CIS 3990

Mobile and IoT Computing

https://penn-waves-lab.github.io/cis3990-24spring

Lecture 9: ML-based Sensing & Through-Wall Pose Estimation

Instructor: Mingmin Zhao (mingminz@cis.upenn.edu)

TA: Haowen Lai (hwlai@cis.upenn.edu)

Objectives of This Module

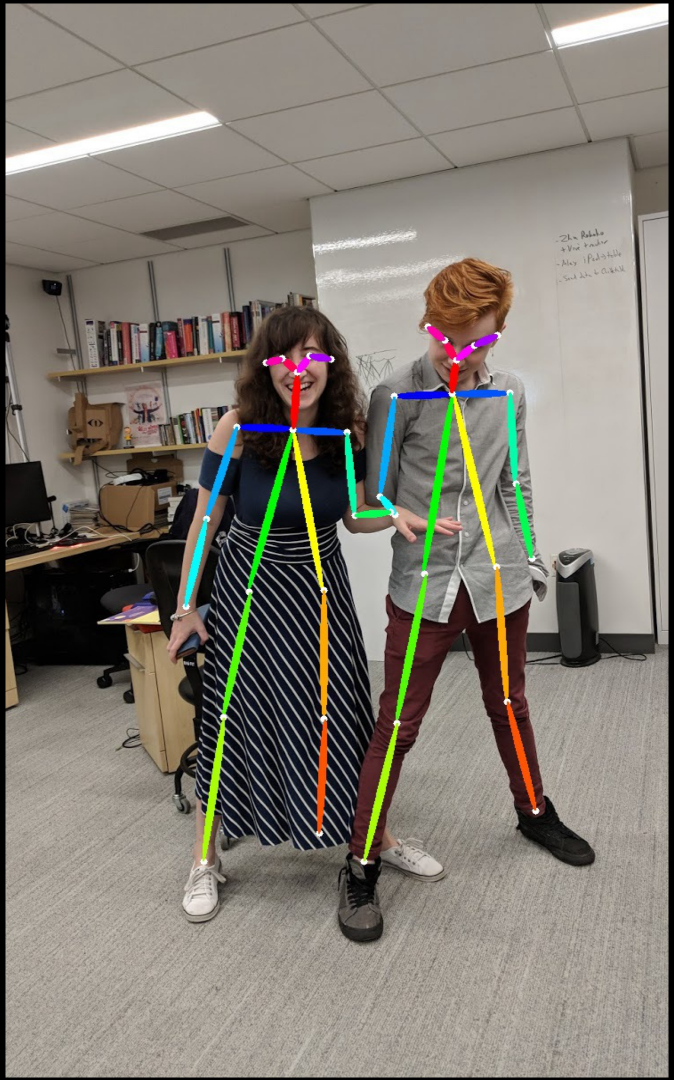
Learn how foundational sensing technologies can be used to extract diverse and meaningful insights

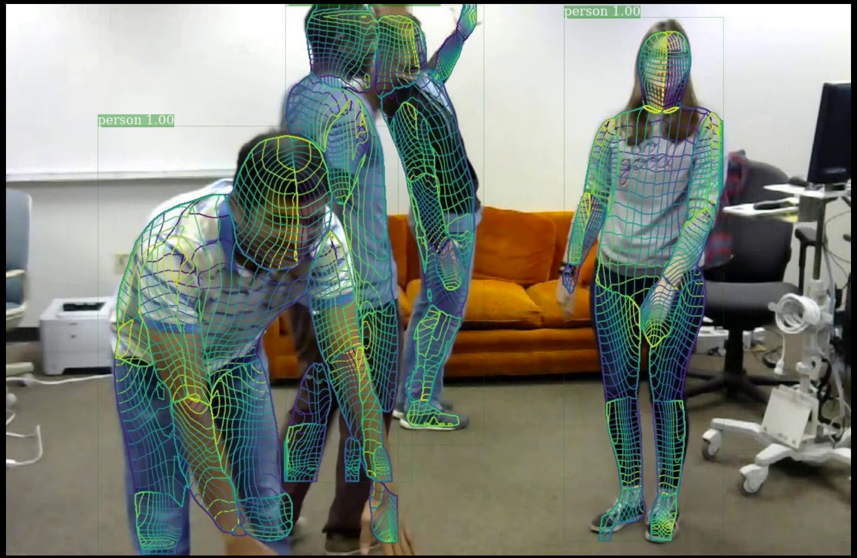
1. What are important application areas of Mobile and IoT sensing?
2. What are the foundational sensing mechanisms and how are they related to localization?
3. What processing algorithms can be used to transform raw sensor data?
4. Example sensing systems/solutions with real-world case studies.

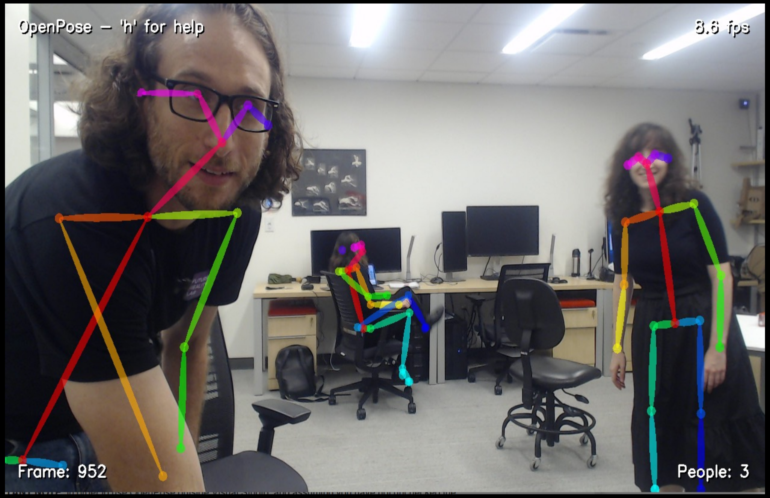
Focus of this lecture:

ML algorithms to extract insights from raw sensor data

Sensing Humans in the Environment

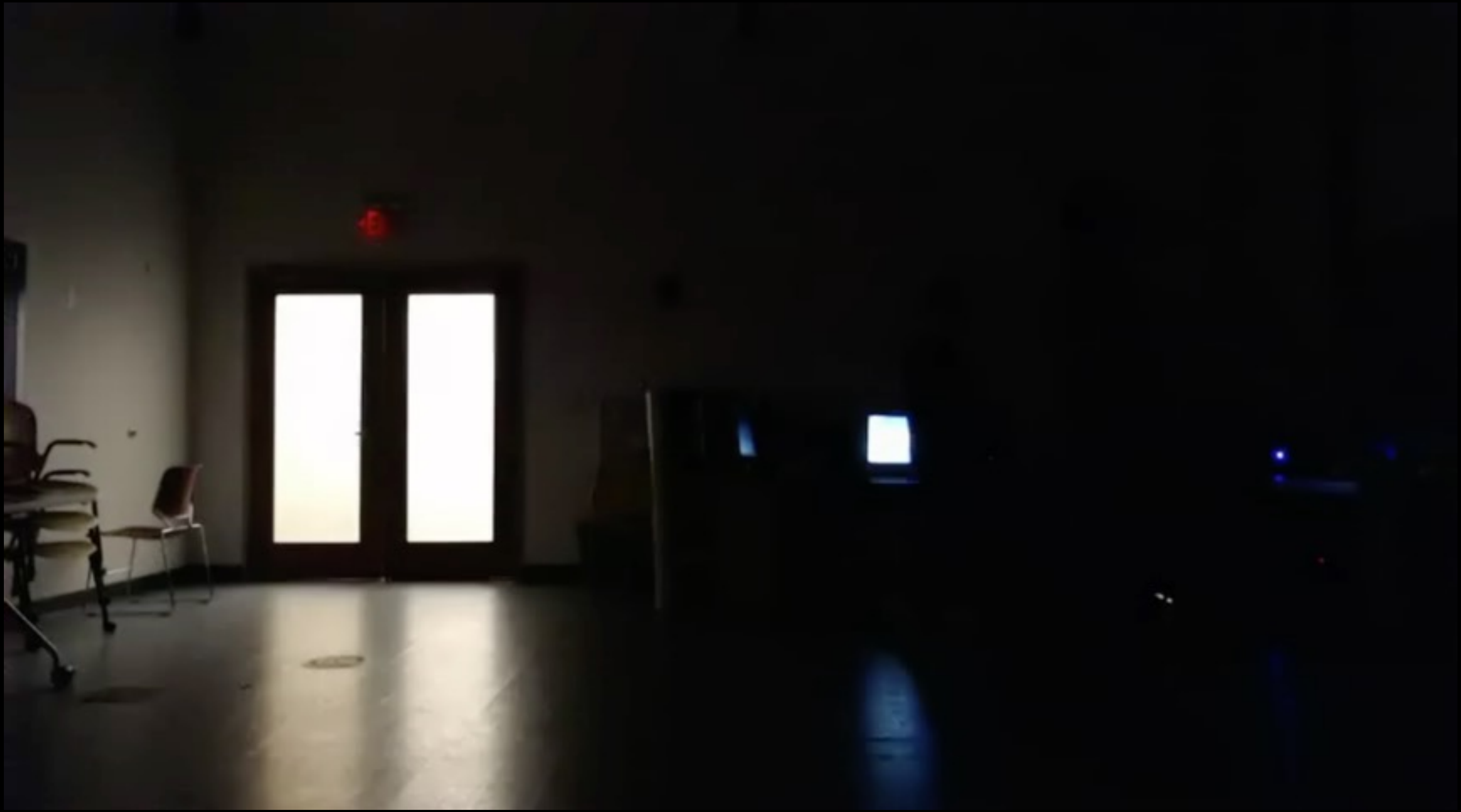




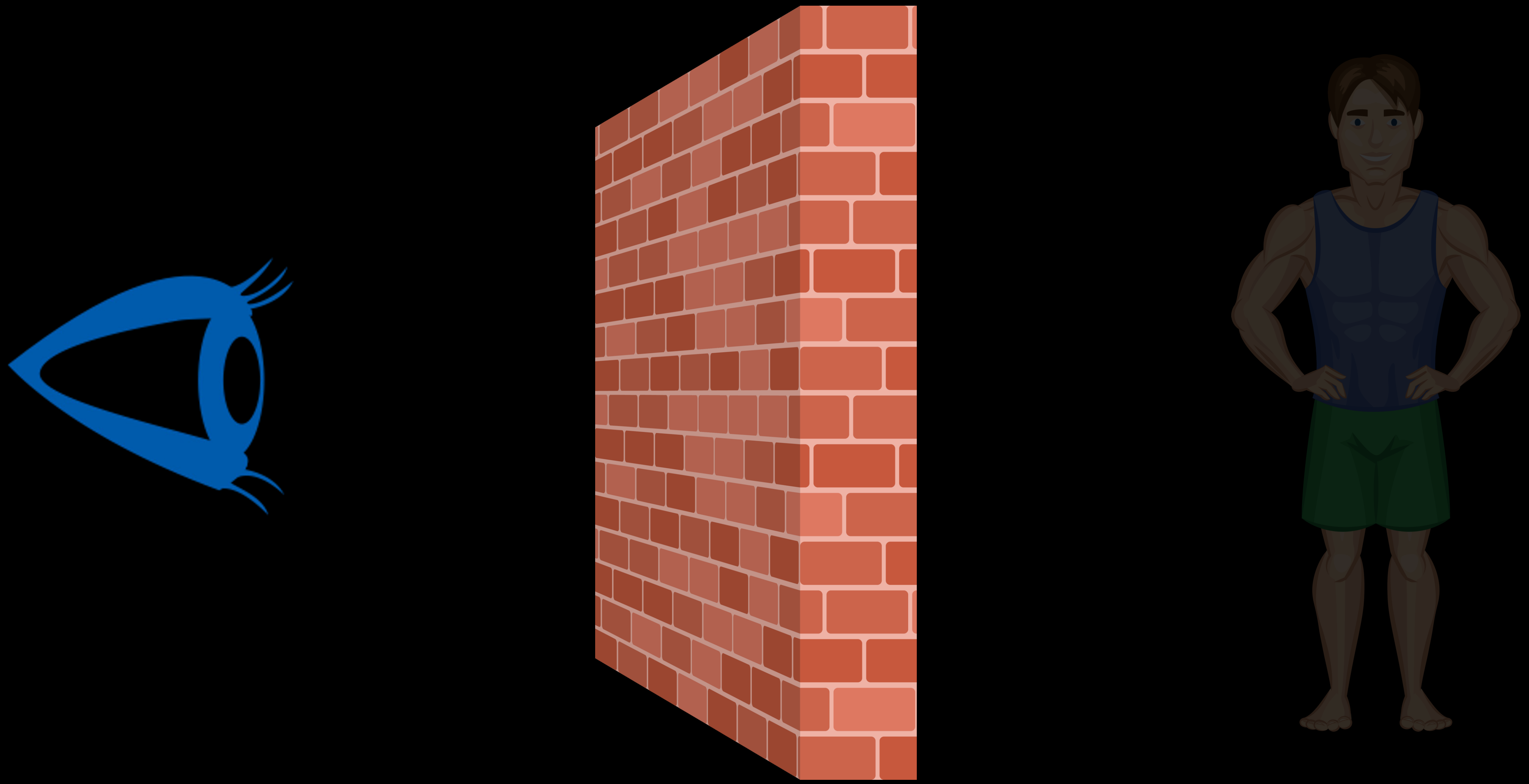


Occlusion is a fundamental challenge for vision

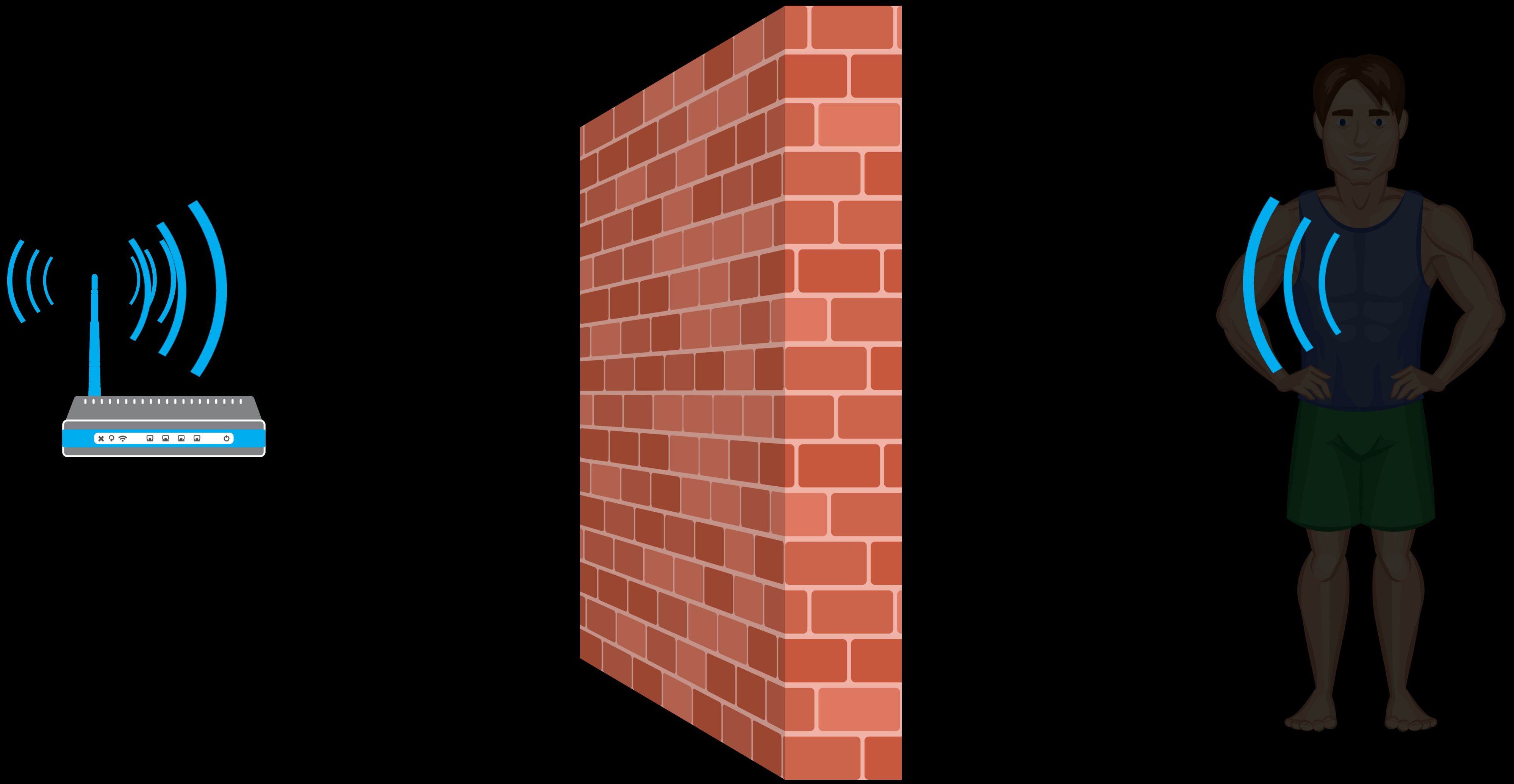
Vision also fails in bad lighting conditions



Want to see the human through walls & in the dark



Want to see the human through walls & in the dark



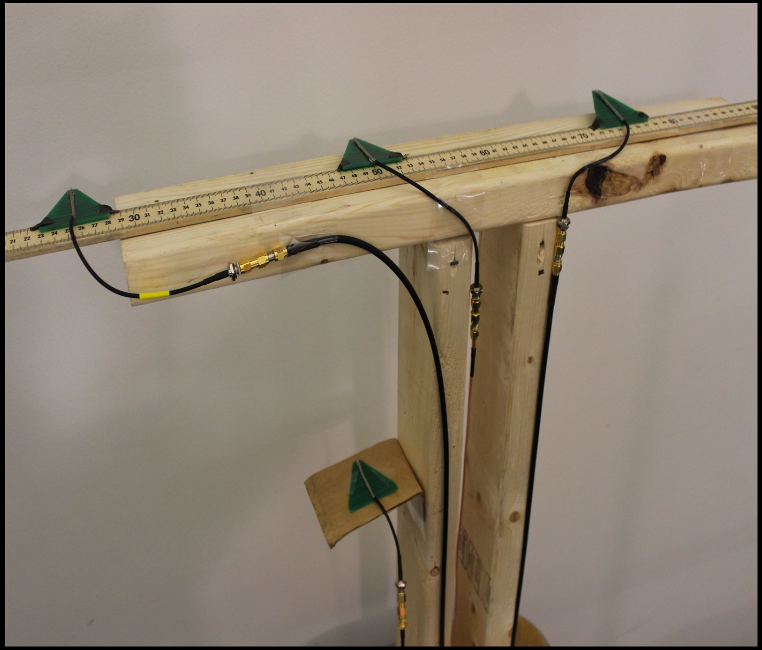
RF-based Approach

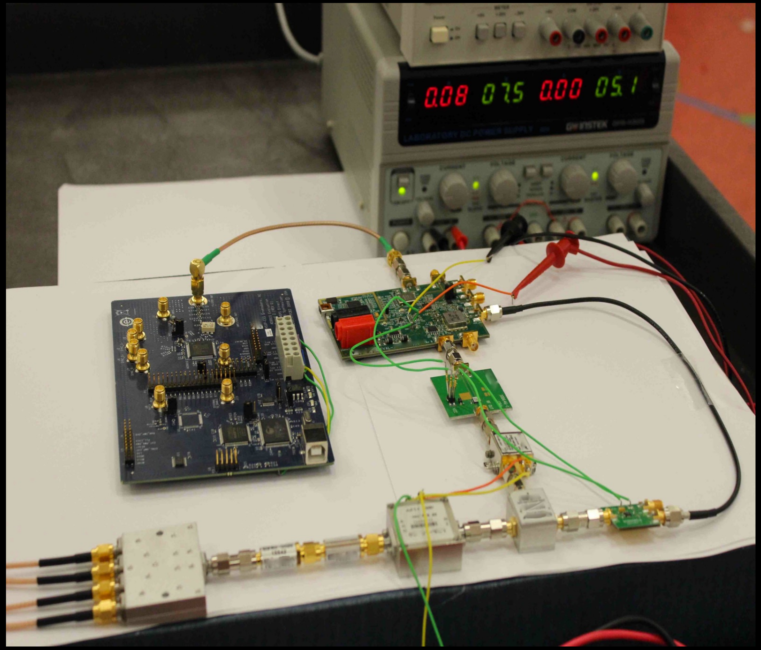
DARPA see-through-wall (mid 2000)

MIT Lincoln Lab (2011)

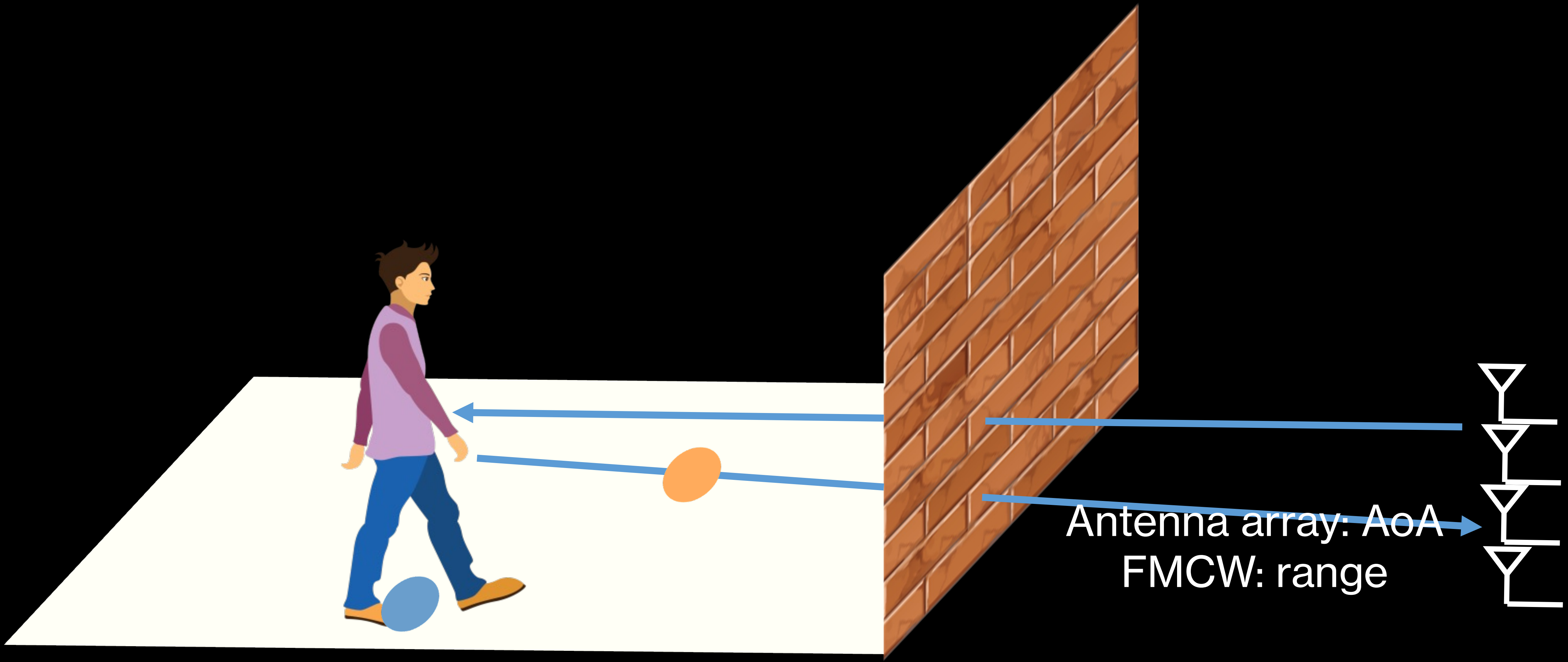
RF-based Approach

Wi-Vi and WiTrack from MIT
(2013)

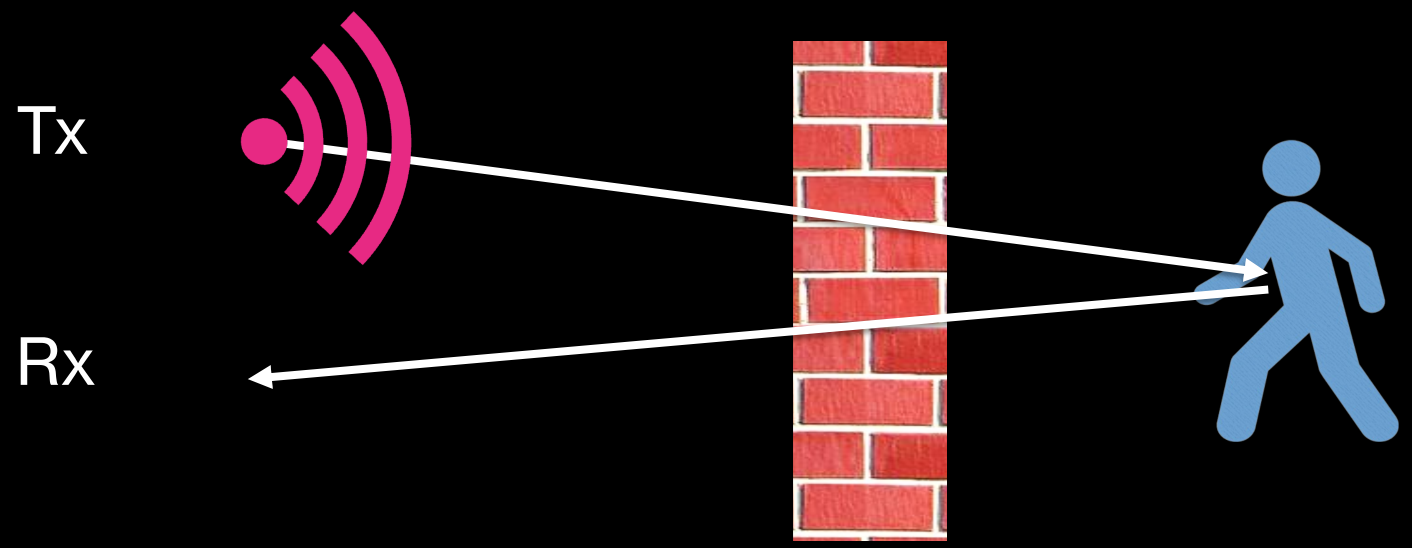




RF-based Approach

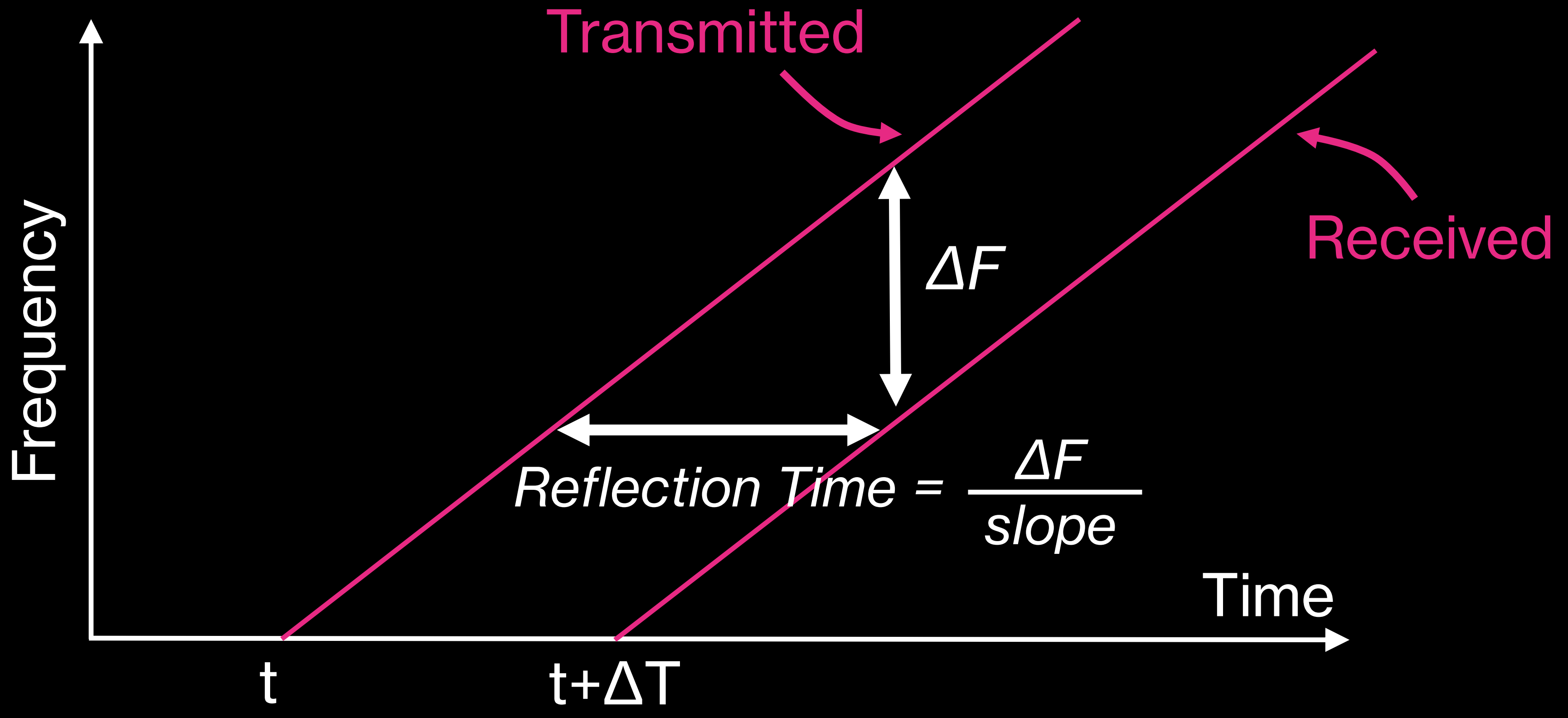


Measuring Distances



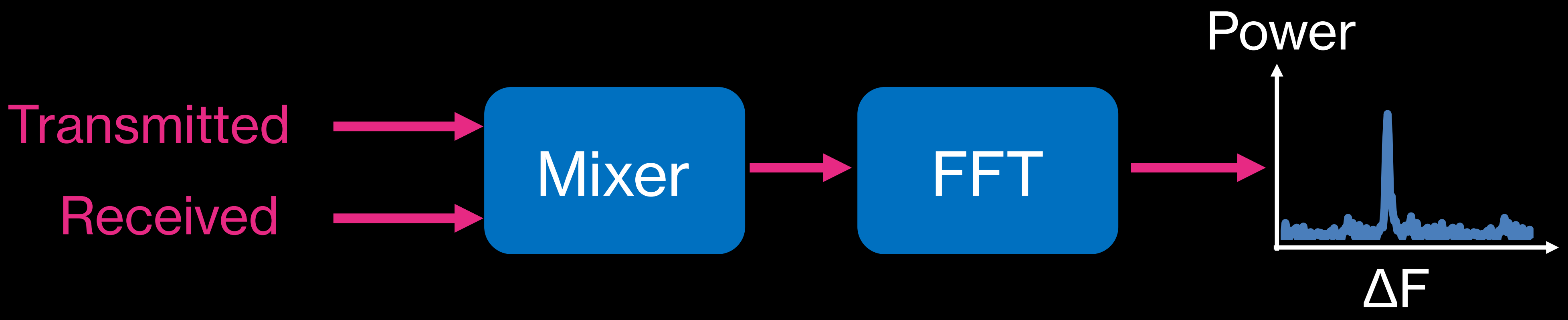
Round-Trip Distance = Reflection time x Speed of light

FMCW: Measure time by measuring frequency

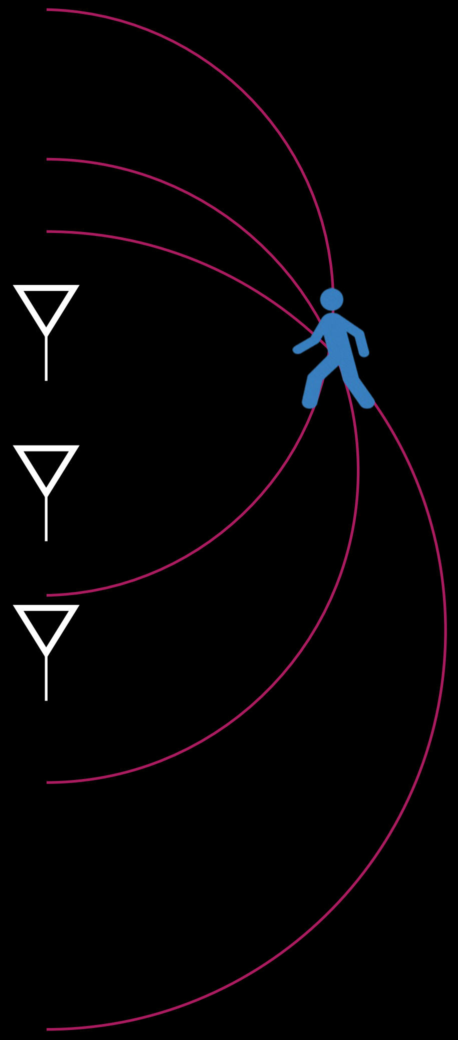


Measuring ΔF

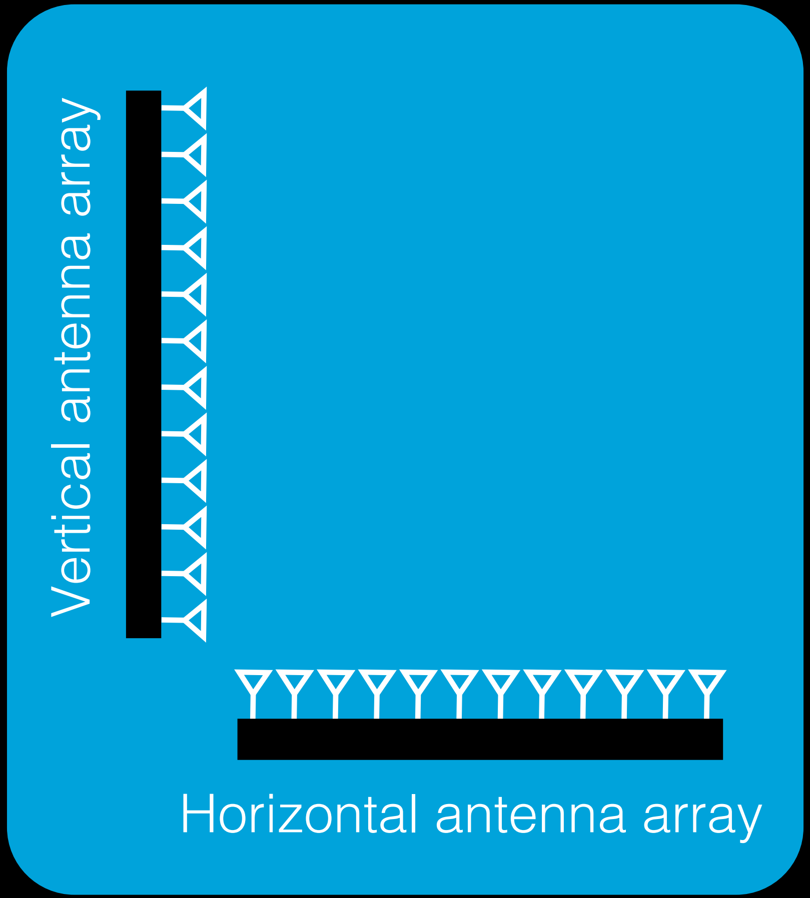
- Subtracting frequencies is easy (e.g., removing carrier in WiFi)
- Done using a mixer (low-power; cheap)



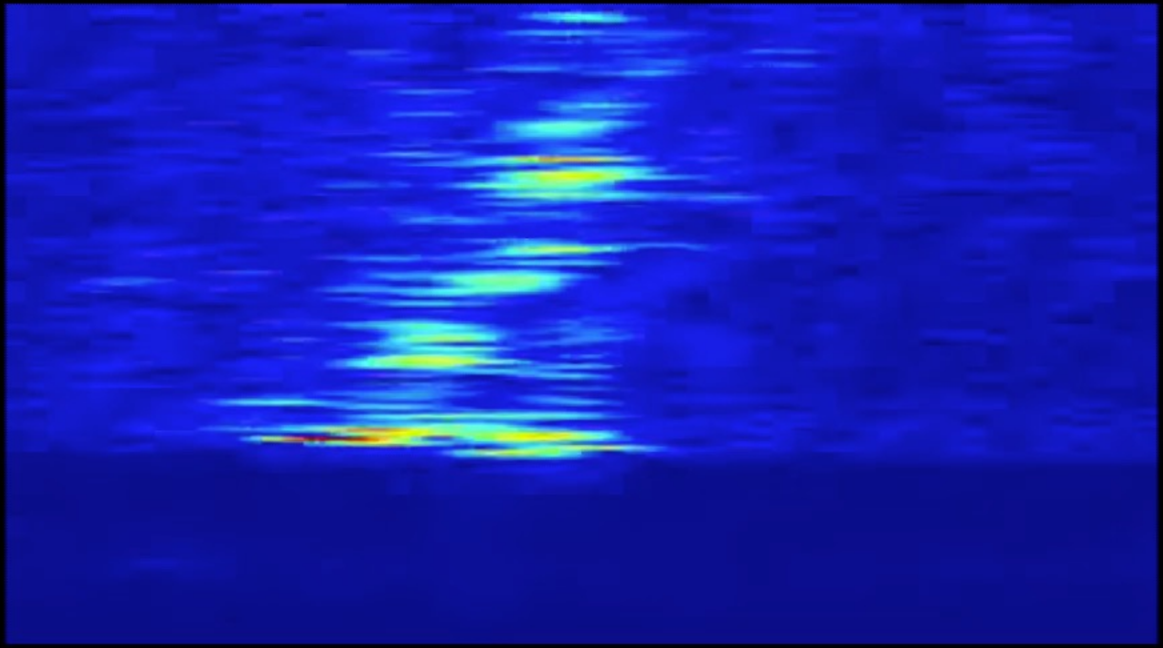
(Ideal) Multilateration

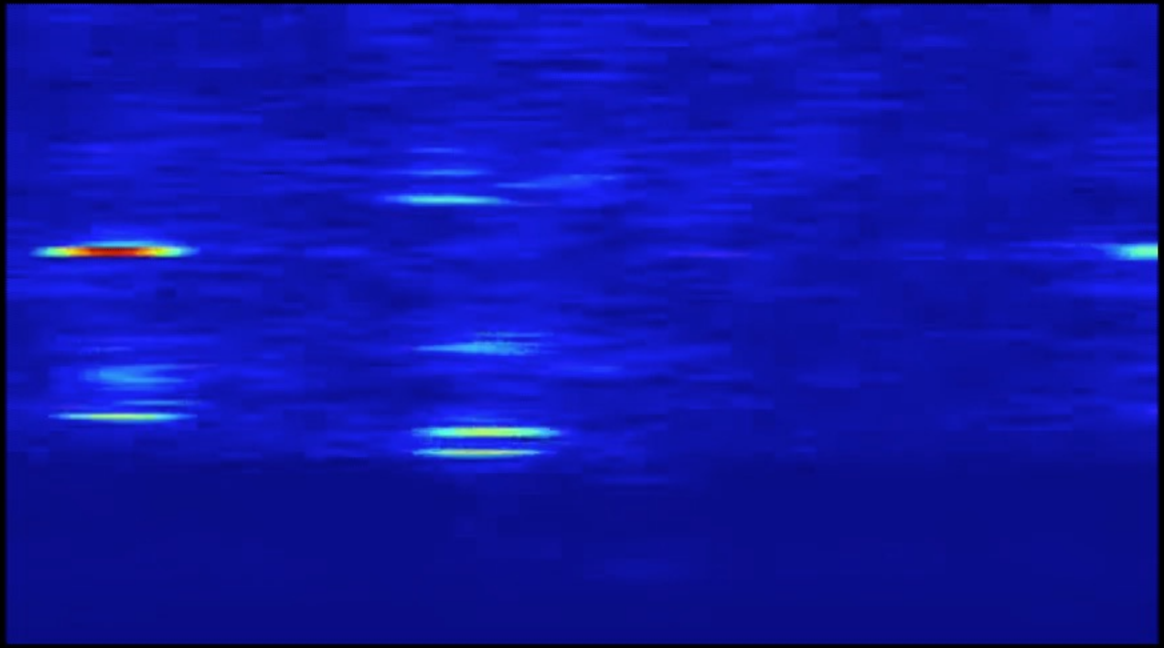


Antenna Arrays for AoA Estimation



How to train a model to estimate pose from RF?





Estimate Pose from RF

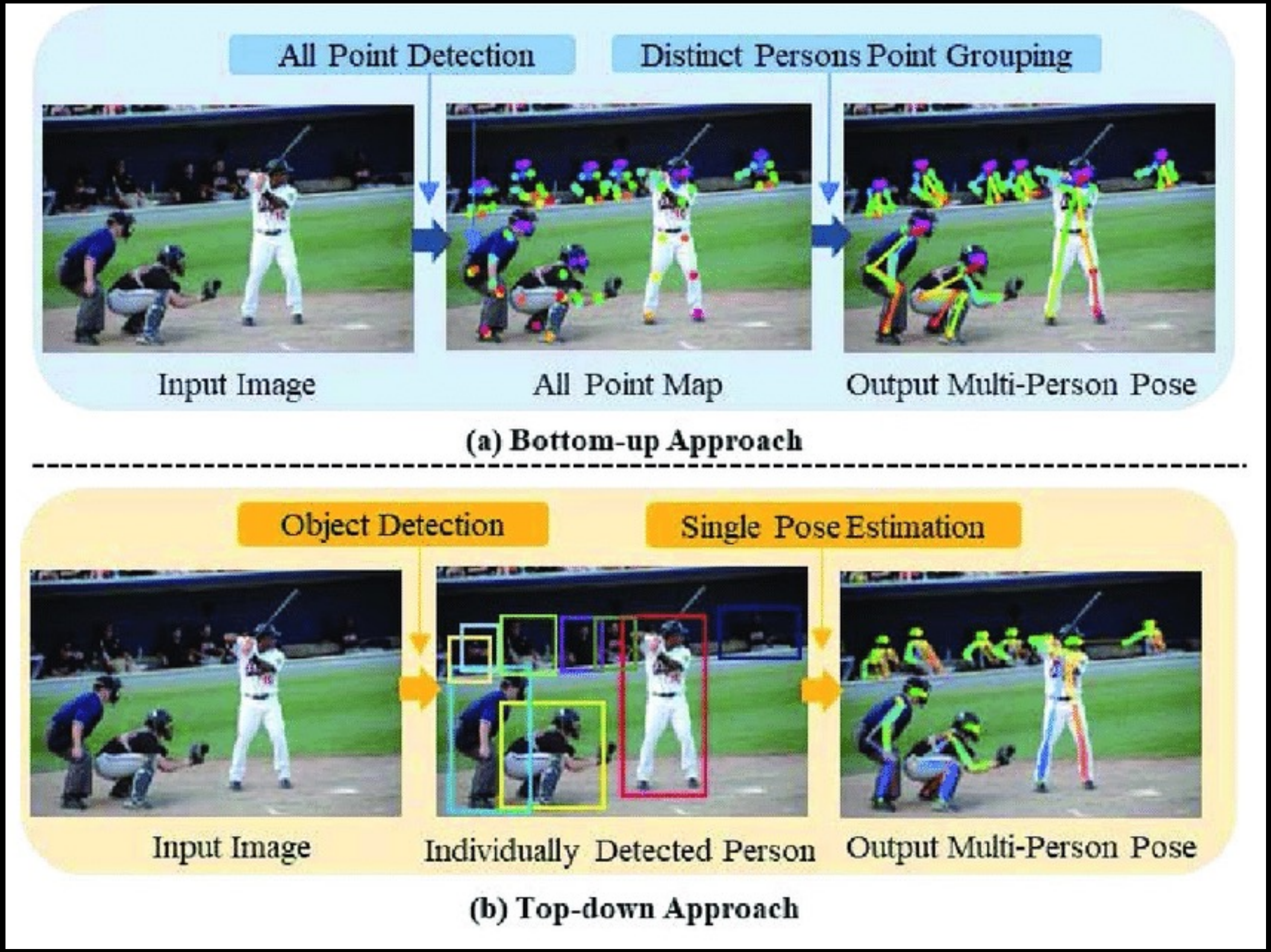
14 body keypoints:

- Head
- Neck
- Shoulders
- Elbows
- etc.

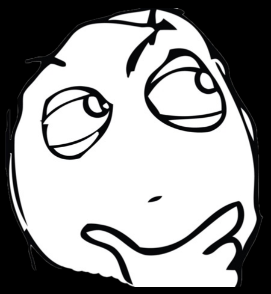
Outline

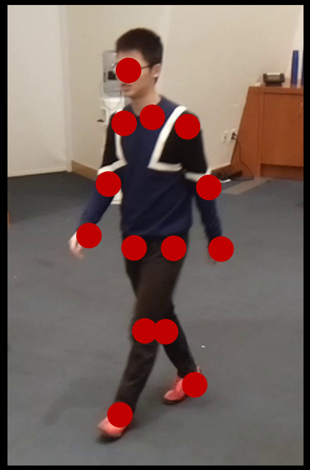
1. Two paradigms for multi-instance pose estimation
2. Challenges in RF-based pose estimation
3. Through-Wall human pose estimation (2D & 3D)

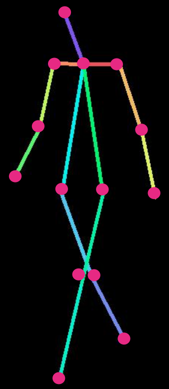
Pose Estimation: Bottom-up and Top-down Approaches

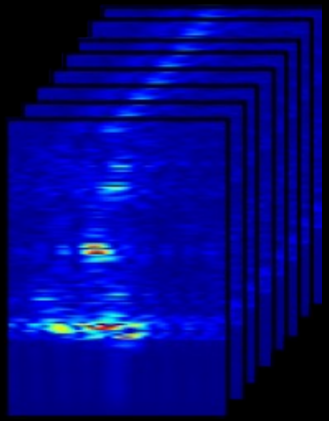


Challenge: How to obtain labeled data?

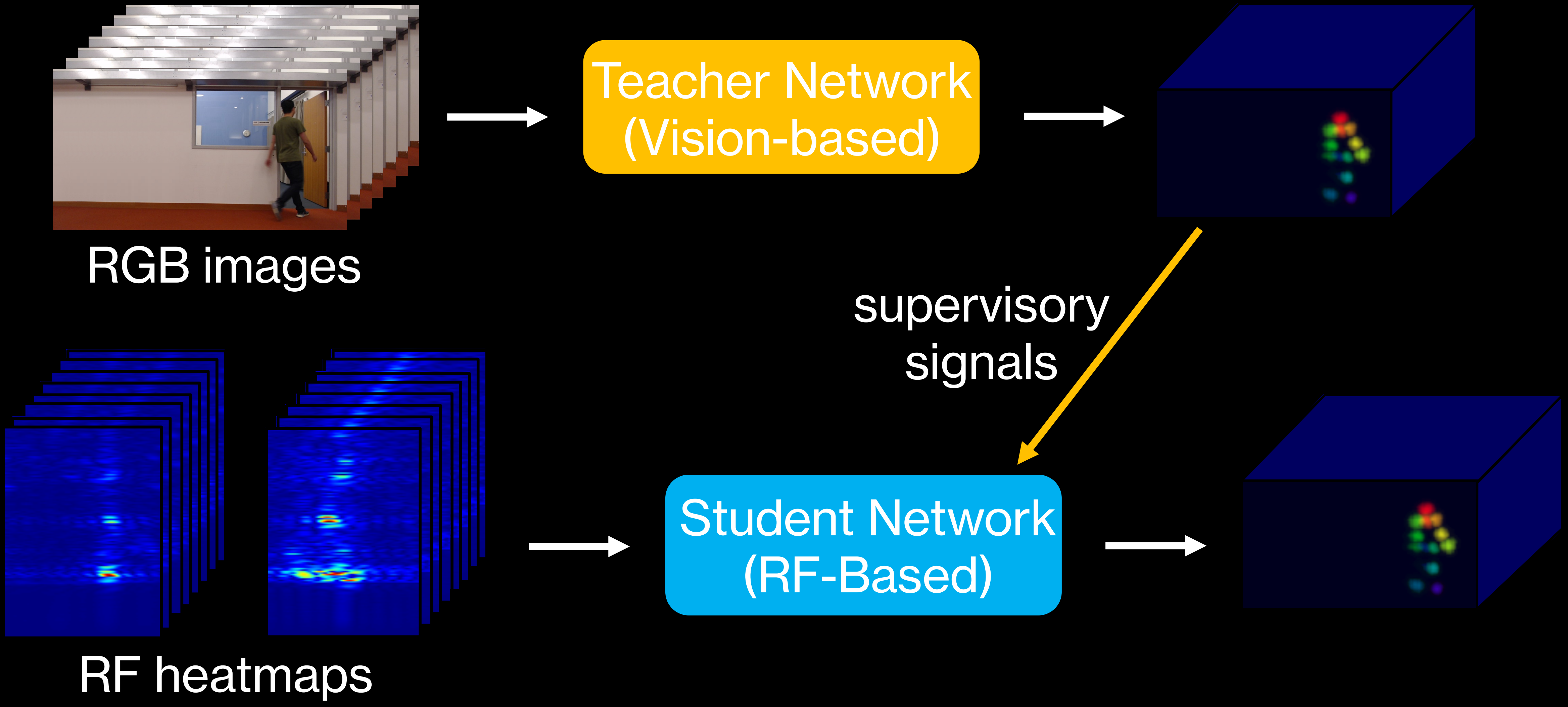




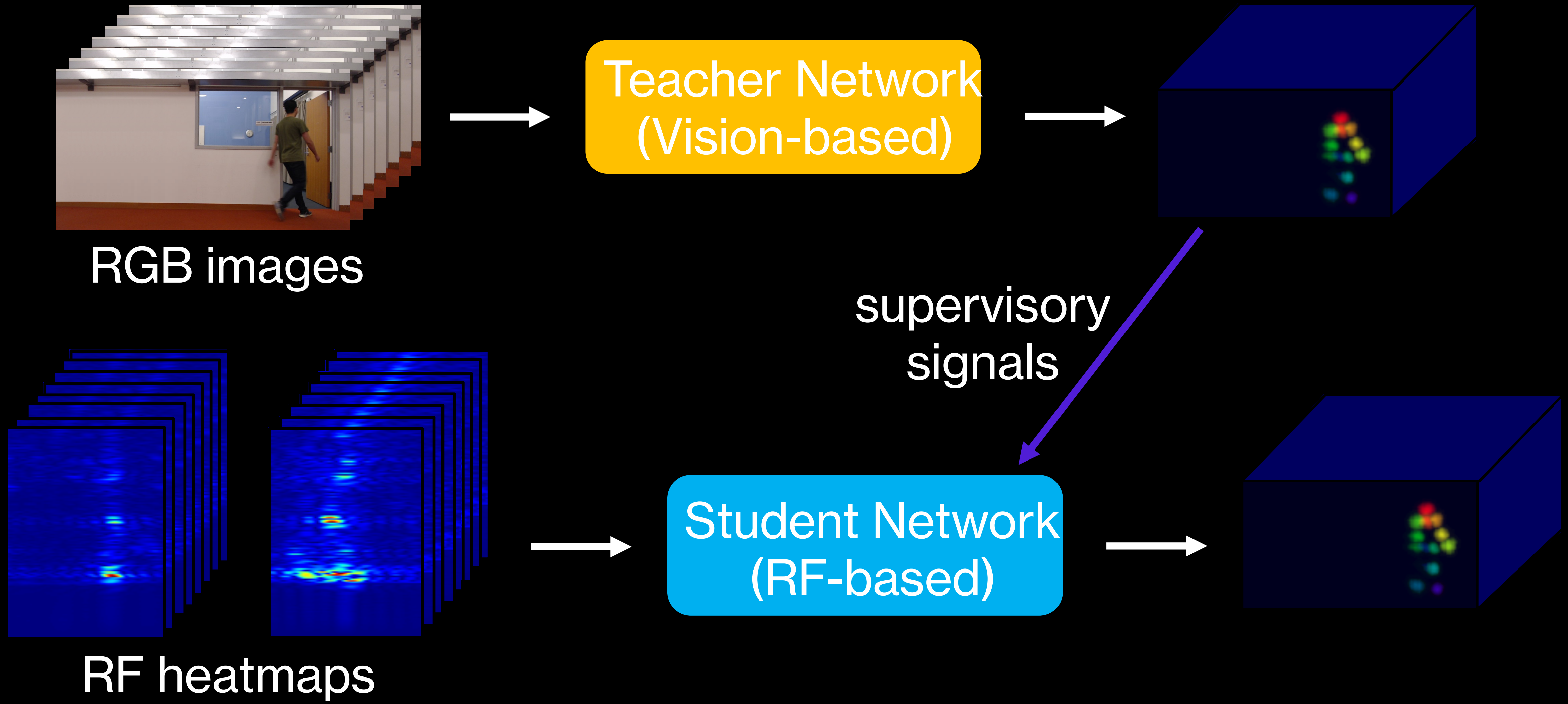




Idea: Cross-Modal Supervision

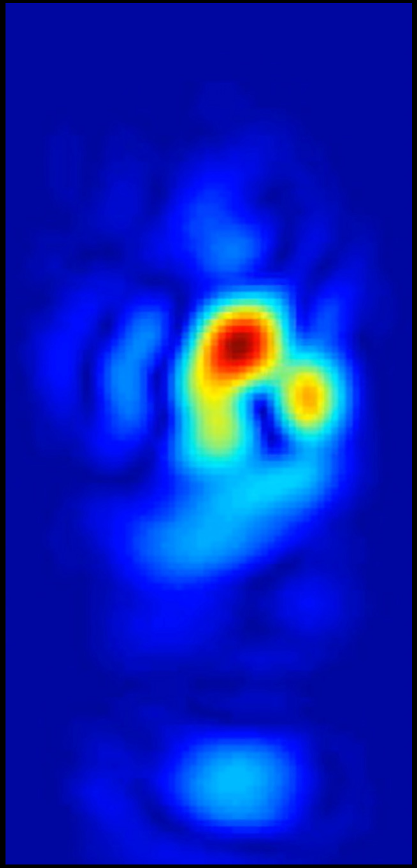


During inference

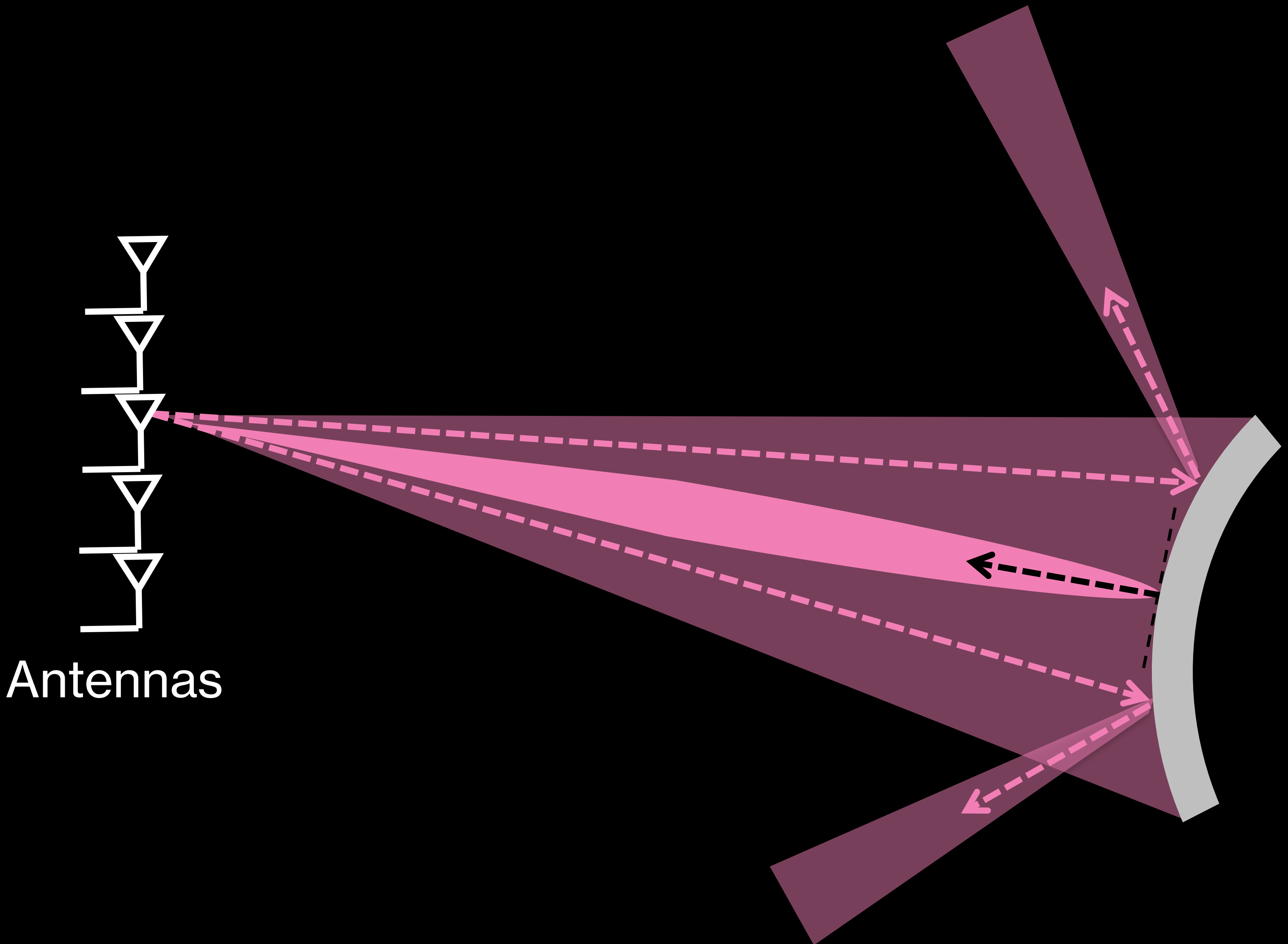


Challenge: Specularity of Human Body

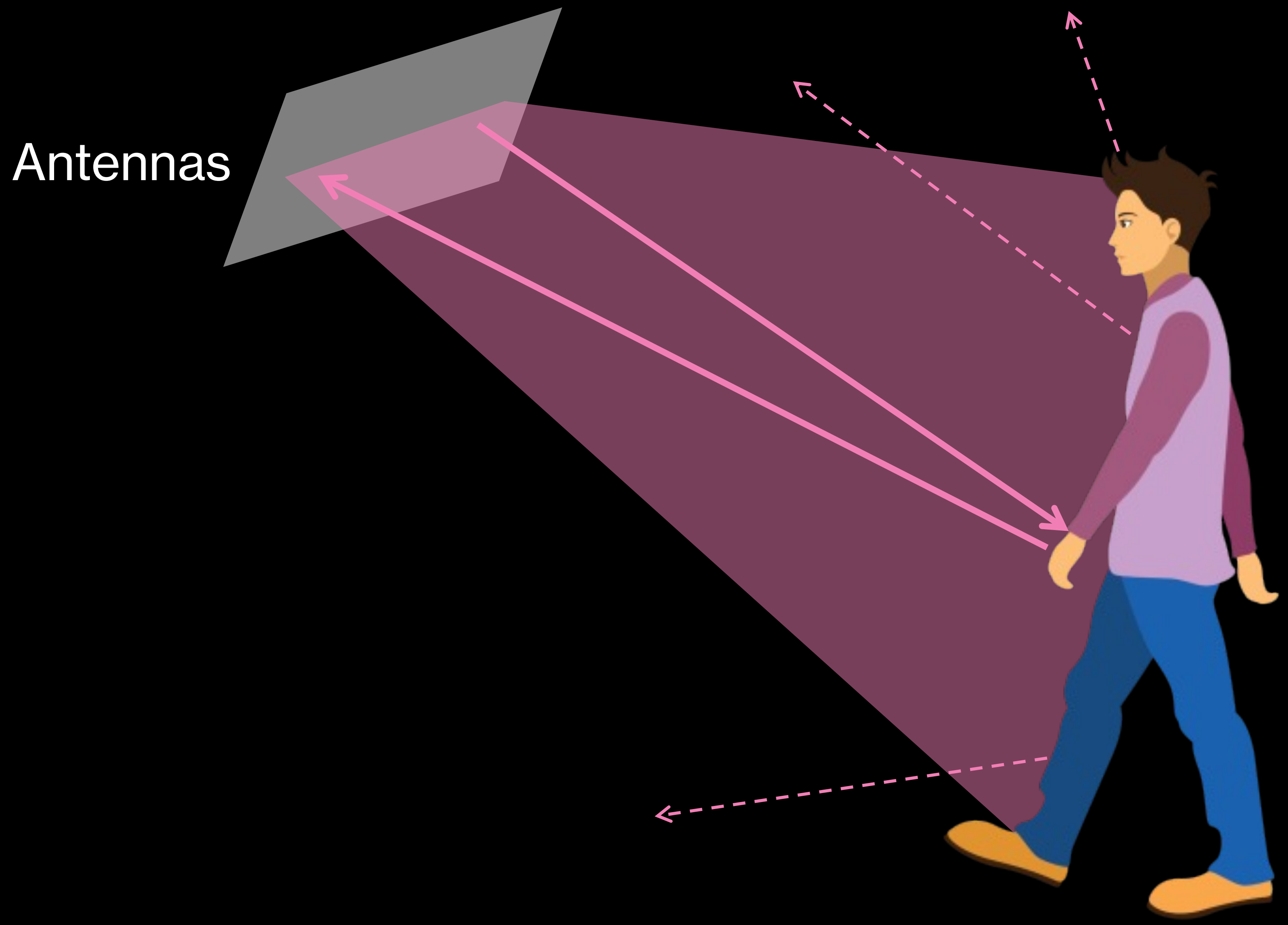
At frequencies that traverse walls, human body is specular (pure mirror)



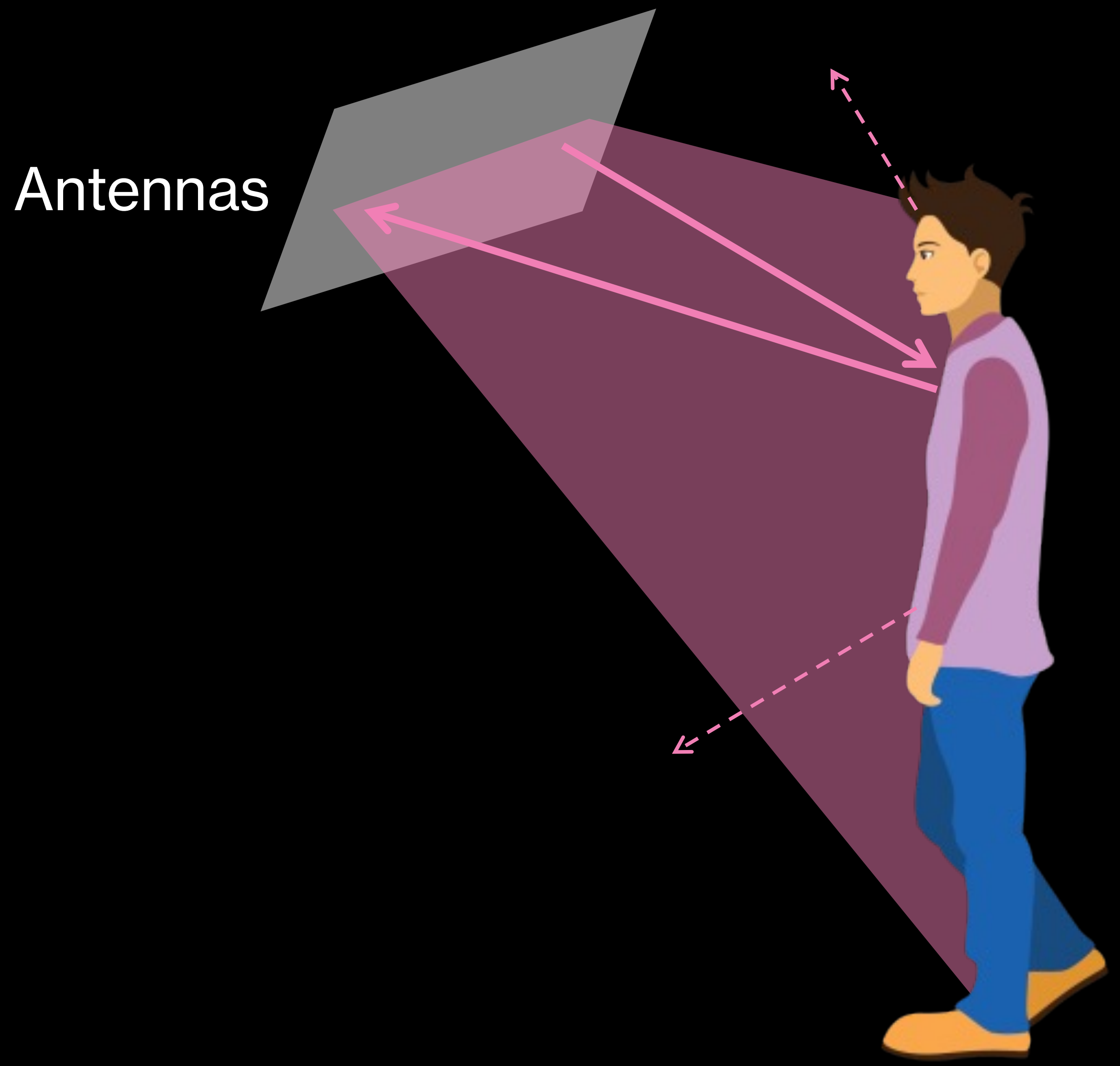
Challenge: Specularity of Human Body



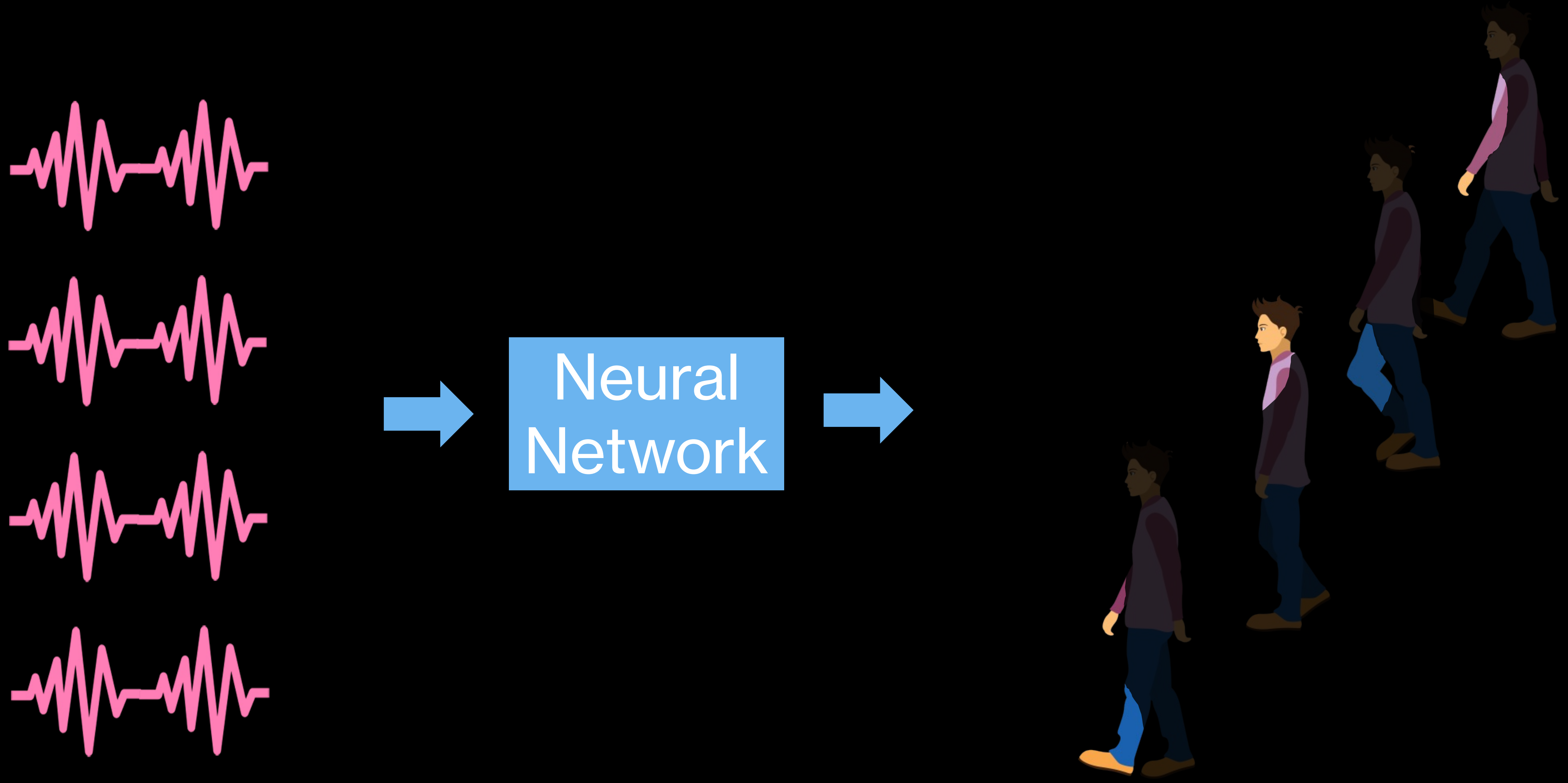
A Snapshot Doesn't Capture Full Skeleton

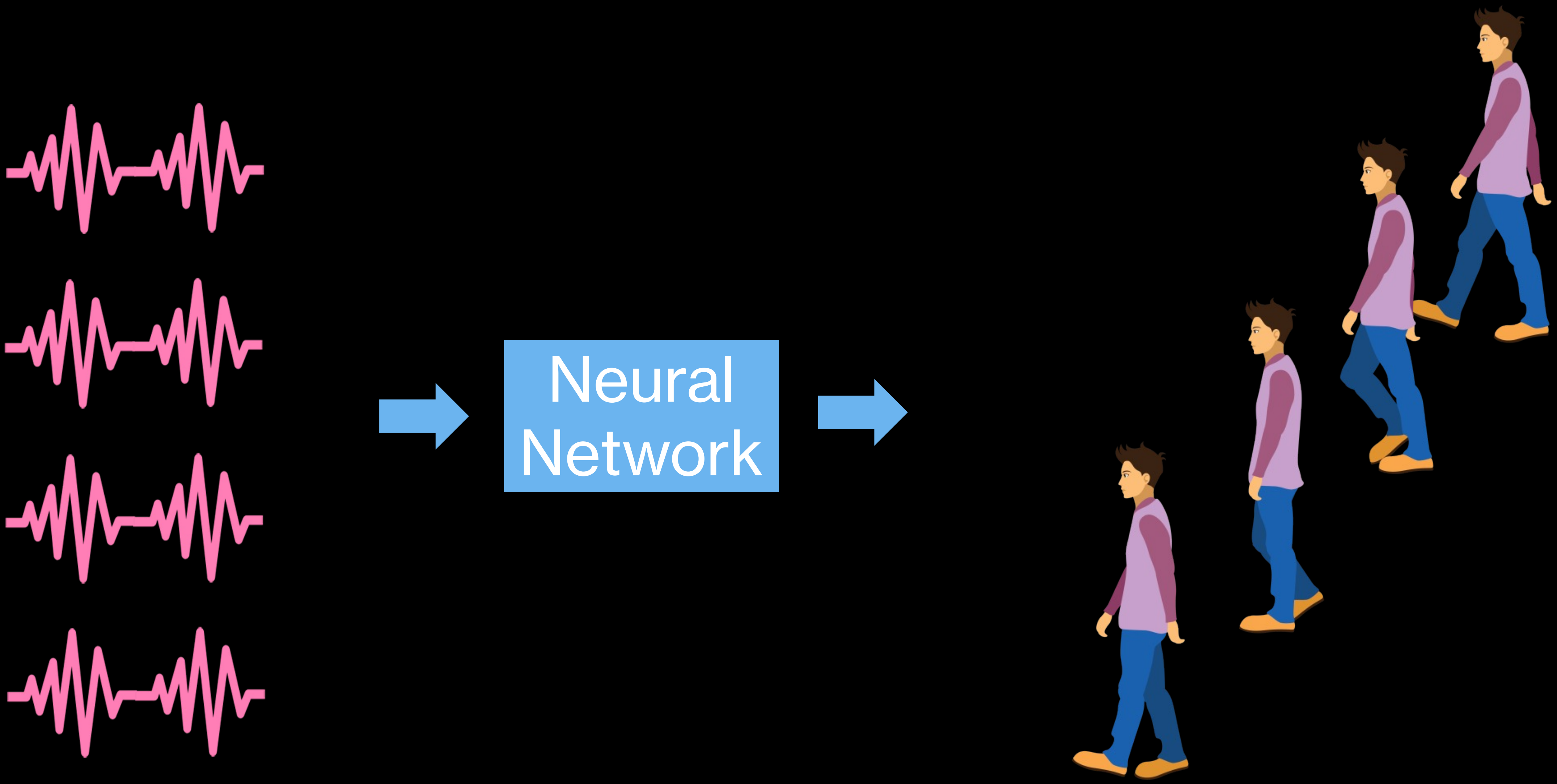


<u>Solution:</u> Use Human Motion Across Time

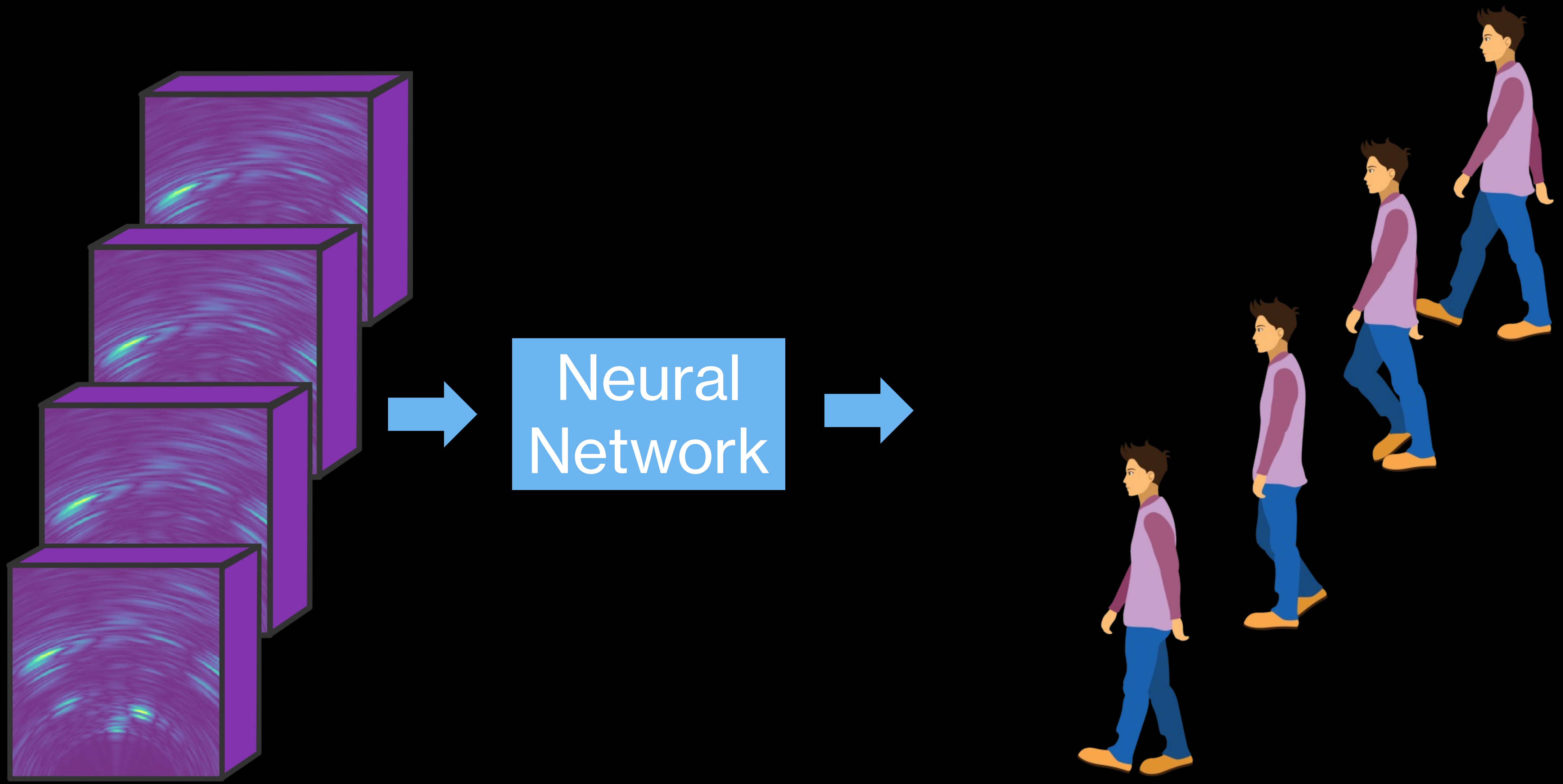


Solution: Use a series of RF snapshots




Neural
Network

Challenge: 4D signals are too large for NN!

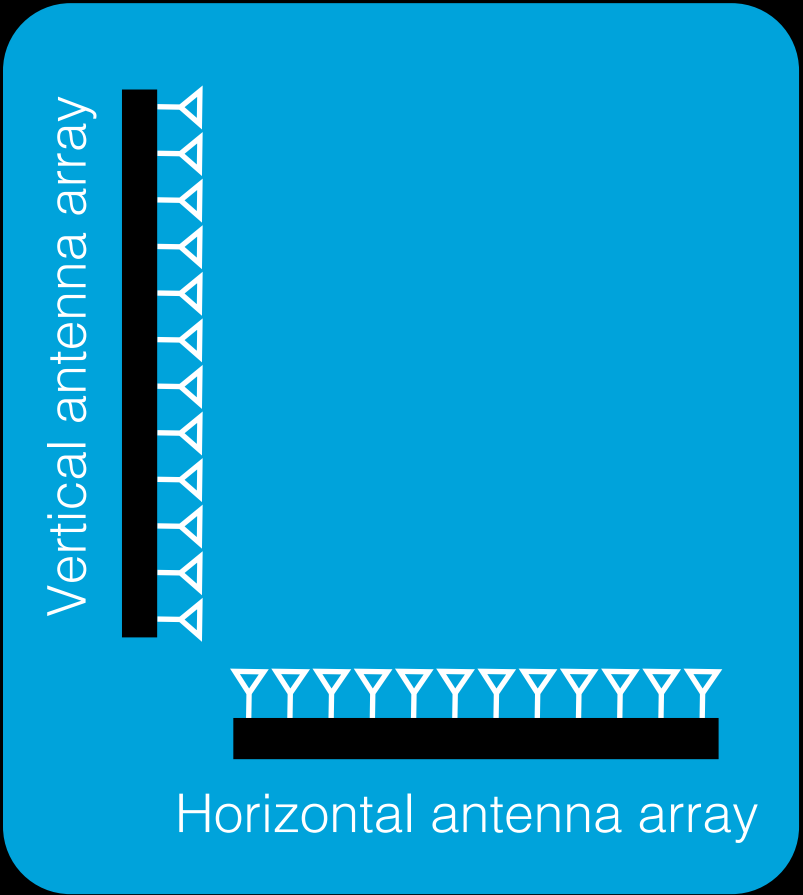


<u>Solution:</u> Neural Network Decomposition

Idea: leverage the sparsity of RF signals to decompose computation.

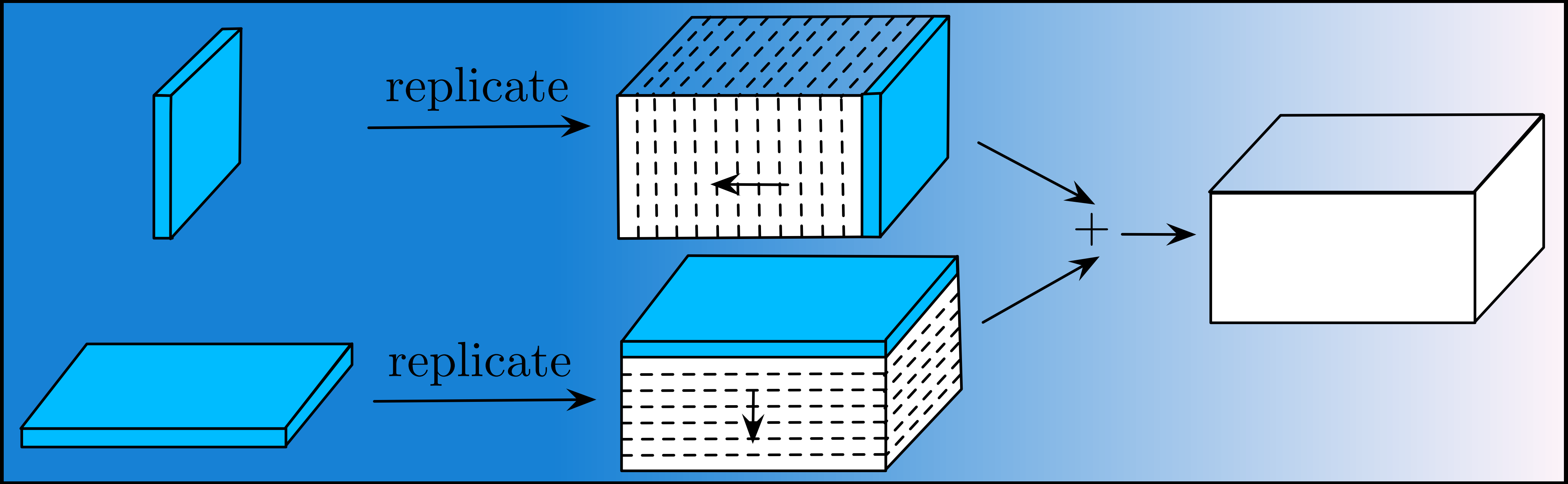
Theorem (informal): A 4D CNN for L-shaped antenna array is equivalent to a combination of two 3D

220x speedup during training

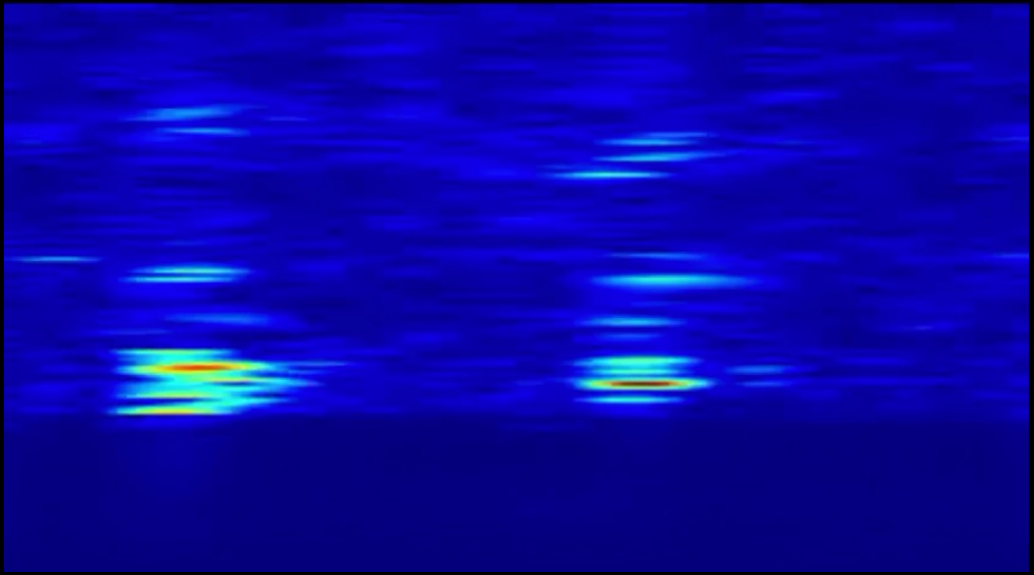


<u>Solution:</u> Neural Network Decomposition

Theorem (informal): An RF-based 4D Neural Networks is equivalent to a combination of two 3D Neural Networks

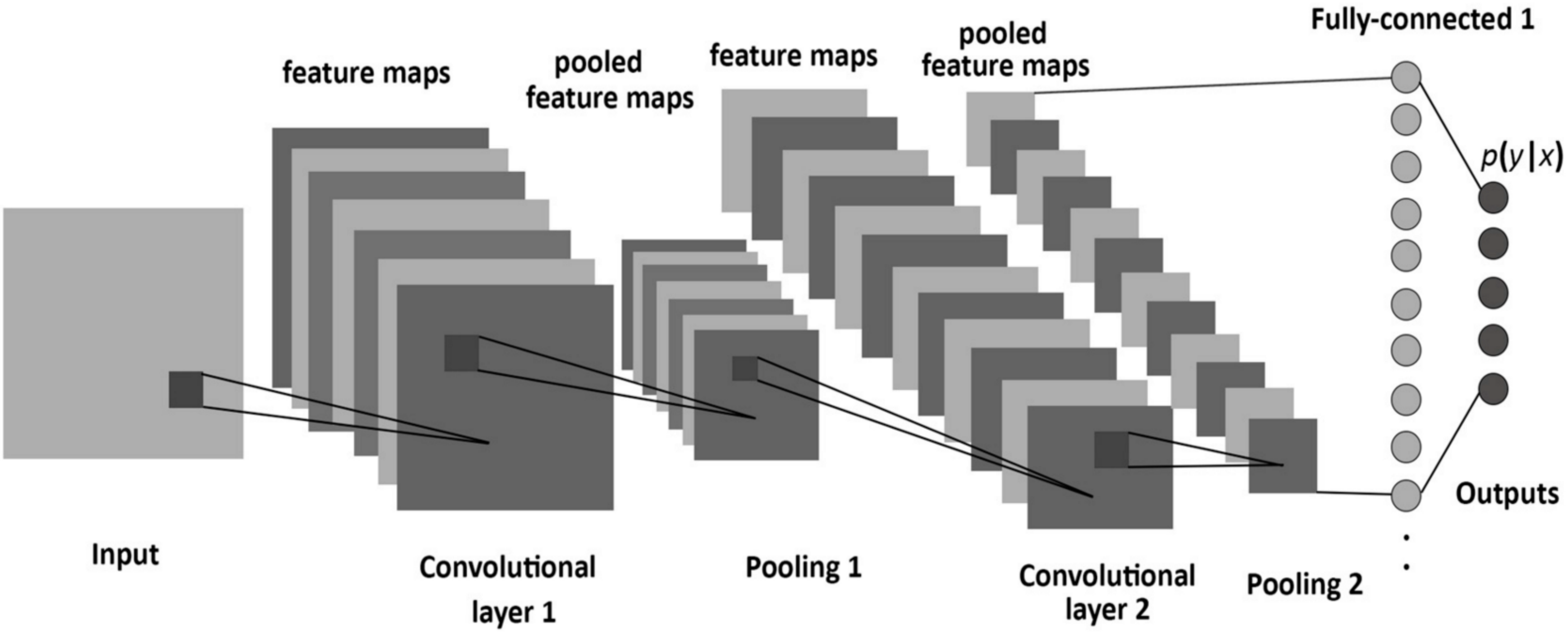


How about Multipath?

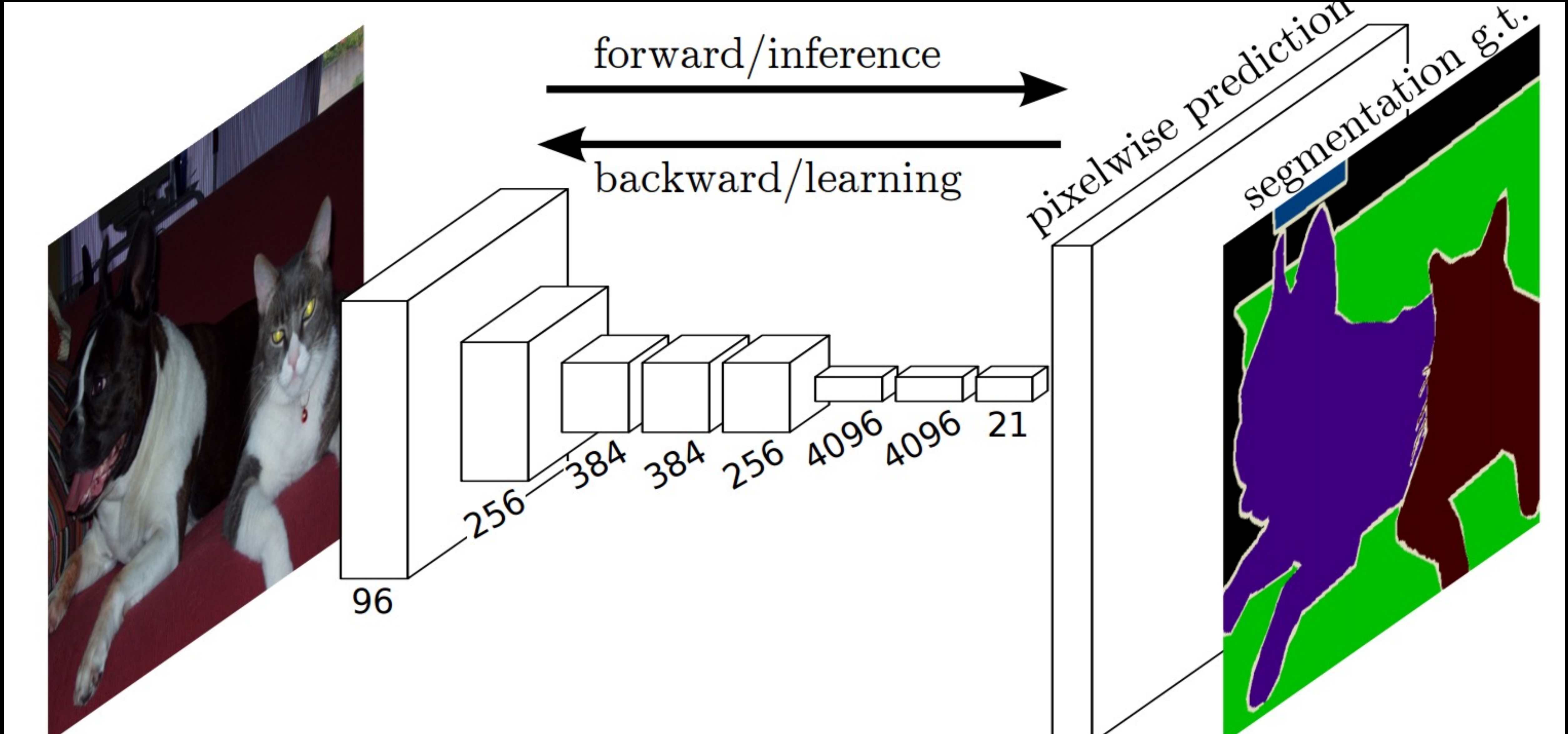


RF-Pose

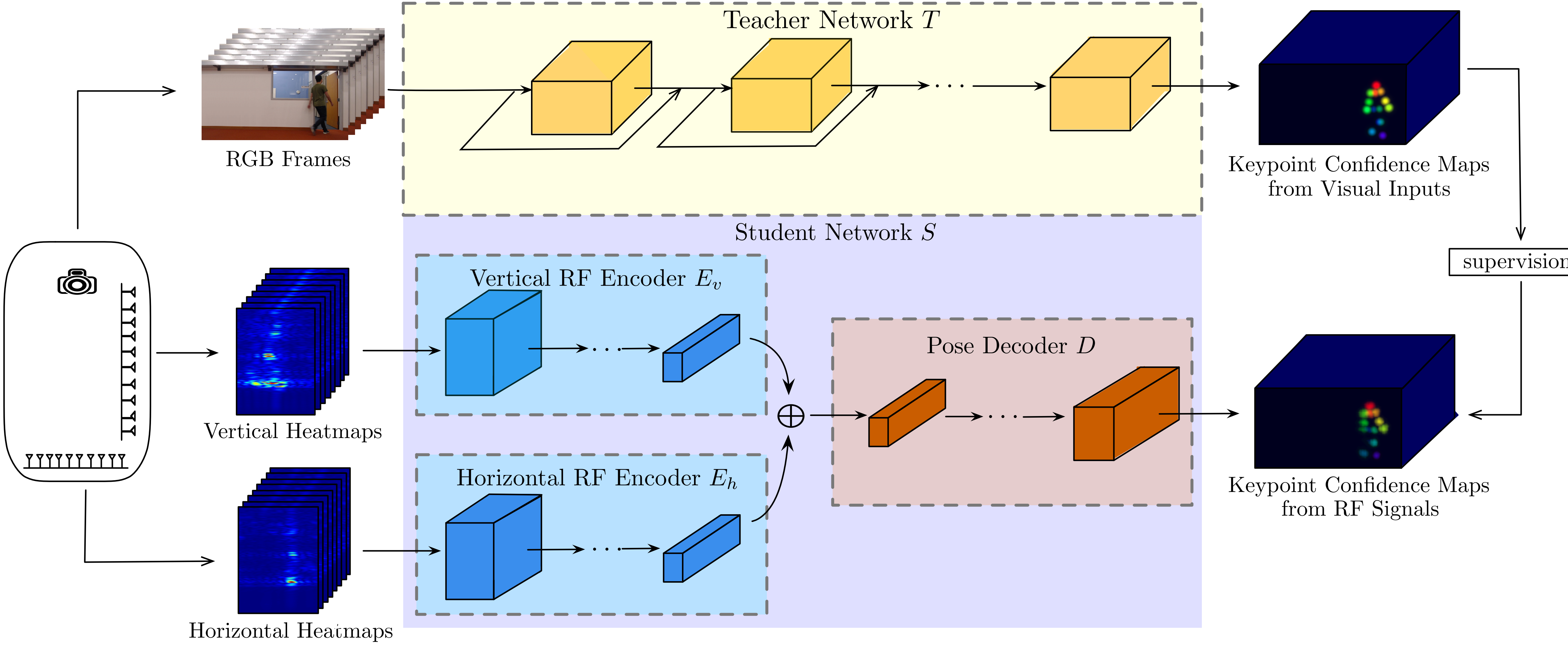
Convolution Neural Network (CNN):



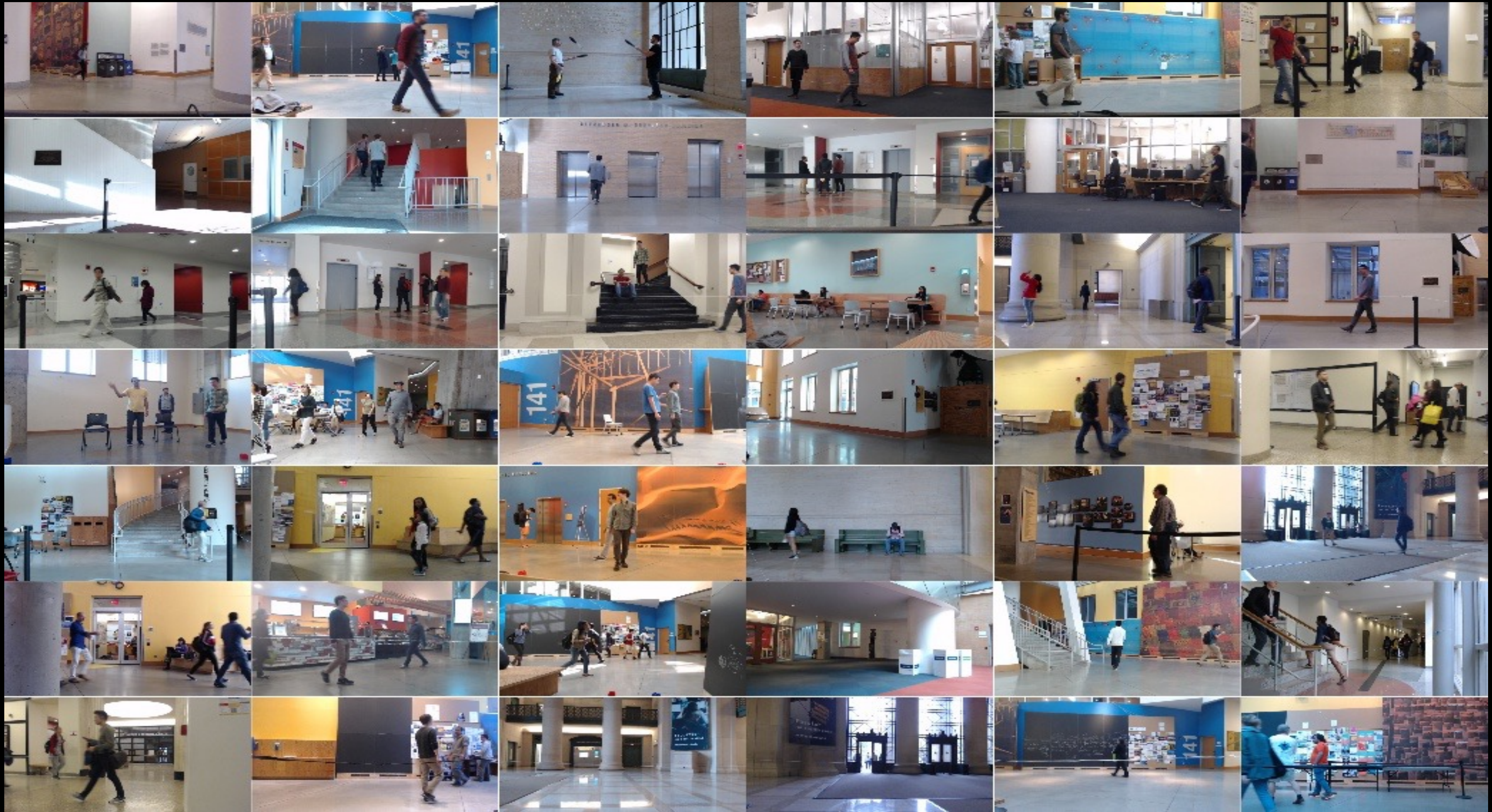
Fully Convolution Network (FCN):



Model Architecture of RF-Pose:



Dataset: 50 hours, 50 locations, daily activities



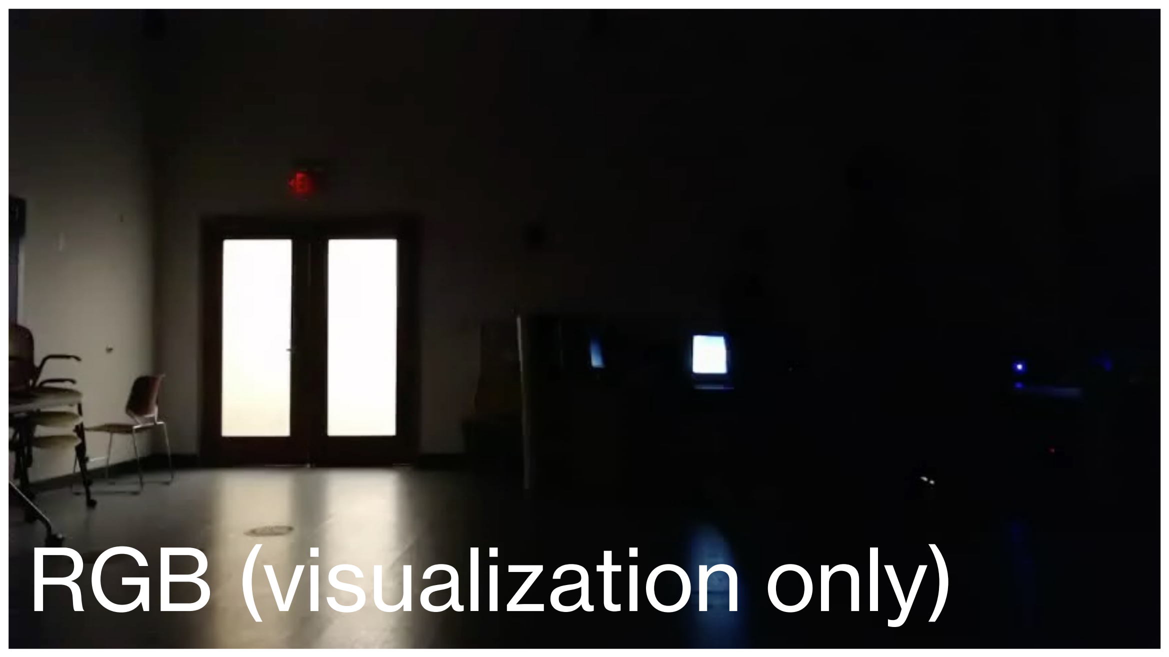
Through-wall poses
using **only** RF

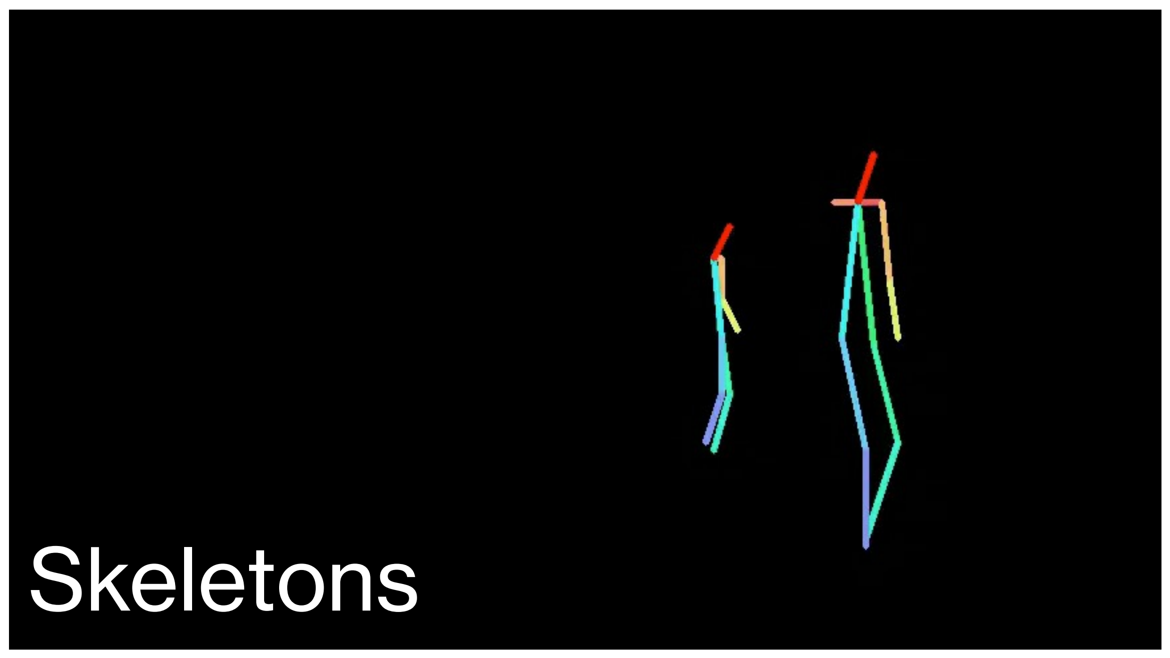

RGB (visualization only)

Skeletons

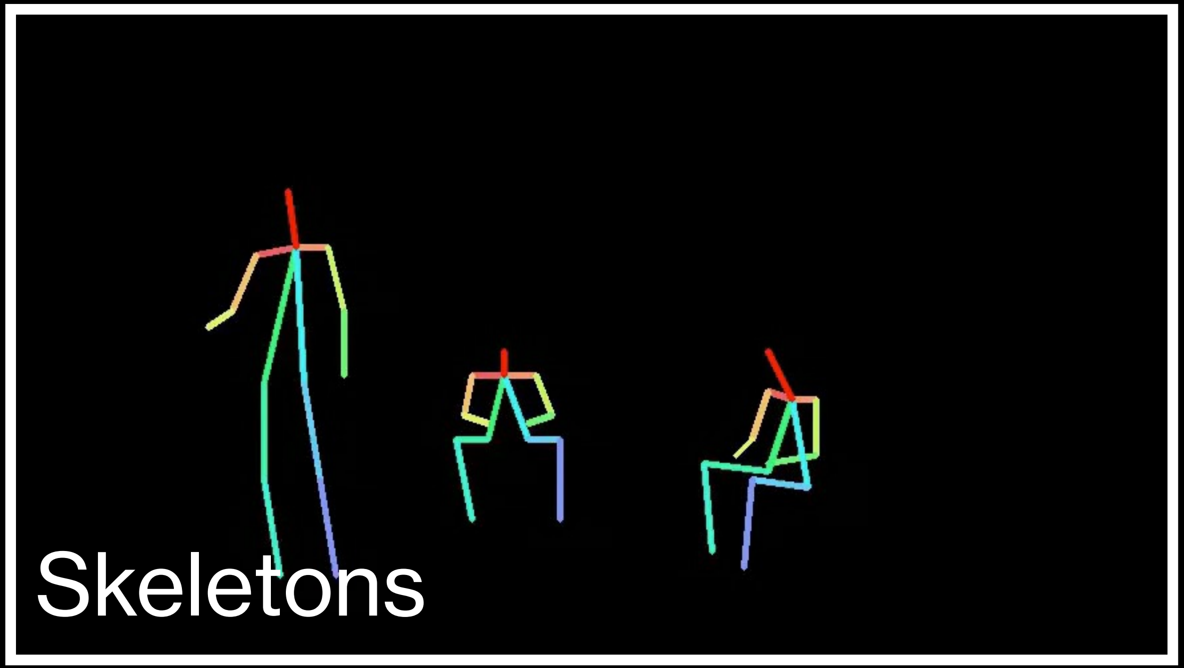
Confidence Maps

RF-Pose also works
in bad lighting



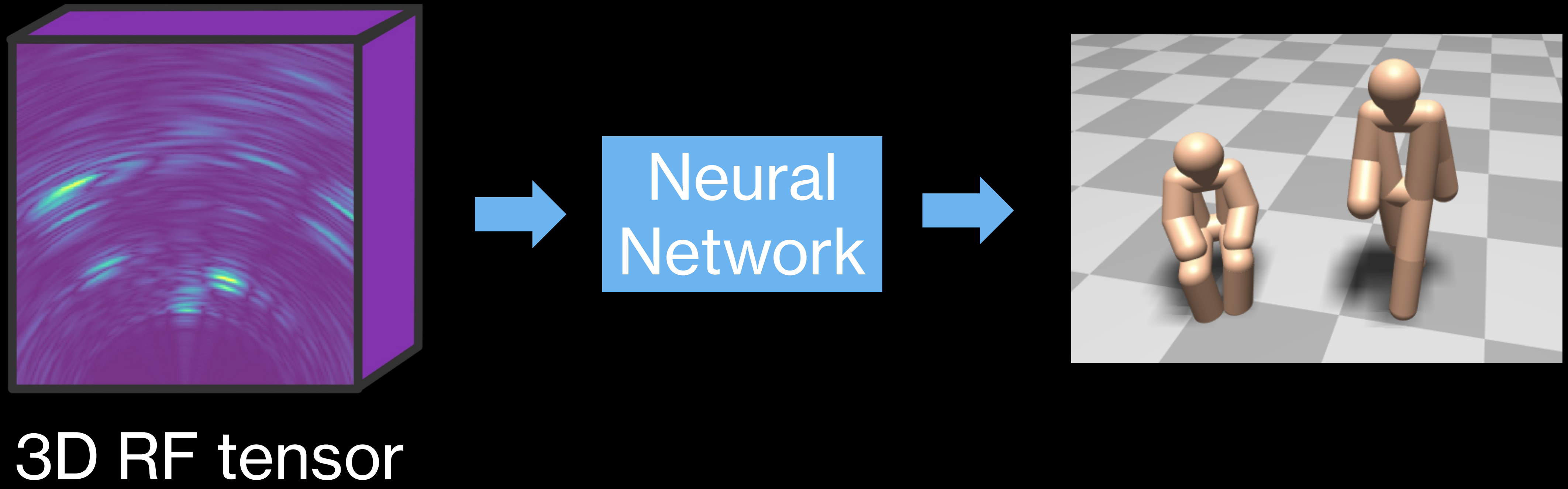


RF-Pose works with different environment and daily activities

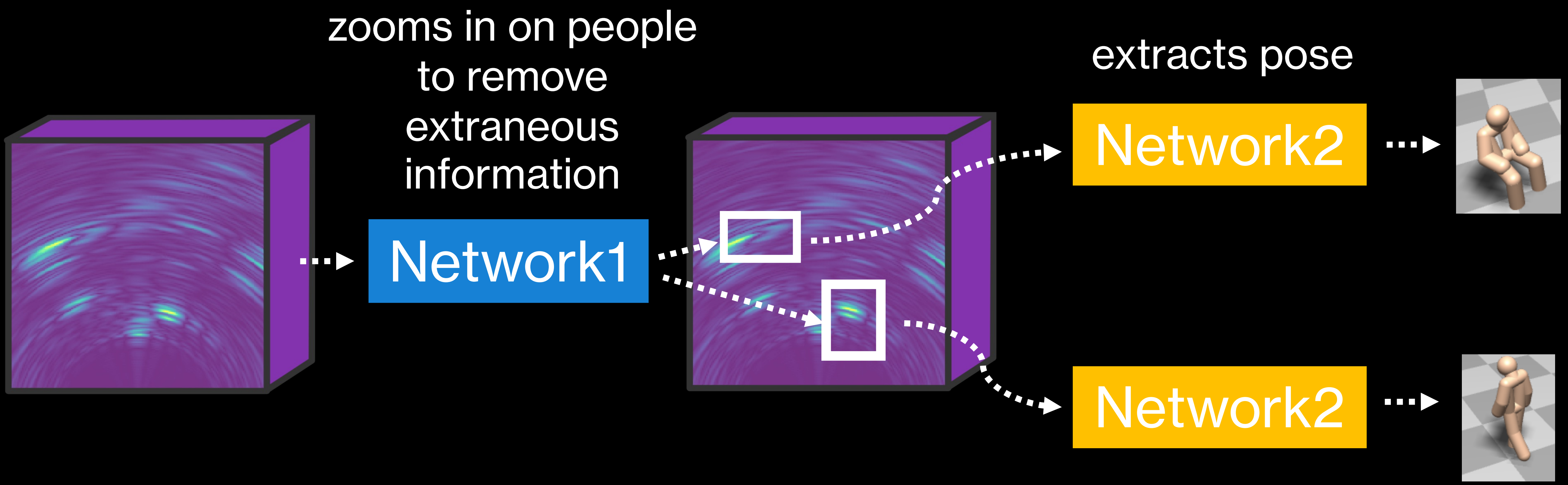


RF-Pose3D

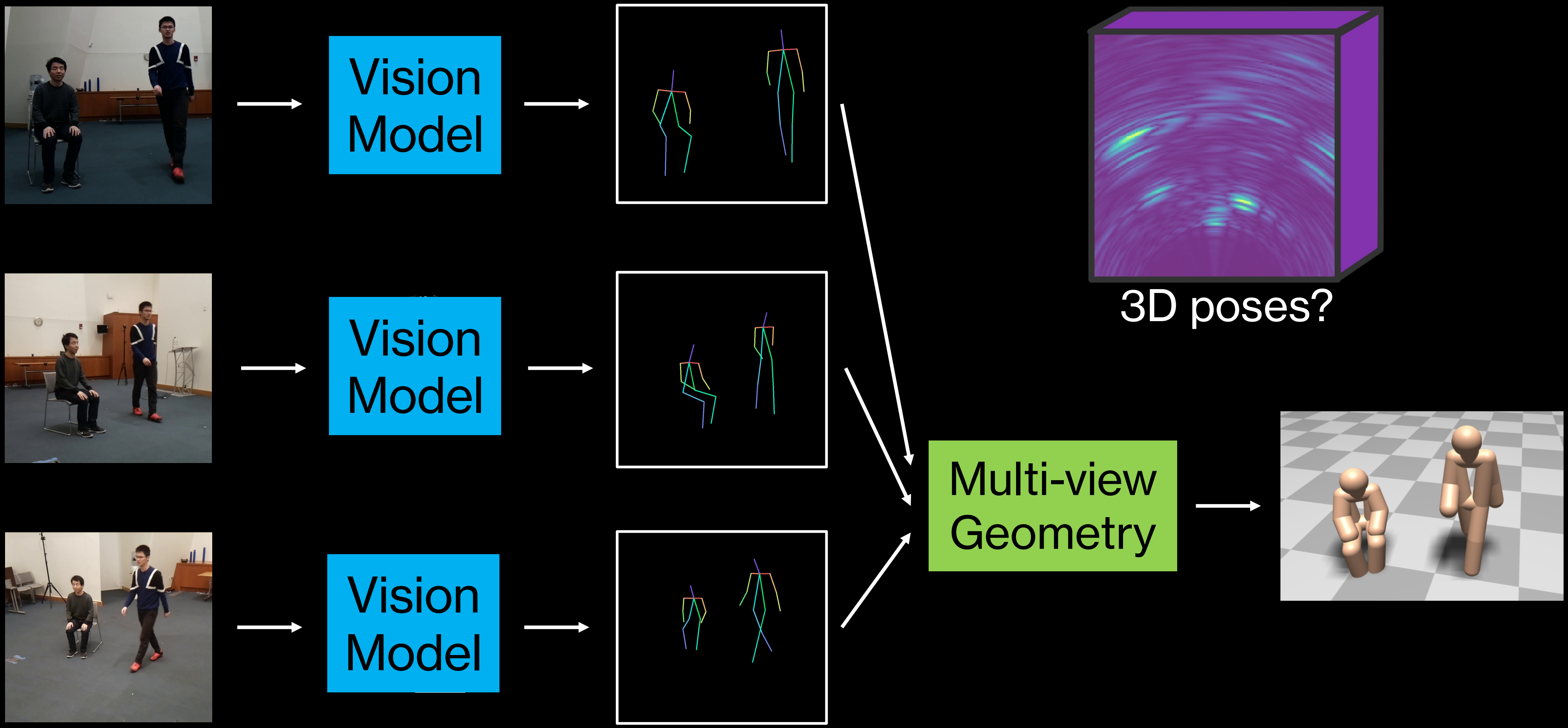
Model Design: Complexity

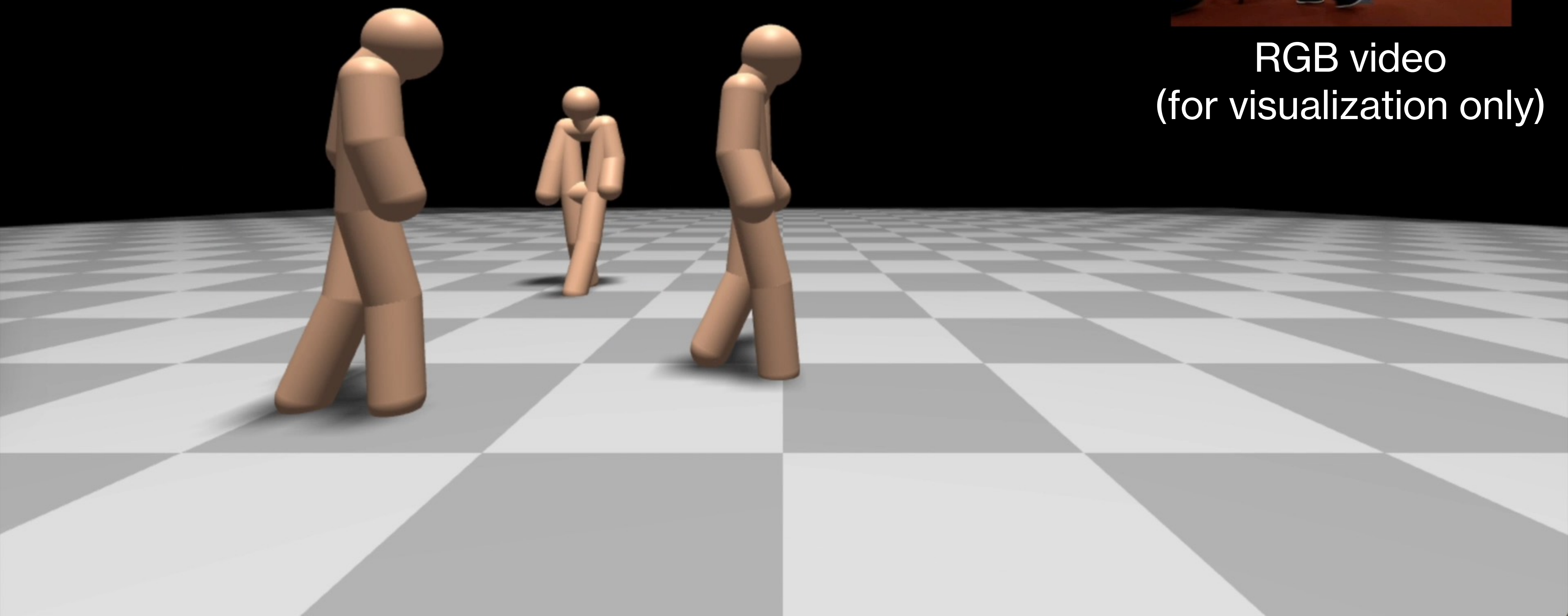


Solution: Two-Stage Model for Task Separation

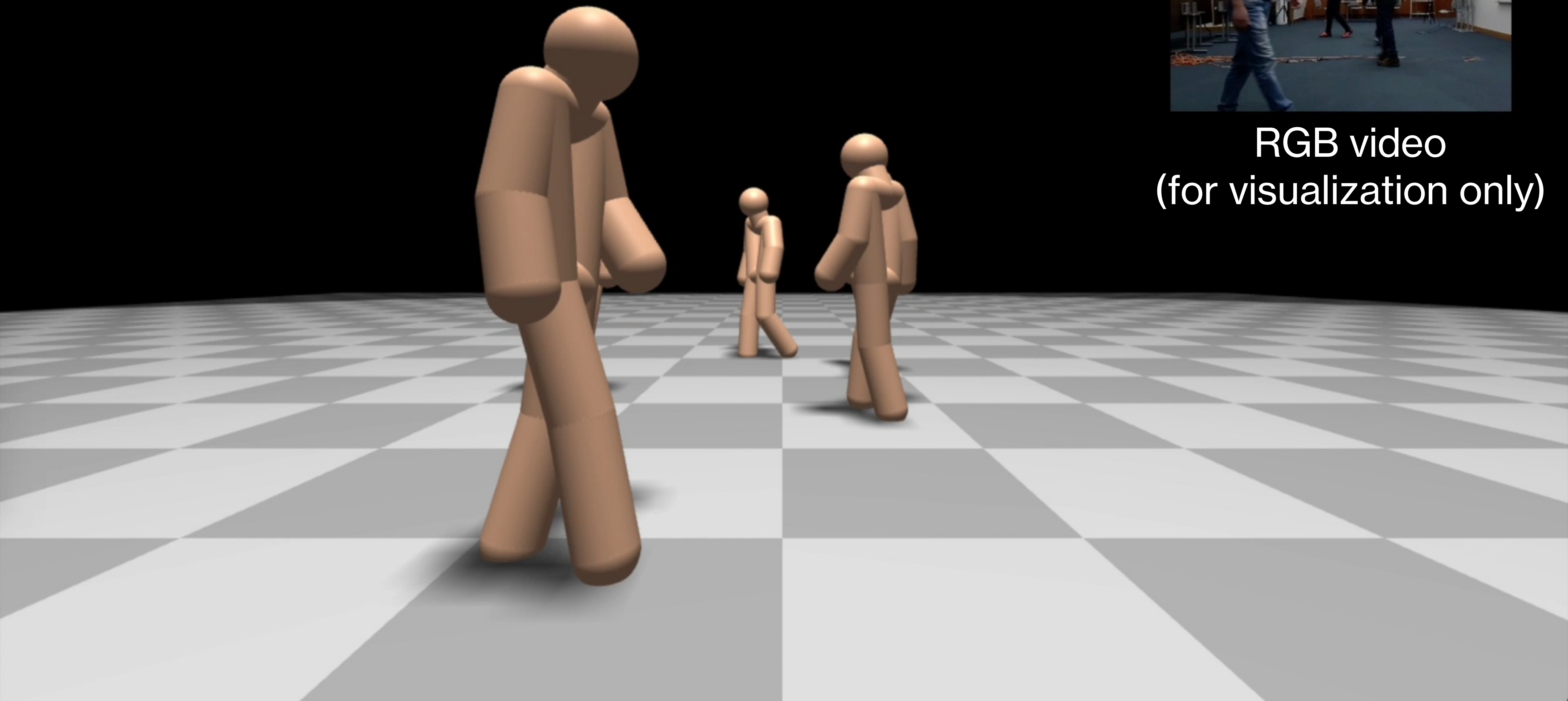


Automatic labeling of 3D poses with cameras





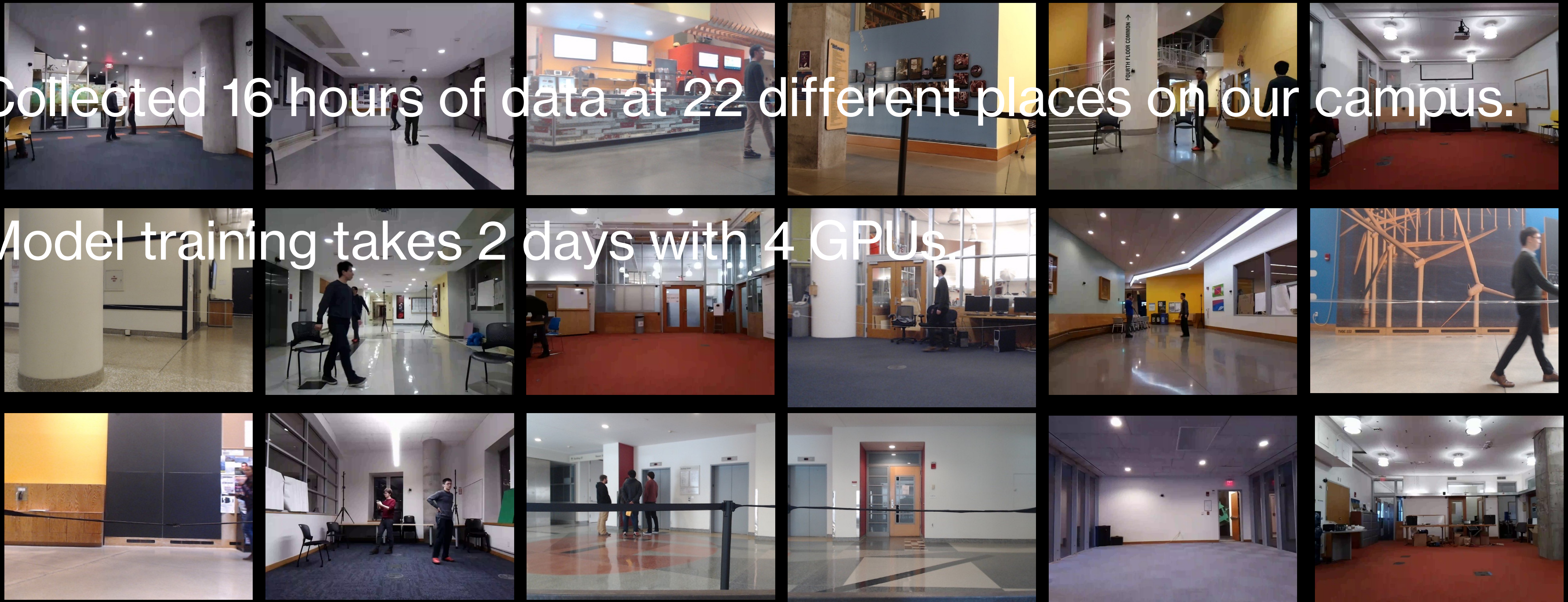
RGB video
(for visualization only)



RGB video
(for visualization only)

Implementation

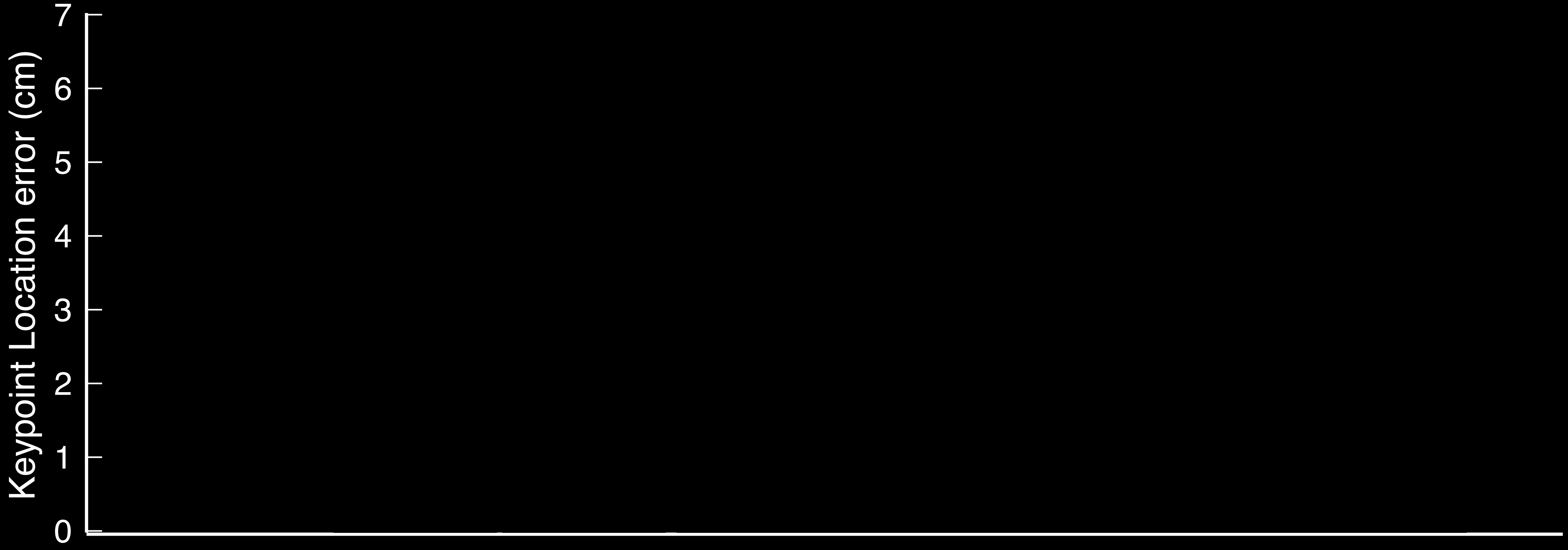
- HW is similar to past work; uses an FMCW radio with antenna arrays
- Model implemented with decomposition in PyTorch.
- Collected 16 hours of data at 22 different places on our campus.
- Model training takes 2 days with 4 GPUs.



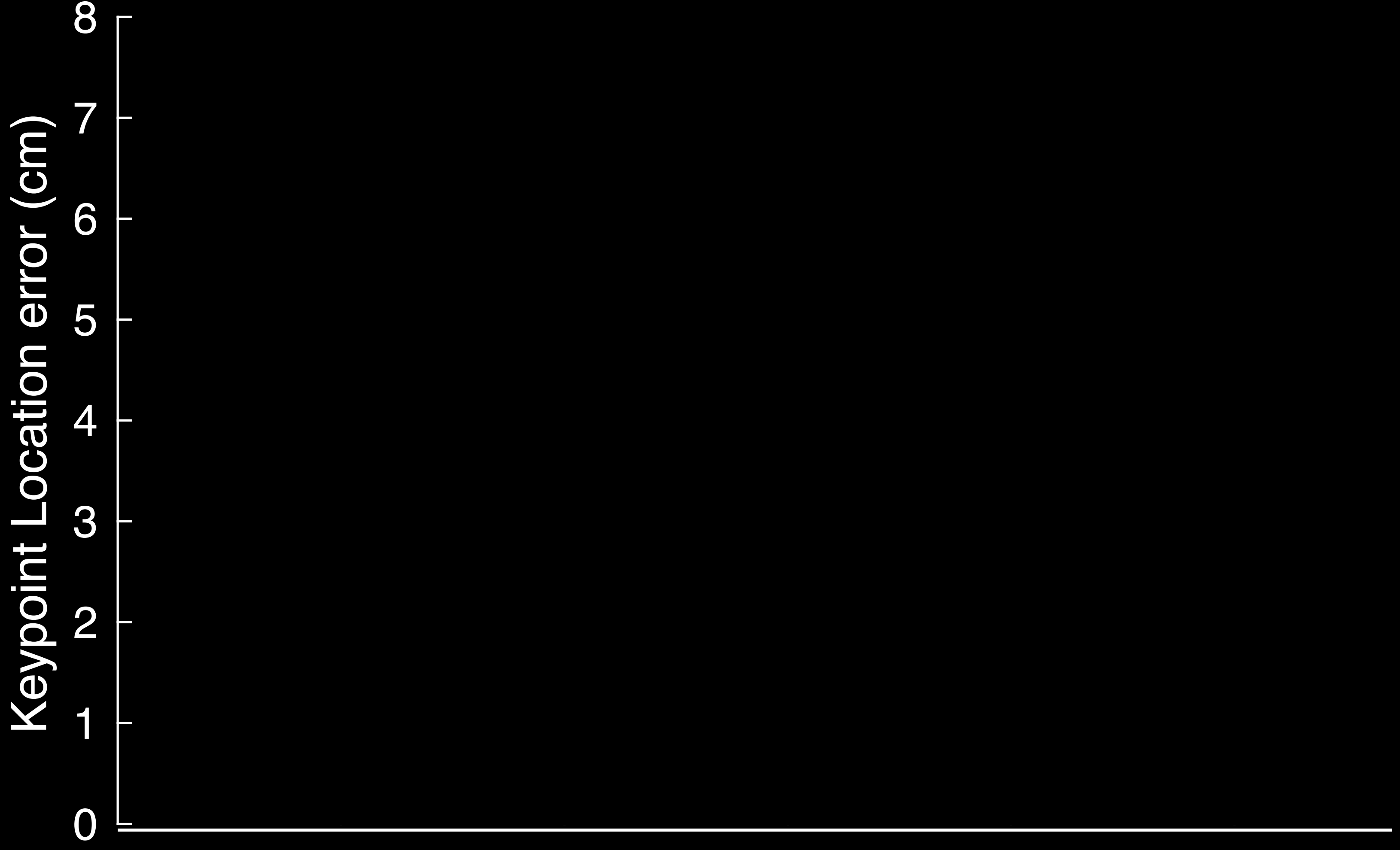
How Accurate is the Skeleton?

Metric: Keypoint localization error

Skeleton Accuracy for Different Keypoints



Skeleton Accuracy vs. Number of People



Objectives of This Module

Learn how foundational sensing technologies can be used to extract diverse and meaningful insights

1. What are important application areas of Mobile and IoT sensing?
2. What are the foundational sensing mechanisms and how are they related to localization?
3. What processing algorithms can be used to transform raw sensor data?
4. Example sensing systems/solutions with real-world case studies.

Review for Module 2:

Due Feb 28^{th} , 11:59 pm

iOS Lab 2 is out

- Topic: Gesture detection and breathing monitoring with acoustic signals
- Due: Fri Mar 1st, 11:59 pm

Course Project

Next Lecture

- Time: Mon Feb 26th
- Topic: Connectivity