ESE 650 - Learning in Robotics
Image Segmentation

Vishnu Purushothaman Sreenivasan

Introduction

Problem Statement

The objective of the project is to use image segmentation to detect an object in an image. In this project the object was a bright pinkish-red barrel. Once the object is detected we had to identify the depth at which the barrel is located using the information from the training images.

Description of approach

I trained a Gaussian model with full covariance for each of the color classes. Using the connected object information I removed spurious noise and detected the depth. Depth was detected using the closeness of the detected major or minor axis length to the mean of the major or the minor axis length for the barrel of different depths.

Methodology

The methodology can be split into the training and the testing part:

Training

a) In the training, I hand selected 6 different color classes i.e. barrel red, white, floor red, yellow, the red color of the soda vending machine and the red color of the reflection of the barrel on wet surfaces.
b) I converted the RGB images to YCbCr for luminance resilience and converted the training images into a 3 dimensional vector of values, picking the corresponding pixels for each color using the data from the masks generated in the previous step.
c) I computed the mean (μ), the covariance matrix ∑ whose inversion gave me the A matrix and I also computed the prior of each color class by finding the ratio of the total no of pixels in all the images corresponding to a class to the total number pixels in all the training images.
d) I wrote a function to determine the log likelihood of each pixel belonging to a color class using the following equation

$$P(\boldsymbol{x}) = \sqrt{\frac{\det(\boldsymbol{A})}{(2\pi)^D}} \times e^{\left\{\frac{-1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T \boldsymbol{A}(\boldsymbol{x}-\boldsymbol{\mu})\right\}}$$

e) Applying Bayes rule, I obtained the probability of each color class given a pixel value i.e. the probability that a pixel belongs to a color class.

$$P(C|\boldsymbol{x}) = \frac{P(\boldsymbol{x}|C)\sum_{\{x'\}} P(C, \boldsymbol{x'})}{\sum_{\{c'\}} P(C', \boldsymbol{x})}$$

f) I built a lookup table which is a cell matrix containing the probability of all of the color classes given a pixel value. We apply something similar to a snap to grid technique where I round of the pixel values to the nearest multiple of 3 from (0 to 255).
g) Using the masks built for the barrel red color, the masks around the barrel, I obtained the properties of the barrel at different distances such as major axis length, minor axis length, convex and filled area, eccentricity, equivalent diameter and the ratio of the major and minor axis length.

Testing

a) The input image is resized and an image of 0.1 times the size of the original image is obtained. This provides for a significant speed up by 300 times for the detection algorithm.

b) The image is converted to YCbCr and the pixel values are rounded off to the nearest multiple of three. This allows for quick table lookup for the probability of barrel red.

c) Then probability response map is thresholded using an adaptive threshold. The threshold is chosen as 0.77 times the maximum probability as it performed the best.

d) Then the binary image is analyzed using connected component analysis using bwconncomp. Spurious noises are removed by checking the number of connected pixels; components with less than 6 pixels or less than 40 percent of the largest component, are removed.

e) Now the components are analyzed for shape factors for depth detection. Objects which have an eccentricity close to that of the mean eccentricity in the training data are selected. Learning algorithms like regression and logistic regression produced unsatisfactory results in situations where the barrel is occluded.

f) The best technique to depth detection was the nearest neighbor, where the chosen neighbors are the centroids of the shape properties of the barrels at different depths. This hard classification produced better results than, soft classification on the training set. The closest neighbor to the major and minor axes is found and the axis with the minimum Euclidean distance is chosen as the norm for the hard classification for the distance.

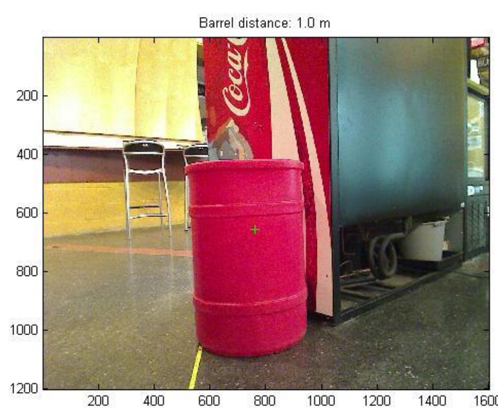
g) In order to account for images which are half cut like the ones where there is a pillar in front of the barrel at the middle. I find the find connected components which are at least 80 percent of the largest component and if they are a major axis distance away, I combine the two as a single component.

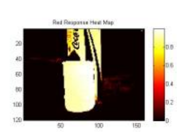
Results

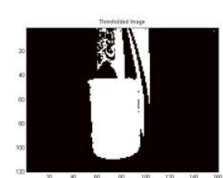
The algorithm works very well for most of the images. The detection is excellent in most of the situation but it fails in a few cases where objects of similar colors are present right next to the image or when we have reflections. We cannot eliminate the regions of unnatural shape as it would affect the response to occluded objects.

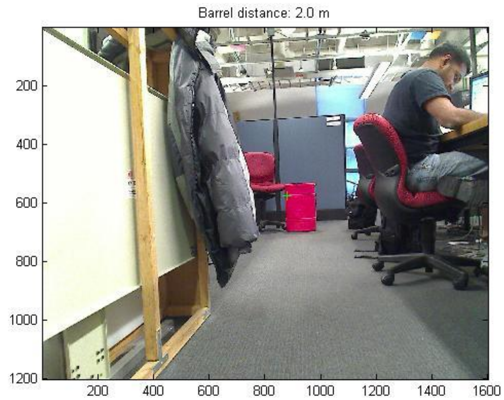
The depth detection works exceedingly well in most cases except a few cases where the chosen bounding box is wrong as the image may contain similar colored objects right next to the barrel.

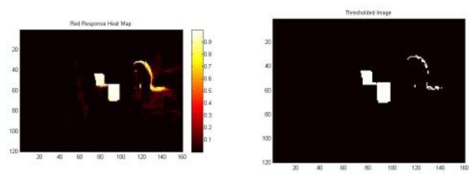
Below are some cases where it fails.

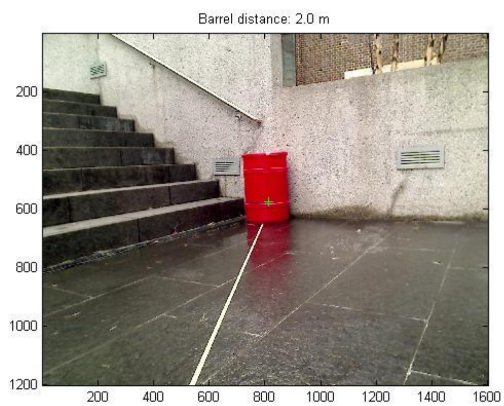


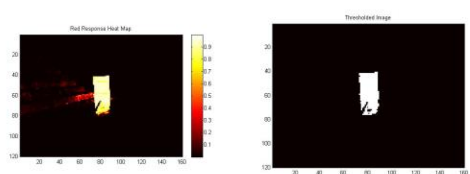






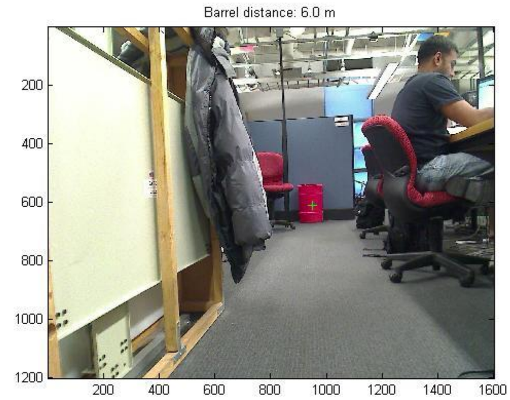


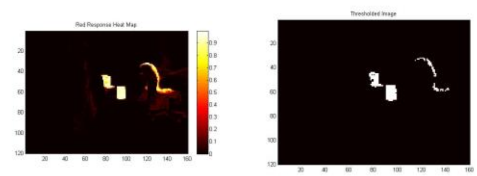


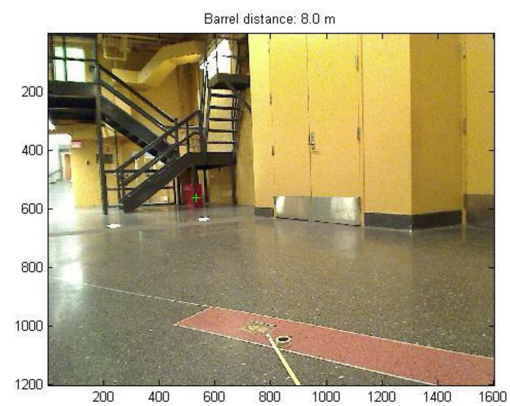


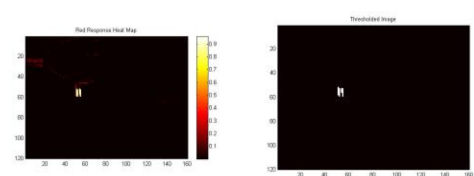
So the algorithm fails in scenarios where there is an object of the same color right next to the barrel. Even if there is a slight gap between the objects, algorithm can differentiate correctly.

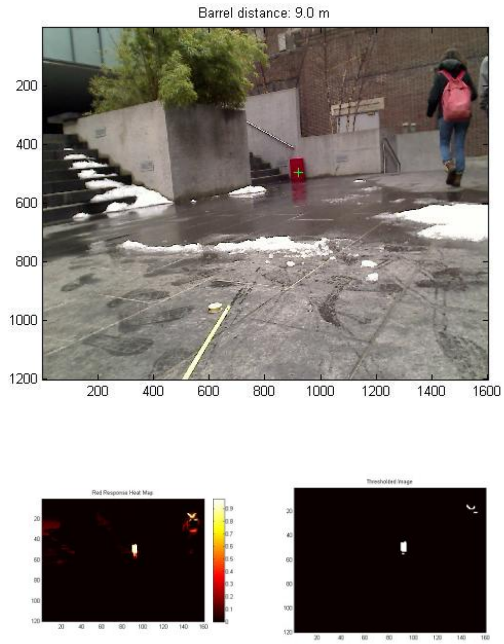
Here are some results where the algorithm works,

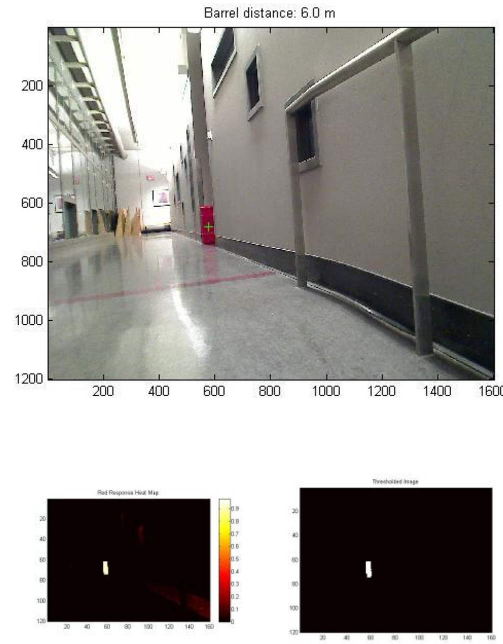


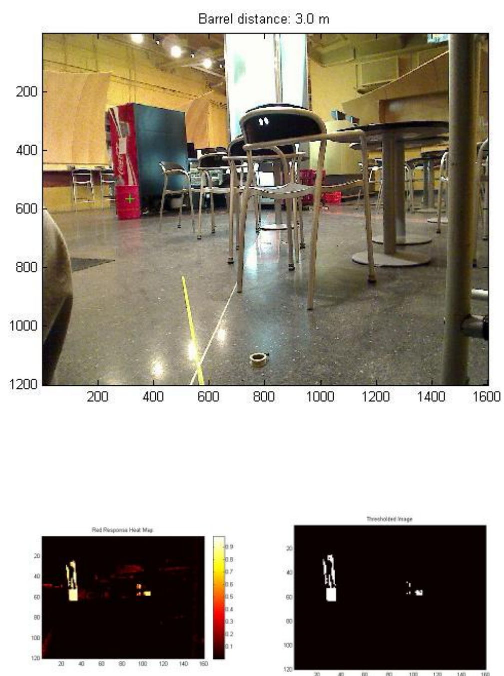


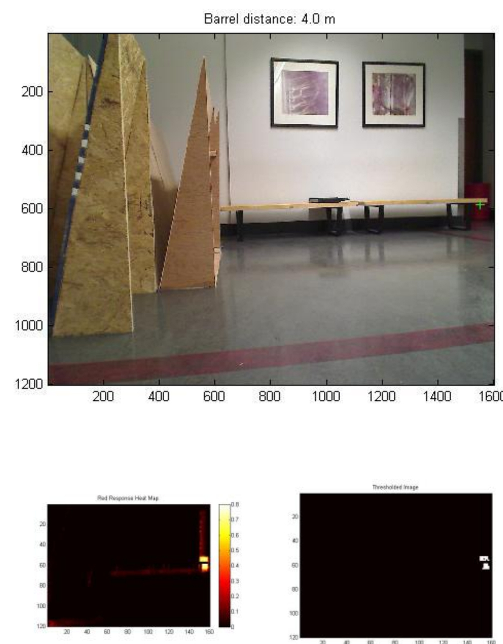


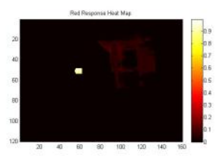


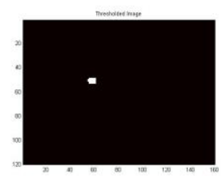

Barrel distance: 9.0 m
Red Response Heat Map
Thresholded Image

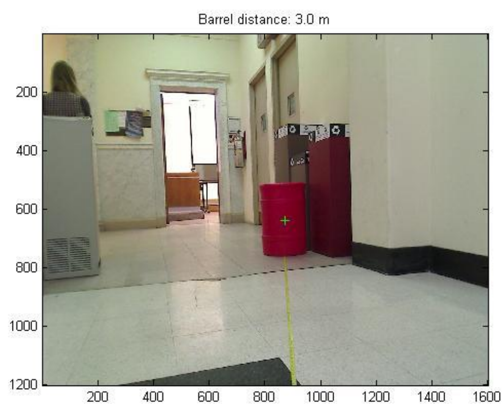

Barrel distance: 6.0 m
Red Response Heat Map
Thresholded Image

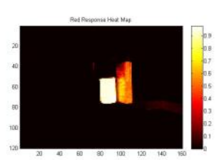

Barrel distance: 3.0 m
Red Response Heat Map
Thresholded Image

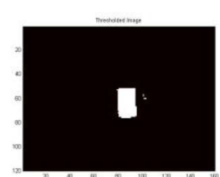

Barrel distance: 4.0 m
Red Response Heat Map
Thresholded Image











We see that even in the cases where the barrel is occluded or half visible and in cases where the barrel obstructed by a pillar or a table the algorithm detects the depth perfectly.