ECE 276A HW1 Report

Blue Barrel Detection

Zhaoliang Zheng
PID: A53254665
E-mail: zhz503@eng.ucsd.

***Abstract*—In this assignment, we are going to use hand-label examples of different colors. We are given a set of training images, in this assignment, we are given 46 images. We should build color classifiers for several colors and finally build a blue barrel detector. My solution for this assignment is to build a GMM algorithm to classify my barrel blue and then contour it. Besides that, I will used some filter to enhance my barrel areas for better detection and draw a bounding box.**

Keywords—Logistic Regression, K-mean, EM algorithm, Gaussian Mixture Model, Color Space

I. Introduction

Artificial intelligence is becoming more and more popular all over the world. This advanced technique has been wildly applied in natural language processing, strategic games, autonomously operating vehicles and so on.

Machine learning, as the main scientific approach, play a very crucial role in all these applications. Machine learning is a machine capable of miming "cognitive" functions and solving problems, for example: feature matching, color segmentation and so on.

In this assignment, I'll use Expectation Maximization to fit in Gaussian Mixture Model (GMM) on the training set. I will first specify K classes for this problem and then develop an EM algorithm to solve the GMM model. The input is a bunch of validation set we separate from the training pictures. The output will be a mask binary picture. After we obtain the mask images, we then set some criterion to draw a bounding box for the barrel in the image. Color segmentation and object identification are two crucial aspects of the machine learning process which could benefit a lot in our future work as well as in daily life.

II. Problem Formulation

A. Color Segmentation

In order to do color segmentation, we need to first specify the K classes for our pictures. We can divide classes K as follows:

$$K = \{normal - barrel\ blue, bright - barrel\ blue, dark - barrel\ blue, \ \dots\}$$

The feature in the pictures can be represented as x_i ($i = 1 \dots n$), more specifically, x_i is the color information of each pixel points in the training, validation and testing dataset.

So the whole problem can be simplify as follow: For the given training data set X_i, we want to find its corroding class Y_i to maximize the probability of X given Y with the parameter θ, which is $p(x, y|\theta, \omega)$.

Then, the finial target is to

$$\max_{\theta,\omega} p(x, y|\theta, \omega) = p(y|\theta)p(x|y, \omega)$$

B. Barrel detection

In this assignment, we are required to draw a bounding box based on what we get from the previous assignment. The only information we have is color information with the label true or false.

Under given mask, I used skimage.measure.regionprops to obtain for the further analysis and selection of barrel. In order to identify the barrel. I set some criterion for this mission.

First, the regions are smaller than 600 pixels could be regarded as noise, because they are too small:

$$\text{if}(\text{region. area} \leq \ 600)\text{: continue}$$

Second, to make sure that the bounding box is on the barrels, we also need to specify the ration of height and width. In this case, I use the range of 1.1 to 2 as the bounding box as for the shape of battels:

$$\text{if}\left(\frac{\text{height}}{\text{width}} \geq 2.2 or \frac{\text{height}}{\text{width}} \leq 1.05\right)\text{: continues}$$

Last, the bounding box should be proper edge for barrels, it is better to set the ration of region area and bounding box area:

$$\frac{region\ area}{bbox\ area} > 0.3$$

Consider some part of barrel may be hidden by other objects, so set this ratio to be 0.3

III. Technical Approach

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar.

Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph. Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

A. Hand-label Training data

After we specify For color segmentation purpose, it is better to consider and manipulate pixels in HSV space rather than in RGB space. In most of cases, the color of objects in the pictures can be greatly affected by lighting conditions.

B. Color Space

RGB color space is a very common color representation, we can first try our model by using RGB color space.

For color segmentation purpose, it is better to consider and manipulate pixels in HSV space rather than in RGB space. In most of cases, the color of objects in the pictures can be greatly affected by lighting conditions.

HSV space is actually the cylindrical coordinates of RGB representation. In HSV space, H is hue, angular dimension, S is saturation, and V is Value/Brightness. Since it separates color information from brightness or lightning, it is wise to convert from RGB space to HSV space by using the Opencv function cv2.COLOR_BRG2HSV。

C. *Data proprocessing*

Data preprocessing is necessary to do at the beginning. In this step, we can save our images into RGB file for future use.

Meanwhile, we should separate our training data into two set. One part is training set for training our model, the other part is called validation set. In my example, the ratio between these two sets is:

$$\text{Train}:\text{Validation} = 4:1$$

D. *Logistic Regression*

- Sigmoid Function:

$$\sigma(z) = \frac{1}{1+\exp(-z)}$$

- Logistic regression probability calculation :

$$p(y|x,\omega) = \prod_{i=1}^{n} \sigma(y_i x_i^T \omega)$$

- Use MLE algorithm to calculate global minimum:

$$\omega_{\text{MLE}}^{t+1} = \omega_{\text{MLE}}^{t} + \alpha \sum_{i=1}^{n} y_i x_i (1 - \sigma(y_i x_i^T \omega))$$

E. *GMM*

- Expectation Maximization is an iterative maximization technique based on auxiliary lower bounds. It is an MLE algorithm to find the lower bound estimation method to fit in GMM model. We have to make a initial guess about the model parameters: $\theta_k : \mu_k, \Sigma_k, \alpha_k$. Here, we used K-mean algorithm to obtain initial μ_k.
- K-means algorithm is an iterative clustering algorithm that uses coordinate descent to solve the following optimization:

$$\min_{\mu,r} C(\mu,r) := \sum\nolimits_{i=1}^{n} \sum\nolimits_{j=1}^{j} r_{ij} \left\|\mu_j - x_i\right\|_2^2$$

- In the EM algorithm, I used Estimation (E) step and Maximization (M) step until we reach the convergence of log likelihood $p(x,y|\theta,\omega)$.
- At E-step, for each data point, the membership probability has to be estimated:

$$r_k(j|x) := \frac{\alpha_{kj}\phi(x;\mu_{kj},\Sigma_{kj})}{\sum_{l=i}^{j} \alpha_{kj}\phi(x;\mu_{kj},\Sigma_{kj})}$$

- Then in the M-step, the new parameters would be estimated and updated by using the following equations:

$$\alpha_{kj}^{t+1} = \frac{\sum_{i=1}^{n} 1\{y_i = k\} r_k^t(j|x)}{\sum_{i=1}^{n} 1\{y_i = k\}} = \frac{1}{n}\phi(x;\mu_{kj},\Sigma_{kj})$$

$$\mu_{kj}^{t+1} = \frac{\sum_{i=1}^{n} 1\{y_i = k\} r_k^t(j|x) \cdot x_i}{\sum_{i=1}^{n} 1\{y_i = k\} r_k^t(j|x)}$$

$$\Sigma_{kj}^{t+1} = \frac{\sum_{i=1}^{n} 1\{y_i = k\} r_k^t(j|x)\, diag\left(x_i - \mu_{kj}^{t+1}\right)^2}{\sum_{i=1}^{n} 1\{y_i = k\} r_k^t(j|x)}$$

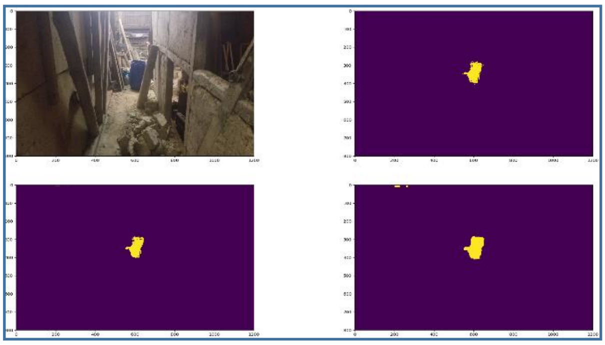
- After we figure out the parameters, we can use log likelihood for a single pixel under GMM model is obtained as:

$$\log \sum_k p(x,y|\theta,\omega) = \log \sum_k \alpha_k \cdot \phi(x;\mu_{kj},\Sigma_{kj})$$

F. *Smart refine mask for barrel detection*

Sometimes the mask we create is not that good in some cases, especially in the environment like there are lots of identical blue background or some part of it are blocked by other objects.

In my algorithm, I design a refine_mask function by using cv2. morphologyEx and cv2. Dilate. As you can see from the performance from the above:

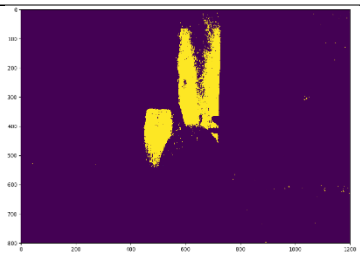


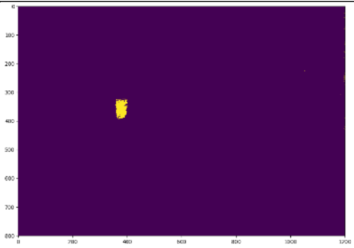
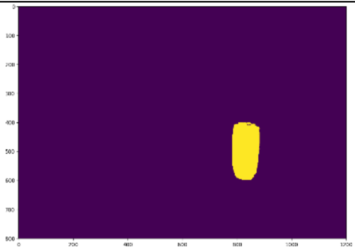
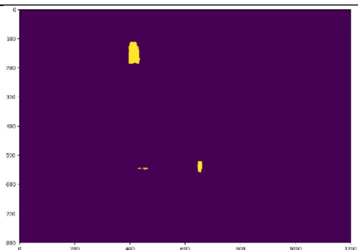
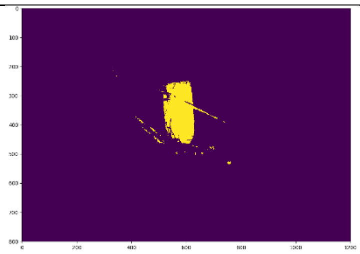
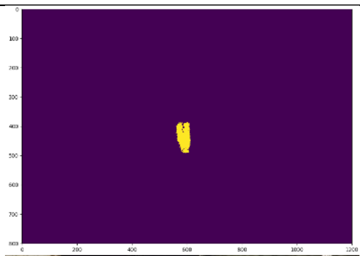
IV. RESULT AND DISCUSSION

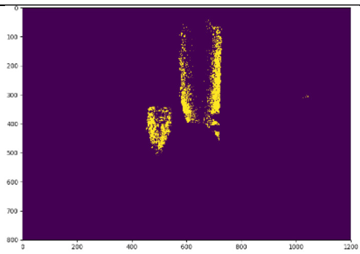
In the result of the barrels detection, I used mask to do barrel detection and used green rectangular boxes to circle the barrel region.

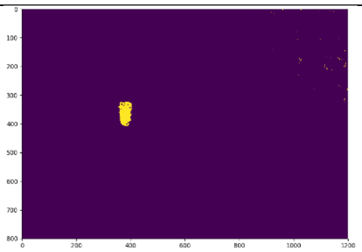
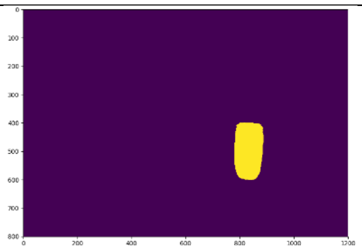
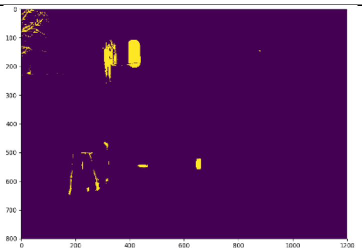
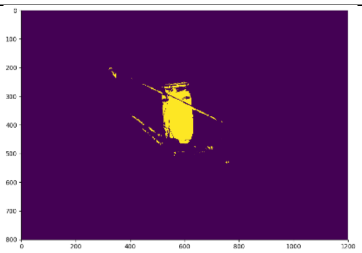
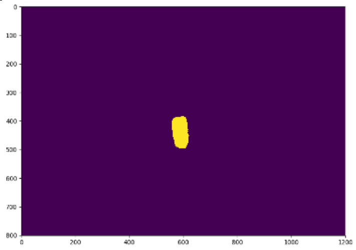
Since I have used two different models, one is Logistic Regression, the other is Gaussian Mixture Model. So I will present my validation result from both models.

A. *Validation set Comparison (LG vs. GMM)*

LG_Barrel detection result		
No.	***Detected images***	***B_box info***
1		$x1 = 453$ $y1 = 343$ $x2 = 565$ $y2 = 545$

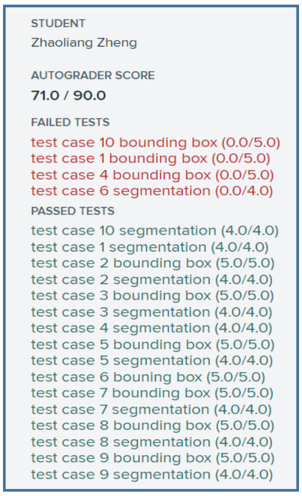
		
2	 	$x1 = 362$ $y1 = 330$ $x2 = 413$ $y2 = 403$
3	 	$x1 = 785$ $y1 = 403$ $x2 = 895$ $y2 = 610$
4	 	$x1 = 399$ $y1 = 114$ $x2 = 447$ $y2 = 199$ And $x1 = 648$ $y1 = 524$ $x2 = 674$ $y2 = 569$
5	 	$x1 = 564$ $y1 = 391$ $x2 = 623$ $y2 = 503$
6	 	$x1 = 564$ $y1 = 391$ $x2 = 623$ $y2 = 503$

GMM_Barrel detection result		
No.	***Detected images***	***B_box info***
1	 	$x1 = 460$ $y1 = 344$ $x2 = 555$ $y2 = 502$

2	 	$x1 = 360$ $y1 = 326$ $x2 = 415$ $y2 = 418$
3	 	$x1 = 783$ $y1 = 401$ $x2 = 899$ $y2 = 612$
4	 	$x1 = 397$ $y1 = 110$ $x2 = 451$ $y2 = 216$ And $x1 = 647$ $y1 = 523$ $x2 = 674$ $y2 = 570$
5	 	Detection failded!
6	 	$x1 = 560$ $y1 = 387$ $x2 = 632$ $y2 = 507$

B. *Test set Comparison (LG vs. GMM)*

LG several Result: (1st and best)

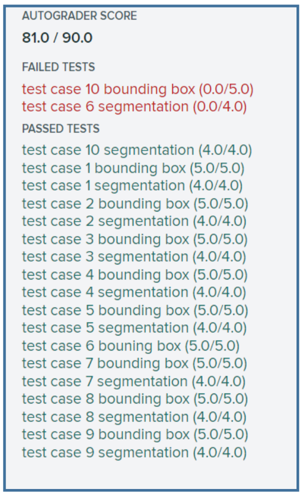


STUDENT
Zhaoliang Zheng

AUTOGRADER SCORE
71.0 / 90.0

FAILED TESTS
test case 10 bounding box (0.0/5.0)
test case 1 bounding box (0.0/5.0)
test case 4 bounding box (0.0/5.0)
test case 6 segmentation (0.0/4.0)

PASSED TESTS
test case 10 segmentation (4.0/4.0)
test case 1 segmentation (4.0/4.0)
test case 2 bounding box (5.0/5.0)
test case 2 segmentation (4.0/4.0)
test case 3 bounding box (5.0/5.0)
test case 3 segmentation (4.0/4.0)
test case 4 segmentation (4.0/4.0)
test case 5 bounding box (5.0/5.0)
test case 5 segmentation (4.0/4.0)
test case 6 bouning box (5.0/5.0)
test case 7 bounding box (5.0/5.0)
test case 7 segmentation (4.0/4.0)
test case 8 bounding box (5.0/5.0)
test case 8 segmentation (4.0/4.0)
test case 9 bounding box (5.0/5.0)
test case 9 segmentation (4.0/4.0)

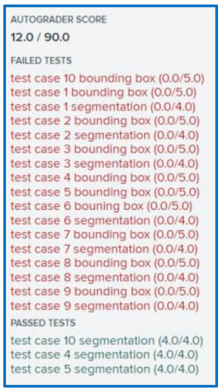


AUTOGRADER SCORE
81.0 / 90.0

FAILED TESTS
test case 10 bounding box (0.0/5.0)
test case 6 segmentation (0.0/4.0)

PASSED TESTS
test case 10 segmentation (4.0/4.0)
test case 1 bounding box (5.0/5.0)
test case 1 segmentation (4.0/4.0)
test case 2 bounding box (5.0/5.0)
test case 2 segmentation (4.0/4.0)
test case 3 bounding box (5.0/5.0)
test case 3 segmentation (4.0/4.0)
test case 4 bounding box (5.0/5.0)
test case 4 segmentation (4.0/4.0)
test case 5 bounding box (5.0/5.0)
test case 5 segmentation (4.0/4.0)
test case 6 bouning box (5.0/5.0)
test case 7 bounding box (5.0/5.0)
test case 7 segmentation (4.0/4.0)
test case 8 bounding box (5.0/5.0)
test case 8 segmentation (4.0/4.0)
test case 9 bounding box (5.0/5.0)
test case 9 segmentation (4.0/4.0)

GMM test Result:



AUTOGRADER SCORE
12.0 / 90.0

FAILED TESTS
test case 10 bounding box (0.0/5.0)
test case 1 bounding box (0.0/5.0)
test case 1 segmentation (0.0/4.0)
test case 2 bounding box (0.0/5.0)
test case 2 segmentation (0.0/4.0)
test case 3 bounding box (0.0/5.0)
test case 3 segmentation (0.0/4.0)
test case 4 bounding box (0.0/5.0)
test case 5 bounding box (0.0/5.0)
test case 6 bouning box (0.0/5.0)
test case 6 segmentation (0.0/4.0)
test case 7 bounding box (0.0/5.0)
test case 7 segmentation (0.0/4.0)
test case 8 bounding box (0.0/5.0)
test case 8 segmentation (0.0/4.0)
test case 9 bounding box (0.0/5.0)
test case 9 segmentation (0.0/4.0)

PASSED TESTS
test case 10 segmentation (4.0/4.0)
test case 4 segmentation (4.0/4.0)
test case 5 segmentation (4.0/4.0)

Discussion:

From the validation set that I set for two different models.

(1) The color segmentation:

Color segmentation ability seems that LR model is slightly better than the GMM model. I think the main reason is set the training labeled images that I choose for LR are purer than GMM. The training labeled images for GMM are over 35 pictures, and the total pixels are more than 700000, and also, I used HSV color space for GMM, in this way, GMM will detect more black colors and preforms better in the dark environment, as you can see from the picture (2) and (6).

(2) The bounding box detection:

Since the bounding box detection is fully based on color segmentation, on this ability, LR is also better than GMM as well.

If I change different parameters, and training set, this will greatly affects the performance of color segmentation as well.

CONCLUSION

In this assignment, I mainly focus on K-mean algorithm, EM algorithm in solving Gaussian Mixture Model. By using this classifier, we used our label images and label color to train our model. After training, we can get the parameters, with this parameters, we can utilize these parameters to do color segmentation on other new images. The bounding box can help us to identify the target region we need.

BIOGRAPHY

Zhaoliang Zheng Received his Bachelor degree of the Process Equipment and Control engineering in Dalian University of Technology. He is currently a UCSD Master student in Mechanical and Aerospace (MAE) Engineering. He is also the lab member of Coordinated Robotics LAB and Drone lab. His research interest are path planning, UAV application development, Embedded systems, Robotics design, control algorithm and SLAM.