

ESE 650 Learning in Robotics, Spring 2016

Project 1: Barrel Detection

Due: Thursday, Jan 28th

Yiren Lu

Introduction

In this project, we are provided with a training set of 51 images, including 52 barrels in them. We are going to hand label the barrels area and use this data set to do color pixels (HSV) classification. And then use connected component analysis to select the barrel area and the center. Finally we use the features of detected region to estimate the distance of the barrel.

1 Bayesian Method

Our aim is to learn the distribution:

$$\mathbf{Pr}(\mathbf{C}|\mathbf{X})$$

Where $C = \{ "Barrel", "Others" \}$. \mathbf{X} are 3 dimensional HSV color pixels. Using Bayes' rule, we have:

$$\mathbf{Pr}(\mathbf{C}|\mathbf{X}) = \frac{\mathbf{Pr}(\mathbf{X}|\mathbf{C})\mathbf{Pr}(\mathbf{C})}{\mathbf{Pr}(\mathbf{X})}$$

Given an input color pixel vector \mathbf{x} , we can find a class $\mathbf{c} \in \mathbf{C}$ that maximizes the probability:

$$c = \operatorname{argmax}_c \mathbf{Pr}(\mathbf{c}|\mathbf{x})$$

To learn the distribution, for each class $\mathbf{c} \in \mathbf{C}$, given the training set \mathbf{X} , which's labels are \mathbf{c} , we want to maximize the likelihood

$$\mathbf{Pr}(\mathbf{X}|\mathbf{c})$$

2 Unimodel Gaussian

First we adapt a single multi-variate Gaussian model. In this case the pixel values are 3 dimensional.

$$\Pr(\mathbf{x}|\mathbf{c}) = (2\pi)^{-\frac{k}{2}} |\Sigma|^{-\frac{1}{2}} \exp(-\frac{1}{2}(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu))$$

where, k is the dimension of the color pixel. In our case, we use HSV values, thus, $k = 3$.

Maximum Likelihood Estimation

Using MLE (maximum likelihood estimation), we can derive that:

$$\mu = \frac{1}{N_v} \sum_{v=1} \mathbf{x}^v$$
$$\Sigma = \frac{1}{N_v} \sum_v (\mathbf{x}^v - \mu)(\mathbf{x}^v - \mu)^T$$

where N_v is the number of training samples in class \mathbf{c} .

3 Gaussian Mixtures Model

In order to gain more robust performance, instead of using a single multivariate gaussian model, I use Gaussian Mixtures Model (GMM).

$$Pr(\mathbf{x}|c) = \sum_{k=1}^K \pi_k N(\mathbf{x}; \mu_k, \Sigma_k)$$

Where π_k is the probability that a sample is drawn from k th mixture component, and $N(\mathbf{x}; \mu_k, \Sigma_k)$ is the gaussian model parameterized by μ_k and Σ_k . Note that $\sum_{k=1}^K \pi_k = 1$ and $\pi_k > 0$.

Expectation Maximization (EM)

I use Expectation Maximization (EM) to optimize the model. EM algorithm is guaranteed to converge to a local minimum.

First randomly initialize π, μ, Σ

E-Step:

$$\Pr(z_i = k|\mathbf{x}, \mathbf{c}) \propto \pi_k N(\mathbf{x}; \mu_k, \Sigma_k)$$

M-Step: Estimate new μ_k and Σ_k

$$\begin{aligned}\mu_k &= \frac{\sum_i \mathbf{Pr}(z_i = k | \mathbf{x}_i) \mathbf{x}_i}{\sum_i \mathbf{Pr}(z_i = k | \mathbf{x}_i)} \\ \pi_k &= \frac{1}{n} \sum_i \mathbf{Pr}(z_i = k | \mathbf{x}_i) \\ \Sigma_k &= \frac{\sum_i \mathbf{Pr}(z_i = k | \mathbf{x}_i) (\mathbf{x}_i - \mu_k)(\mathbf{x}_i - \mu_k)^T}{\sum_i \mathbf{Pr}(z_i = k | \mathbf{x}_i)}\end{aligned}$$

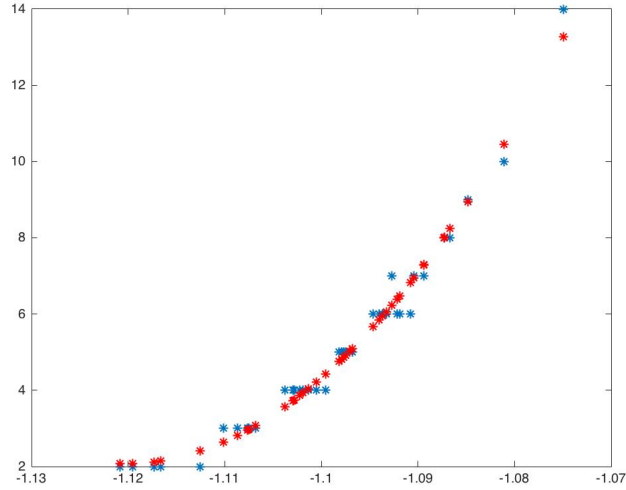
4 Object Detection

After classified the pixels, we have to detect which part of the red pixels is the actual barrel object. To gain robust performance, I first use **imfill** to fill out all the holes so that the detected region pixels are more solid. Then I tried image dilation and erosion. Unfortunately, in my model, dilation and erosion do not work so well. I use **bwlabel** to label detected regions and use **regionprops** to extract regions' centroids and boundingboxes. Then I exclude the regions with incorrect aspect ratio. Among the valid ranges, I pick the one with largest number of pixels as the barrel. And take the **convex hull** of the region as the barrel region.

Due to the time limitation of this project, I only detect **one** barrel in the image. While, my program could be easily extended to detect multiply barrels by selecting the top n largest regions and identify if each of them is actual barrel.

5 Distance Estimation

I model the distance and corresponding number of pixels in training data. By playing around the relationship between distances (*dist*) and number of barrel pixels (*npix*). I model the 2-d polynomial regression between $-npix^{0.01}$ and *dist*. This curve fits really well as shown in the graph below. The blue * are ground truth, and red * consist fitting curve.

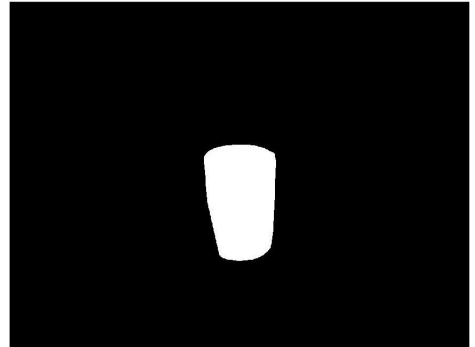
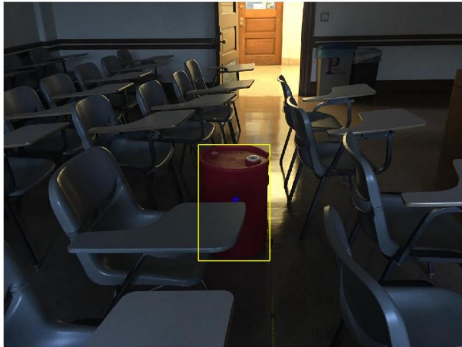


6 Code Execution

Please put test images into **Test_set** folder and execute **code.m** in the submitted project. My algorithm is in **detect_barrel.m**. This function has 5 return values. **[x,y,d,BW,cc,bb]**. **x,y** is the centroid, **d** is the distance. **BW** is the masked region. **cc** and **bb** are the corresponding centroids object and boundingbox objects.

The code will automatically show the detected region and the centroid and print the distance to the matlab **command line**.

The image labeling and model training are in **labelimages.m** and **test_models.m** respectively.



Estimated distance:

d =

|

2.2476