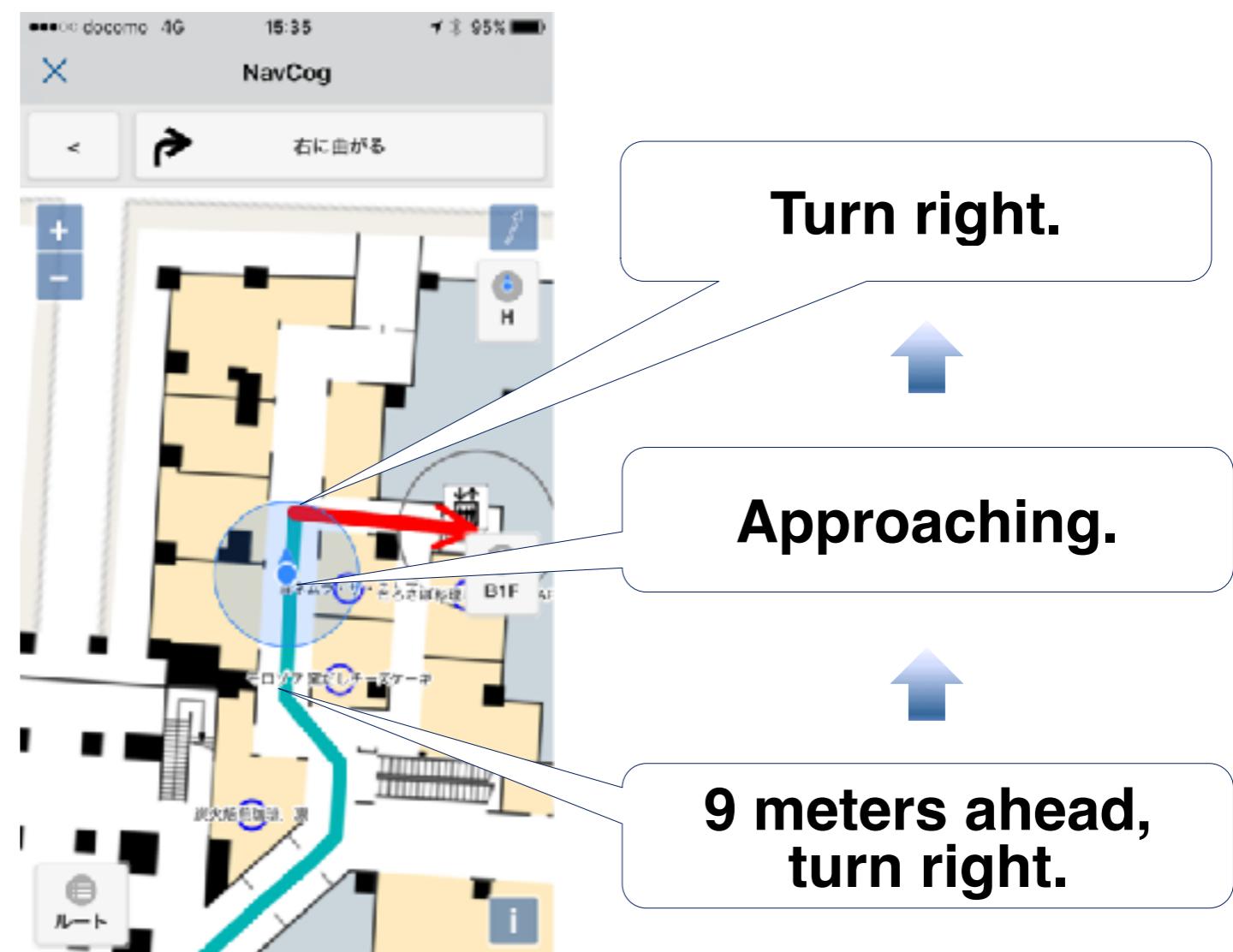


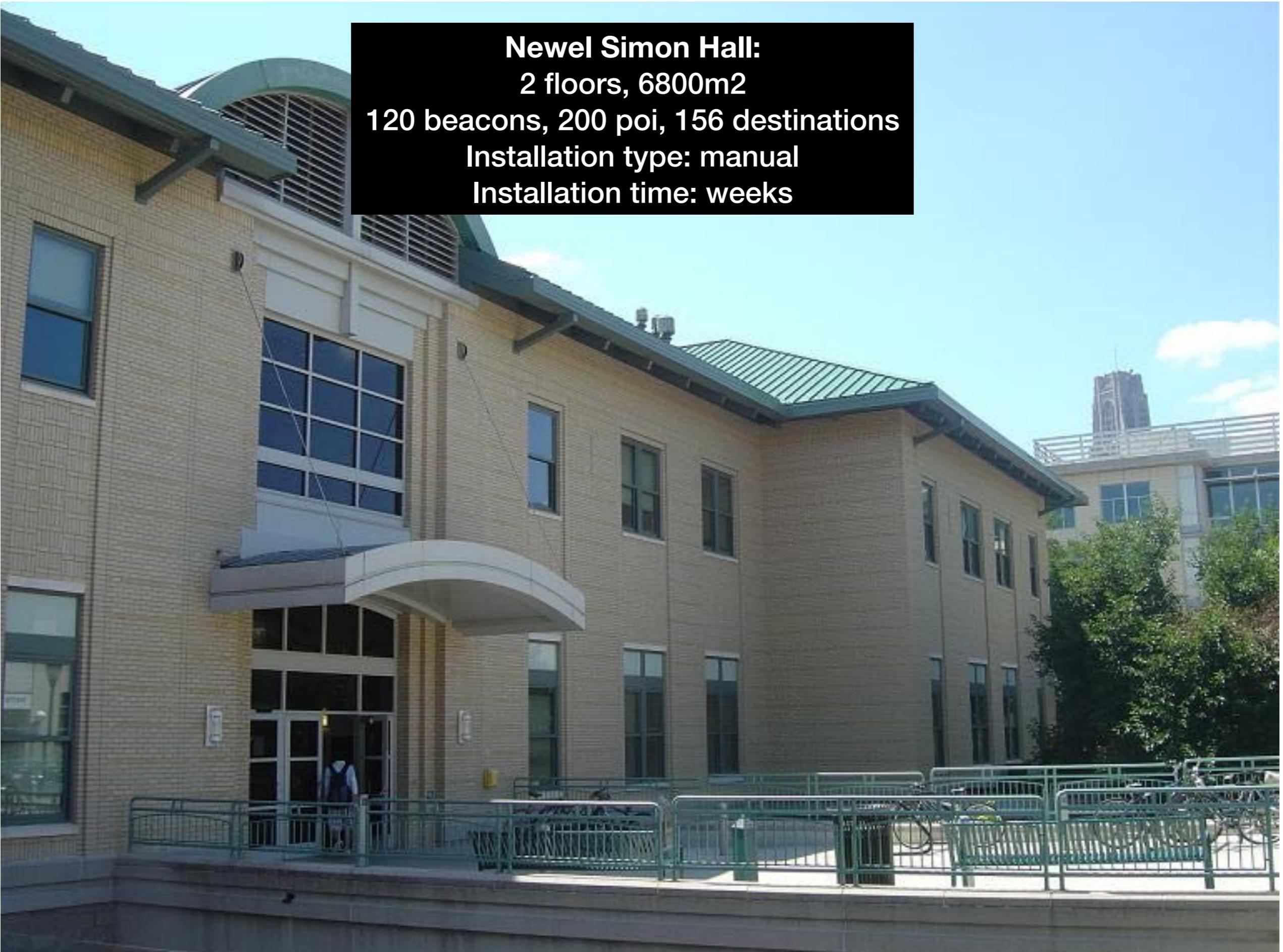
-  **Uran Oh, Daisuke Sato, Kris Kitani, Hironobu Takagi, and Chieko Asakawa.** "NavCog3: An Evaluation of a Smartphone-Based Blind Indoor Navigation Assistant with Semantic Features in a Large-Scale Environment". In: *Proceedings of the International Conference on Computers and Accessibility*. ACM. 2017
-  **J Eduardo Pérez, Myriam Arrue, Masatomo Kobayashi, Hironobu Takagi, and Chieko Asakawa.** "Assessment of Semantic Taxonomies for Blind Indoor Navigation Based on a Shopping Center Use Case". In: *Proceedings of the Web for All Conference*. ACM. 2017
-  **Dragan Ahmetovic, Masayuki Murata, Cole Gleason, Erin Brady, Hironobu Takagi, Kris Kitani, and Chieko Asakawa.** "Achieving Practical and Accurate Indoor Navigation for People with Visual Impairments". In: *Proceedings of the Web for All Conference (2017)*
-  **Dragan Ahmetovic, Cole Gleason, Chengxiong Ruan, Kris M Kitani, Hironobu Takagi, and Chieko Asakawa.** "NavCog: a navigational cognitive assistant for the blind." In: *Proceedings of the International Conference on Human-Computer Interaction with Mobile Devices and Services*. ACM. 2016
-  **Dragan Ahmetovic, Cole Gleason, Kris M Kitani, Hironobu Takagi, and Chieko Asakawa.** "NavCog: turn-by-turn smartphone navigation assistant for people with visual impairments or blindness". In: *Proceedings of the Web for All Conference*. ACM. 2016



Turn-by-turn Navigation

- Providing voice & visual instructions at each corner
- Re-routing capability
- Similar to car navigation systems
- High-precision localization via Bluetooth beacon network and particle filtering





Newel Simon Hall:
2 floors, 6800m²
120 beacons, 200 poi, 156 destinations
Installation type: manual
Installation time: weeks

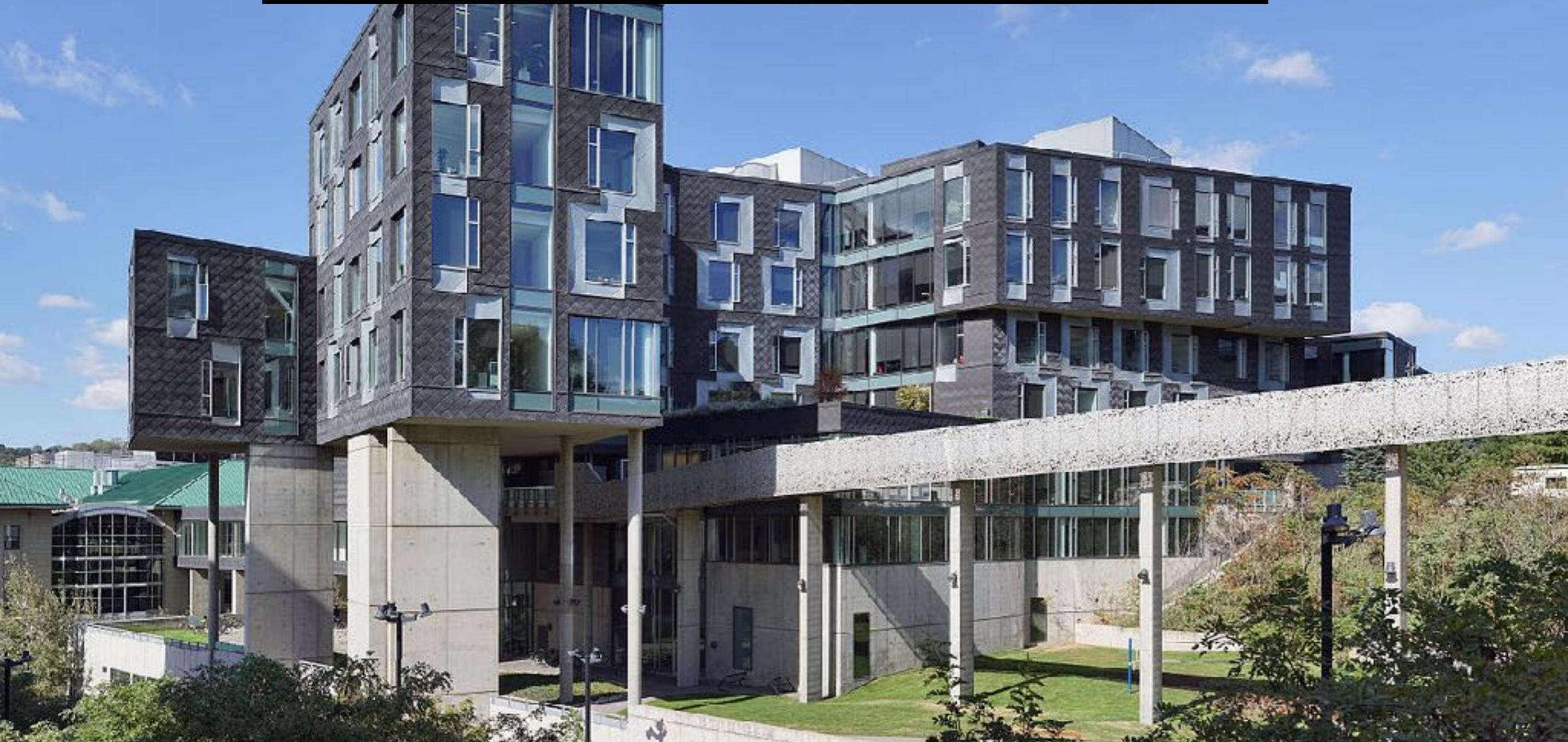
Gate Hillman Center:

6 floors, 22, 000m²

450 beacons, 487 poi, 396 destinations

Installation type: crowdsourcing (beacons), cobot (fingerprints)

Installation time: 1 week (beacons) + 2 days (fingerprints)



Wean Hall:

8 floors, 30, 000m²

300 beacons, 668 poi, 626 destinations

Installation type: admin app (beacons), lidar (fingerprints)

Installation time: 1 day (beacons) + 1 day (fingerprints)



Coredo Muromachi:
3 buildings, 4 floors, 21, 000m²
218 beacons, 200 POI, 92 destinations
Installation type: manual* (beacons), lidar (fingerprints)
Installation time: 2 weeks



Greentree DoubleTree Hotel 2017



Beacons here

OR TAKE OUT.

pinkberry

Fresh Deli

The Strip Ma

AJ | ARMANI
JEANS

Allegheny General Hospital 2018



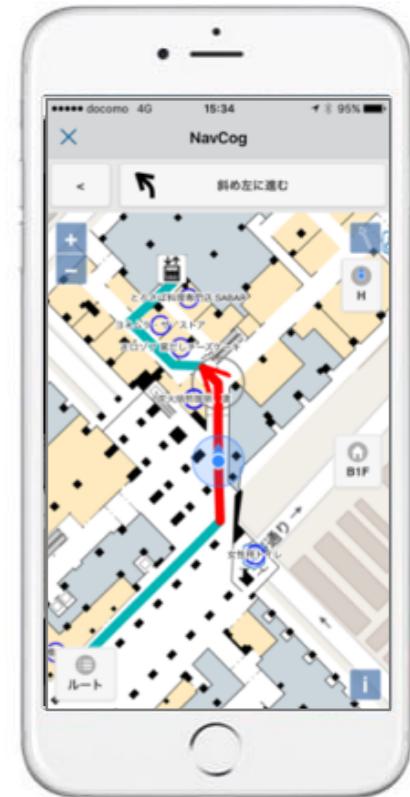
Nihonbashi Muromachi Pilot

(Feb. 1 – Feb. 28, 2017)



226,000ft² (21,000m²)

92 stores



Available on the
App Store

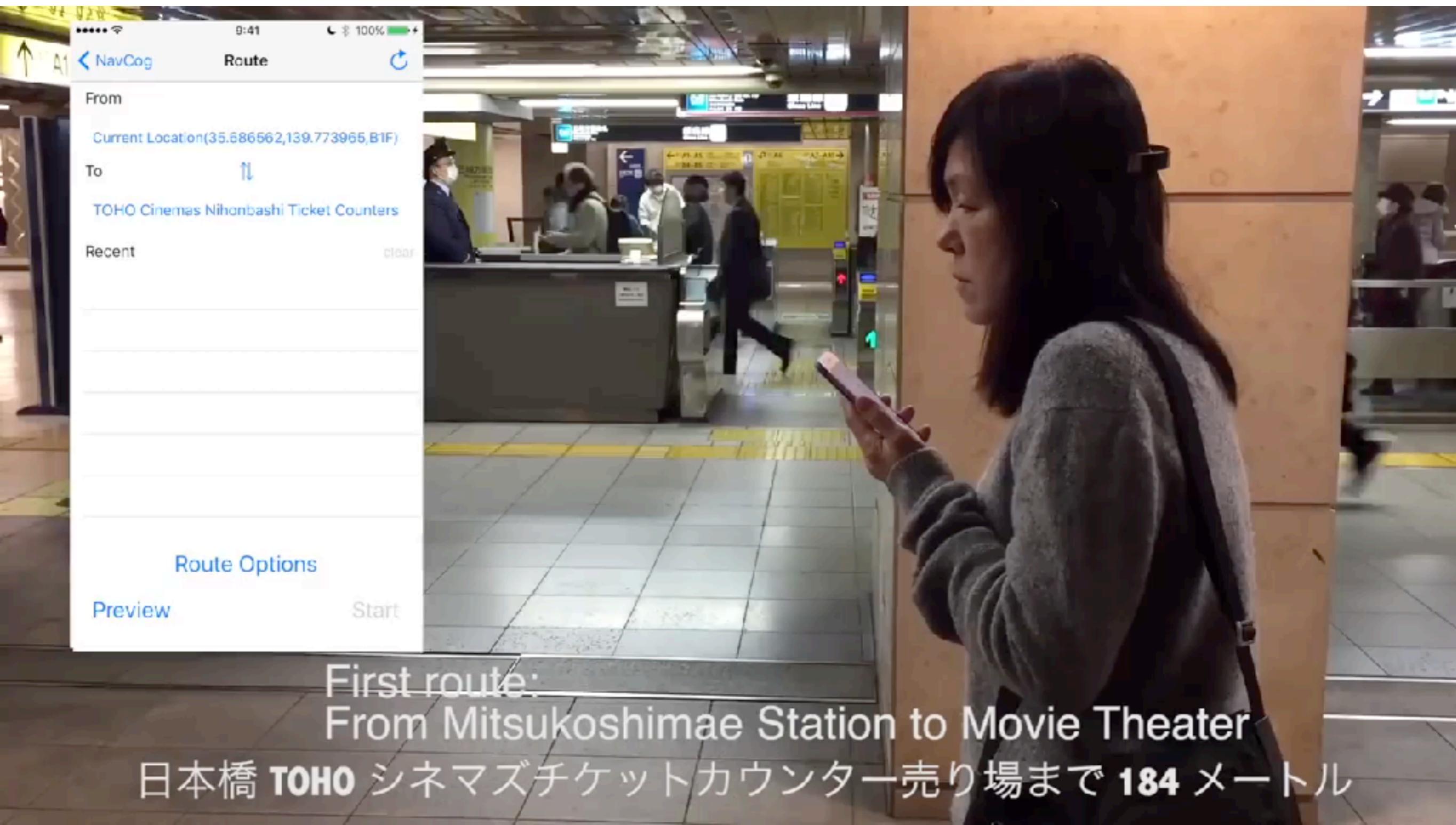
Target users:

- Blind
- Wheelchair
- Foreign visitors

Source: MITSUI FUDOSAN (<http://www.mitsufudosan.co.jp>), Mitsui Shopping Park Urban (<https://31urban.jp/>)

Conversational Concierge + Navigation





Wearable Sensing for Understanding, Forecasting and Assisting Human Activity

Kris Kitani

Assistant Research Professor
Carnegie Mellon University

