Advanced Vision Assignment 2 Report

Mindaugas Dabulskis, Marat Subkhankulov

26th February 2013

1 Final Touch

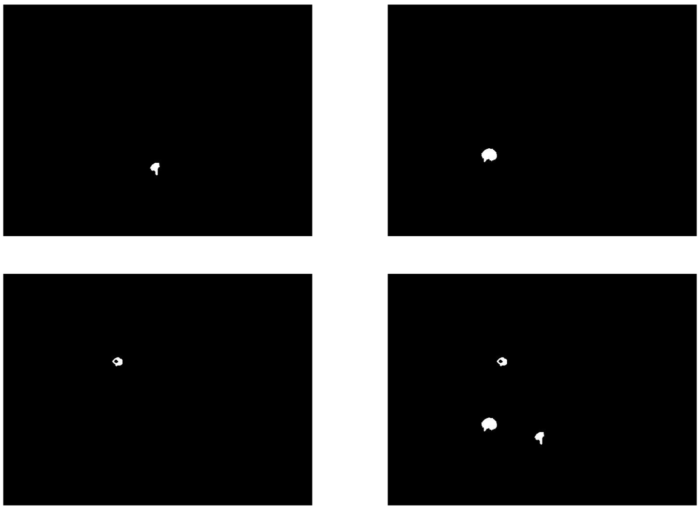

Figure 1: Cleaned images

Final clean up and ball identification was made using matlab **bwlabel** and **regionprops** functions. The function **bwlabel** takes our intermediate file, which we got after applying all the masks, and labels all the connected objects in the file. Later, this file is supplied to **regionprops** function, which measures and returns set of the properties for each connected object in the file. We only needed three properties:

1. *Area* - number of actual pixels in the region,
2. *PixelIdxList* - vector containing the linear indices of the pixels in region,
3. *PixelList* - matrix specifying the locations of pixels in the region.

The idea was:

1. to find maximum connected object in the file by the pixels area,
2. set all other connected object areas pixel to 0 (delete these areas),

3. find the middle point of the largest connected objects area.

Second point cleaned the image of all the noise, because the object with biggest area was the ball itself. Third point was needed for evaluation, to evaluate how much our detected ball differs from ground truth. We had four different approaches how to calculate the mass of the ball:

1. Use the mean of the area pixels,

2. use the median of the area pixels,

3. use the maximum and minimum pixels in **y** and **x** coordinates and find the mean between them. This creates a bounding box of the area and finds it centre.

The evaluation of these approaches are reported in the section below. We decided to use just normal mean, but left an option to change if needed and found that other approaches with different data sets works better. We stored all found centres in the matrix. The format of that matrix is exactly the same as the format for the ground truth. In next stage we just plot the found mass centres on each image with the centre from ground truth and draw the trajectory at the end by connecting all the points.

2 Experiments and results

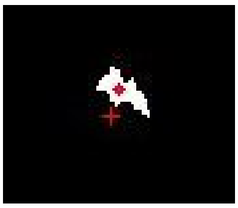

Figure 2: Image in which the distance between ground truth (+) and found mass centre (.) is bigger than 10 pixels

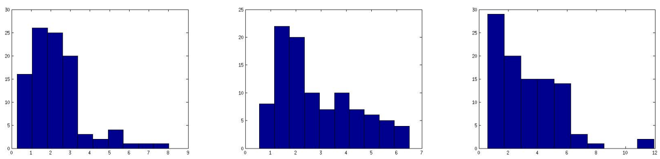

Figure 3: Cleaned images

The experiments shows that in all the image balls were detected. However, using the criterion that the detected ball mass centre should be more further than 10 pixels, we found out that in two images it is not the case. The main reason while the centre is further is because of the jugglers fingers. The fingers cover half the ball, thus it is not even hard to recognize the ball, but also impossible to get whole ball, because it is partly covered. Automatically, we only detec half the ball, resulting in shifted mass centre. Potential remedies would be somehow identify that the ball is in jugglers

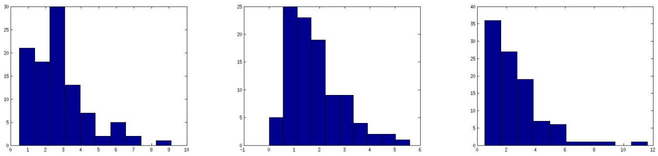

Figure 4: Cleaned images

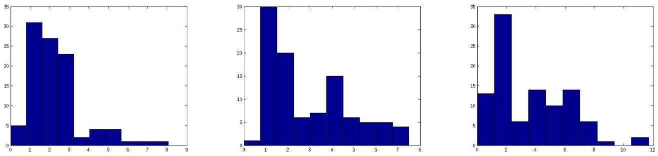

Figure 5: Cleaned images

Figure 6: Cleaned images

Figure 7: Cleaned images

hand and make bigger bounding box for the ball at those points. This way we would increase the area of the ball artificially, which would be more similar to actual ball size than what is recognized.

We have also evaluated different approaches to the mass centre calculation. From statistics we can see that bounding box approach has the smaller total error and have only one image in which found mass centre in yellow ball is further than 10 pixels, compared to other approaches, which have two images. This is a case, because in this approach we take account that half the ball might be covered by hand, thus we creating artificial area in which we predict that the ball is located. If we look into individual balls error, we would noticed that with bounding box approach red ball has the largest error compared with other approaches. This is because red ball has very distinct color and it is very easy to recognize, thus usually the ball is easily recognized. That is why in mean and median approach the red balls error is smallest. However, in bounding approach we creating artificial area, which might not be correct area of the ball, thus we make more mistakes. In conclusion, bounding box approach works great n balls, which can be hard to recognize, but does not work well with easily recognizable balls.

We decided not to use bounding box approach, because the drawn path is quite choppy and not as smooth as other two approaches.

Table 1: *Mean*

	Red	Green	Yellow
ball found in images	98/98	98/98	98/98
mass centre is less than 10 pixels from ground truth	98/98	98/98	96/98
maximum error	8.0093	6.5060	11.9423
minimum error	0.2799	0.5622	0.5995
one ball error	224.0428	276.9718	328.6480
total error	829.6625		

Table 2: *Bound*

	Red	Green	Yellow
ball found in images	98/98	98/98	98/98
mass centre is less than 10 pixels from ground truth	98/98	98/98	97/98
maximum error	9.1241	5.5902	11.6297
minimum error	0.5000 20	0	0.5000
one ball error	264.7848	189.4427	258.1291
total error	712.3566		

Table 3: *Median*

	Red	Green	Yellow
ball found in images	98/98	98/98	98/98
mass centre is less than 10 pixels from ground truth	98/98	98/98	96/98
maximum error	8.0623	7.5000	11.7047
minimum error	0	0	0
one ball error	227.6602	298.2528	363.9370
total error	889.8500		

2.1 References

References should be numbered in order of appearance, for example [1], [2], and [3]. You *can* use `bibtex` to prepare references, or do it by hand if there are very few.

References

[1] Smith, J. O. and Abel, J. S., "Bark and ERB Bilinear Transforms", IEEE Trans. Speech and Audio Proc., 7(6):697–708, 1999.

[2] Lee, K.-F., Automatic Speech Recognition: The Development of the SPHINX SYSTEM, Kluwer Academic Publishers, Boston, 1989.

[3] Rudnicky, A. I., Polifroni, Thayer, E. H., and Brennan, R. A. ”Interactive problem solving with speech”, J. Acoust. Soc. Amer., Vol. 84, 1988, p S213(A).