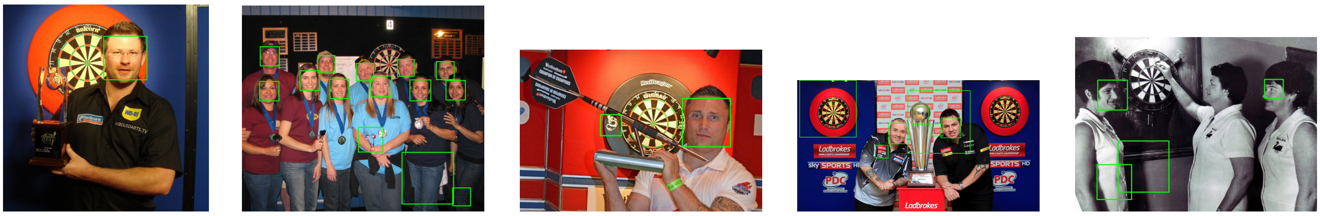
# 1 The Viola-Jones Object Detector

## 1.1 The Performance of the Detector for Human Faces

Five images were processed using the given Viola-Jones object detector to match human faces. The results below show detected face candidates, denoted by green bounding boxes.



(a) dart4.jpg (b) dart5.jpg (c) dart13.jpg (d) dart14.jpg (e) dart15.jpg

Figure 1: Result of the given algorithm detecting faces

## 1.2 Evaluating the Performance of the Face Detector

The True Positive Rate (TPR) on dart5 is 100%. The TPR on dart15 is less clear-cut. The decision comes down to two main factors - which parts of the image qualify as faces, and how well the bounding boxes fit the chosen faces. Given that the detector we are working with is looking for frontal faces, we decided to qualify a "face" in this context as only being acceptable when both of the eyes and the mouth are fully visible. As such, we ruled that the dart15 image didn't have any actual faces, and thus got a 100% TPR by default.

This problem we had assessing dart15 is one you must tackle when attempting to calculate the TPR of a detector - both assessing which objects count as positives in the first place, and how to determine if your detector has correctly identified them.

Of course, the TPR isn't the be-all and end-all when evaluating detector effectiveness - any TPR test can be "aced", no matter how stringent, by simply providing every possible group of pixels in the image as a detected positive. Among the countless results, you are guaranteed to include every positive. A better measure, then, is the F score. This uses both the detector's precision (correct positive results divided by all positive results) and recall (the TPR) as parameters. These are then weighted to decide which you consider more important: written as an Fn score, where n is the relative importance of recall to precision (F2 seeing recall as twice as important, for example).

To calculate the F1 score of our face detector, we created several rules to help determine when a positive result is correct. They are as follows:

1. A face is only considered valid if both of the eyes, as well as the mouth, are visible. It must be a human face. This defines what we consider a face.
2. The bounding box must encompass at least the eyes and the mouth fully, and cannot be more than 1.5x the dimensions of the tightest ideal box that does this. This is to give the correct positives a minimum and maximum size, to prevent overly large boxes detecting faces by "accident".

Written by: Team 82, Tom M. & Harry M-T.

# 2 Building & Testing the Dartboard Detector

## 2.1 TPR vs FPR on 3 Stages of Training Data

Figure 5 shows the three stages of training for our dartboard detector. Before training we can see that the FPR and TPR were 1, meaning that all objects were detected. We can see that our method of training was to detect everything first (represented with an FPR and TPR of 1), then as training progresses, detect less objects (represented with a lower FPR score). Rather than starting by detecting nothing and then trying to detect objects (increasing the FPR score). The graph shows that at each stage of training the FPR is reduced while keeping the TPR at 1.

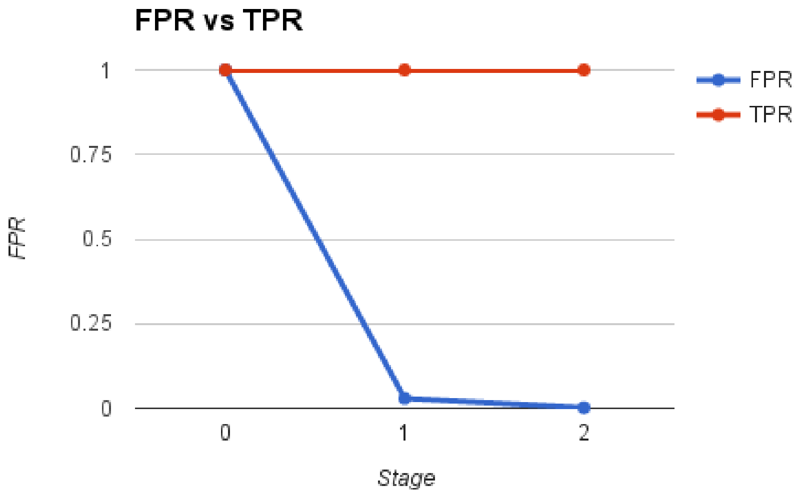



Figure 2: FPR vs TPR of the dartboard detector at each stage of training

## 2.2 Dartboard Detectors Performance

The F1 score was calculated for all 16 images (see figure 6), resulted in 0.224. The graph we plotted in figure 5 measures performance looking only at FPR and TPR. The F1 score looks at more factors, that are more important when measuring performance, such as the Precision and Recall. These are more important factors when predicting system performance because the F1 score shows the balance between precision and recall on actual data rather than the graph, that shows test data with other factors that affect the results.

(a) dart4.jpg (b) dart5.jpg (c) dart14.jpg (d) dart15.jpg

Figure 3: Result of the trained algorithm detecting dartboards

# 3 Integration with Shape Detectors

## 3.1 Two Example Dart Images showing Merit and Limitations of our implementation

Two example images were used (see figure 4) describing the Thresholded gradient magnitude image used as input to the Hough Transform, a 2D representation of the Hough space and the final result images showing the final detections using bounding boxes.

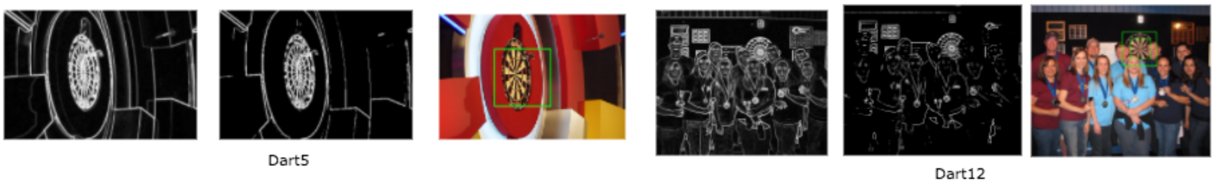



Figure 4: The Thresholded Gradient Magnitude, a 2D representation of the Hough space and the result of the final detection for twoDartboard images

## 3.2 Evaluation of our Detector on all example images

In the process of this coursework we ended up with two different configurations of the Viola Jones detector. Earlier in this report, we used the F1 score of one of these, but when the time came to compose our final detector, we ended up using the other, as although its F1 score was roughly half (.106 compared to .224), its recall was much greater (.95 vs .7). We made this decision due to our detector aiming to trim false positives away to leave true positives; so we wanted the maximum number of true positives to begin with.

The F1 score of our final detector was .586, a marked improvement on both original Viola Jones configurations. It achieved this by means of a far greater precision (.447, up from the .056 of the Viola Jones configuration we used as a base). It unfortunately regressed a little from the base in terms of recall (.85 vs the original .95), which is something we would aim to remedy when improving the detector further (although the precision score of .447 still has much greater room for improvement than the recall score).

Specific shortcomings of the final detector were that it had no way to remove overlapping boxes, and detected some false positives where there were only line endings or circles, but not both; both of these factors limit its precision. It would also be wise to attempt to find a Viola Jones configuration with a perfect recall score, as any true positives missed by the Viola Jones detector cannot be found by our refinement of it.

## 3.3 Flow Diagram Showing Evidence from Hough and Viola-Jones Detector

- We used the Viola-Jones detector to generate boxes around dartboards and other similar images. This produced lots of boxes and a high FPR.
- We combined this data with the Hough Circle Transform to remove boxes that didn't contain a centre of a circle, because we found dartboards were usually round.
- However, to catch slanted dartboards (dartboards that were not a perfect circle) and dartboards that were partly visible we needed to combine evidence from the Hough Line Transform because we found line endings usually clustered around the centre of a dartboard.
- We kept a box if it contained a centre circle, or if it contained more than 5 clustered line endings, to get a lower FPR.

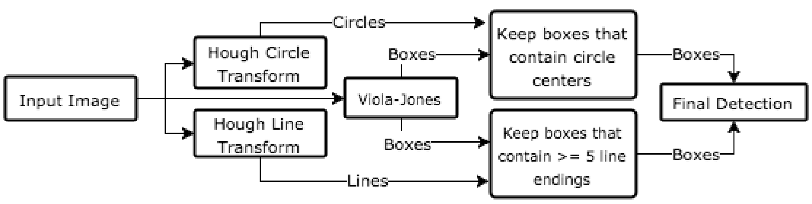



Figure 5: Flow diagram showing how we combined evidence from the Hough Transform and Viola-Jones detector.

Written by: Team 82, Tom M. & Harry M-T.

# 4 Improving your Detector

There were lots of methods discussed to improve the detector, but unfortunately we did not end with the time to implement any of them.

- Looking at choosing the best fitting box from any overlapping boxes to improve overall precision.
- Find a way to identify circles and ellipses (rather than just circles) and combine this information with an AND rather than an OR to pick up oval and squashed dartboards, so we are not just detecting just lines, or just circles. This would reduce the False Positive Rate.
- Find a Viola Jones detector configuration that has a perfect recall rate to use as our initial detector.
- Tweak settings on line and circle detection such that we are not losing any recall in our final detector.

# 5 References

## 5.1 F1 Score for all dartboard bounding box images

| Image | Recall | Precision | F1 Score |
|---|---|---|---|
| dart0.jpg | 1 | 0.25 | 0.4 |
| dart1.jpg | 1 | 0.5 | 0.67 |
| dart2.jpg | 1 | 0.05 | 0.09 |
| dart3.jpg | 1 | 0.33 | 0.5 |
| dart4.jpg | 0 | 0 | 0 |
| dart5.jpg | 1 | 0.17 | 0.29 |
| dart6.jpg | 0 | 0 | 0 |
| dart7.jpg | 1 | 0.07 | 0.13 |
| dart8.jpg | 1 | 0.10 | 0.17 |
| dart9.jpg | 1 | 0.25 | 0.4 |
| dart10.jpg | 1 | 0.08 | 0.15 |
| dart11.jpg | 1 | 0.33 | 0.5 |
| dart12.jpg | 1 | 1 | 1 |
| dart13.jpg | 0 | 0 | 0 |
| dart14.jpg | 1 | 0.07 | 0.13 |
| dart15.jpg | 1 | 1 | 1 |

Figure 6: The F1 Score for each Dartboard image, calculated using the Recall and Precision