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

David Fouhey et al.

Experiment Presentation by Edward Banner

Outline

- Dataset
- Overview of model
  - Which components I experimented with
- Experiments
  - Where are gains coming from?
  - Noisy functional surface estimates?
  - Use different heatmap information?
  - When does the model perform poorly?

Thanks David Fouhey!

Dataset

- 141 timelapse YouTube videos (versus 40 in the paper)
- Larger and more challenging

- http://www.di.ens.fr/willow/research/scenesemantics/

Example Room

Pipeline

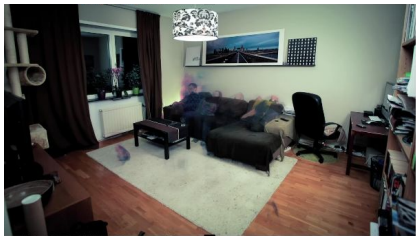
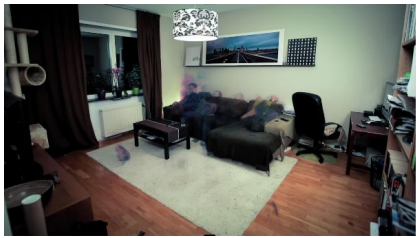


Image credit: Fouhey et al.

Pipeline

Room proposals

- Headau et al.



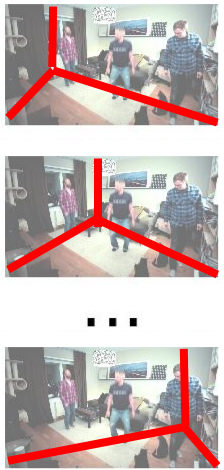
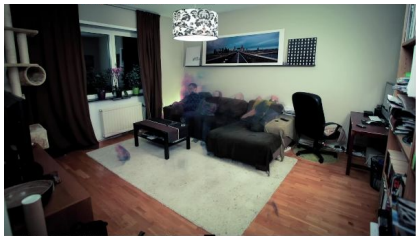


Image credit: Fouhey et al.

Pipeline

Room proposals

- Headau et al.



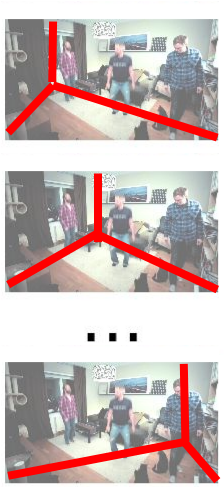
Action and Pose Detection

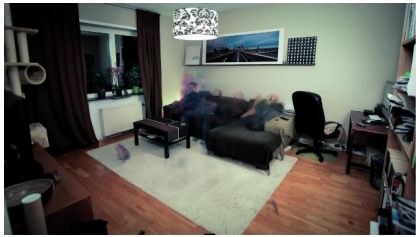
- Felzenswab et al.
- Yang and Ramanan

Image credit: Fouhey et al.

Room proposals

- Headau et al.





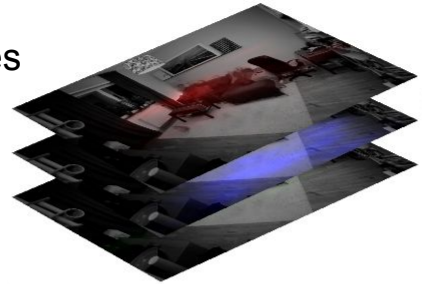
Action and Pose Detection

- Felzenswab et al.
- Yang and Ramanan

Image credit: Fouhey et al.

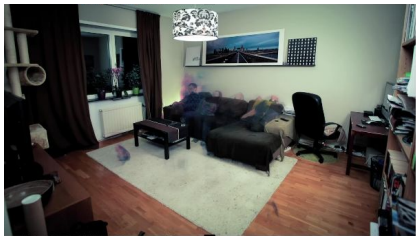
Estimates of
Functional Surfaces

- Standing
- Sitting
- Reaching



Pipeline

Room proposals
- Headau et al.



Re-ranker

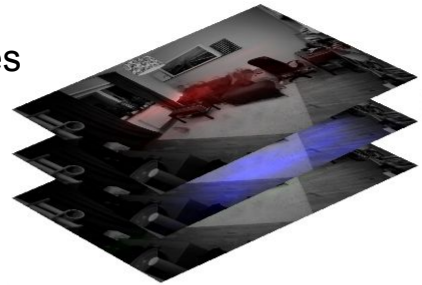
Action and Pose Detection

- Felzenswab et al.
- Yang and Ramanan

Image credit: Fouhey et al.

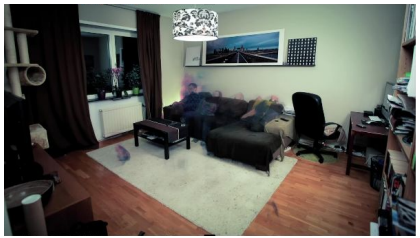
Estimates of
Functional Surfaces

- Standing
- Sitting
- Reaching



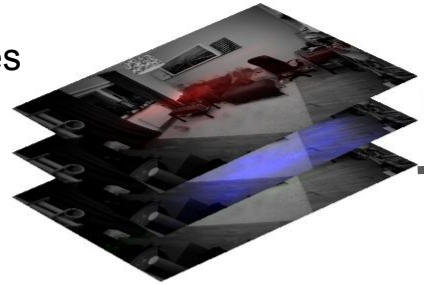
Pipeline

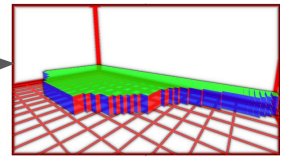
Room proposals
- Headau et al.



Re-ranker

Free Space Estimation





Action and Pose Detection

- Felzenswab et al.
- Yang and Ramanan

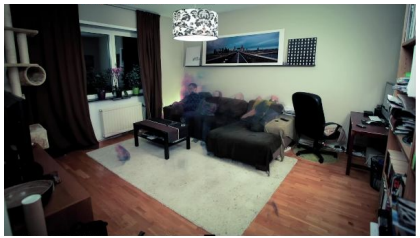
Estimates of Functional Surfaces

- Standing
- Sitting
- Reaching

Image credit: Fouhey et al.

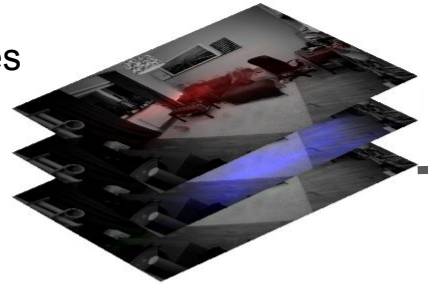
Pipeline

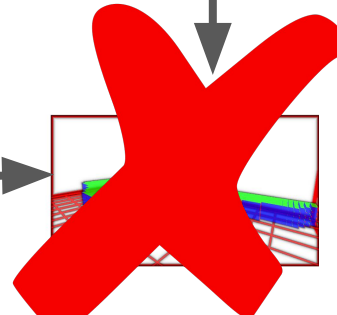
Room proposals
- Headau et al.



Re-ranker

Free Space Estimation





Action and Pose Detection

- Felzenswab et al.
- Yang and Ramanan

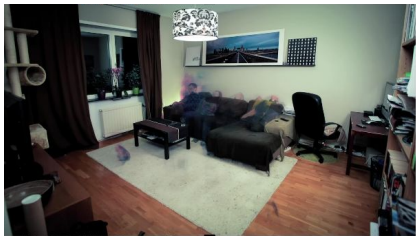
Estimates of Functional Surfaces

- Standing
- Sitting
- Reaching

Image credit: Fouhey et al.

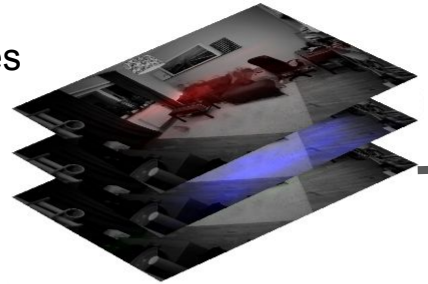
Pipeline

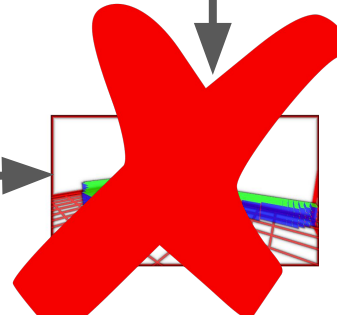
Room proposals
- Headau et al.



Re-ranker

Free Space Estimation





Action and Pose Detection

- Felzenswab et al.
- Yang and Ramanan

Estimates of Functional Surfaces

- Standing
- Sitting
- Reaching

Image credit: Fouhey et al.

Experiment 1: Ignoring functional surfaces

Experiment 1: Ignoring functional surfaces

$$f(x, H, y) = \psi(x, y) + \phi(H, y) + \rho(y)$$

Image features

Human actors (walkable area)

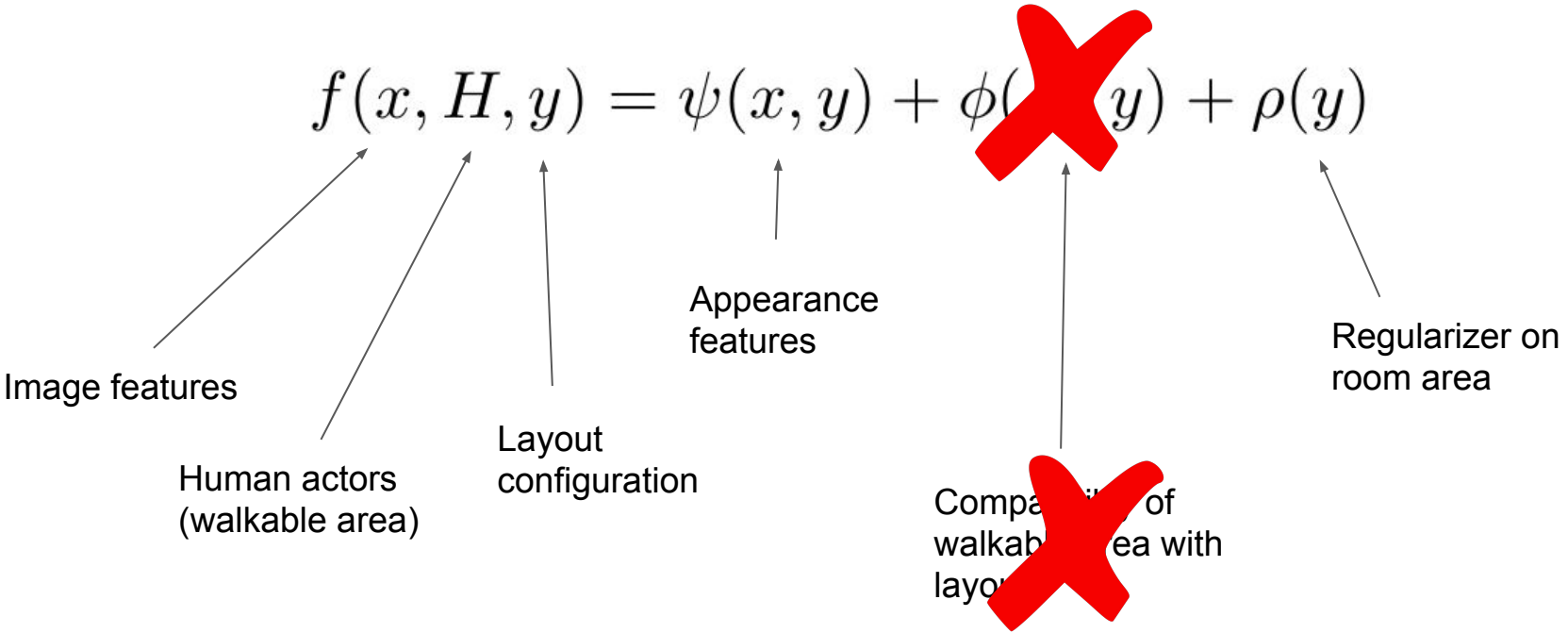
Layout configuration

Appearance features

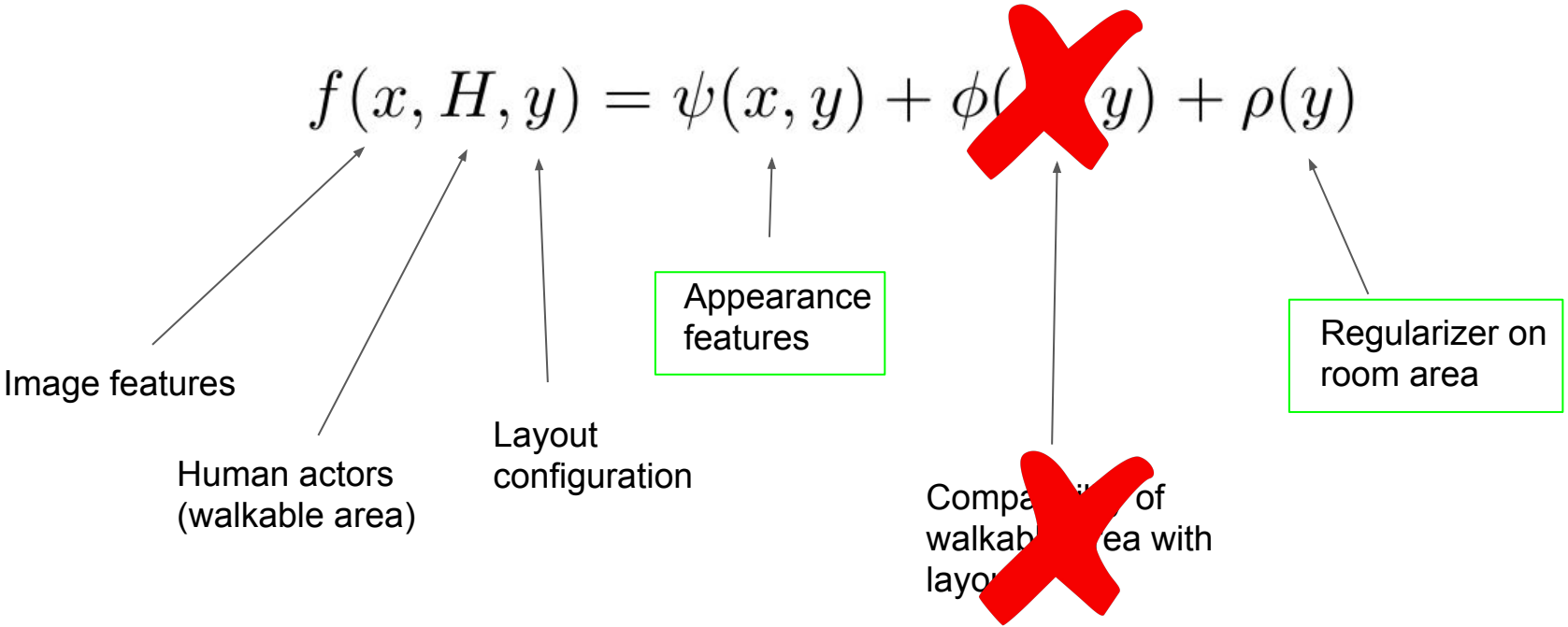
Compatibility of walkable area with layout

Regularizer on room area

Experiment 1: Ignoring functional surfaces

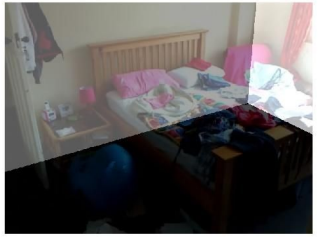


Experiment 1: Ignoring functional surfaces



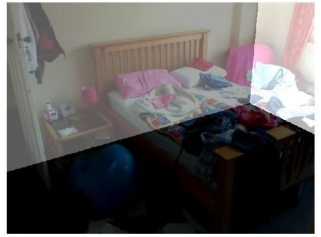
Encouraging Larger Rooms: **Successes**

Encouraging Larger Rooms: **Successes**



Appearances Feature Only

Encouraging Larger Rooms: **Successes**

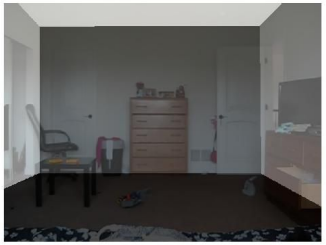


Appearances Feature Only

Low Penalty

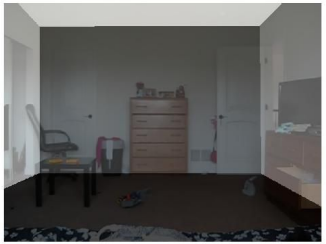
Low Area Penalty: **Failures**

Encouraging Larger Rooms: **Failures**



Appearances Feature Only

Encouraging Larger Rooms: **Failures**

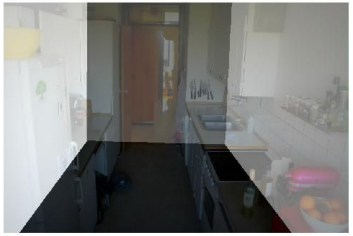


Appearances Feature Only

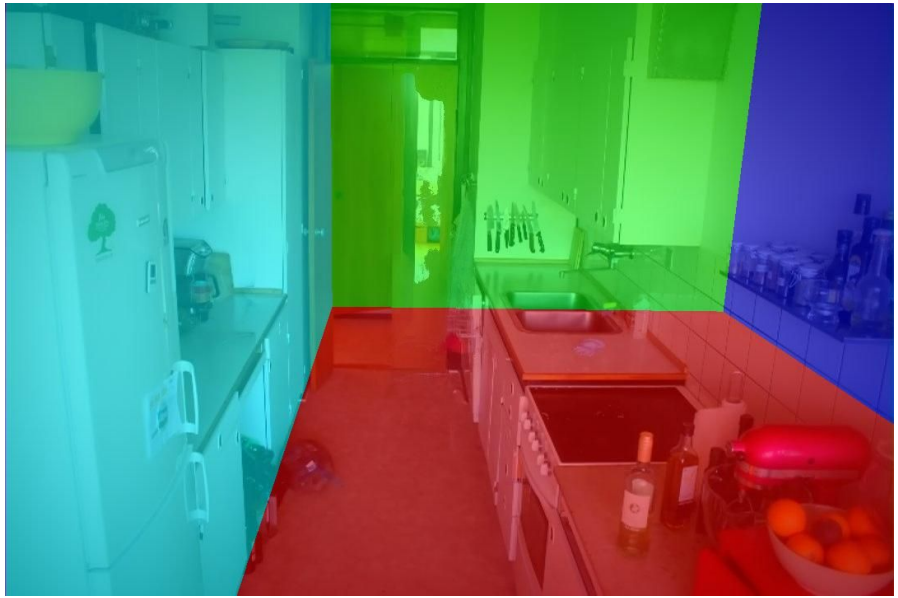
Low Penalty

Encouraging Smaller Rooms: **Successes**

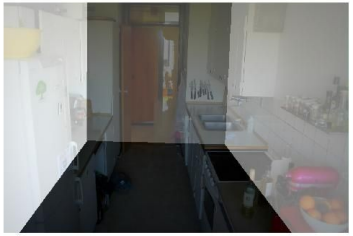
Encouraging Smaller Rooms: **Successes**



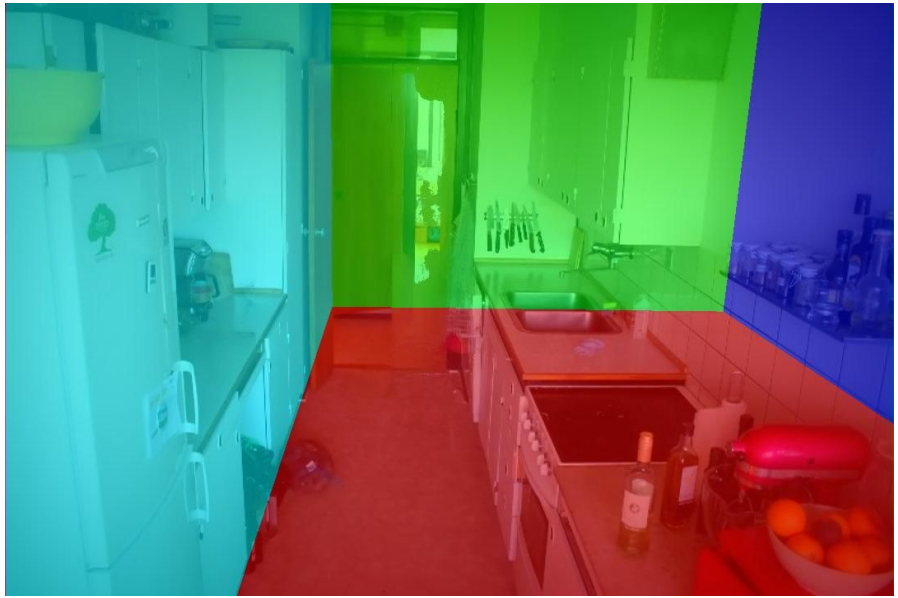
Appearances Feature Only



Encouraging Smaller Rooms: **Successes**



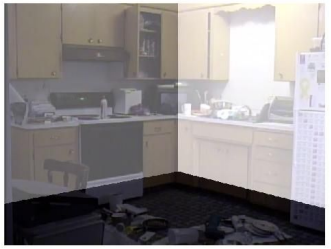
Appearances Feature Only



High Penalty

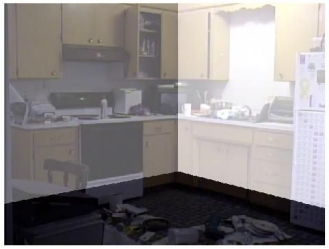
Encouraging Smaller Rooms: **Failures**

Encouraging Smaller Rooms: **Failures**



Appearances Feature Only

Encouraging Smaller Rooms: **Failures**

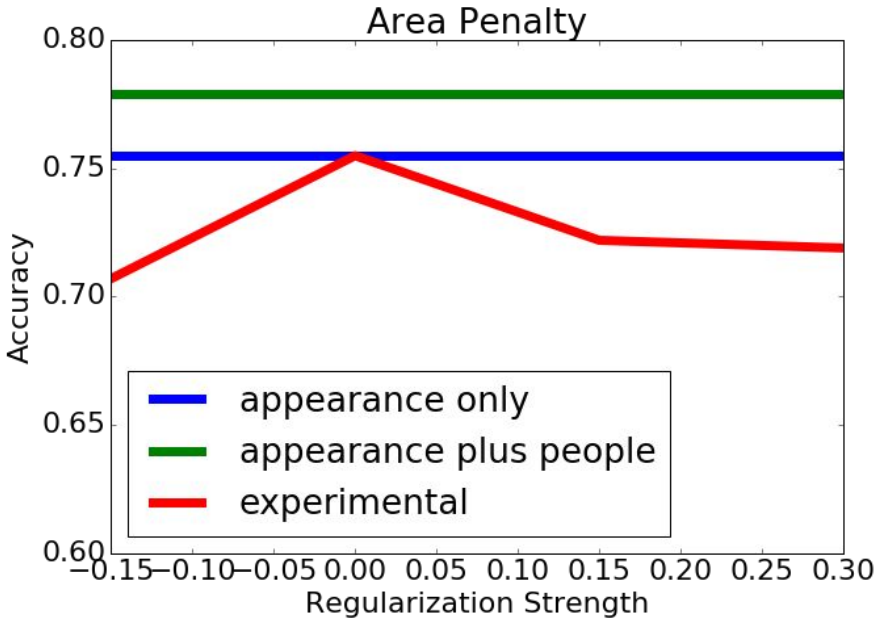


Appearances Feature Only

High Penalty

Experiment 1: Ignoring functional surfaces

20 timelapses

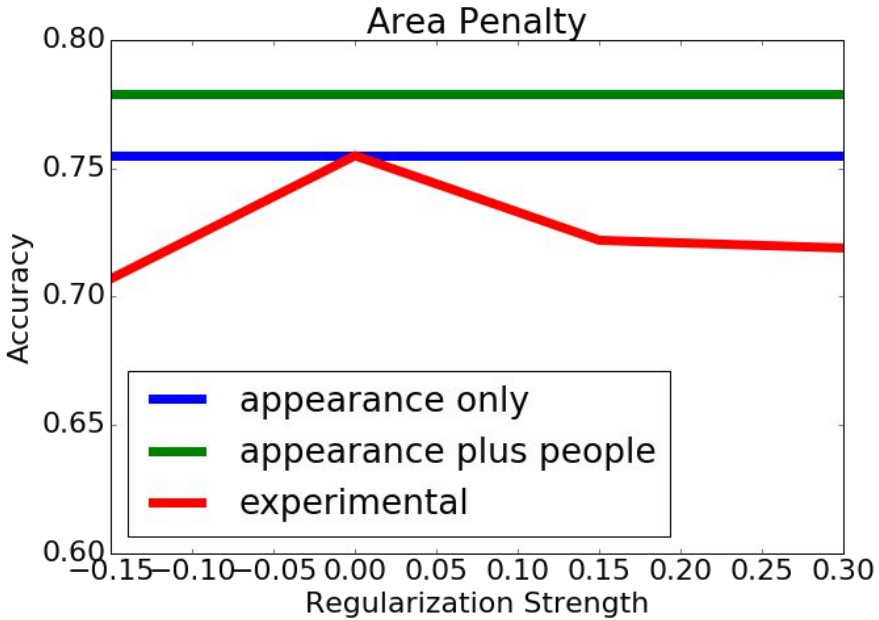


Experiment 1: Ignoring functional surfaces

20 timelapses

Conclusions:

- System ***is*** getting gains from human surface estimates!



Experiment 2: Robustness to pose estimation error

Experiment 2: Robustness to pose estimation error

- Sparsify estimates
- Diffuse estimates

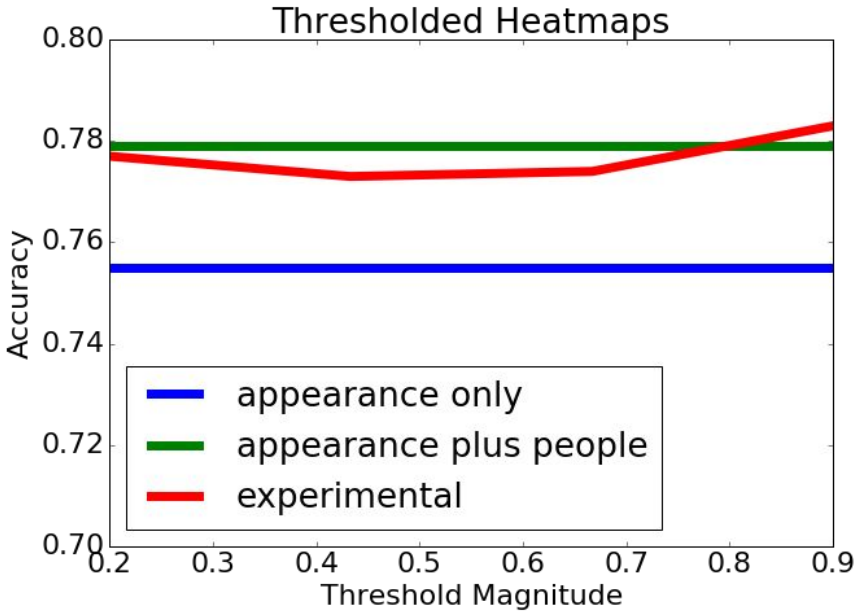
Experiment 2: Robustness to pose estimation error

- **Sparsify estimates**
- Diffuse estimates

Experiment: Sparsifying Functional Surface Estimates

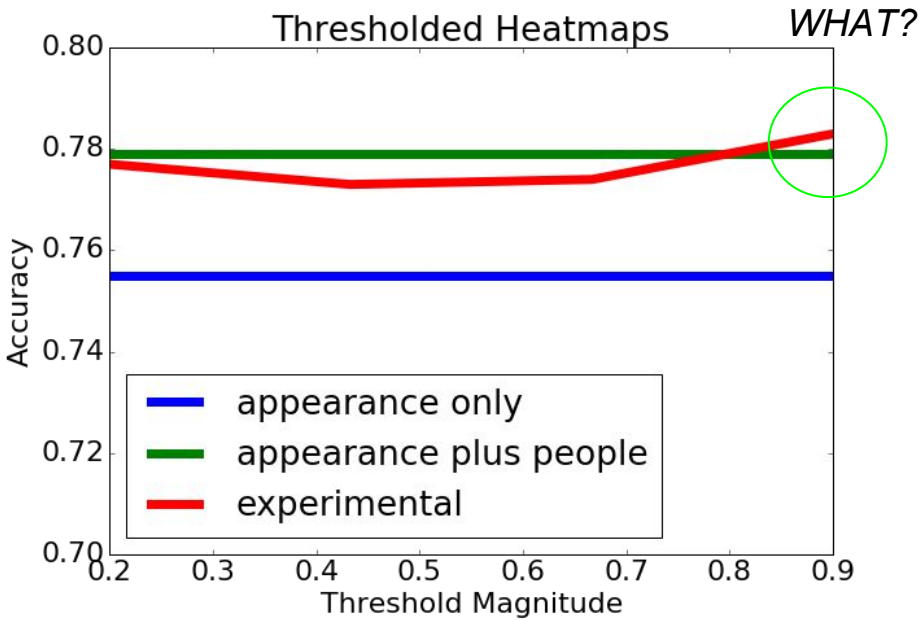
Experiment: Sparsifying Functional Surface Estimates

20 timelapses

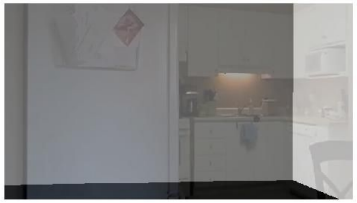


Experiment: Sparsifying Functional Surface Estimates

20 timelapses

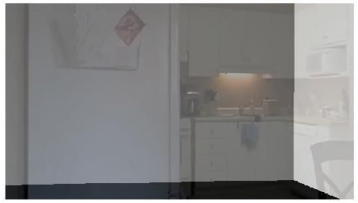


Sparsifying Walkable Area: **Successes**



Full Walkable Estimates

Sparsifying Walkable Area: **Successes**



Full Walkable Estimates

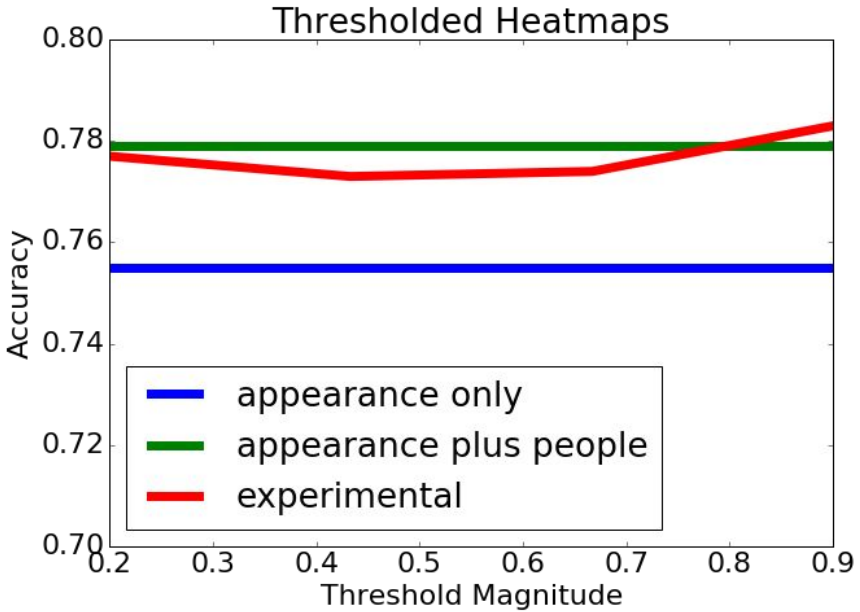
90% Thresholded Heatmap

Experiment: Sparsifying Functional Surface Estimates

20 timelapses

Conclusion: Spurious example. More data needed!

Model is surprisingly resilient!



Experiment: Robustness to pose estimation error

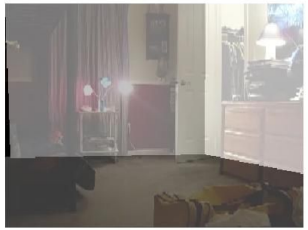
- Sparsify estimates
- Diffuse estimates

Experiment: Robustness to pose estimation error

- Sparsify estimates
- **Diffuse estimates**

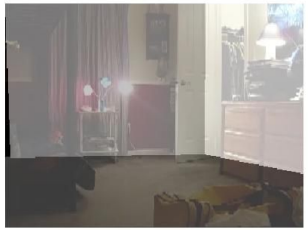
Diffusing Walkable Areas: **Failures**

Diffusing Walkable Areas: **Failures**



Appearance Feature Only

Failures!

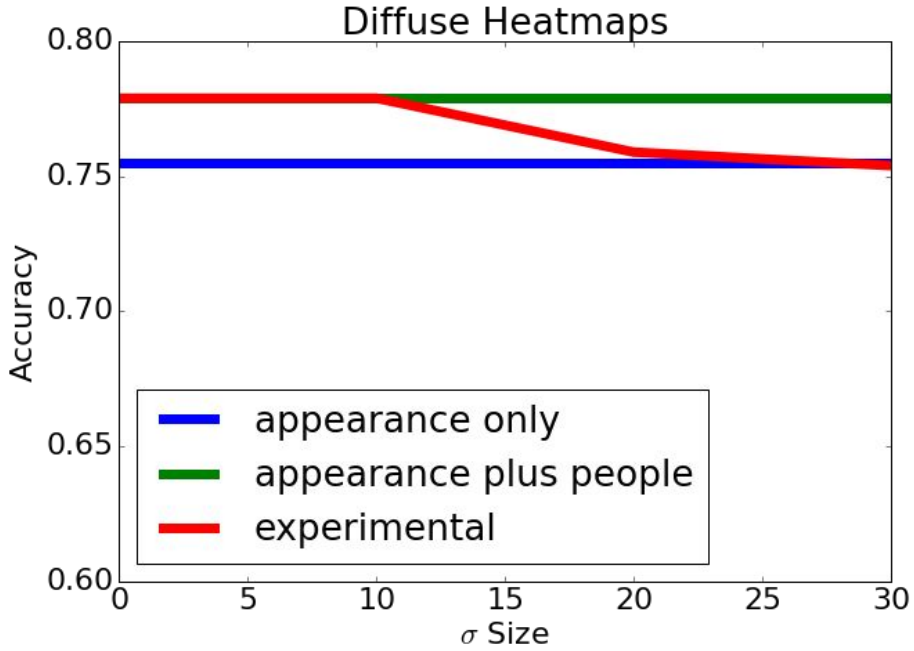


Appearance Feature Only

High Diffuse Level

Experiment: Making Heatmaps more Diffuse

20 timelapses

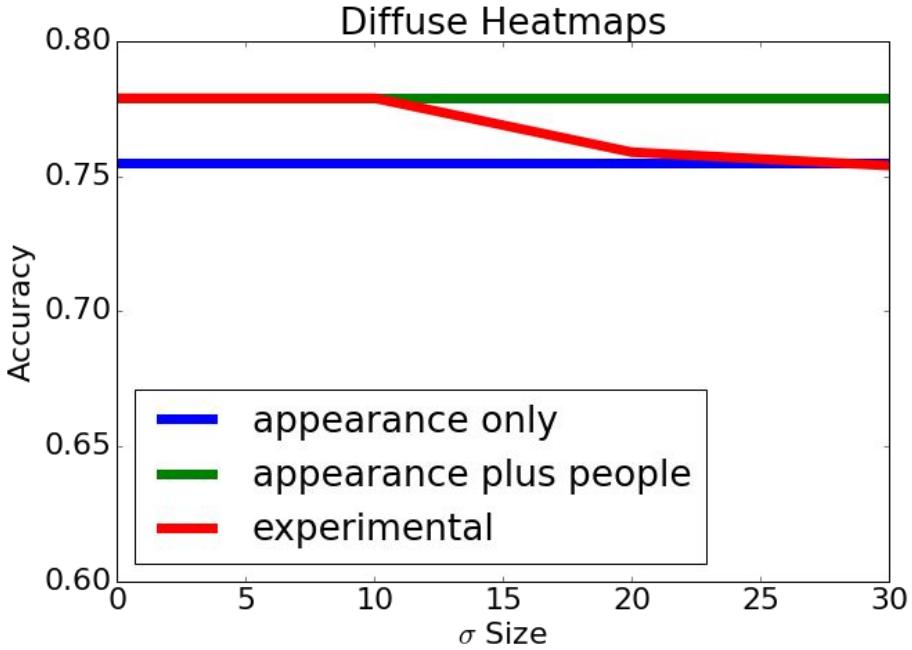


Experiment: Making Heatmaps more Diffuse

20 timelapses

Conclusion: Making walkable areas more diffuse has similar effect to encouraging larger rooms.

Model is fairly robust!



Experiment: Using only sittable regions

Containment constraint: Walkable area must be within the proposed room

Sittable area obeys this constraint!

Using Only Sittable Regions: **Succeses**

Using Only Sittable Regions: **Succeses**

Just Standable

Using Only Sittable Regions: **Succeses**

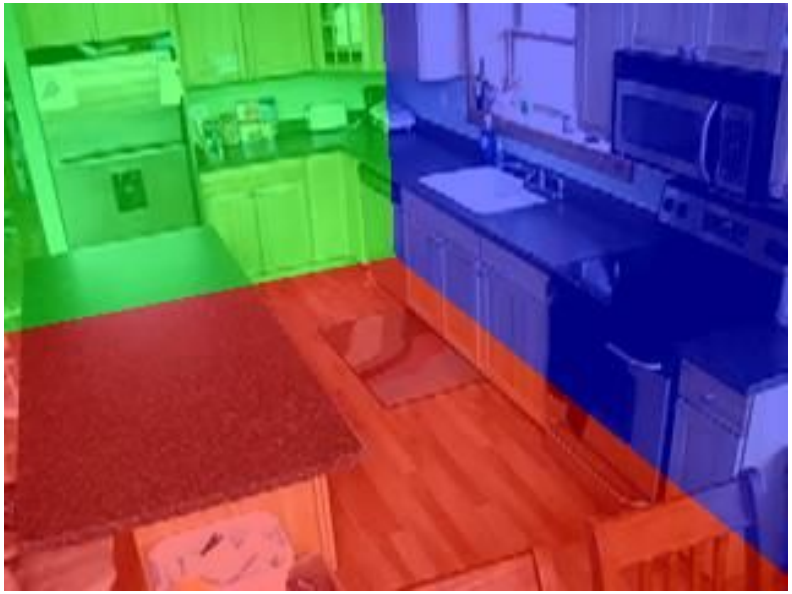
Just Standable

Just Sittable

Using Only Sittable Regions: **Failures**

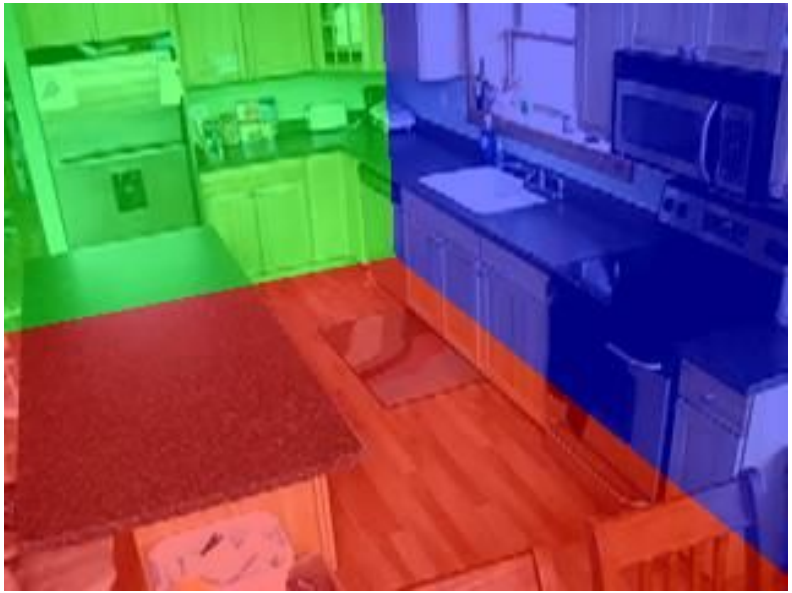
Using Only Sittable Regions: **Failures**

Just Standable

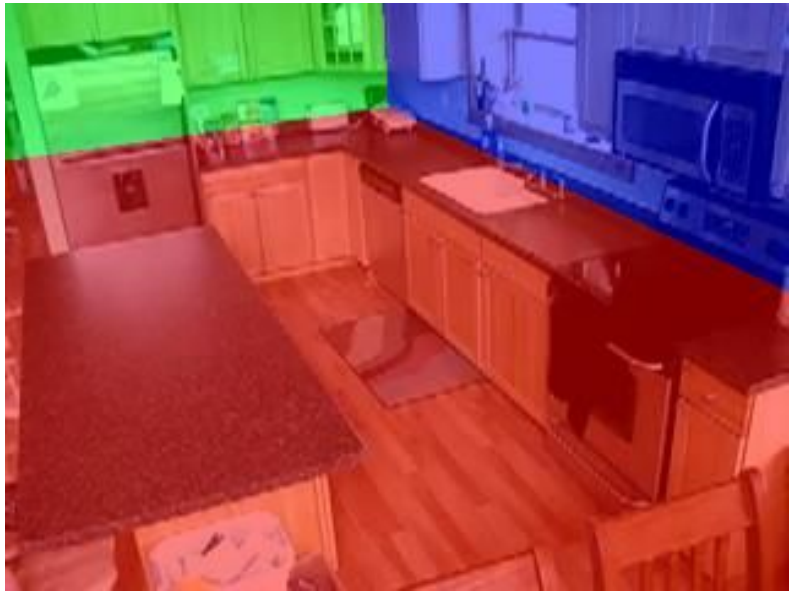


Using Only Sittable Regions: **Failures**

Just Standable

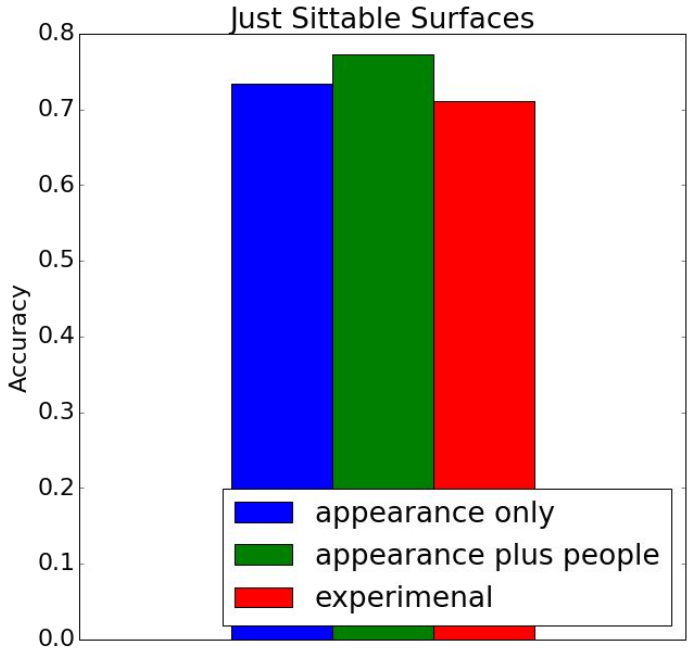


Just Sittable



Experiment 4: Using sittable regions instead of standable regions

80 timelapses

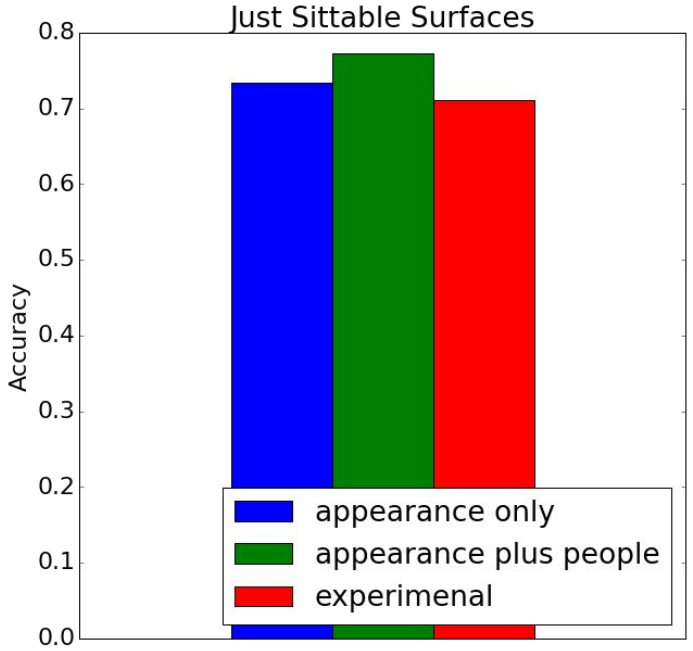


Experiment 4: Using sittable regions instead of standable regions

80 timelapses

Conclusion: Sittable regions tend to “push up” floor to avoid floating humans.

Misconception about the containment constraint!



Experiment 5: Relationship between detected floor space and actual floor space

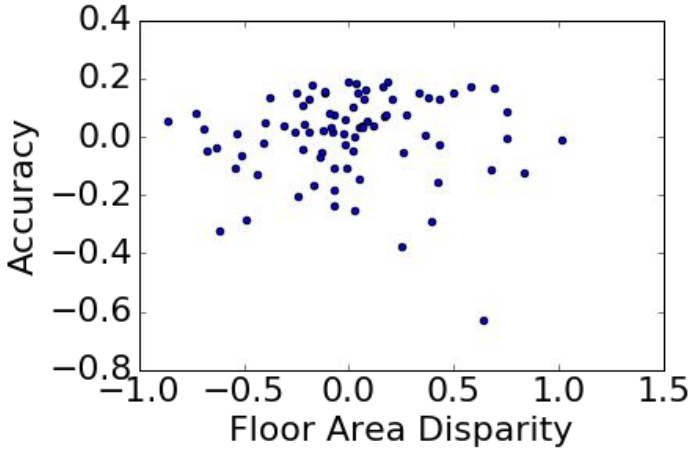
Experiment 5: Relationship between detected floor space and actual floor space

Hypothesis: Accuracy will suffer when there is high disparity between actual floor space and detected floor space

Experiment 5: Relationship between detected floor space and actual floor space

Hypothesis: Accuracy will suffer when there is high disparity between actual floor space and detected floor space

80 timelapses



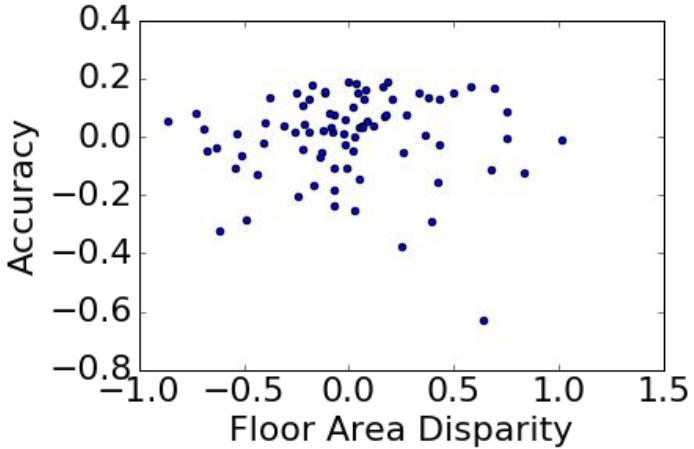
Experiment 5: Relationship between detected floor space and actual floor space

Hypothesis: Accuracy will suffer when there is high disparity between actual floor space and detected floor space

80 timelapses

Conclusion: The two are ***uncorrelated***!

Takeaway: ***Location*** of walkable area more important than magnitude



Computation Time Breakdown

- ~200-2000 room proposals per room
- ~7-20 seconds per room
- ~2 hours for 80 rooms