ESE 650 Project 1: Color Segmentation

James Yang

January 29, 2015

1 Introduction

Color segmentation is a very useful method for identifying objects in robotics. While basic color thresholding is capable of partitioning a digital image, this method for color segmentation is not robust to varying lighting conditions or relatively homogeneous color scenes. Employing machine learning to color segmentation offers a far more robust method for segmenting colors in different lighting conditions. For this project, we were tasked with implementing machine learning and computer vision algorithms capable of partitioning RGB images and extracting the pixel location and actual distance of a red barrel. An ensemble of Gaussian Mixture Models was implemented for this particular task.

2 Data Collection

A series of training images were provided, each depicting a red barrel in various lighting conditions and juxtaposed with similarly shaded red objects. The images were converted to the YCbCr color space so as to separate color and luminance information. Barrel colors were hand-masked from each image to create two color classes: "barrel" and "not barrel".

3 Gaussian Mixture Model Training and Barrel Detection

Three separate GMMs were trained as binary pixel classifiers. Each GMM modeled a unimodal 3D Gaussian for each cluster. Using the EM algorithm, pixels of interest were clustered by color and subsequently labeled as "barrel" or "not barrel. Then, based on the classifications, we checked connected components on a binary image to observe all the potential objects in the image that could be a barrel. From there, we had two metrics to check if an object was a barrel or not. The first was a minimum pixel area within a bounding box to ensure that small or sparsely filled objects would not show up as false positives. The second was a restriction that the width and height ratio of each bounding box should come within 25% of the actual width to height ratio of a barrel which is $\frac{40}{57}$.

The first GMM trained on four clusters for barrel colors and two clusters for non-barrel colors. The second GMM trained on two clusters for barrel colors and two clusters for non-barrel colors. Finally, the third GMM trained on three clusters for barrel colors and one cluster for non-barrel colors. Based on the type of training error, it was empirically determined that the first GMM performed well with medium lighting conditions, though was lacking in poorly and excessively exposed images. The second performed well in low lighting conditions, but not so much in bright or saturated images. Finally, the third model performed well in bright images, but not as well in darker images. Combined, these models form a best-first ensemble method, meaning that if a barrel was not found using the first GMM's classification scheme, barrel detection is attempted with the second GMM's classification scheme. Likewise, the second GMM's classification scheme failed to produce a barrel location, the third GMM is used. The checks always occur in this order by virtue of each GMM's error rate. Combined, these GMMs were expected to cover a broad spectrum of lighting conditions.

4 Depth Estimation

Depth estimation was accomplished by estimating the focal length and sensor size of the camera. Since many of the pictures in the training set contained wet floors and windows, reflections caused all three GMMs to return false positives on pixel classification, more so on reflections from water than windows.

As a result, depth estimation was accomplished using the horizontal dimensions of the bounding box on each potential barrel object. By averaging labeled distances and the horizontal dimension of each barrel, distinguishing between the actual height and width of the barrel, we were able to produce a fairly accurate scaling factor converting horizontal width and distance.

5 Future Work

While the detection worked on more than 80% of the training images, two issues have yet to be addressed by this algorithm. The first is consolidating occlusions, and the second is eliminating false positives. While more robust color segmentation can always be achieved with more data, the vision portion of the detection algorithm can ultimately make a large difference. For instance, if two objects very close to each other would, combined, match the dimensions of a barrel, it would be reasonable to assume the presence of a barrel as well as an obstacle in front of the barrel. Furthermore, in the presence of many parallel lines, it would be nonsensical for a barrel to appear above the horizon and not directly above a parallel line, implying that the barrel is sitting on a ledge. Finally, it is not unreasonable to find a barrel that is rotated about an axis normal to the image plane of the camera, meaning that the barrel could reasonably be in any orientation as well, perhaps leaning against a wall or the picture was taken at such an angle. These are several issues that, if addressed, can make an enormous difference in robustness.

6 Running the Code

Place images inside working directory and run es$_{d}etection.m$.