CV2015Spring—Assignment #2

Due: Thursday, Apr 30 10:00 AM

Zhao Haiwei

Apr 28, 2015

1. The arrangement of experiment

The whole framework of the implementation for object detection is shown in Figure 1.

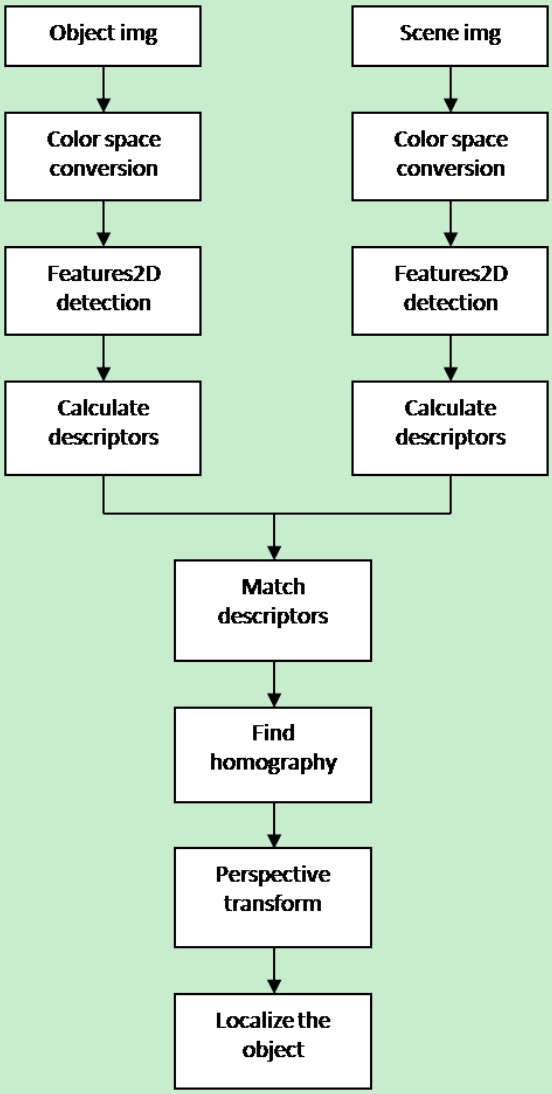


Figure 1: Framework of the implementation for object detection.

1.1 Step 1: Input image

The object image and scene image is given. Firstly, I input the two images.

(a) (b)

Figure 2: Input image.

1.2 Step 2: Color space conversion (5 points)

Input The input color image ($m \times n \times 3$ matrix), m is the width of the input image, n is the height of the input image.

Output $m \times n$ matrix.

Implementation This step determines the kind of color space I want to use in step 2, and I converse the color space from RGB to GRAY.

(a) (b)

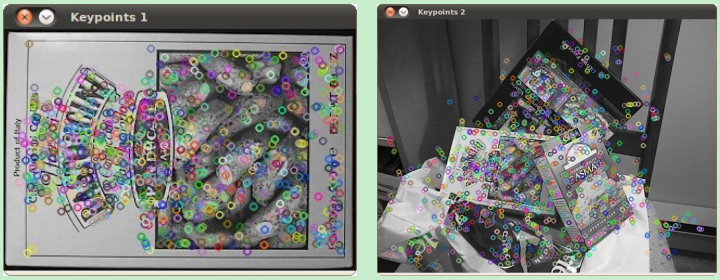
Figure 3: Input image.

1.3 Step 3: Features2D detection (15 points)

Input The gray image ($m \times n \times 1$ matrix), m is the width of the input image, n is the length of the input image.

Output Keypoints of object image and scene image.

Implementation There are lots of Features2D detection algorithms, such as SIFT, SURF, and so on. I choose SIFT which is fast and easy to implement.



(a) (b)

Figure 4: Features2D detection.

1.4 Step 4: Calculate descriptors (15 points)

Input The gray image ($m \times n \times 1$ matrix), m is the width of the input image, n is the length of the input image.The keypoints detected in Step 3.

Output One matrix describes the keypoints in the object image and one matrix describes the keypoints in the scene image.

Implementation Descriptors of the keypoints in the object image are stored in one Matrix and descriptors of the keypoints in the scene image are stored in another Matrix.

1.5 Step 5: Match descriptors (20 points)

Input One matrix describes the keypoints in the object image and one matrix describes the keypoints in the scene image.

Output M good matched keypoints between object image and scene image.

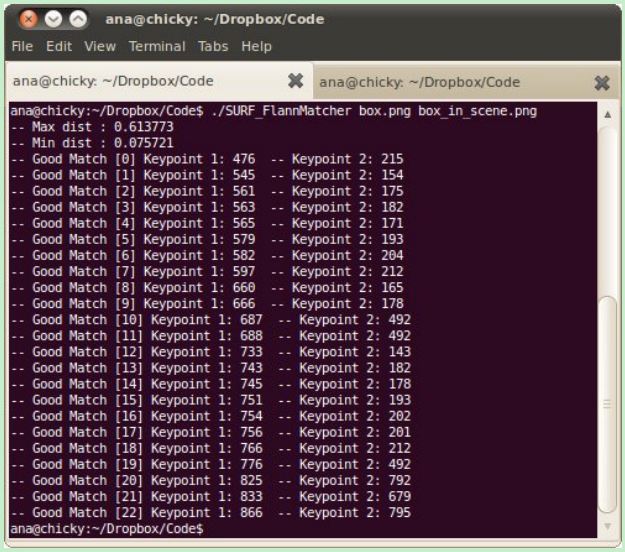
Implementation

- Step 5-1: I quick calculate $x \times y$ (x, y are the number of keypoints) distances (value) between keypoints, each value indicates the similarity of one keypoint from object image and one keypoint from scene image.

- Step 5-2: Fix one keypoint from object image and select the min or max distance between the keypoint and one keypoint from scene image to find the best match,then change the fixed keypoint and do the same.
- Step 5-3: I draw only "good" matches (i.e. whose distance is less than $3 * \min(distance)$).

Hint There are lots of distance calculation algorithms, such as Hausdorff Distance, Euclidean Distance, and so on. I choose Euclidean Distance to calculate the similarity between two keypoints and find the best match.

(a)



(b)

Figure 5: Match descriptors.

1.6 Step 6: Find homography transformation (20 points)

Input Keypoints of object image and scene image.

Output Matrix of homography transformation between two different point sets.

Implementation

- Step 6-1: Get N keypoints from the good matches and store them in two different point sets.
- Step 6-2: Calculate the matrix of homography transformation between the two different point sets.

Hint There are lots of mapping transformation algorithms, I use homography transformation.

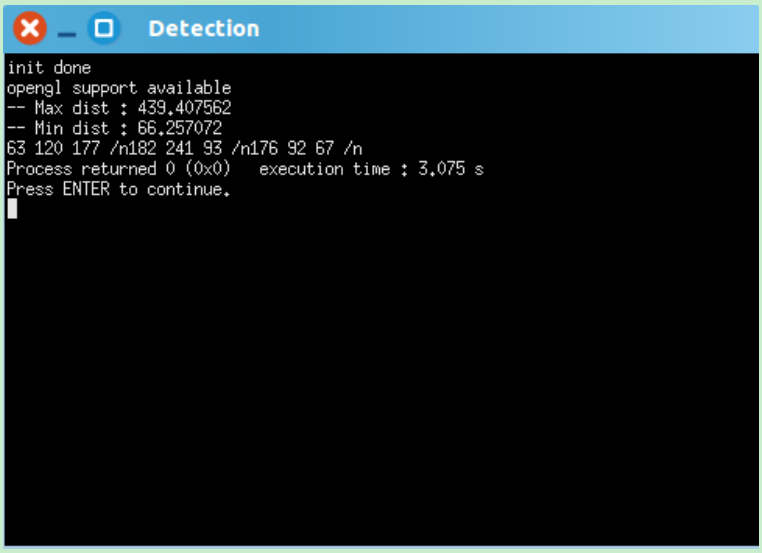


Figure 6: Find homography transformation.

1.7 Step 7: Perspective transform (20 points)

Input Four corners of object image.

Output Four points of scene image which are corresponding to the four corners.

Implementation I use the matrix of homography transformation which was got in Step 6 to perspective transform the four corners of object image to four points of scene image.

Figure 7: Perspective transform.

1.8 Step 8: Localize the object (5 points)

Input The scene image which has four points (the mapped object in the scene image).

Output The matrix box in the scene image which contains the object in object image.

Implementation I draw lines between the four points (the mapped object in the scene image).

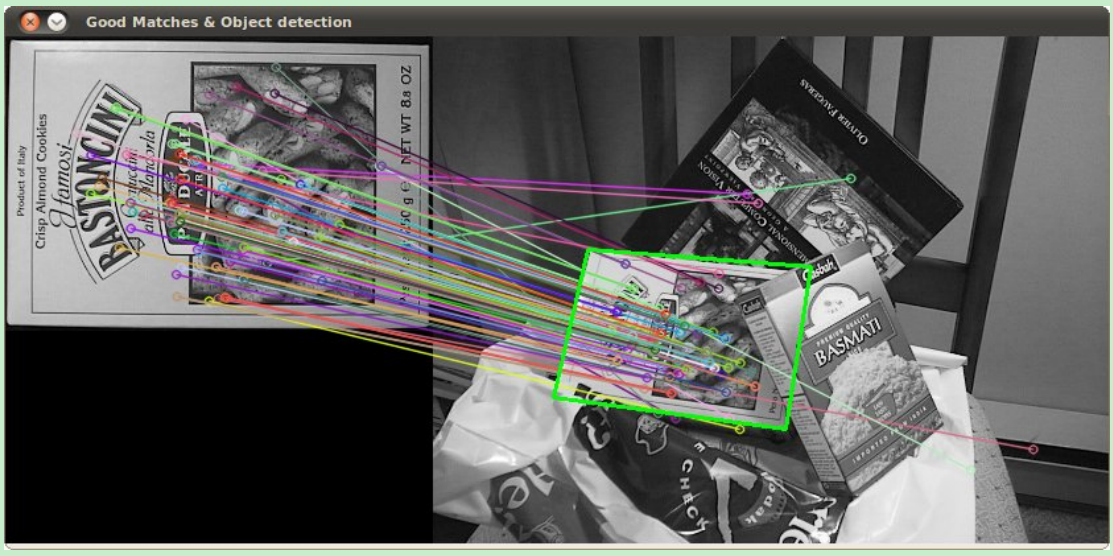


Figure 8: Localize the object.