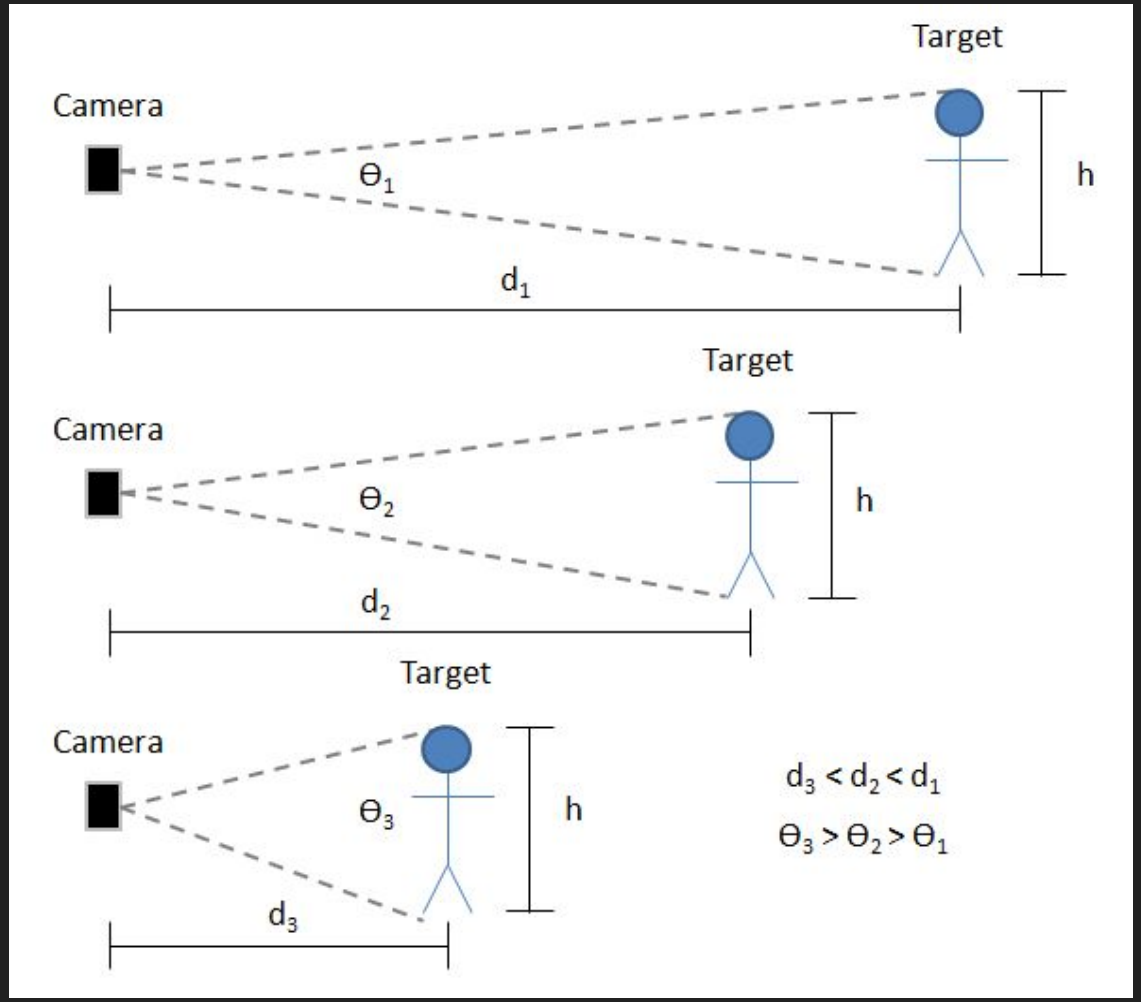
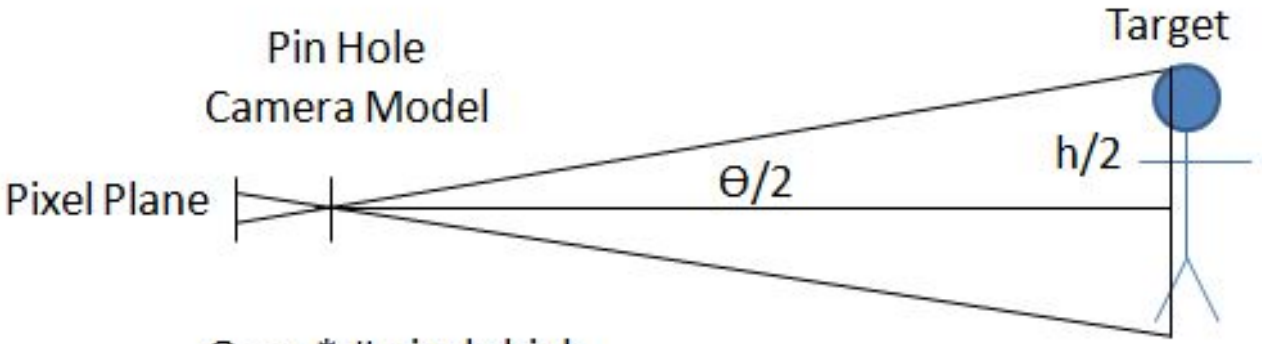
# Predicting Object Depth from Segmented RGB Image

Ahmed AlMutawa, Austin Anderson, Rohit Raje, John Stechschulte, Wade Wu


Target
Camera
Θ1
h
d1
Target
Camera
Θ2
h
d2
Target
Camera
Θ3
h
d3
d3 < d2 < d1
Θ3 > Θ2 > Θ1




Θ = a * # pixels high

a –> scalar conversion from pixels to degrees

a ≈ Camera Field of View/Total Pixel Height

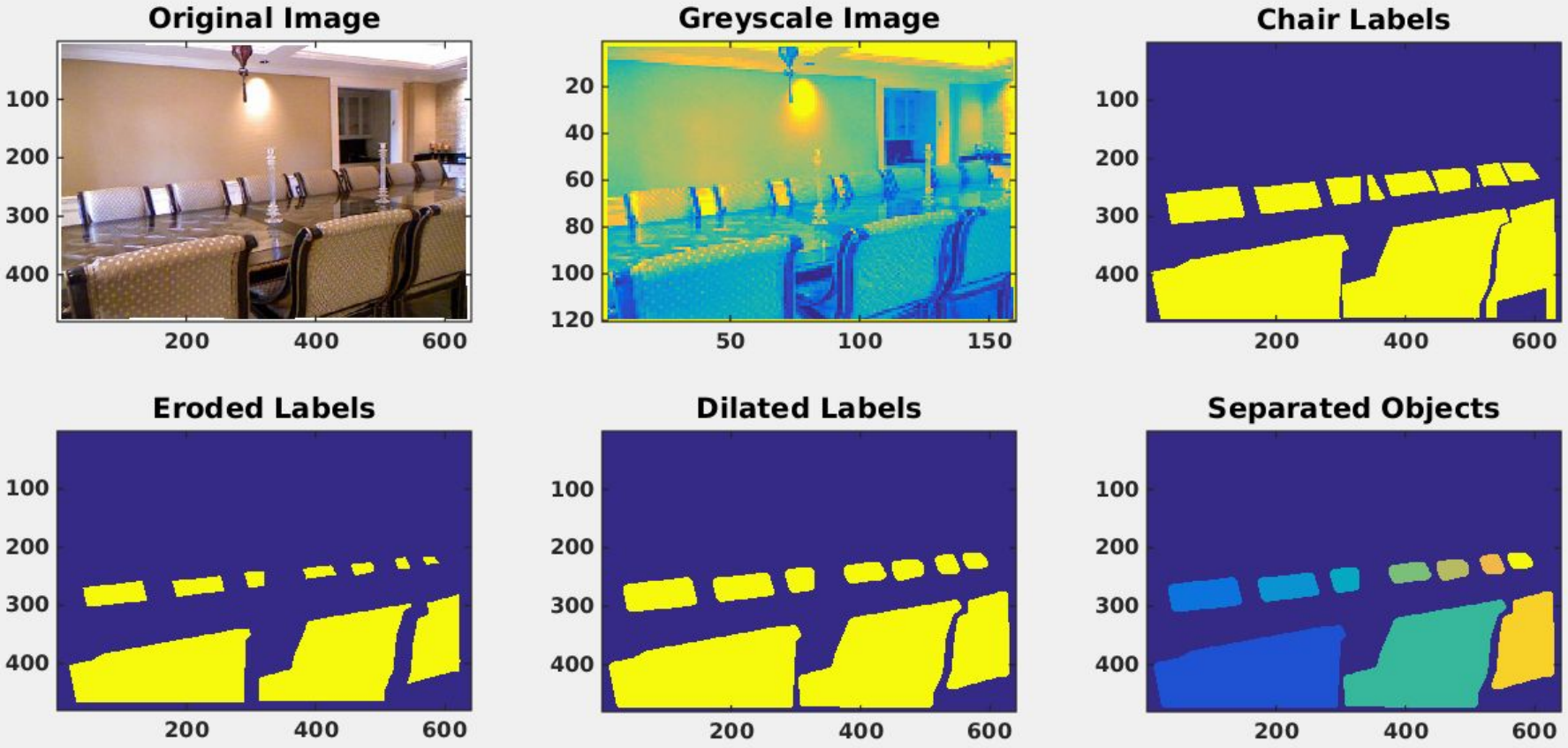
For Kinect:

$a_{vert}$ ≈ 43°/480 pixels

$a_{horiz}$ ≈ 57°/640 pixels

d ≈ h/$a_{vert}$* # pixels high

From: https://msdn.microsoft.com/en-us/library/jj131033.aspx


Original Image
100
200
300
400
200
400
600
Greyscale Image
20
40
60
80
100
120
50
100
150
Chair Labels
100
200
300
400
200
400
600
Eroded Labels
100
200
300
400
200
400
600
Dilated Labels
100
200
300
400
200
400
600
Separated Objects
100
200
300
400
200
400
600

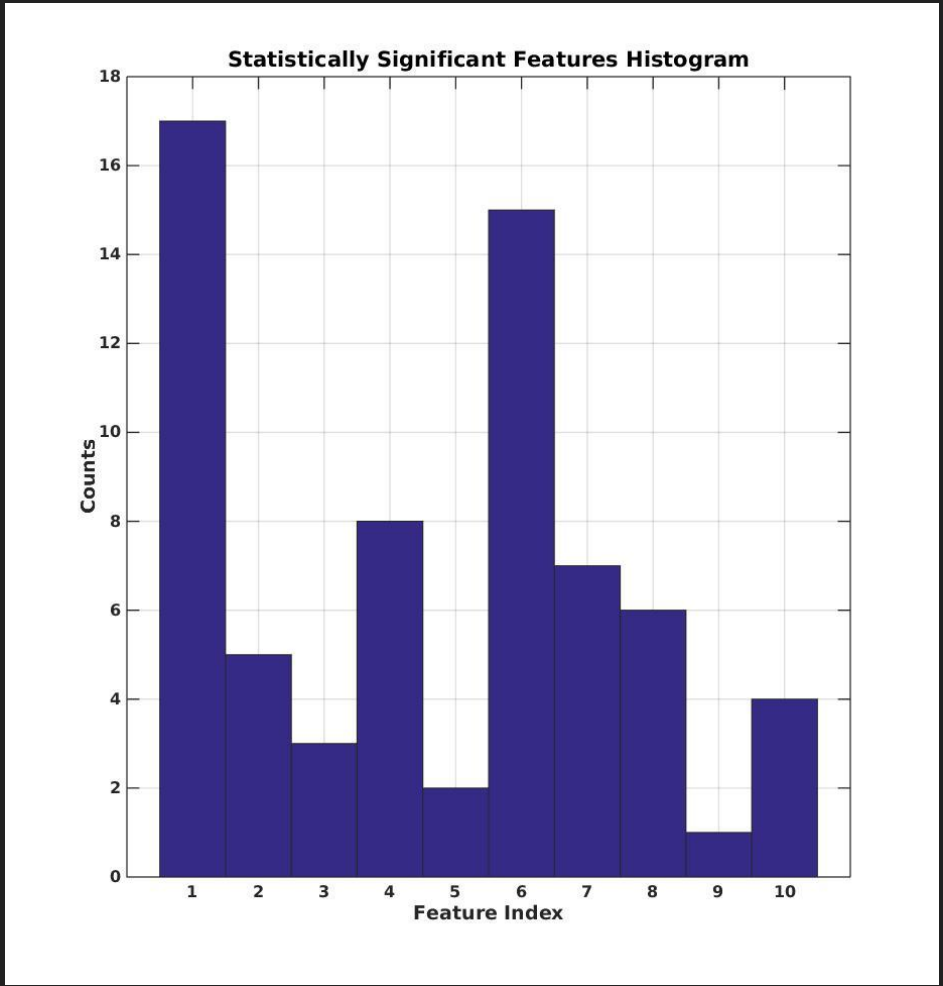
## Basic Features

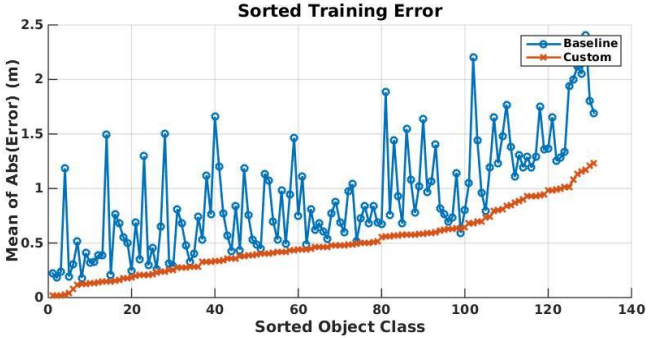
- Num. pixels (and inverse sqrt)
- Pixel width (and inverse)
- Pixel height (and inverse)
- Object location (x and y)

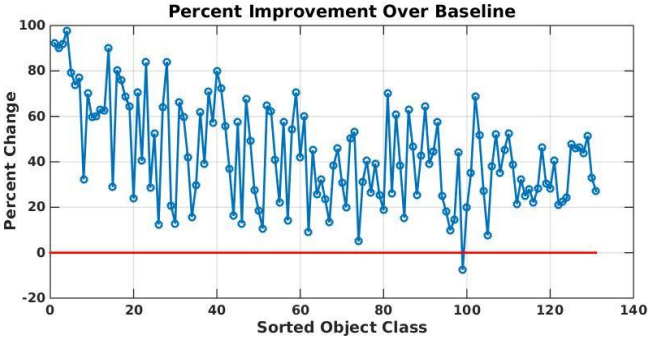
## Engineered Features

- Corner detection
- Sparse filtering
- Principal Component Analysis
- Hough Transform
- Water filling algorithm

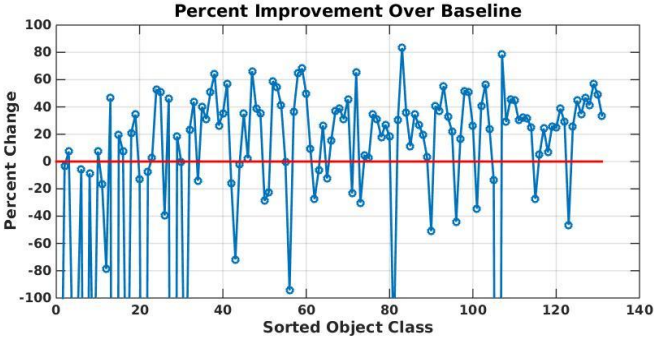
| | |
|---|---|
| 1 | Bias |
| 2 | 1/sqrt(# pixels) |
| 3 | Width (dx) |
| 4 | Height (dy) |
| 5 | x-position |
| 6 | y-position |
| 7 | 1/dx |
| 8 | 1/dy |
| 9 | Number of corners |
| 10 | PCA eig. 1 |




Sorted Training Error
Baseline
Custom
Mean of Abs(Error) (m)
Sorted Object Class


Percent Improvement Over Baseline
Percent Change
Sorted Object Class

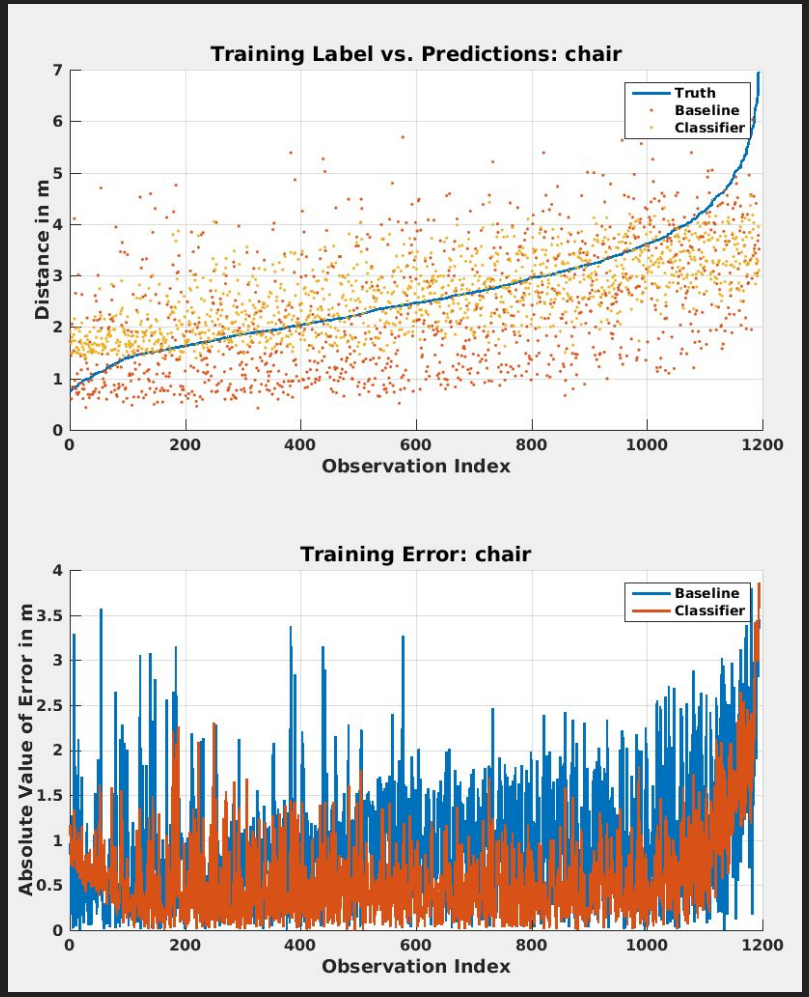

Testing Error Sorted by Training Error
Baseline
Custom
Mean of Abs(Error) (m)
Sorted Object Class

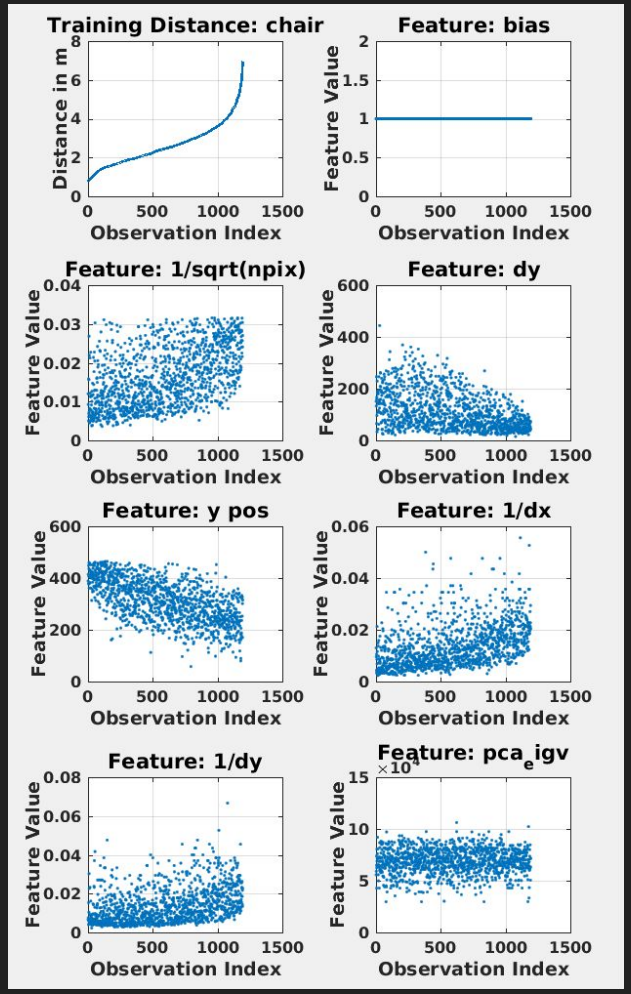

Percent Improvement Over Baseline
Percent Change
Sorted Object Class

# Best and Worst Object Classes

| Class | Test error (m) |
|---|---|
| ‘toilet paper’ | 0.14 |
| ‘tea kettle’ | 0.15 |
| ‘placemat’ | 0.19 |
| ‘sink’ | 0.21 |
| ‘paper towel’ | 0.24 |

| Class | Test error (m) |
|---|---|
| ‘pipe’ | 11.25 |
| ‘flower pot’ | 10.12 |
| ‘classroom board’ | 8.44 |
| ‘mail shelf’ | 7.21 |
| ‘bicycle’ | 6.66 |


Training Label vs. Predictions: chair
Truth
Baseline
Classifier
Distance in m
Observation Index
Training Error: chair
Baseline
Classifier
Absolute Value of Error in m
Observation Index


Training Distance: chair
Distance in m
Observation Index
Feature: bias
Feature Value
Observation Index
Feature: 1/sqrt(npix)
Feature Value
Observation Index
Feature: dy
Feature Value
Observation Index
Feature: y pos
Feature Value
Observation Index
Feature: 1/dx
Feature Value
Observation Index
Feature: 1/dy
Feature Value
Observation Index
Feature: pca_igv
Feature Value
Observation Index

# Conclusions and future work

- Some object classes outperformed the baseline
- One major surprise: effectiveness of y-position feature
- Learned as much about the training data as we did about the problem

- Incorporate segmentation
- Better data augmentation for improved training
- Tune preprocessing parameters for each object class