Image Processing and Computer Vision Coursework

Jeonghyun Kim

Candidate Number: 97485

This report has an aim to handle Image Processing and Computer Vision coursework that focuses on understanding basic theories about hough transform, haar-like features and more other filters or detectors to increase performance of finding objects from image files. This paper is divided into four parts from Subtask 1 to Subtask 4. The report is written on only one page per each section.

1 Subtask1

a) Results of face.cpp for dart4.jpg, dart5.jpg, dart13.jpg and dart14.jpg.

Figure 1: *dart4.jpg*

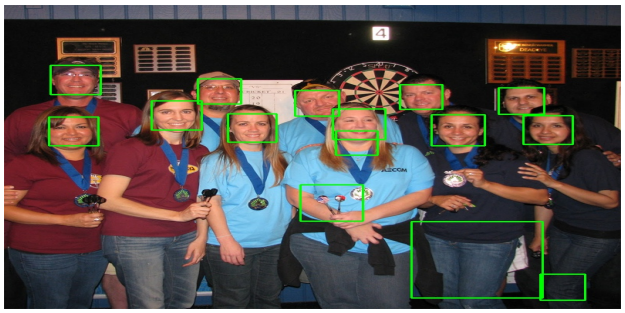

Figure 2: *dart5.jpg*

b) To compare dart5.jpg and dart15.jpg with True Positive Rate (TRP) called sensitivity, it is required to know such as True Positive (TP) and False Negative (FN). With dart5.jpg and dart15.jpg, TPR is 100% for the dart5 and 66.7% for the dart15, because the formula of TRP is,

$$\mathbf{TPR} = \frac{\mathbf{TP}}{\mathbf{TP} + \mathbf{FN}}$$

Figure 3: *dart14.jpg*

Figure 4: *dart15.jpg*

1) When the object detector guesses all faces in an image file, TPR will be shown with high percentage no matter what the number of boxes the machine predicts. It means that TRP never considers False Positive (FP), so this cannot represent precise accuracy as an indicator.

2) If the machine predicts faces a lot, not considering whether the prediction is right or not, TPR can be 100% without any difficulties.

3) Dart5.jpg and dart15.jpg are selected to get F1-score. F1-score is,

$$\mathbf{F_1 score} = \frac{\mathbf{2TP}}{\mathbf{2TP} + \mathbf{FP} + \mathbf{FN}}$$

which means that it is a harmonic mean of precision and sensitivity. Here all of the measurement are based on human annotation. For dart5.jpg, F1-score is 0.846154, and dart15.jpg is 0.571429. It shows better accuracy than use TRP only, because it includes mean of sensitivity and precision, including FP.

2 Subtask2

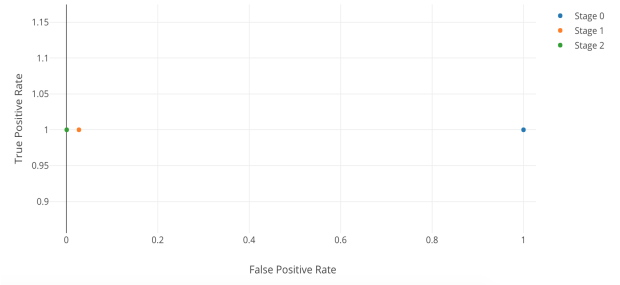

Figure 5: *TPR vs FPR scatter graph*

a) While the number of stages progresses, error rate is decreased as above the graph. TPR is always 1. In other words, False Negative (FN) is 0 because finding FN will stop. This means, there are not the cases that positive sample is mistakenly classified as negative during the whole stages progress. From a FPR's point of view, FP is decreased as the stages progresses. Therefore, this result shows the higher the stage, the better performance, which avoids the cases that a negative sample is classified as positive.

b) Four test images are below.

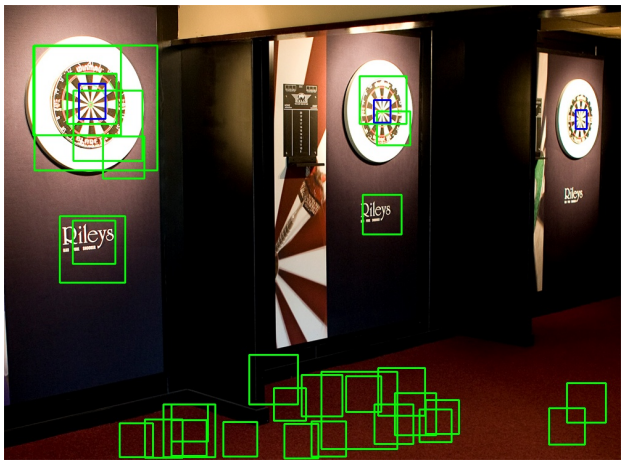

Figure 6: *dart10.jpg*

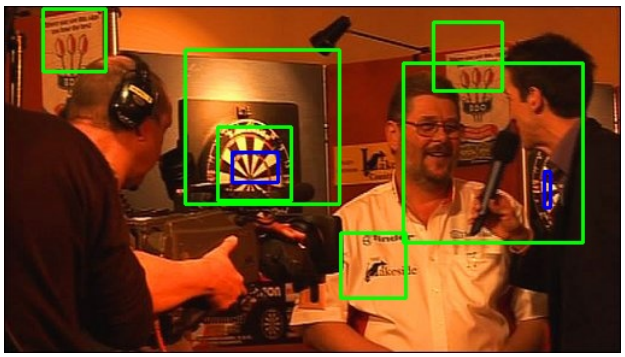

Figure 7: *dart11.jpg*

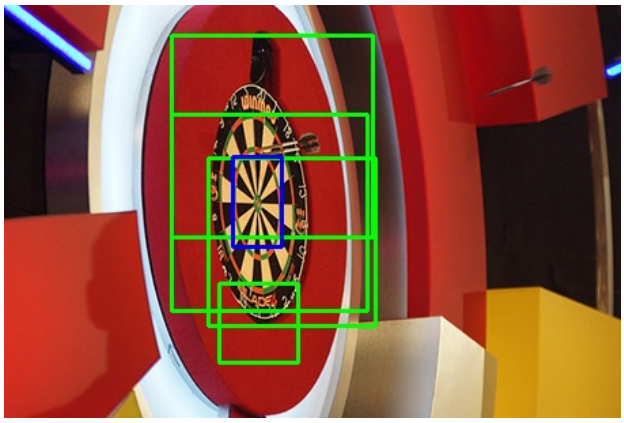

Figure 8: *dart12.jpg*

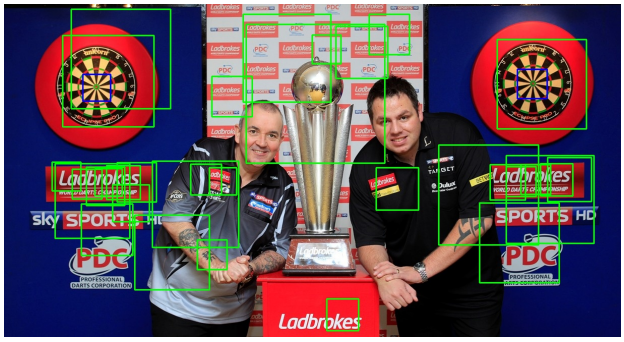

Figure 9: *dart14.jpg*

Green boxes are detected by the machine, and blue boxes are drown based on data "DartDataCommaBased.csv", which is collected answers from perspective of a person. All F1-score from Figure 10 is computed grounded the answer data.

Image	dart0	dart1	dart2	dart3	dart4	dart5	dart6	dart7
F1-score	0.117647	0.4	0.166667	0.285714	0.222222	0.222222	0.181818	0.1
Image	dart8	dart9	dart10	dart11	dart12	dart13	dart14	dart15
F1-score	0.2	0.181818	0.0606061	0.5	0.4	0.153846	0.114286	0.5
Average	0.241715							

Figure 10: *F1-score for all images*

F1-score looks low around 0.24. Even if the machine can predict dartboard correctly, case of FP is many as above results. The performance of the new images, which is not given during training, is worse comparing with task a) because it is not same traning image data. In fact, performance of task a) for TPR is high as it used data for both of traning and test.

3 Subtask3

a) Sobel Filter and Hough Transform.

1) Results of Sobel magnitude with gradient.

Figure 11: *sobel magnitude x gradient of dart8.jpg*

Figure 12: *sobel magnitude x gradient of dart9.jpg*

2) Results of Hough Transform.

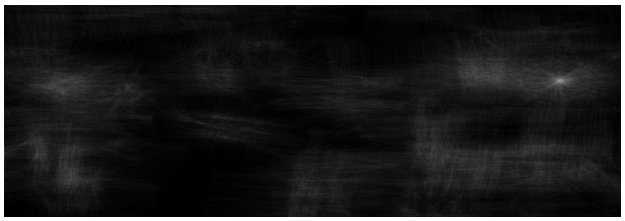

Figure 13: *hough transform of dart8.jpg*

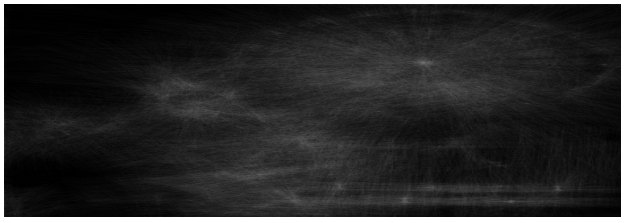

Figure 14: *sobel magnitude x gradient of dart9.jpg*

3) Results of detection.

Figure 15: *result of dart8.jpg*

Figure 16: *result of dart9.jpg*

b) Scores

Scores	Average
Precision	0.38
Recall	0.77
F1-score	0.52

Figure 17: *Final result of scores*

Merits:

- It has lower cases for FN as Recall score is high.
- Can distinguish circles from images.

Shortcomings:

- Still have many FP cases as Precision is low.
- Not considering density of brighter points from Hough space.
- Cannot detect ellipses and crushed circle.

c) Evidences

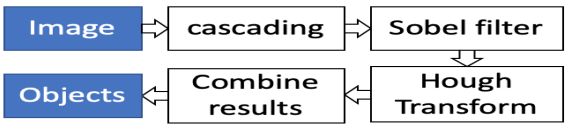

Figure 18: *Final result of scores*

There are two kinds of data to detect dart boards. First is prediction boxes from the result of cascading. It is decided by trained machine that have specific haar-like features. Second is Hough Transform (HT) data, which can detect circle. If combine this two evidences, it shows better performance than use each alone. Here are rationale that I combined,

- Even though Viola-Jones detector predicts wrong objects without any circles, HT can delete those wrong boxes, which means FP will be decreased.

- If HT is used, it is possible to know centre of circle. It means that it can distinguish between right and wrong detections near a dart board.

4 Subtask4

a) Approaches

- Image Enhancement - To get clear edge.
- Centre matching with HT and the detector - To get high performance of Precision score.

b) Merits

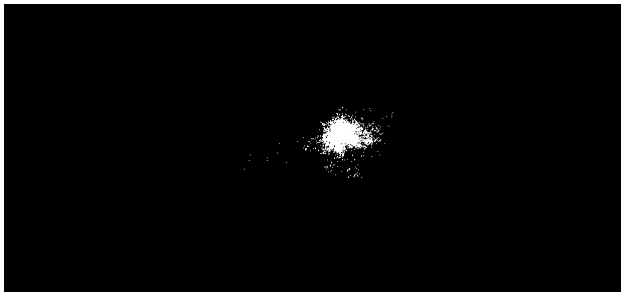

Figure 19: *hough transform of dart13.jpg with threshold*

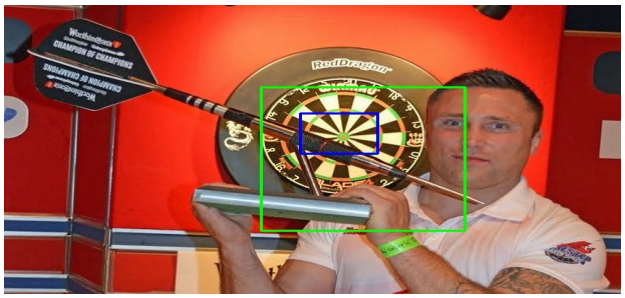

Figure 20: *result of dart13.jpg*

Figure 21: *hough transform of dart15.jpg with threshold*

Figure 22: *result of dart15.jpg*

Scores	Average
Precision	0.49
Recall	0.66
F1-score	0.52

Figure 23: *Final result of scores*

c) Evaluation
Merits

- Increase Precision, focusing on centre of hough space.
- High performance for dim or dark images.

Shortcomings

- Influenced by image lightening.
- Still cannot detect ellipses and semicircles.

Overall result is not changed before I had. The reason why results are similar is that I have tried to increase Precision score no matter what Recall is. Furthermore, image enhancement has considerable performance only for dark and dim images, but less accuracy for brighter images.
In conclusion, Precision and Recall score are inter related, so it needs to consider at a time for both of them. Moreover, if it is possible to choose coefficient of image enhancement depends on image quality, it will show better performance.