Image Mining

Object Tracking in Videos

Xiaoxuan HEI & Yuru HE

Q1 Experiment the tracking performed by the provided code *Tracking_MeanShift.py* that uses the basic Mean Shift algorithm, on the marginal density of the hue component H. Remind the principle of this algorithm and illustrate its advantages and limits by your experiments.

Advantages:

1) The calculation amount of the algorithm is not large, and the real-time tracking can be completely achieved when the target area is known.
2) The kernel function histogram model is adopted, which is insensitive to edge occlusion, target rotation, deformation and background motion.

In Figure 1, the ROI can be well tracked.

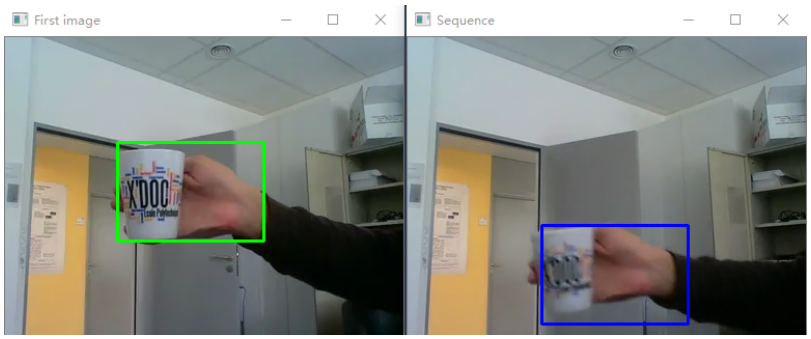

Figure 1 - Q1_1

Limits:

1) During the tracking process, since the window width remains unchanged, when the target scale changes, the tracking will fail.
2) When the target speed is fast, the tracking is not very good.
3) Histogram features are slightly lacking in the description of target color features, lacking spatial information.

In Figure 2 and Figure 3, the tracking areas change a lot.

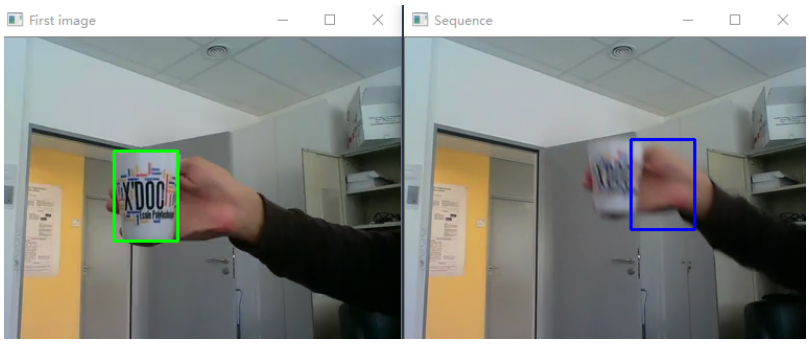

Figure 2 - Q1_2

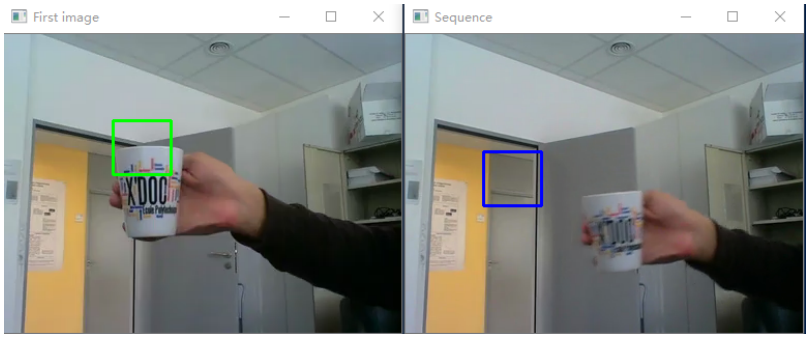

Figure 3 - Q1_3

Q2 Analyse more in-depth the result by displaying the sequences of hue images, and also the weight images corresponding to the back-projection of the hue histogram. Propose and program improvements, by changing the computed density and/or updating the model histogram.

- Improvement 1: At first, when we selected the character's sleeve, the ROI automatically jumped to other places. It's because Pixels with V < 20 are ignored. We can change this threshold in the mask function:

  mask = cv2.inRange(hsv_roi, np.array((0.,30.,20.)), np.array((180.,255.,235.)))

- Improvement 2: roi_hist = cv2.calcHist([hsv_roi],[0],mask,[256],[0,256])

  In this function, we can change the second parameter from [0] to [1], that's to say we use the distribution of saturation instead of hue.

- Improvement 3: we can give a filter to smoothen the histogram, for example, we can use Gaussian filter: gaussian = cv2.GaussianBlur(roi_hist, (5, 5), 1.1, 1.1)

Q3 Calculate for each frame, the local orientation, i.e. the gradient argument of pixels, and also the gradient magnitude. Use a threshold on the gradient magnitude to mask pixels whose orientation is not significant. Display the sequence of orientations, where the masked pixels appear in red.

For this question, first we calculate the coordinates of the gradient ***(Ix, Iy)***. Then the argument of the gradient corresponds to the angle of the gradient vector, that is given by ***arctan(Iy/Ix).*** The Pixels with gradient magnitude beneath 40 are not significant are in red. Here are two pictures for Antoine_Mug.mp4 and VOT-Ball.mp4.

Figure 4 - Antoine_Mug

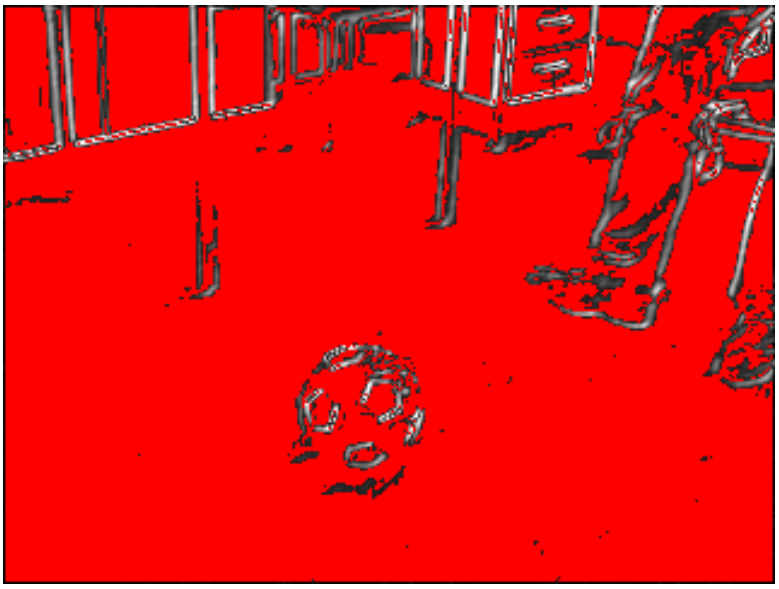

Figure 5 - VOT-Ball

Q4 Build a model of the initial object under the form of an implicit model indexed on the orientation (R-Table). Then, calculate the associated Hough transform on all the images of the sequence. Calculate the straightforward tracking, corresponding to the maximal value of the Hough transform at each

image. Comment and criticise the obtained results.

After calculating the Hough transform, the best detection function tries to find the highest value of the Hough space. In the detection window which shows the Hough space state, we can find the brightest point is the position where we have the object that we want to track.

From the result we can find that, the Mug can be tracked very well.

Figure 6 - Hough Transformation

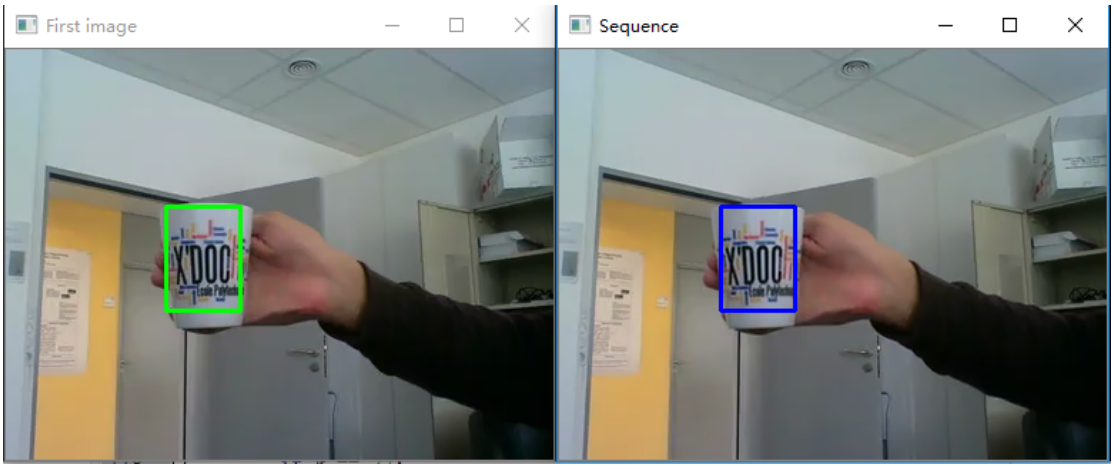

Figure 7 - Result of Argmax

Q5 Replace the computation of the maximal value by the application of the Mean Shift on the Hough transform (i.e. by replacing the back-projection of the Hue

histogram by the output of the Hough transform in the first argument of the Mean Shift). Interpret the result and compare it with the previous one. Propose an update strategy of the model that would allow to be robust to aspect changes of the object.

The ROI in question 1 changes very fast and once the video starts, it's easy to be at the wrong place, like the picture below:

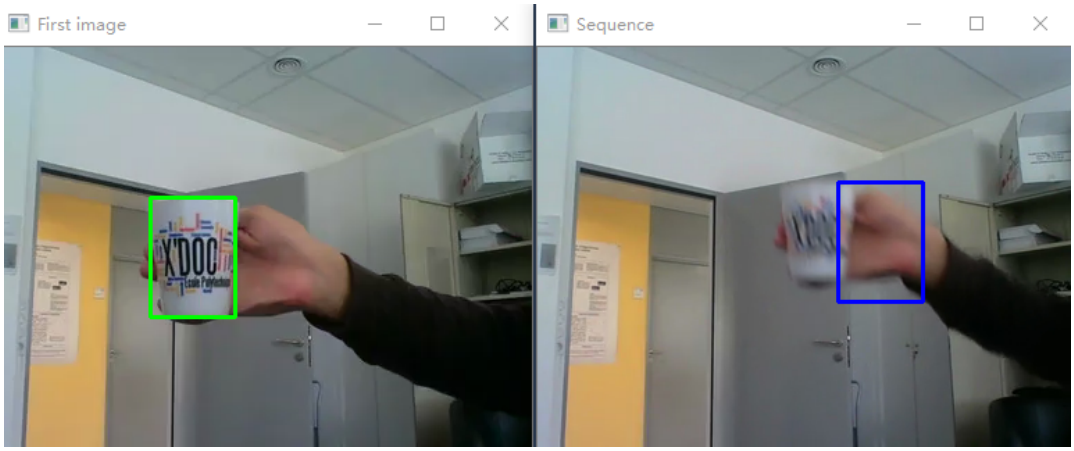

Figure 8 - weight map "dst"

But if we replace the weight map "dst" by the Hough Transform "H", the result is more accurate. But the tracking speed is very low.

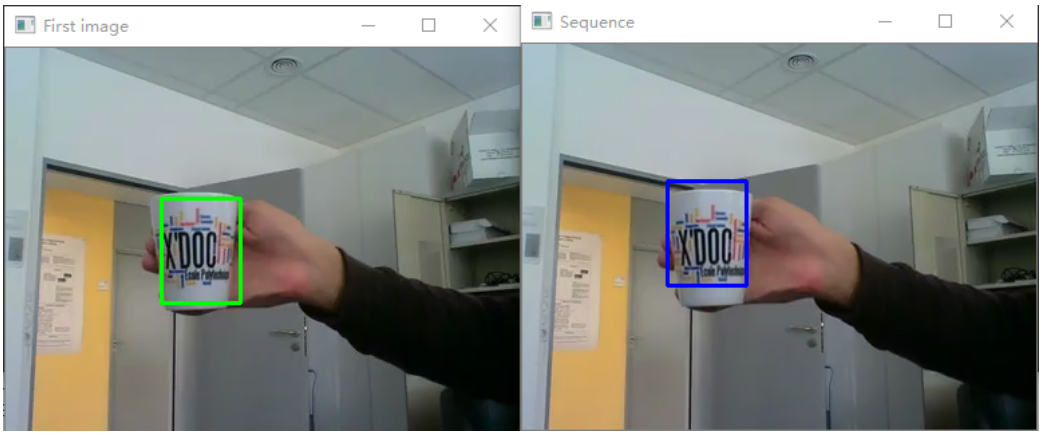

Figure 9 - Q5