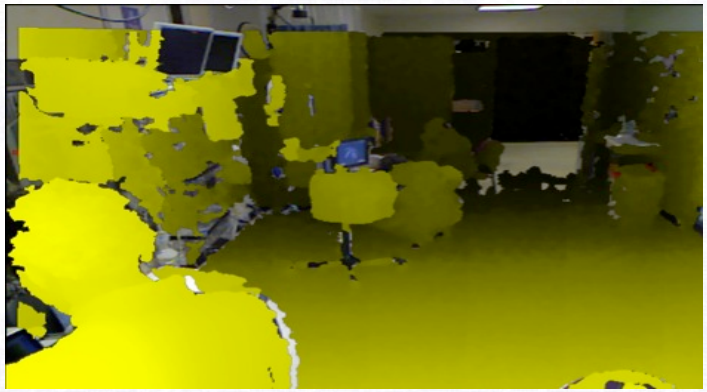
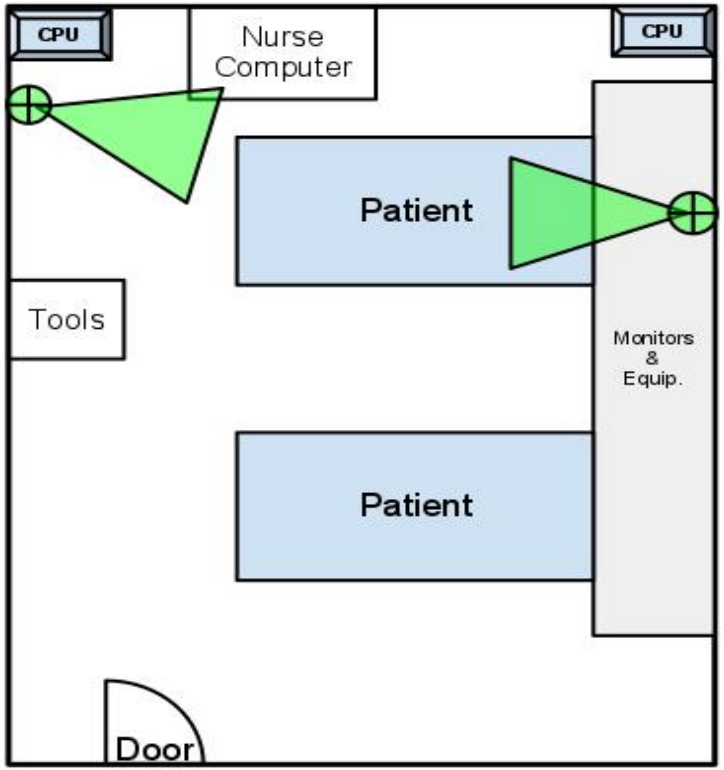
AWARE@ICU

Colin Lea

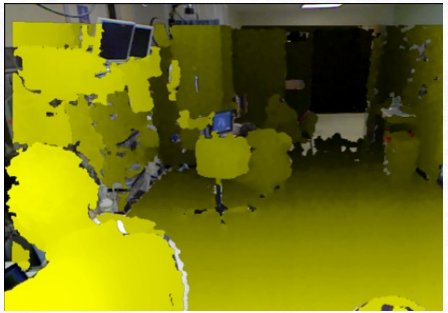
Automated **W**orkflow and **A**ctivity **Re**cognition
CIS2 Checkpoint Presentation #10

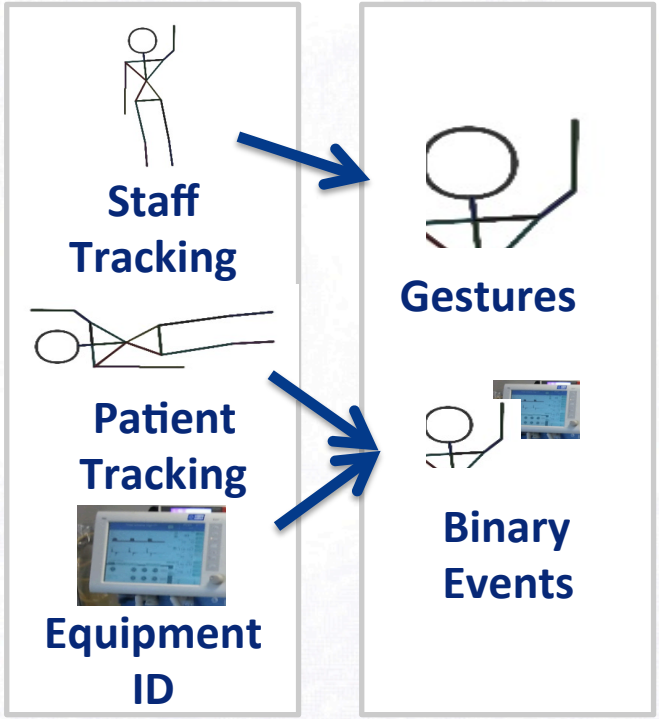
overview

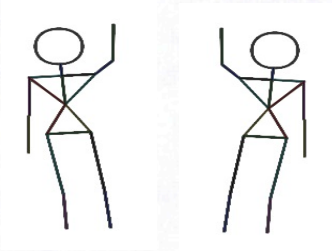



CPU
Nurse Computer
CPU
Patient
Tools
Monitors & Equip.
Patient
Door

pipeline







i.e. "Greeting"

Raw Signal **Derived Signals** **Action**

milestones

Milestone	Mid Feb	End Feb	Early March	Mid March	End March	Early April	Mid April	End April	Early May
Recorder									
Data									
S-Gestures									
S-Locations									
S-Skeleton									
S-Scene									
A-Windows									
A-Recog									
Docs									

Today

Recorder

Task	Difficulty	Notes
General Purpose Recorder	Easy	
Face removal	Medium	
Blackout stretcher	Easy	
Anonymization	Medium	

Experimentation

Task	Difficulty	Notes
Get IRB Approval	Easy	Waiting
Record preliminary data	Easy	Non-publishable
Ensure de-identification	Easy	
Hand annotate activities	Easy	4/4/2012
Additional recording	Easy	

Derived Data

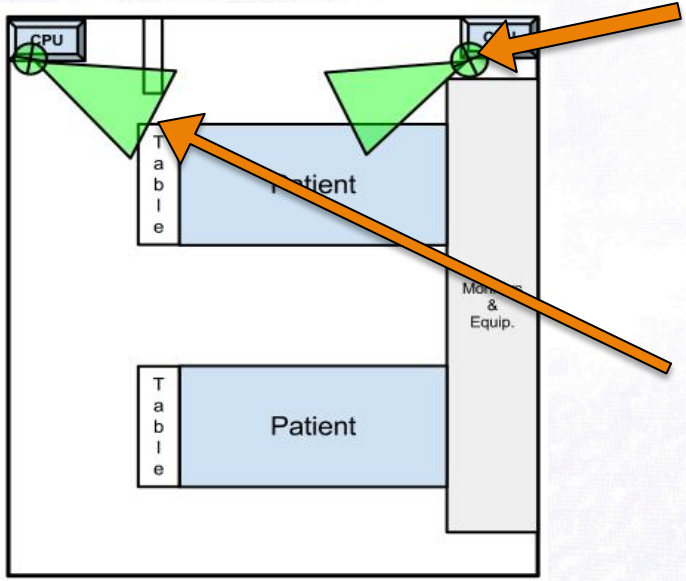
Task	Difficulty	Notes
Location tracker	Medium	
Face direction tracker	Easy	CSIRO Software
Staff body pose estimation	Hard	In Process
Patient tracker	Medium	
Equipment Identification	Hard	

Activity recognition

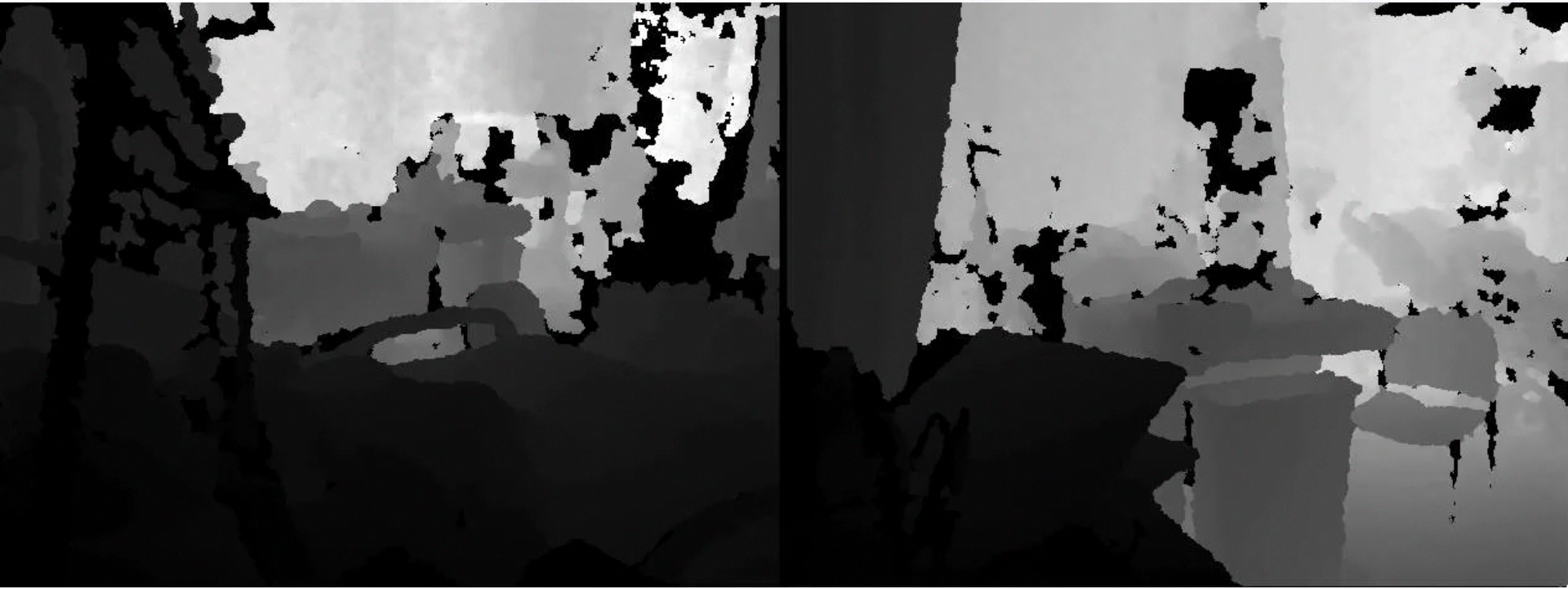
Task	Difficulty	Notes
Gesture recognition	Hard	
Retrospective analysis	Hard	
Patient Tracking	Medium	
Multi-camera models	Medium	
Evaluation	Easy	

data collection

- 3 hours of depth/RGB/skeleton footage
- 2 Kinects
- Pediatric ICU



ICU depth video



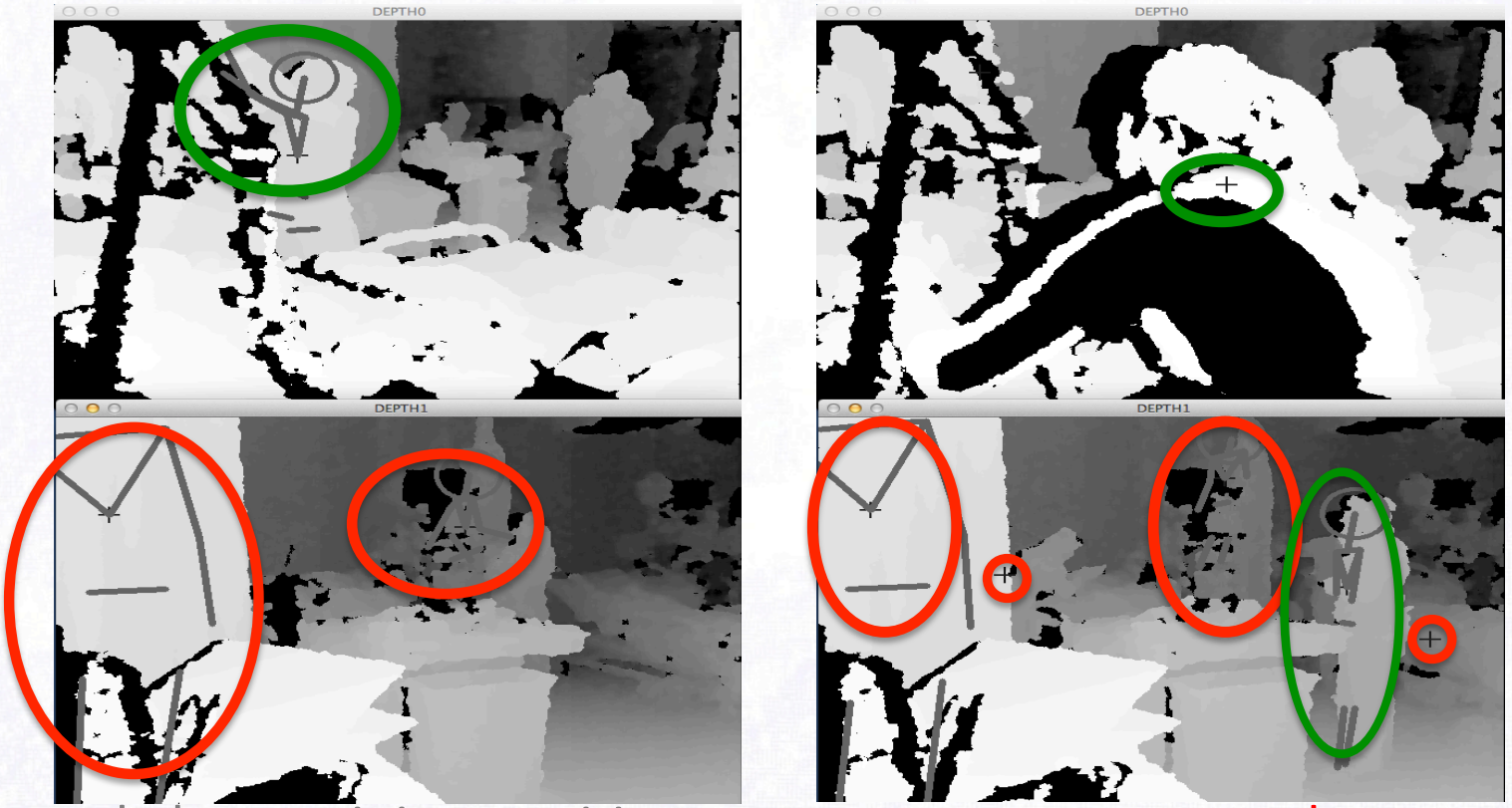
Head view

Foot view

kinect skeletons

Camera 1:

Camera 2:

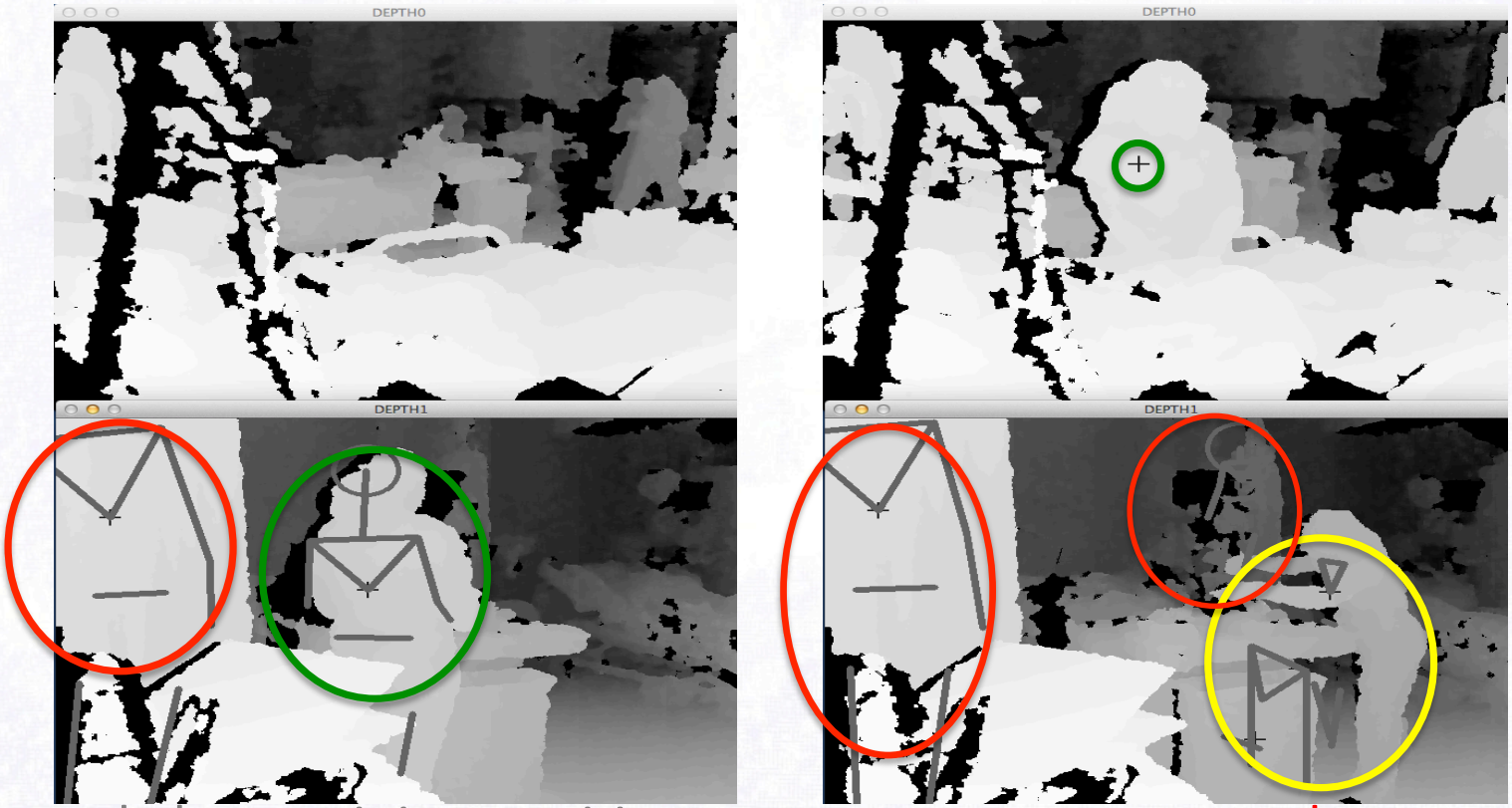


(+) = location :: skeleton = joints positions :: green = correct :: red = wrong

kinect skeletons

Camera 1:

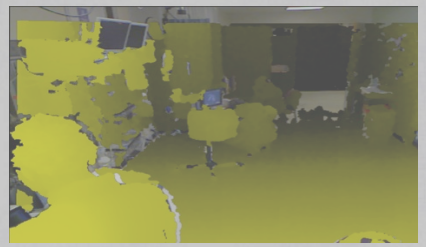
Camera 2:



(+) = location :: skeleton = joints positions :: green = correct :: red = wrong

Can we do better?

activity pipeline



Skeleton Tracking

Gestures

Patient Tracking

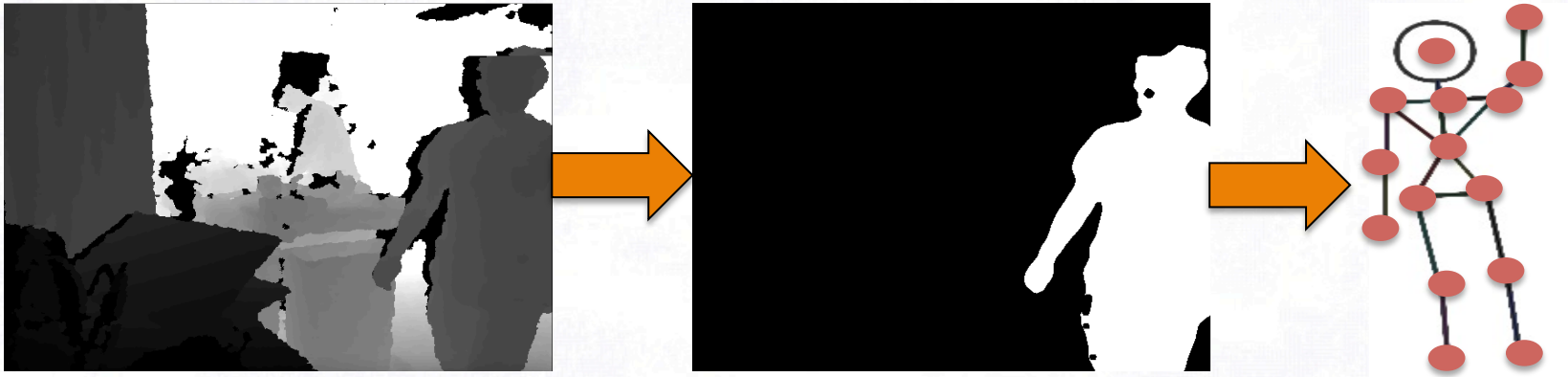
Equipment ID

Binary Events

i.e. "Greeting"

Raw Signal **Derived Signals** **Action**

skeletal tracking



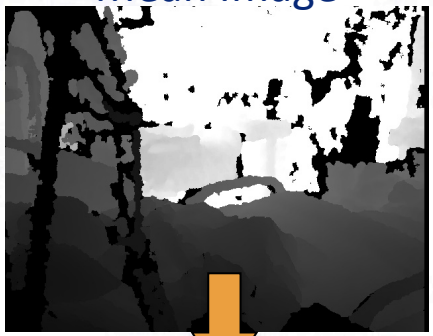
Depth Image

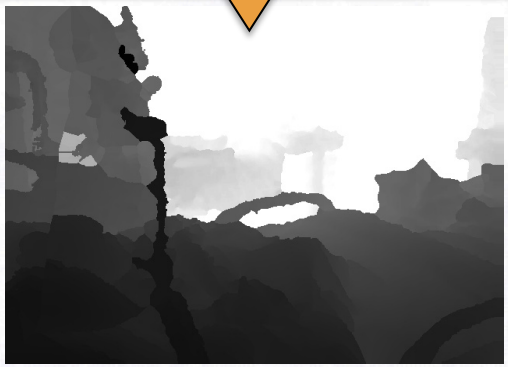
Segmentation

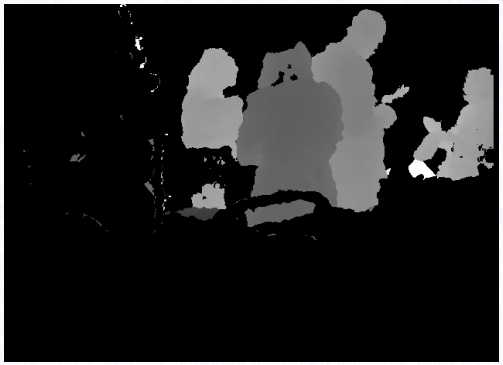
Skeleton

segmentation (step 1)

Mean Image







Depth Image

Background Subtraction

segmentation (step 2)

Problem: people are shades of gray

- Iterative solutions too slow

Solution:

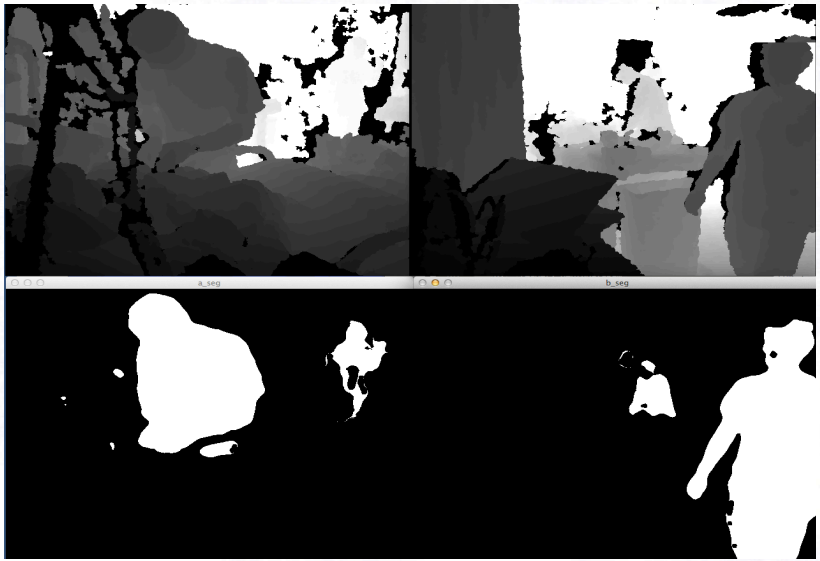
- Method 1: Cluster
- Method 2: Gradient/edge removal

Background Subtraction

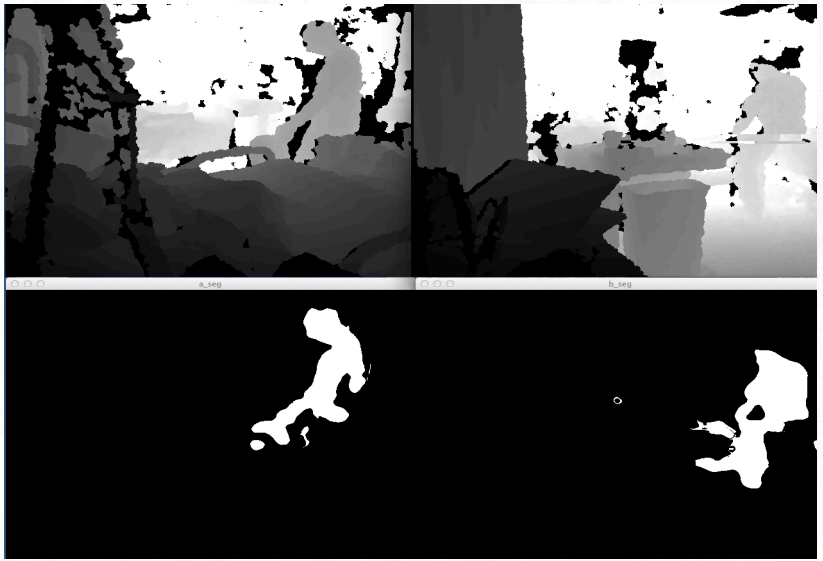
(Method 1) Clustering (DBSCAN)

(Method 2) Gradient-based

segmentation

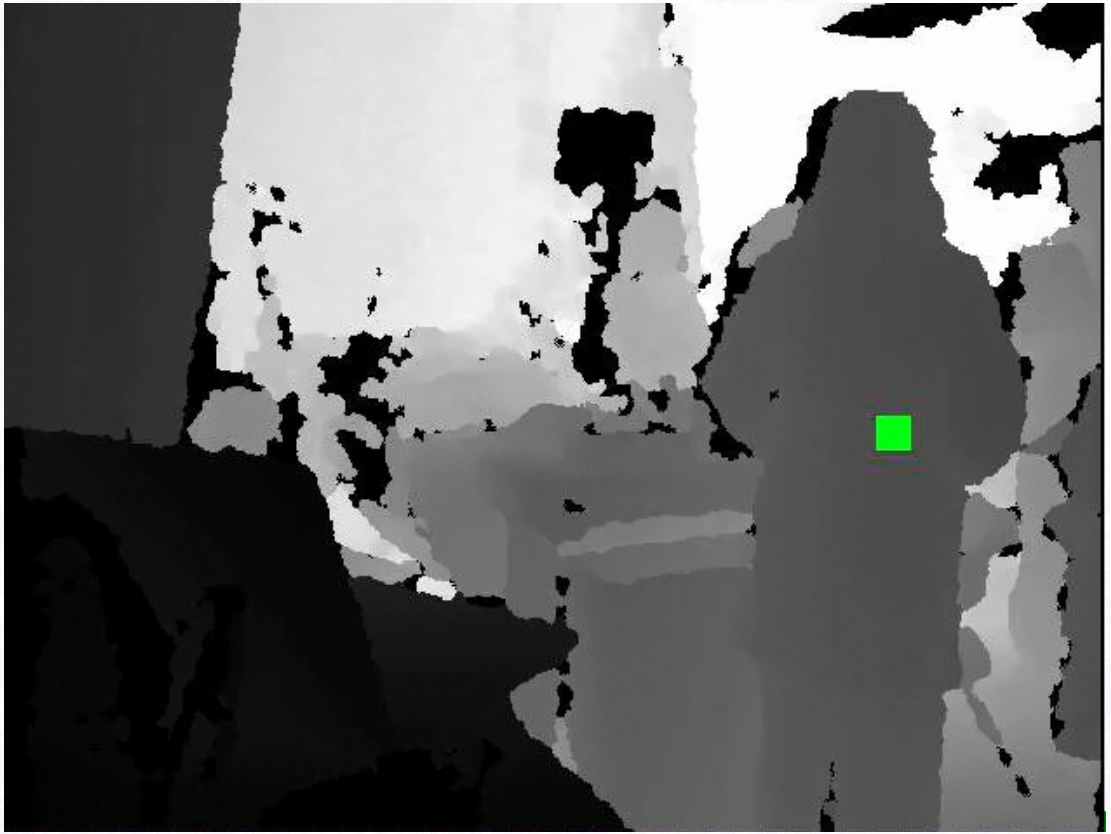


(1a) (1b)



(2a) (2b)

tracking



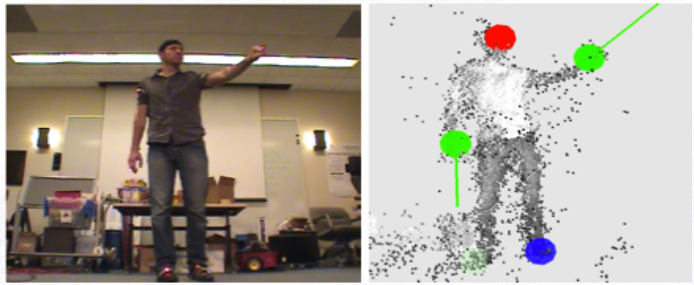
skeletal tracking (in progress)

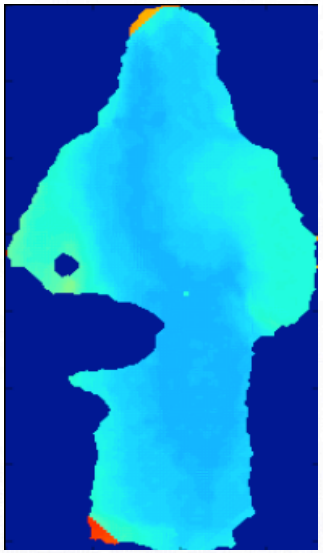
Find extremities (ie. arms, head)

- Floyd-Warshall Algorithm (like Dijkstra's)
- Think flow along surface!

Parts-based model

- All people have similar structure
- Initialize with head + general direction (SVD)





Flow from centroid

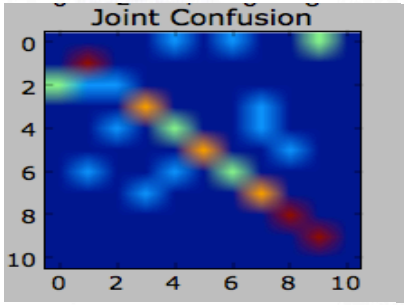
C Plagemann, V Ganapathi, D Koller, S Thrun. "Real-time Identification and Localization of Body Parts from Depth Images." ICRA 2010.

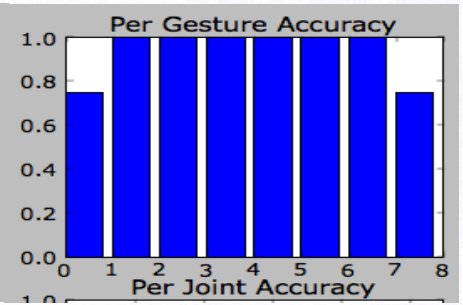
gesture models

Action Recognition

Staff gestures

- Determine sub-actions
  - i.e. Inserting tube
  - Giving medicine



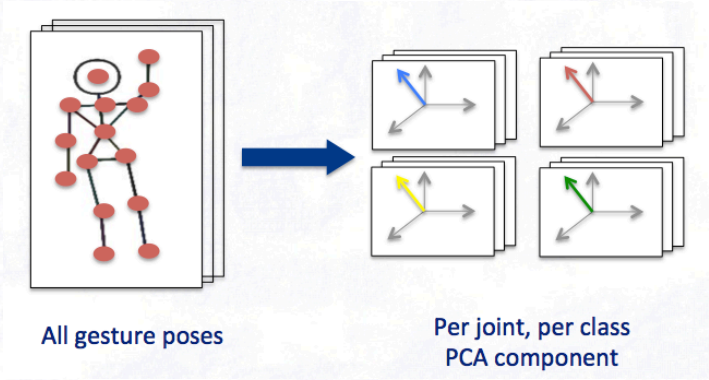


Recent results

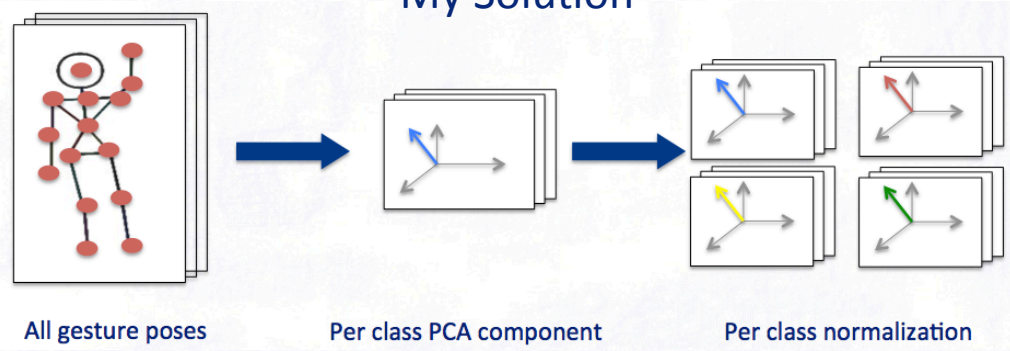
Algorithm

1. Calculate PCA per-joint, per-class
   - 1 [x,,y,z] basis * 13 joints * n classes
2. Correlate input w/ each known gesture's PCA basis
3. Find similarity between test and training (per joint, per class)
4. Vote (Optional: Add prior)

dimensionality reduction



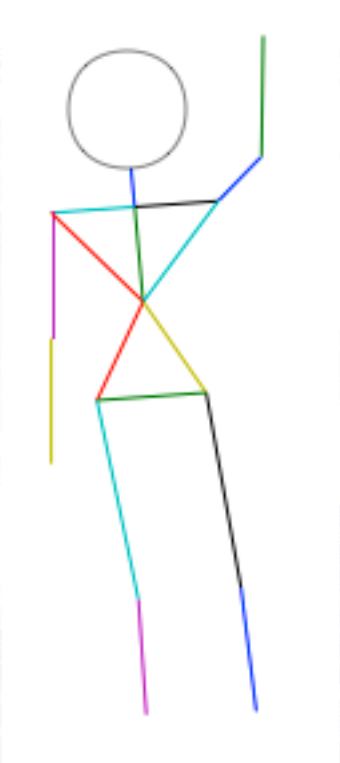
My Solution



[Bigdelou '12]

experiment

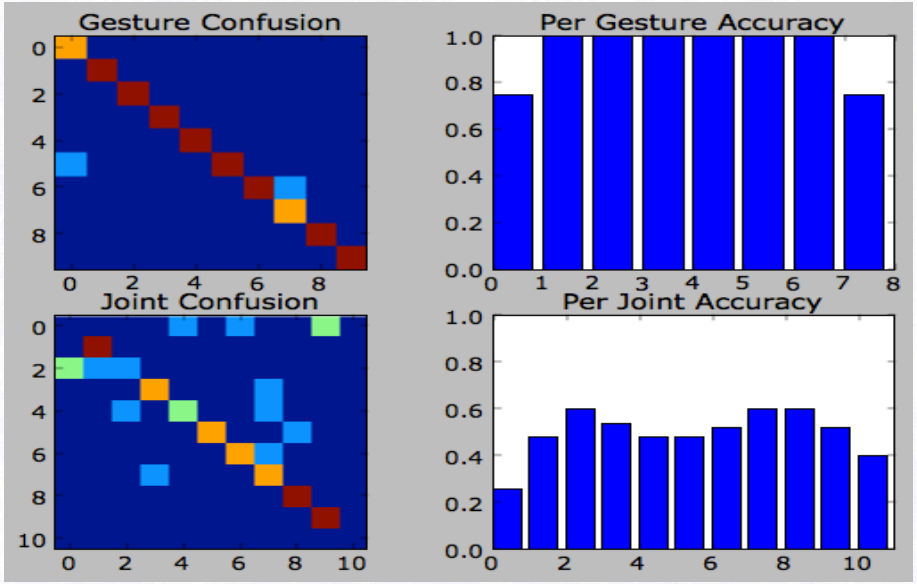
10 gestures
5 trials/gesture
2 people



Wave
Circle (counter-clockwise)
Circle (clockwise)
Push forward
Push left
Push right
Swoosh right
Reach up
Duck
Kick

results

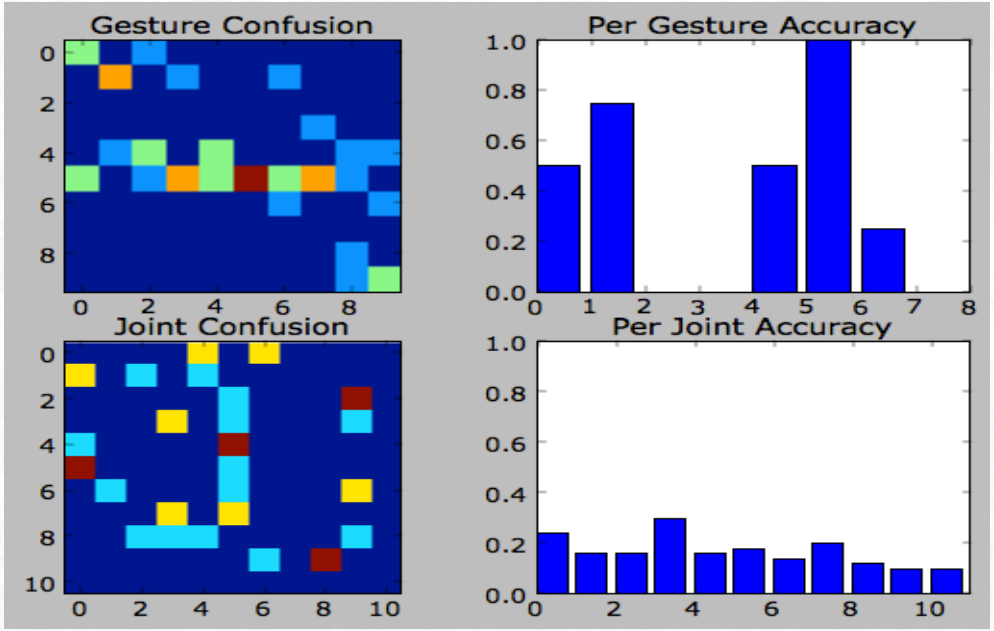
5 samples w/ leave one out



Avg: 94.5% [0.95, 0.95, 0.95, 0.925, 0.95]

results

Train on one person, test on another

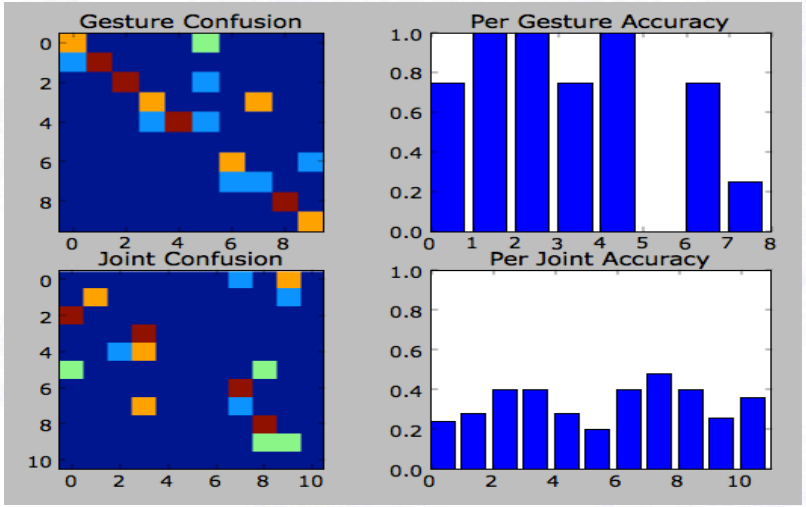


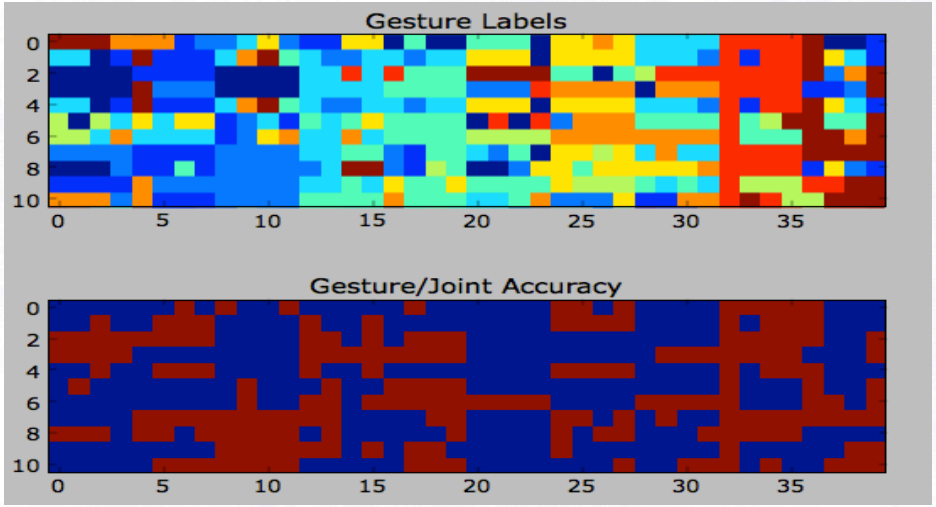
Train A, Test B: **44%** [0.425, 0.475, 0.425, 0.4, 0.475]

Train B, Test A: **61%** [0.575, 0.5, 0.7, 0.65, 0.625]

results

One-shot learning



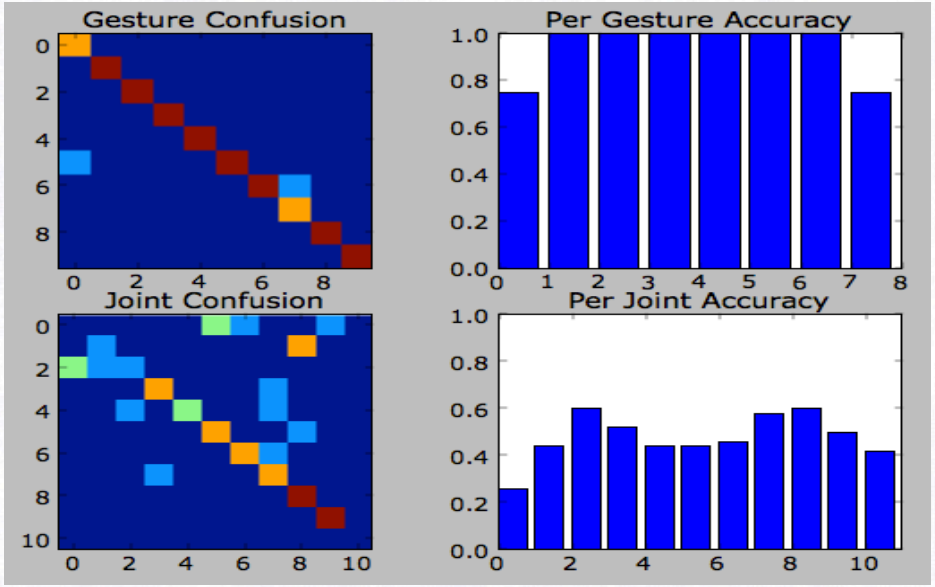


Avg: 77% [0.8, 0.75, 0.75, 0.825, 0.725]

results

Sliding window

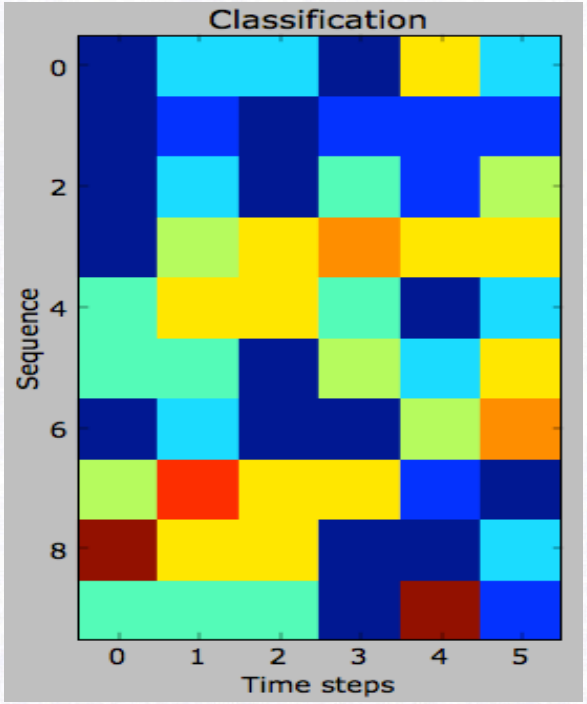
At one time step

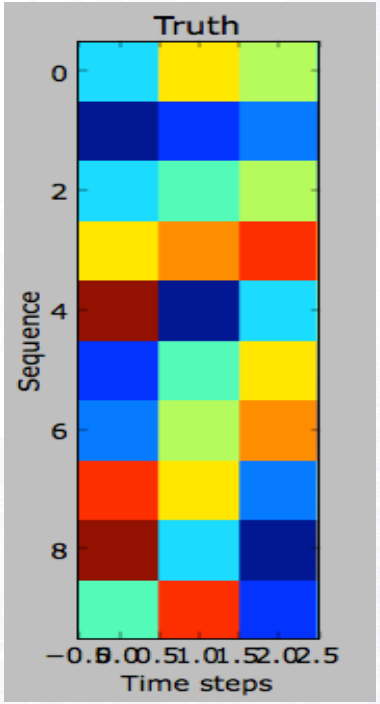


Avg: 95.5% [0.95, 0.95, 0.975, 0.95, 0.95]

results

Sequences





other directions

Action Recognition

Time-series graphical models

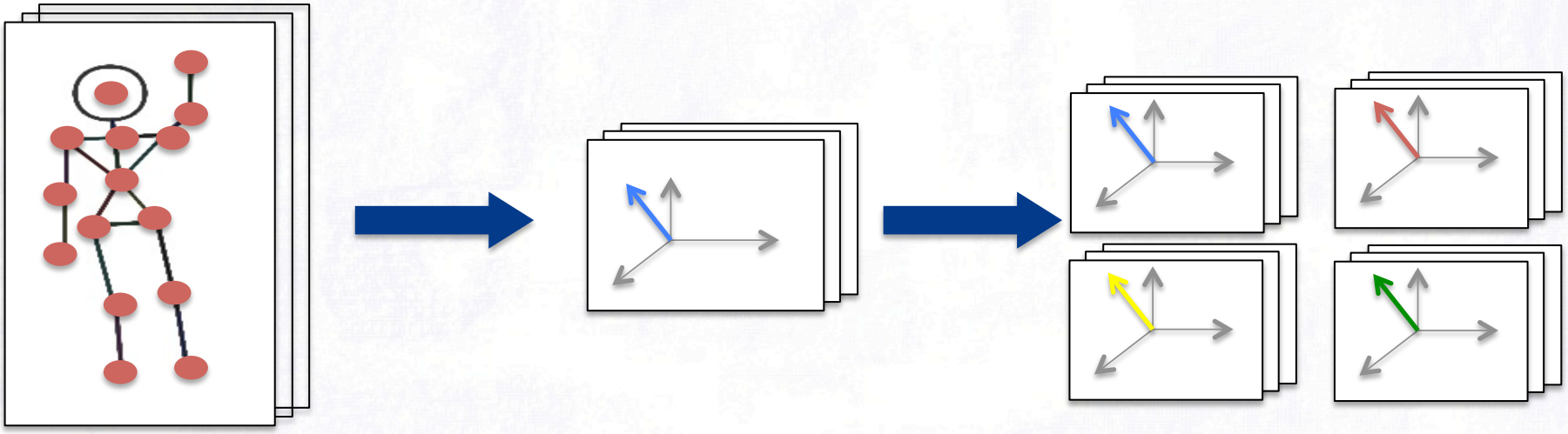
- Switching Linear Dynamical System (code available)
- CRFs

Multiple Instance Learning

- Split into bags -> Iterative SVM

Questions?

dimensionality reduction

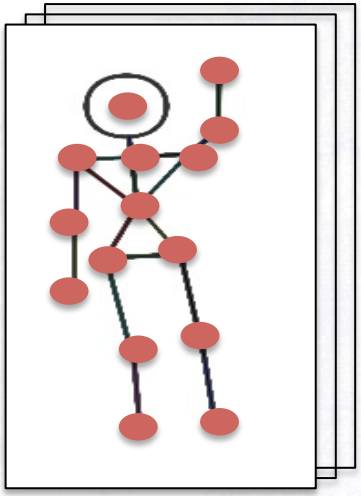


All gesture poses

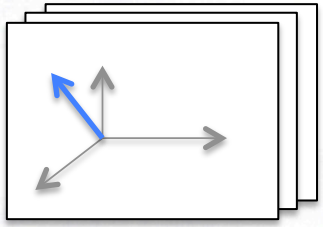
Per class PCA component

Per class normalization

dimensionality reduction



All gesture poses



Per joint, per class
PCA component