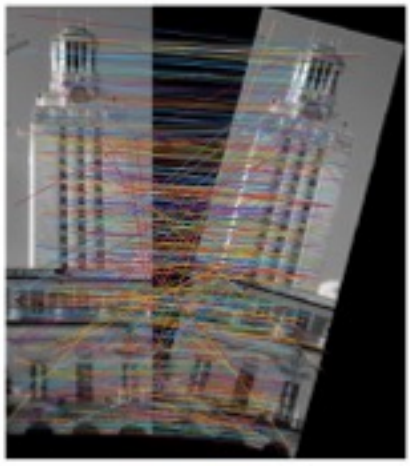
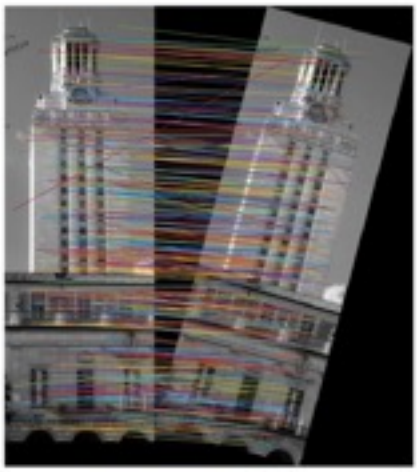
CS 381V Visual Recognition, Spring 2016
Assignment 1: Recognizing specific objects with local feature matching

Abhishek Sinha

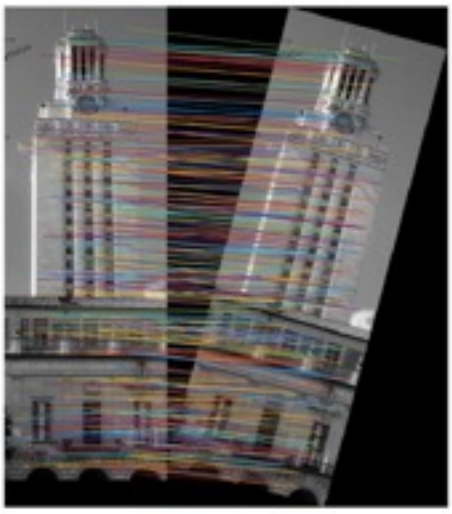
1. Match Comparison
    a. object-template.jpg and object-template-rotated.jpg



After Nearest Neighbors



After Lowe's Test



After RANSAC

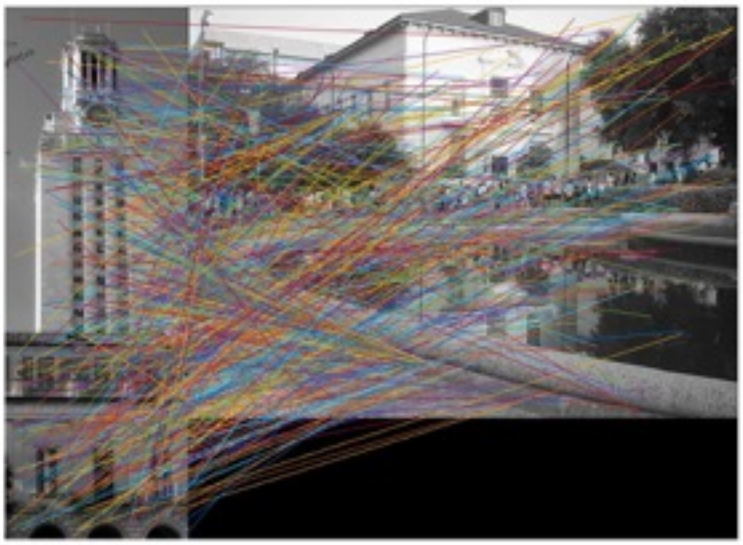
b. object-template.jpg and scene-1.jpg

After Nearest Neighbors

After Lowe’s Test

After RANSAC

c. object-template.jpg and scene-2.jpg



After Nearest Neighbors

After Lowe's Test

After RANSAC

General facts that we observe is that nearest-neighbor with threshold still retains a very large number of matches. We get rid of most of the spurious matches after the Lowe's test and RANSAC further reduces the number of matches by giving us only those matches that can be explained by an Affine transform.

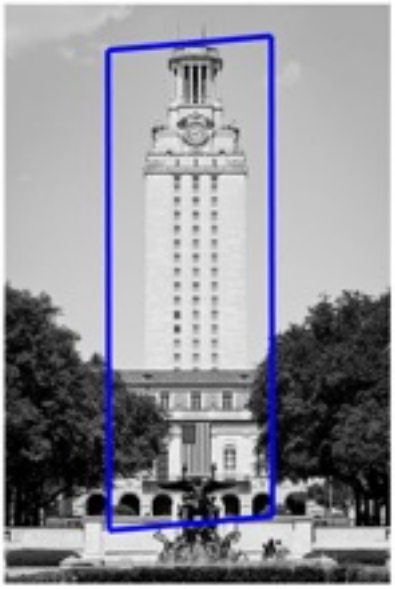
In part (a), since the scene image is simply a in-plane rotation of the template image, there is almost a one-one map between the descriptors of the template image and the scene image. RANSAC is able to discover the simple rotation as the most probable hypothesis and almost all the matches after the Lowe's test also survive the RANSAC verification.

In part(b), there is considerable intensity and perspective (out-of-plane rotation) difference between the 2 images. After Lowe's test, only 10 matches remain and most of these (7) are preserved by RANSAC and we get a suitable bounding box. But what is interesting is the fact that it is not the windows or the top portion that get matched (which intuitively seem to be more descriptive in nature), but rather the textural brick features on the right side that get matched. This does not seem to be a very good match, since other buildings (with completely different general appearance) with similar textual brick features are also likely to get matched.

In part (c), only 7 matches remain after Lowe's test. But still it isn't absolutely clear if we can consider the 2 images a match or not. After RANSAC, only 3 matches remain which clearly shows that it is not able to find any plausible affine transformation relating the 2 objects with enough inliers. Thus RANSAC helps in identifying this negative case in a more definite way.

2. Object Detection

Bounding box for object-template-rotated.jpg



Bounding box for scene-1.jpg

For object-template-rotated.jpg, we get almost a perfect match and RANSAC is able to exactly identify the region where the building is present and the fact that it is just a simple in-plane rotation of object-template.jpg

For the scene-1.jpg, we also get a reasonable bounding box. The background on the right and left gets removed, and the box mostly covers the building since in the original template the building almost entirely covered the image.

For scene-2.jpg, object is not detected because there are only 3 inlier matches found.

Detection Decision

OBJECT_DETECTION_INLIER_THRESHHOLD = 6
EIGEN_VALUE_THRESHHOLD = 1e-4

no_of_matching_inliers >= OBJECT_DETECTION_INLIER_THRESHHOLD & min(abs(eigenValues)) >= EIGEN_VALUE_THRESHHOLD

Thus we make the detection decision based on a threshold on the number of inliers found by the best RANSAC model. The second check is to ensure that the eigenvalues of the rotation matrix corresponding to the affine transformation are not too small. This is to avoid the case where many features in the template end up matching to a point or a line in the scene image (more about this in the extra credit part).

Note: We ran the RANSAC algorithm for 300 iterations and the inlier distance threshold was kept at a value of sqrt(5).

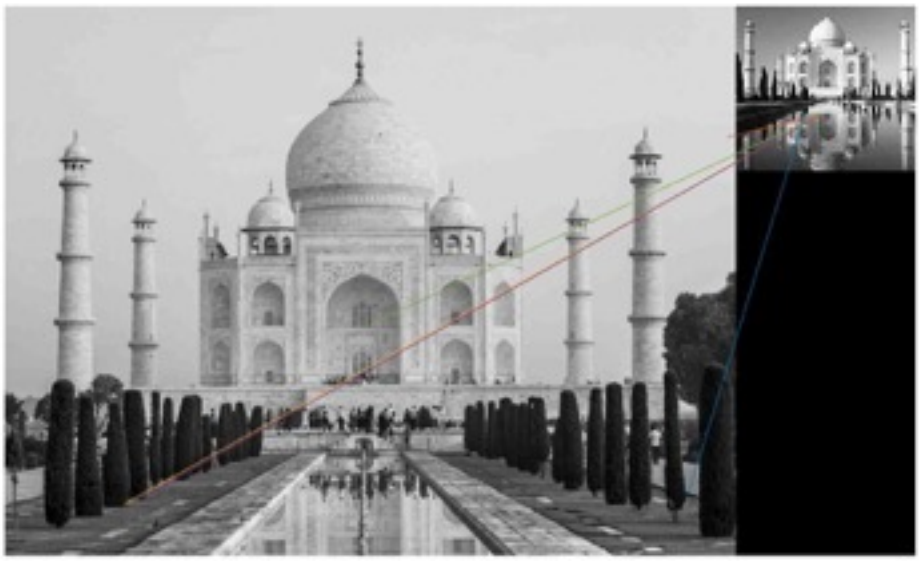
Extra Credit

1. Original Template Image

Example where match is possible

RANSAC Matches

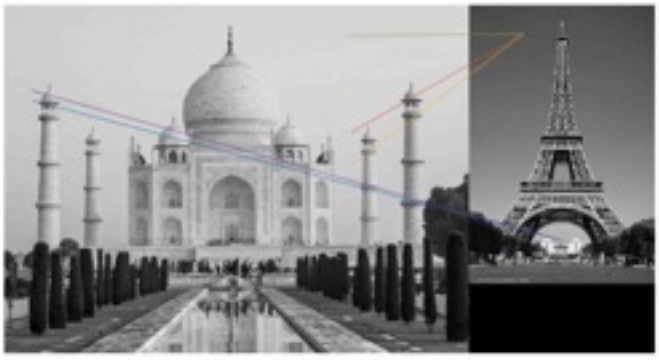
Example where Match is Not Possible



In the first, case there is variation in perspective and scale between the template and the scene image which the algorithm is able to detect.
However, in this case, the variation in perspective and scale is too much and the algorithm is not able to detect a match. Interestingly, some of the matches are between a feature in the left image and the corresponding feature in the reflection in the right image. This reflection might be confusing the RANSAC since some of the descriptors might be getting matched to the corresponding descriptors in the actual building while some others may be getting matched to the reflection. Thus RANSAC will not be able to discover one consistent affine transformation.

2. Examples where Lowe's Test gave a false match which were eliminated after Spatial Verification

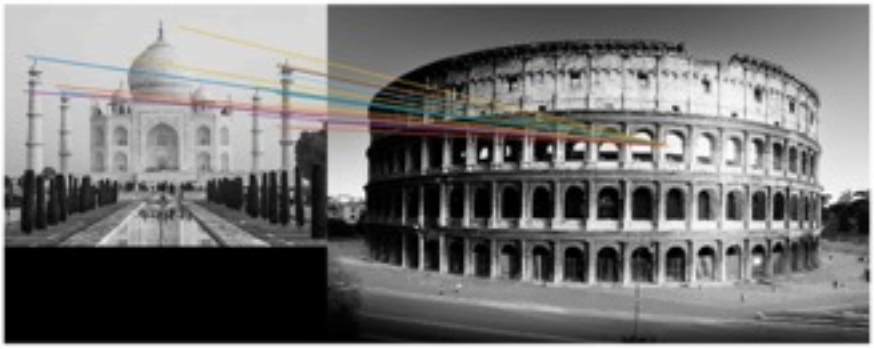
After Lowe's Test (above). After RANSAC (below)



Final Bounding Box

We can see clearly that there are a very large number of matches after the Lowe's Test. The number of matches comes down drastically after RANSAC but still if we made the decision only based on an inlier threshold we would have classified the above case as a match. However, if we set a minimum threshold on the eigenvalue of the rotation matrix, then we are able to avoid the case where there is a point or linear match. Hence, spatial verification, if performed by setting a threshold on both the number of inliers as well as the smallest eigenvalue of the rotation matrix is able to eliminate the match that Lowe's test gives. However, if we only go by the number of inliers, we will still get a match.

A similar scenario is also seen if we try to compare the image of Taj Mahal with Colosseum.



Thus the threshold on the eigenvalues seems to be crucial for correct matching.