People Watching:
Human Actions as a Cue for Single-View Geometry

David Fouhey, Vincent Delaitre,
Abhinav Gupta, Alexei Efros, Ivan Laptev, Josef Sivic
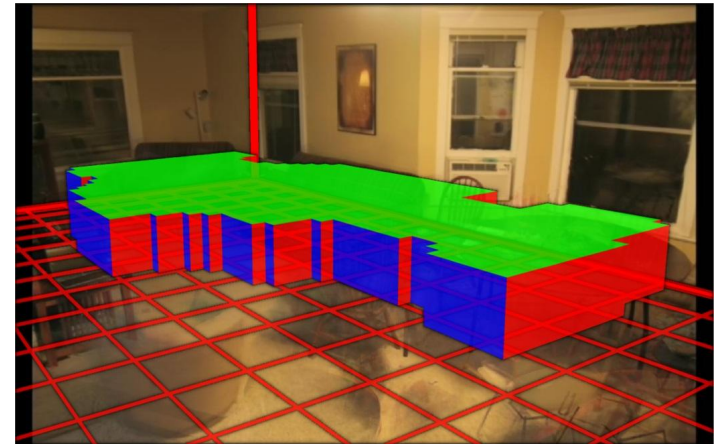
Presented by Ashwini Venkatesh

Where are the people?

People – Cues not Clutter

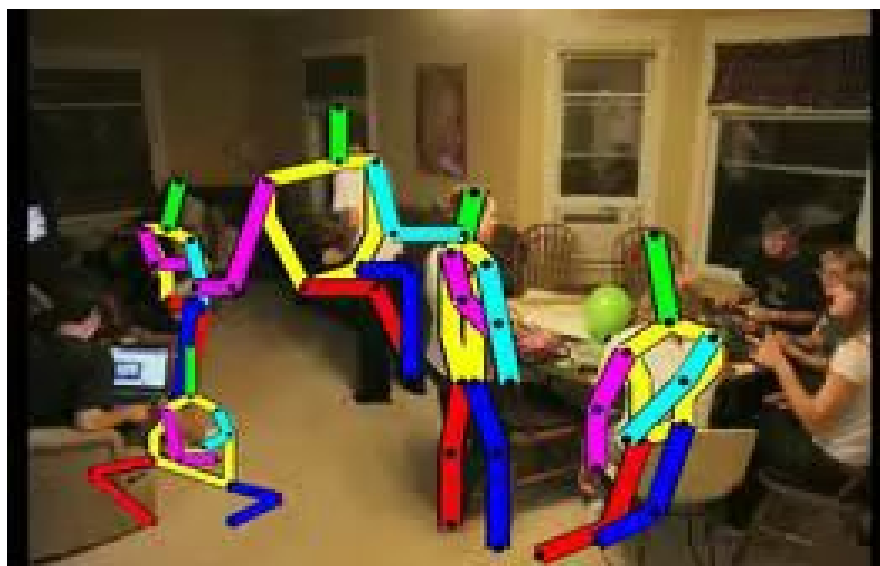# Goal – Inverse Problem

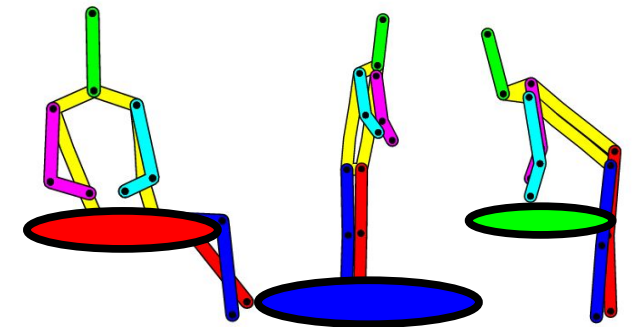Clutter ➞ Sensors !

Approach

Timelapse

Pose Detections

Approach

Timelapse → Pose Detections
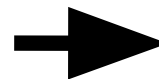
Estimate Functional Regions from Poses

Approach

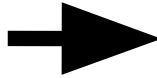Timelapse → Pose Detections → Functional Regions

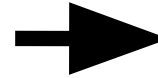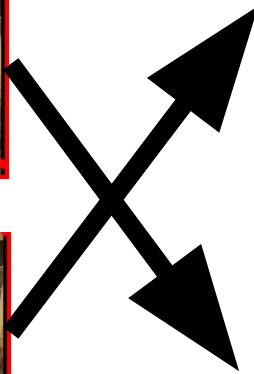# 3D Room Hypotheses From Appearance

Approach

Timelapse → Pose Detections → Functional Regions
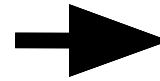
#1

#49

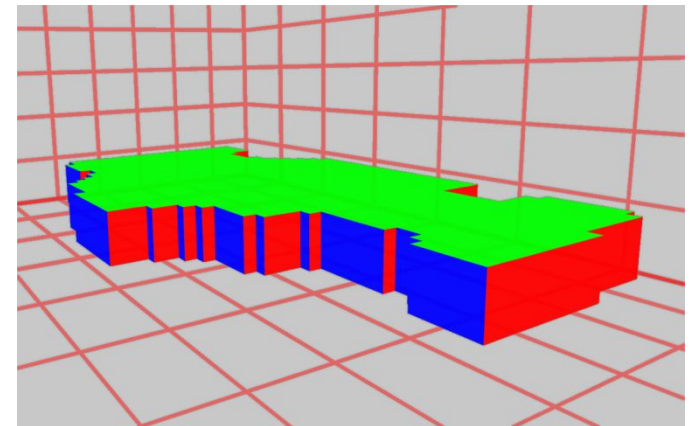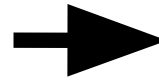Score 3D Room Hypotheses With Appearances + Affordances
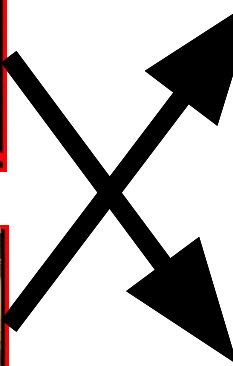
Approach

Timelapse

Pose Detections

Functional Regions
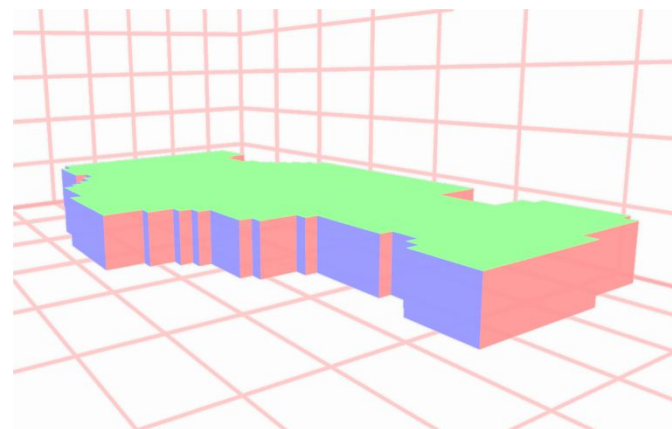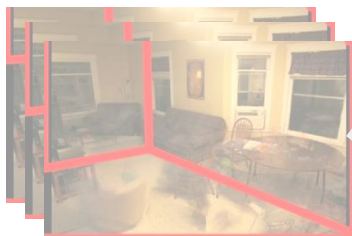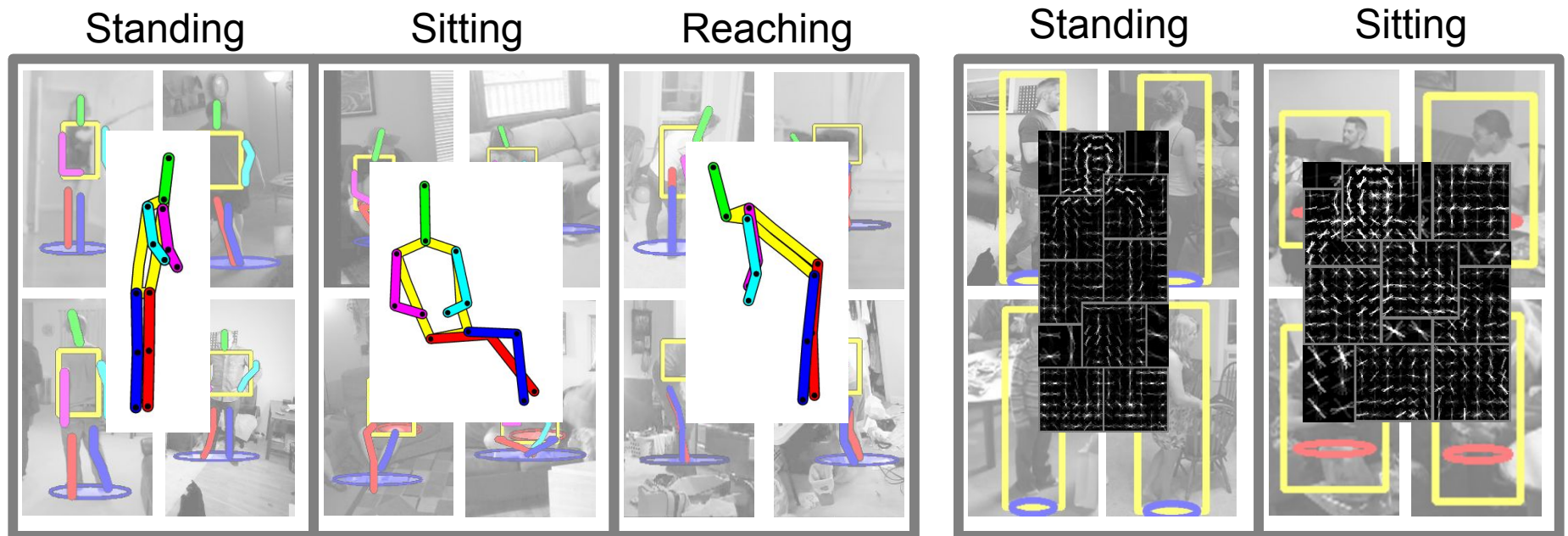
Estimate
Free-Space

Approach

Timelapse

Pose Detections

Functional Regions

Detecting Human Actions

Articulated Pose Estimator

Deformable Parts Model

Train Separate Detectors for Each Pose

Approach

Timelapse

Pose Detections

Functional Regions

From Poses to Functional Regions

Sittable Regions at Pelvic Joint

From Poses to Functional Regions

Walkable Regions at Feet

From Poses to Functional Regions
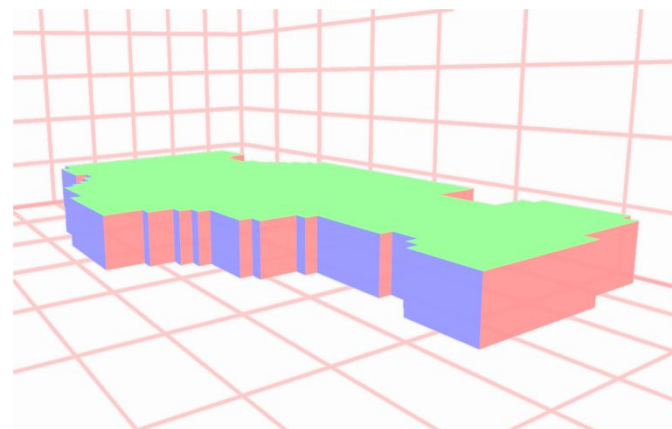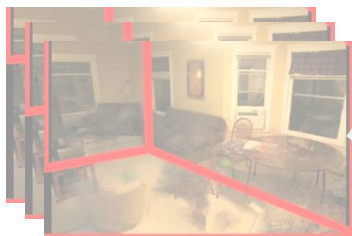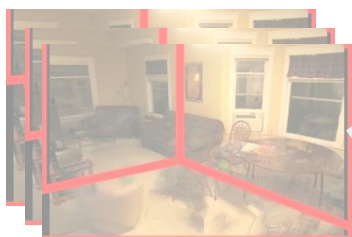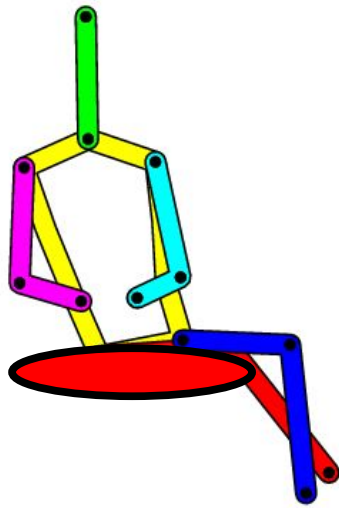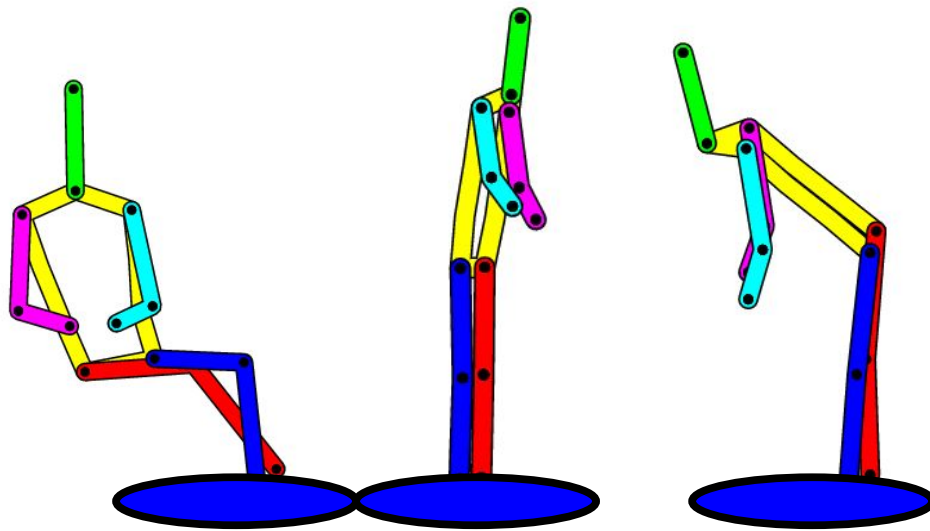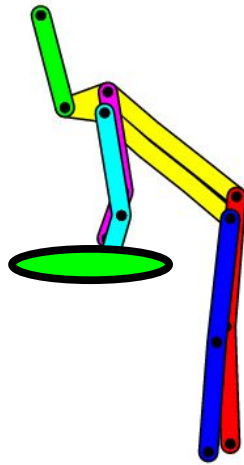

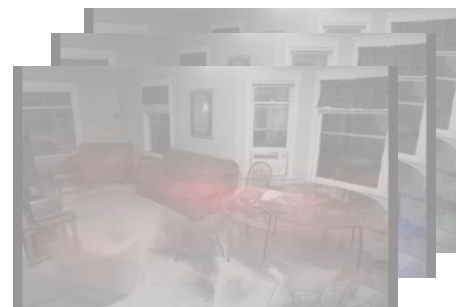Reachable Regions at Hands

Approach
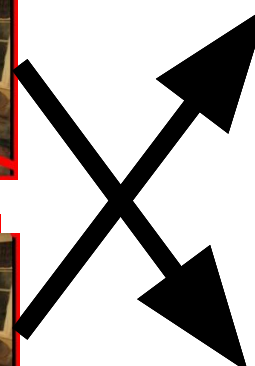

Timelapse

Pose Detections

Functional Regions

#1

#49
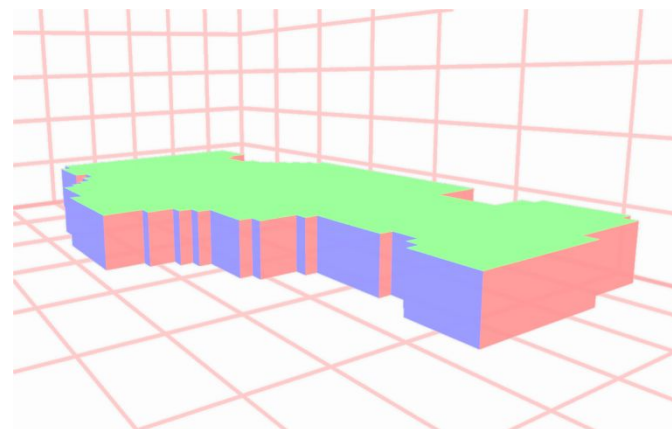
3D Room Hypotheses
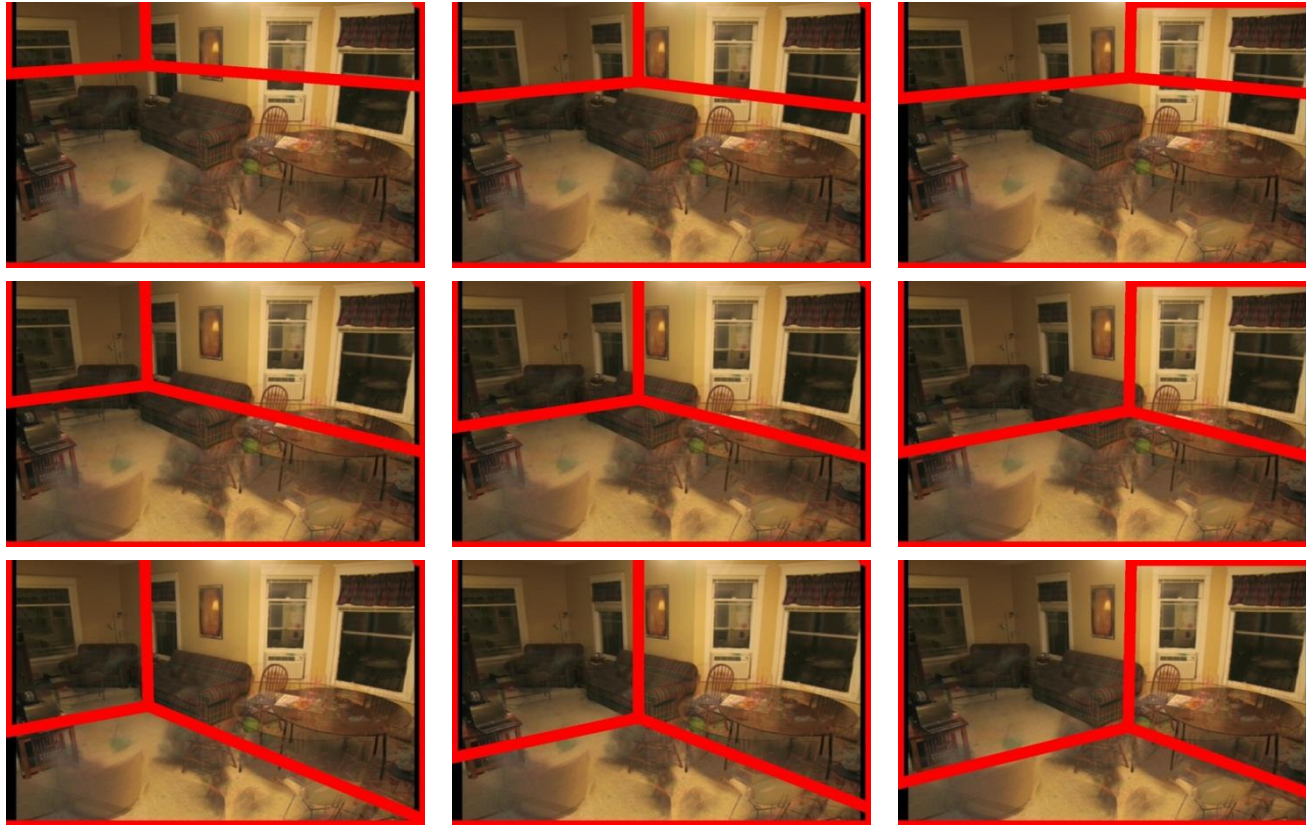
Vanishing-point aligned hypotheses

Constraints

Containment	Free Space	Support
Volume occupied by human should be inside a room	Volume occupied by human cannot intersect any object in the room	Object surfaces which can make the pose physically stable

Putting it together – picking a room

φ ( ) + ψ ( ) + ρ ( )

Compatibility of room layout with surface geometry

Putting it together – picking a room

φ (   )$+\psi$ (   )$+\rho$(   )
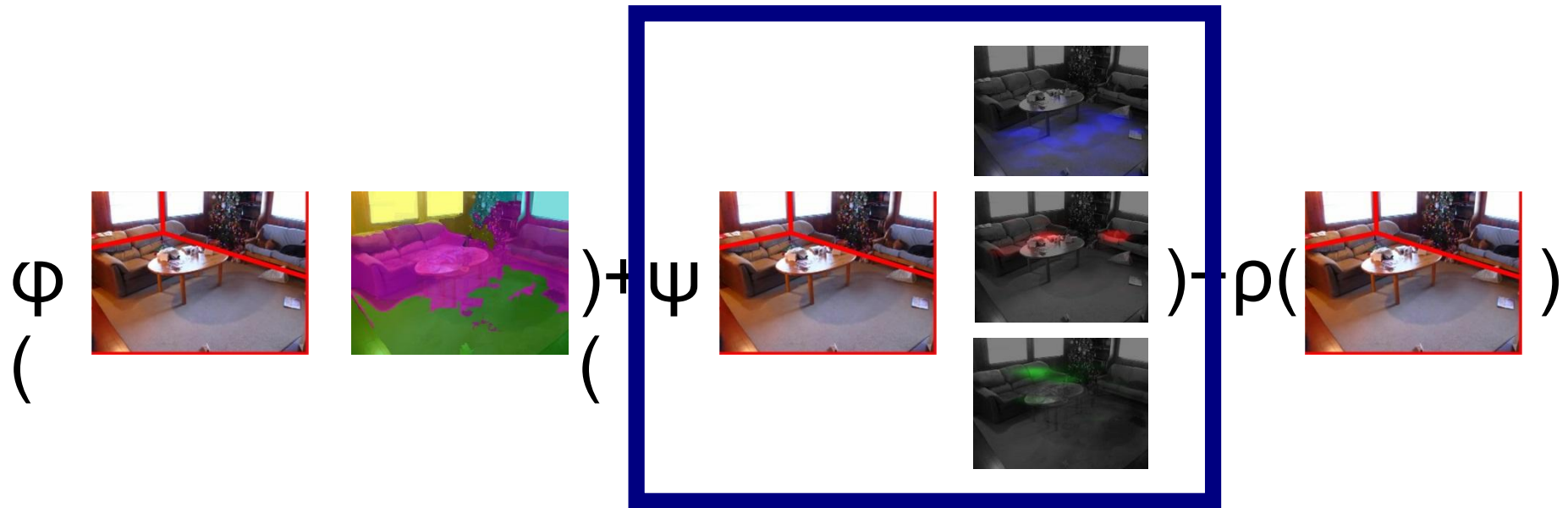
Compatibility of human poses
and room layout

Putting it together – picking a room

$$\varphi(\quad)+\psi(\quad)+\rho(\quad)$$
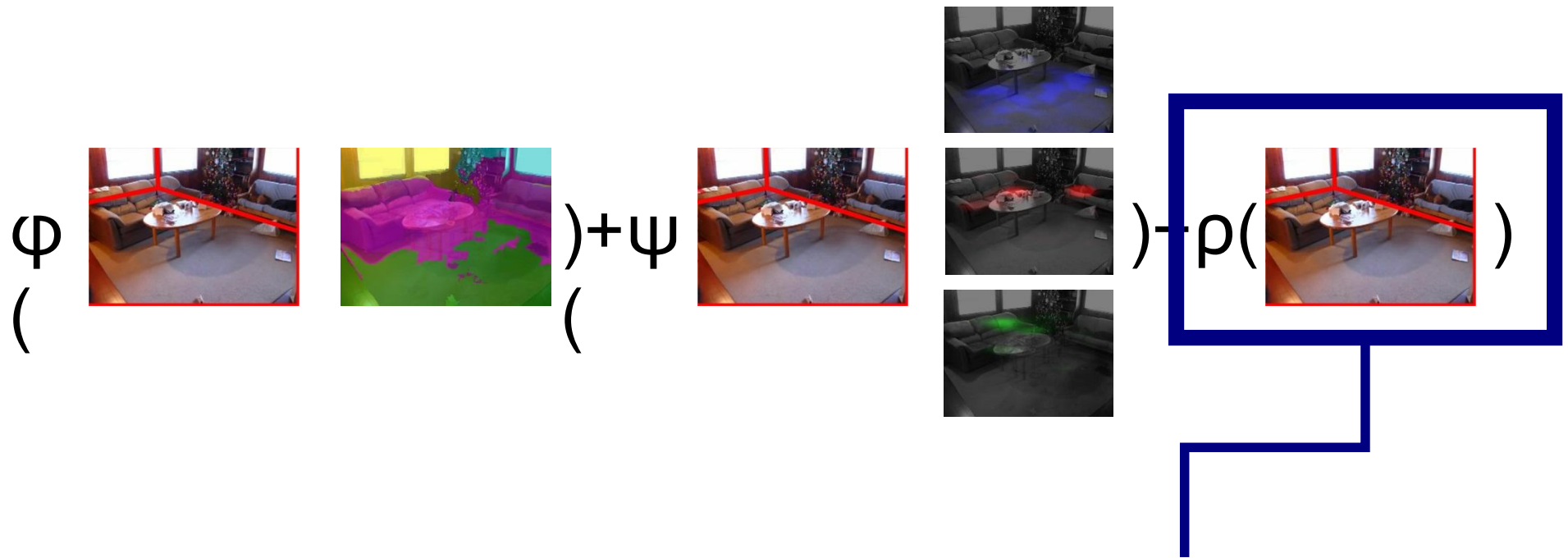
Relative room size regularizer

Reranking Results

Appearance Alone

Appearance + People

#1

Score = -1.7754

...

#82

Score = -1.8859

#1

Score = -1.8865

...
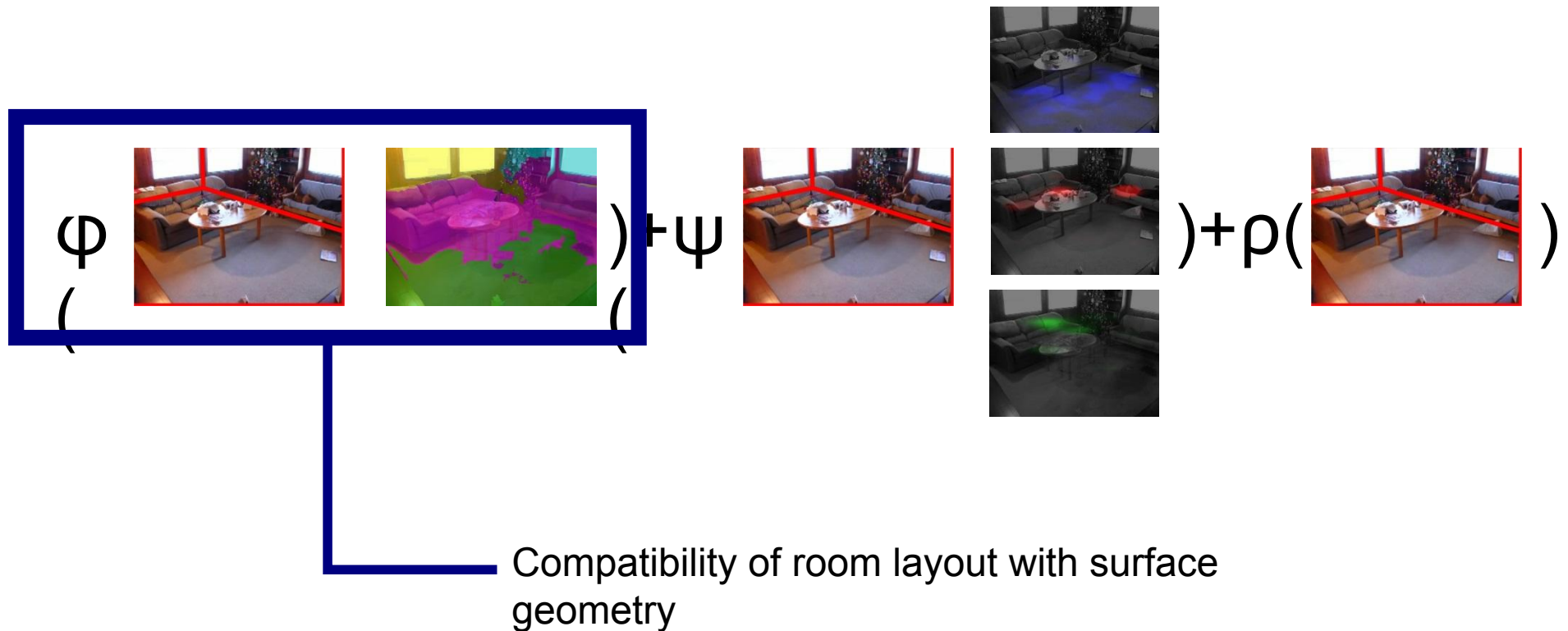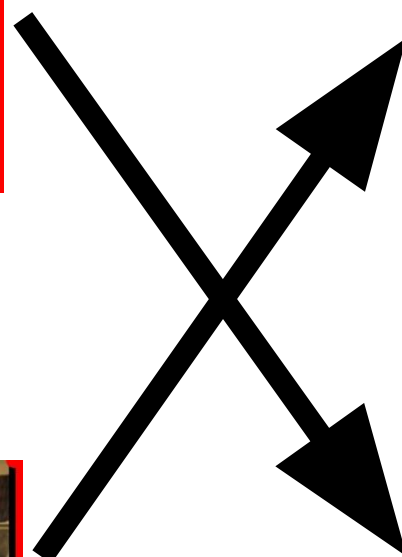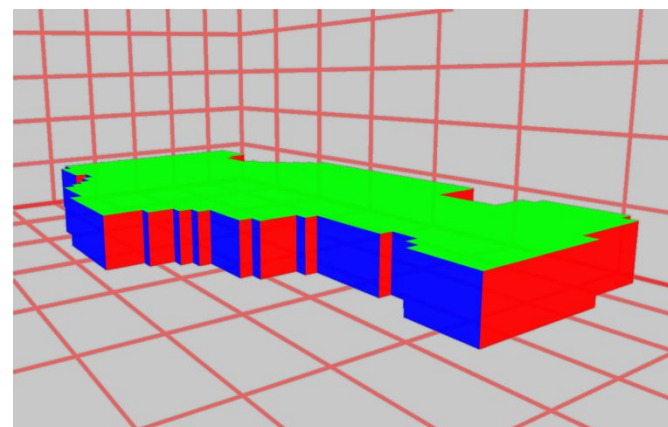
#49

Score = -2.0319

Approach

Timelapse

Pose Detections

Functional Regions

Estimate
Free-Space

Slide Credit: Fouhey et al

Estimating free space

LEGEND
Floor **Wall 1** **Wall 2** **Wall 3** **Ceiling** **Clutter**

Hedau et al. '09
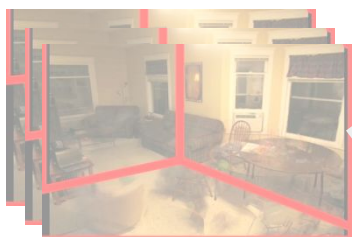
Estimating Free Space
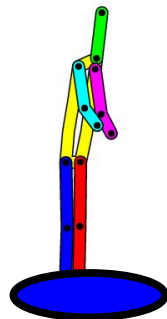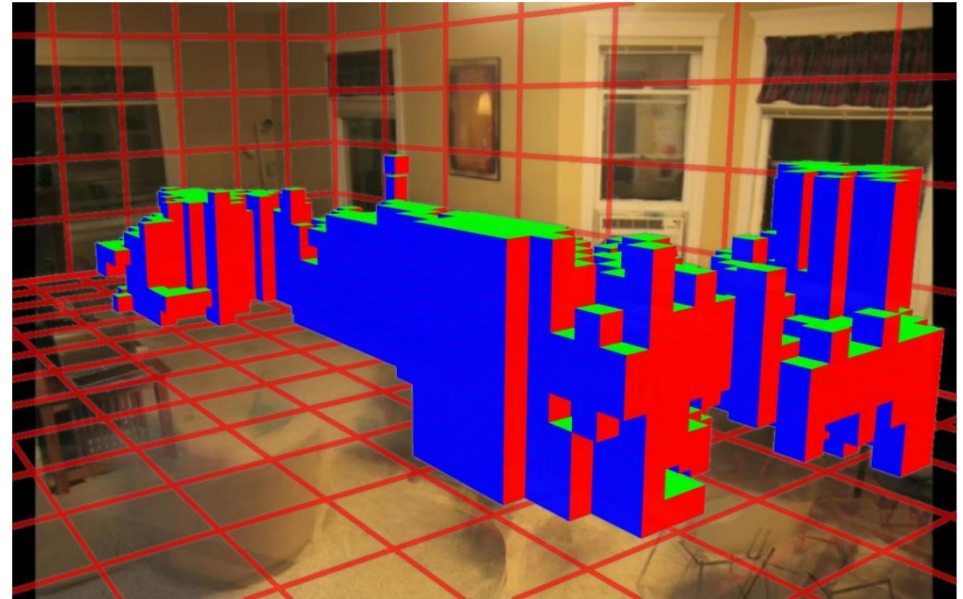
Estimating Free Space

Results

Qualitative Example

Qualitative Results

Appearance Alone

Qualitative Results
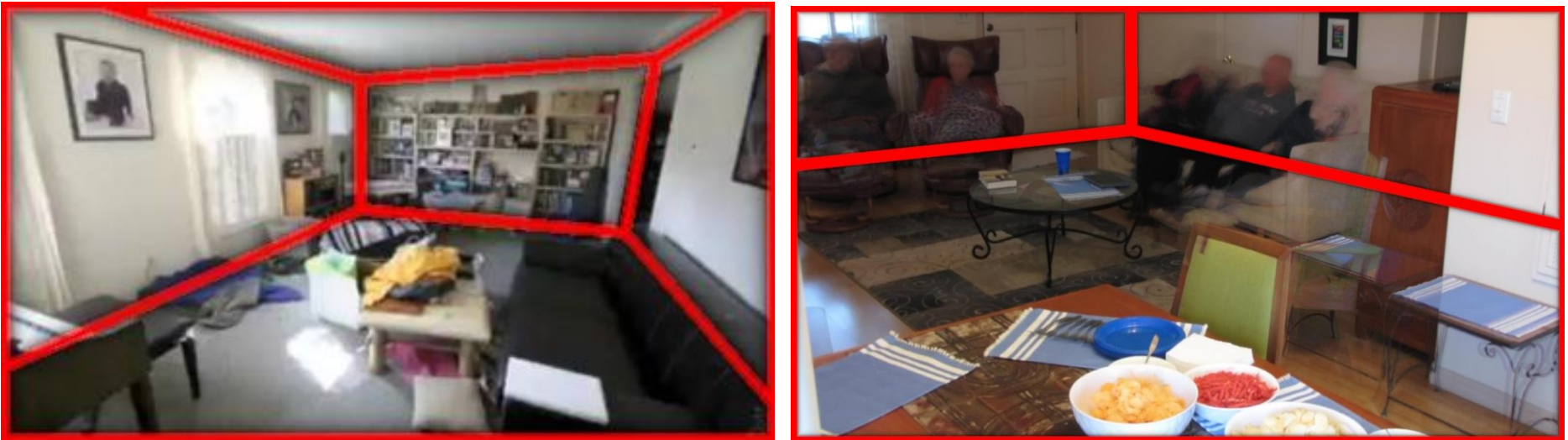
Appearance + People

Single Images with People

Appearance Alone

Single Images with People

Appearance + People
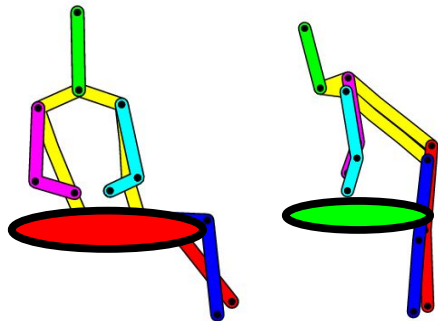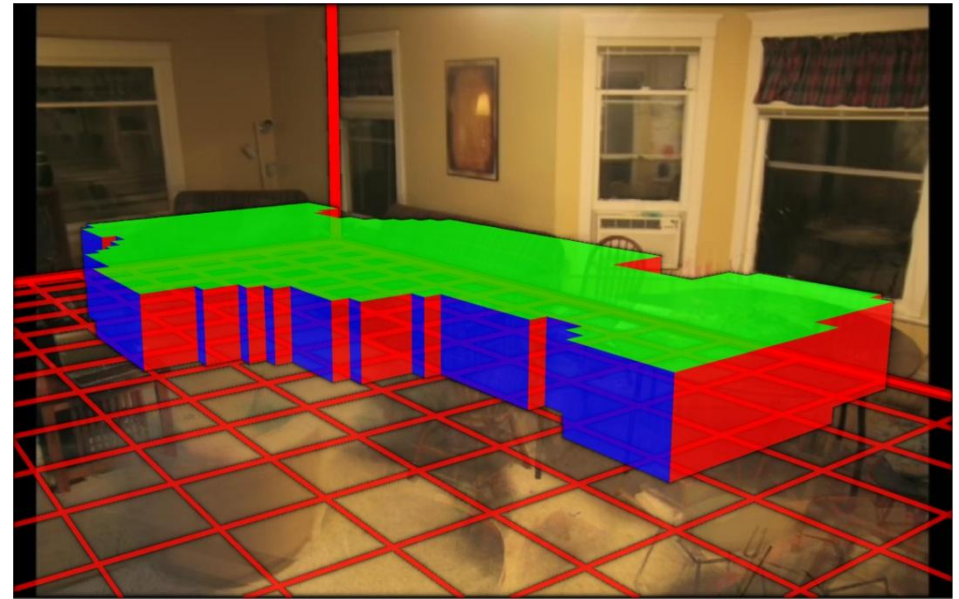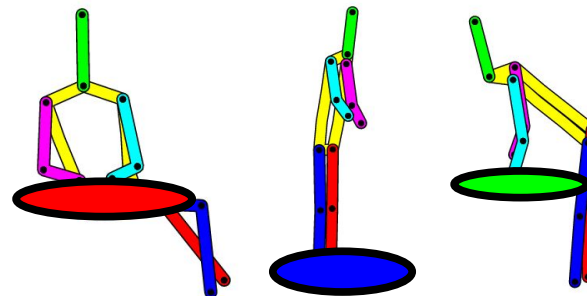
Quantitative Results

	Location	Appearance Only		People Only	Appearance + People
		Lee et al. '09	Hedau et al. '09		
Timelapses	64.1%	70.4%	74.9%	70.8%	**82.5%**
Single Image	66.4%	71.3%	77.0%		**79.6%**

Discussion Points

1. Do some human action exceptions such as can sitting on table, standing on a sofa cause trouble for the algorithm?
2. Role of background subtraction during testing
3. Semantic relationships between humans and objects
4. Performance on odd-shaped rooms and outdoor scenes
5. Evaluation metric used - only evaluates the 3D room scene and the free space estimate