
The University of Texas
at Austin

CS381V Visual Recognition

# Coding Assignment 01

Edited by LaTeX

Department of Computer Science

STUDENT

**Jimmy Lin**
xl5224

INSTRUCTOR

**Prof. Kristen Grauman**

TASSISTANT

**Kai-Yang Chiang**

RELEASE DATE

**Sep. 01 2016**

DUE DATE

**Sep. 16 2016**

TIME SPENT

**15 hours**

September 16, 2016

# Problem 1

## i) Thresholded Nearest Neighours (TNN): threshold = 0.8 * mean$_dist$

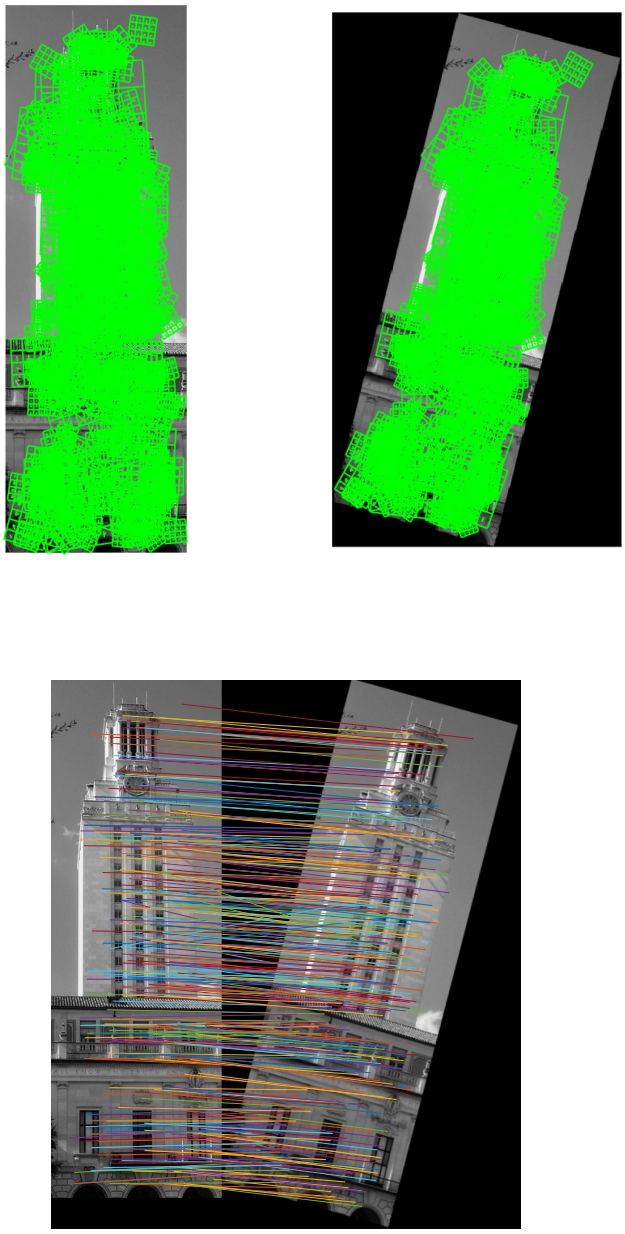


Figure 1: Patch Matches and Line Connections for TNN: object-template-rotated.jpg.

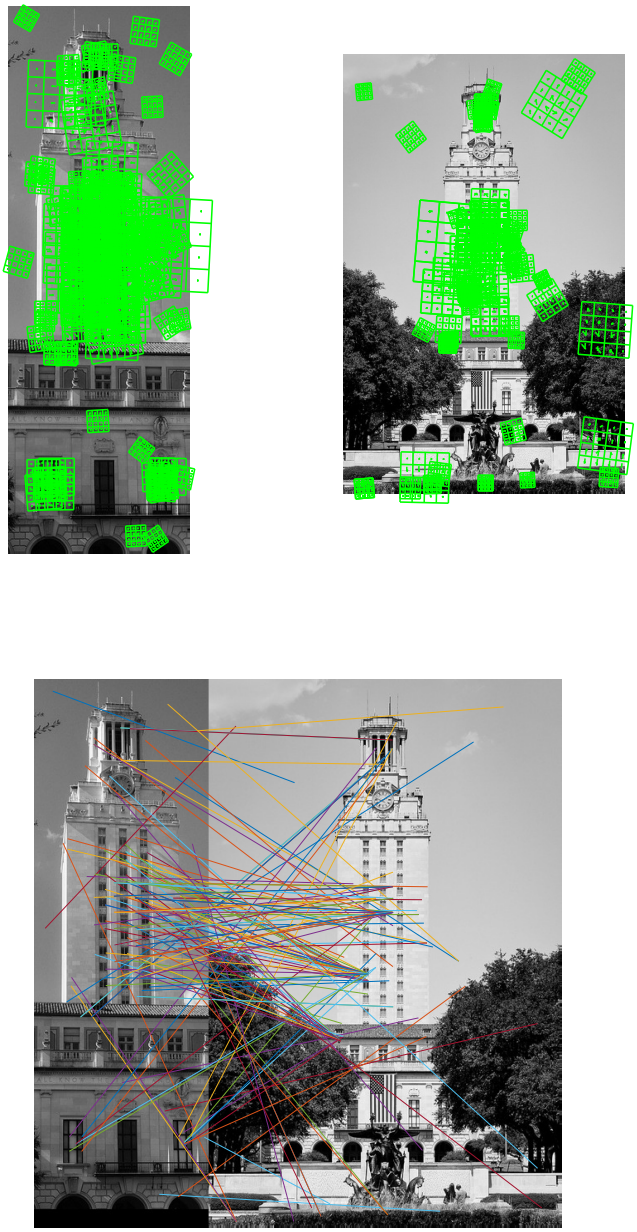


Figure 2: Patch Matches and Line Connections for TNN: scene1.jpg.

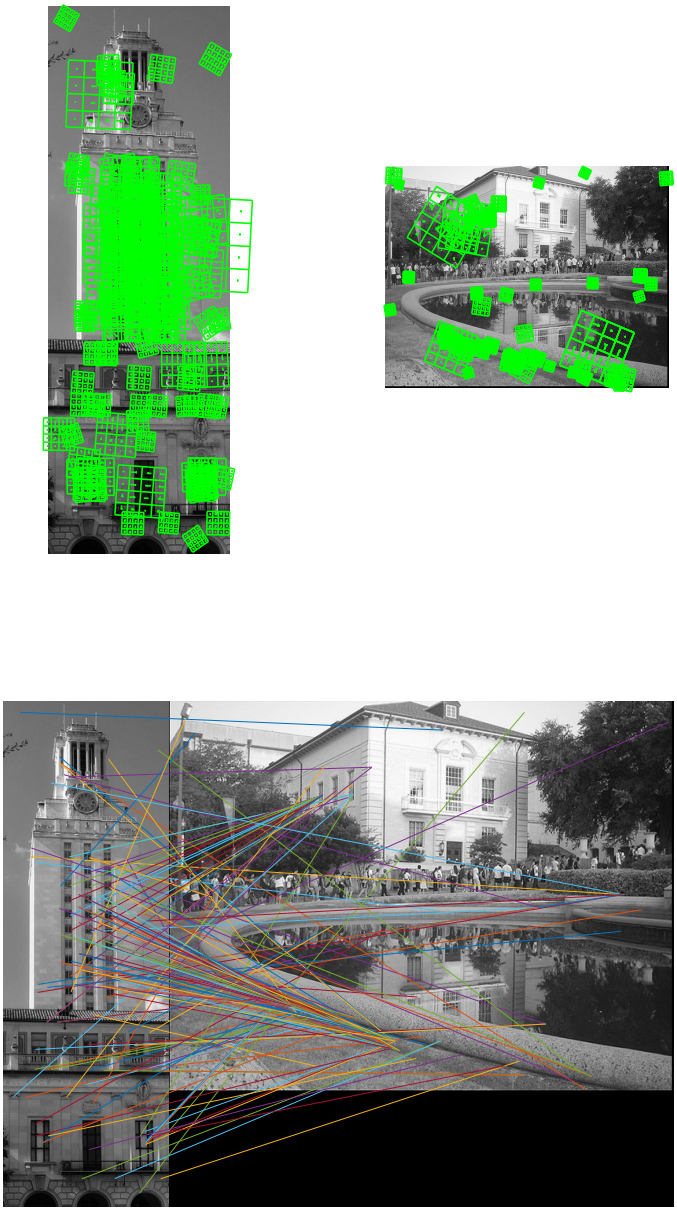


Figure 3: Patch Matches and Line Connections for TNN: scene2.jpg.

### ii) Thresholded Ratio Test (TRT): ratio = 0.6

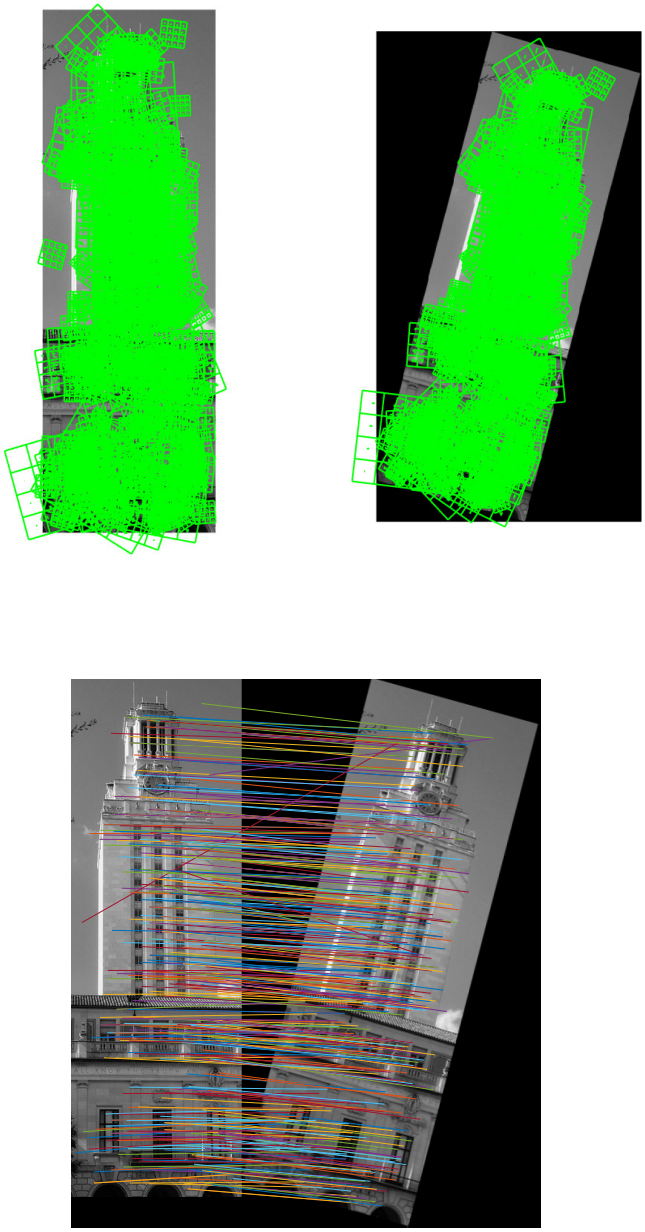
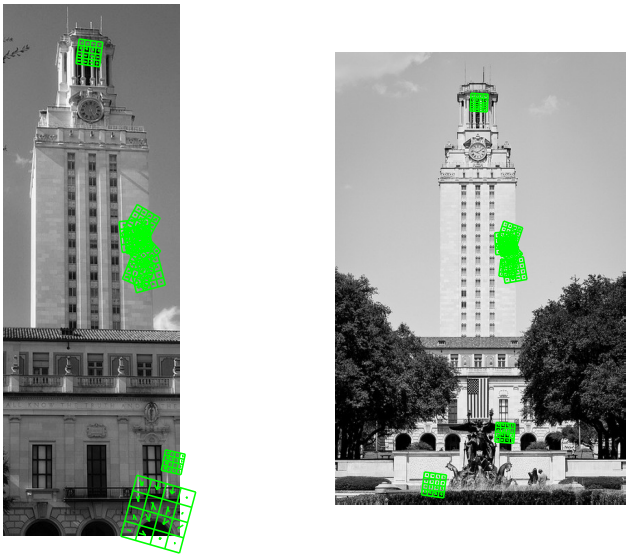


Figure 4: Patch Matches and Line Connections for TRT: object-template-rotated.jpg.



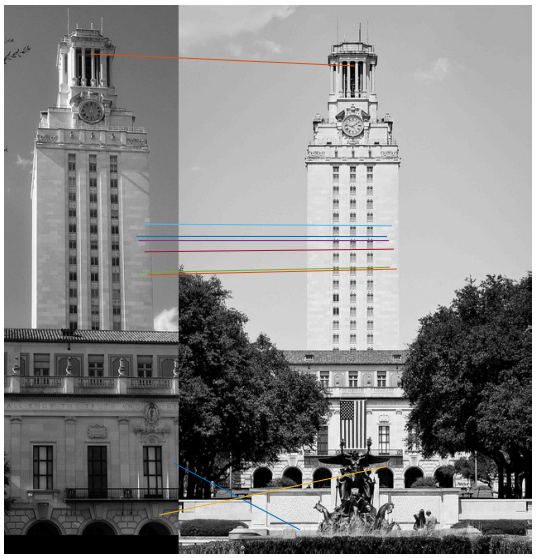


Figure 5: Patch Matches and Line Connections for TRT: scene1.jpg.

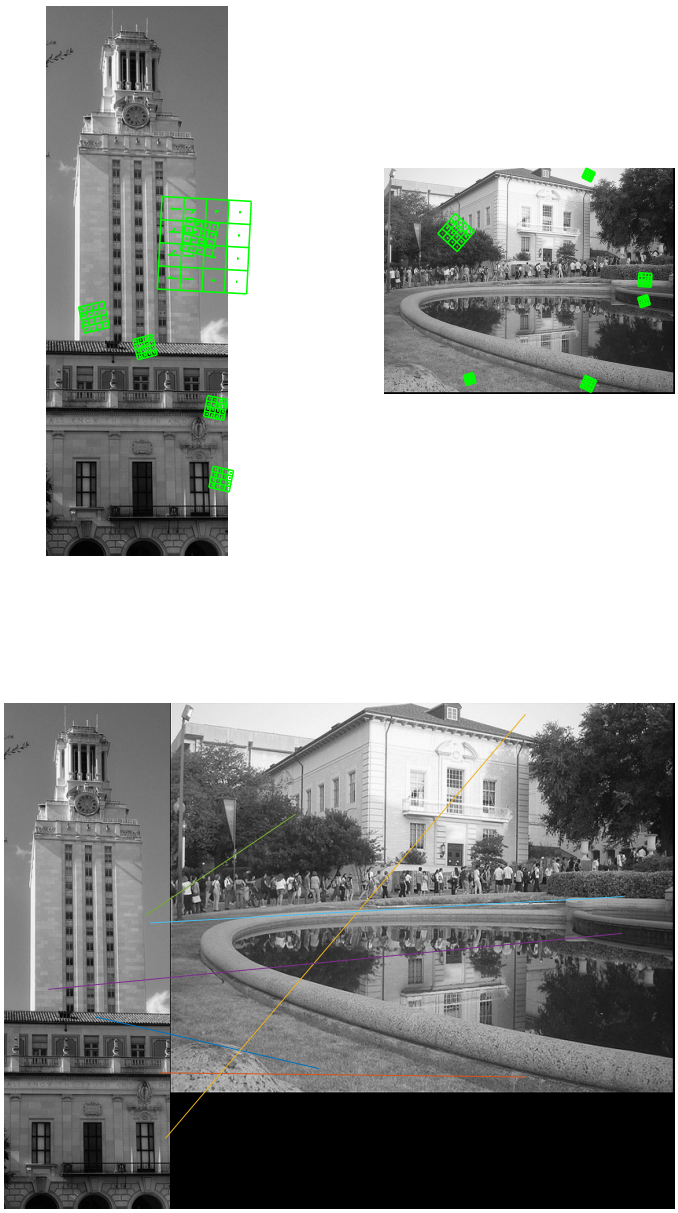


Figure 6: Patch Matches and Line Connections for TRT: scene2.jpg.

### iii) Inliers RANSAC

- Number of iterations to attempt random descriptors for fitting affine transformations:

$$INL\_ITERATIONS = 300$$

- Inliers Thresholds:

$$INL\_THRESHOLD = 50$$

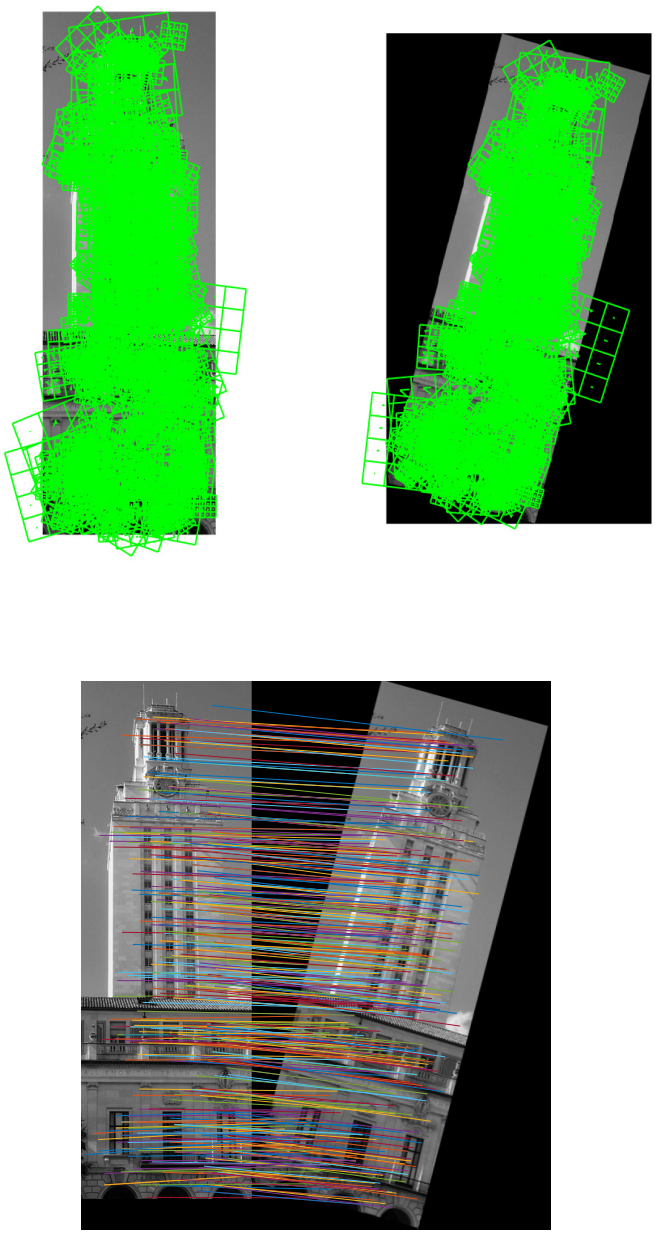


Figure 7: Patch Matches and Line Connections for RANSAC: object-template-rotate.jpg.

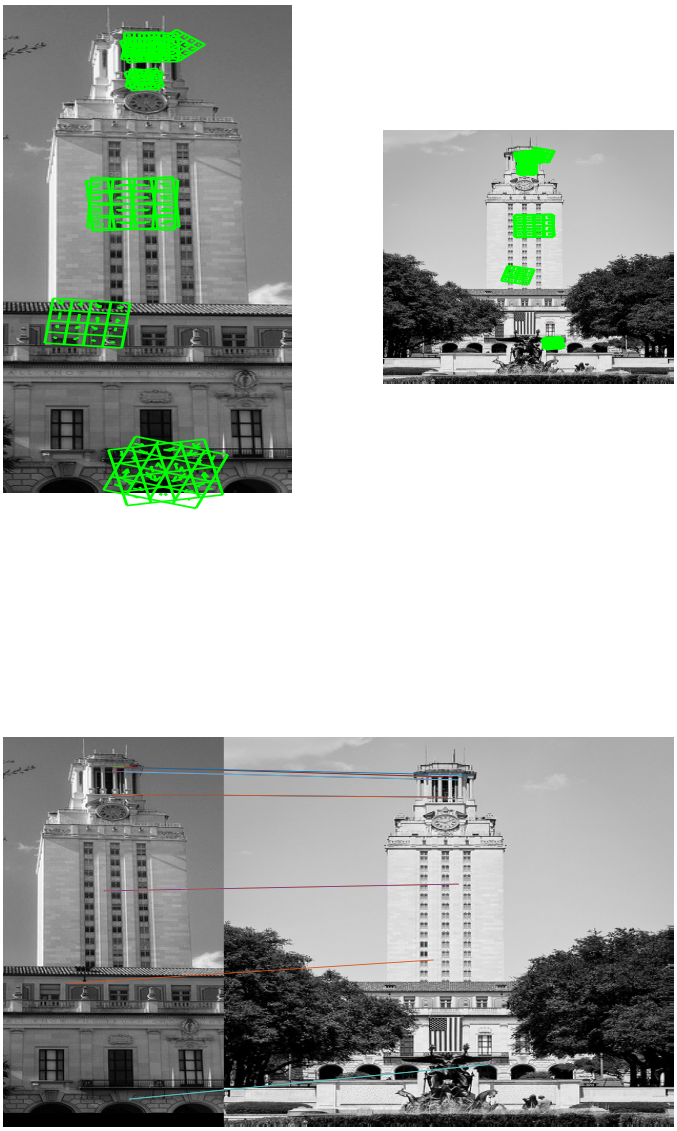


Figure 8: Patch Matches and Line Connections for RANSAC: scene1.jpg.

Figure 9: Patch Matches and Line Connections for RANSAC: scene2.jpg.

**Remark** Overally, inlier RANSAC provides the most robust descriptor matching in that it filters most quantity of affine variant descriptors. For the rotated image, both TNN and TRT provides many more descriptors matching than Inliers but most of those matchings are inaccurate (most descriptors in template and scene are not exactly corresponding to each other). Instead, those descriptors that match in Inlier RANSAC are very accurate at least for the rotated image (see fig. 7). And more importantly, the Inlier RANSAC approach successfully captures descriptor matchings over illumination change (see fig. 8) and filters almost all of descriptor matchings for totally different images (see fig. 9). Thus, Inlier RANSAC is the most preferable approach among all three for the purpose of detection.

## Problem 2 Detection Decision

Rules for Detection Decisions:

- If the number of surviving descriptor matchings (over RANSAC affine verification) is greater than
$$DECISION\_THRESHOLD = 10$$
then it is decided that the desired object is in the scene and we affine transform the corners of template image to locate the bounding rectangle in the scene image. See the decision result for object-template-rotate.jpg and scene1.jpg.
- Otherwise, it is decided that the desired object does not appear in the scene image. See the decision result for scene2.jpg.

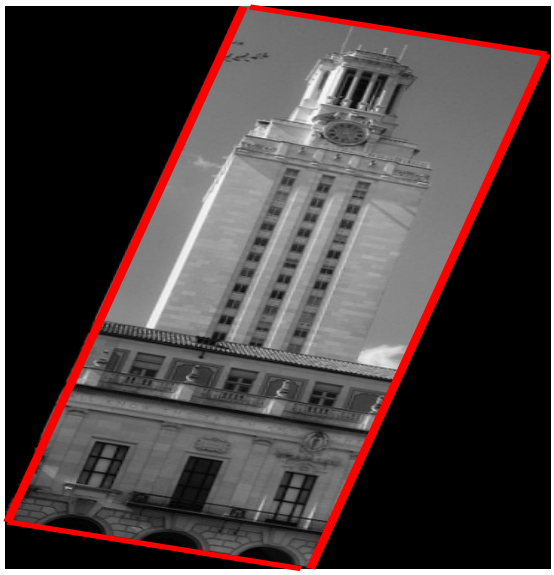


Figure 10: Detection Decision for RANSAC: object-template-rotate.jpg (One perfect detection).

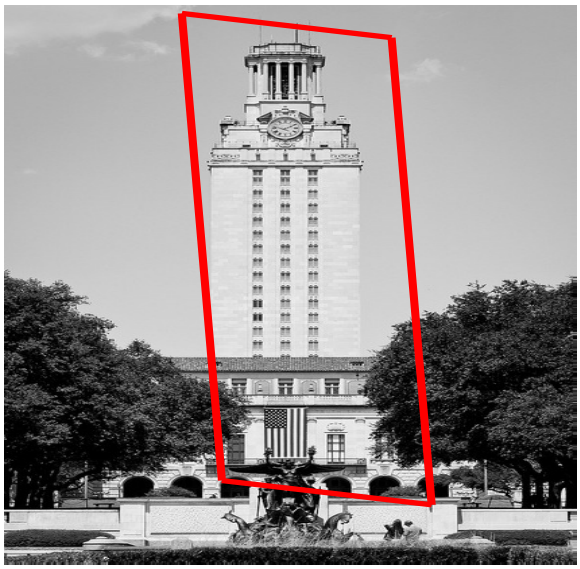


Figure 11: Detection Decision for RANSAC: scene1.jpg (One decent detection).

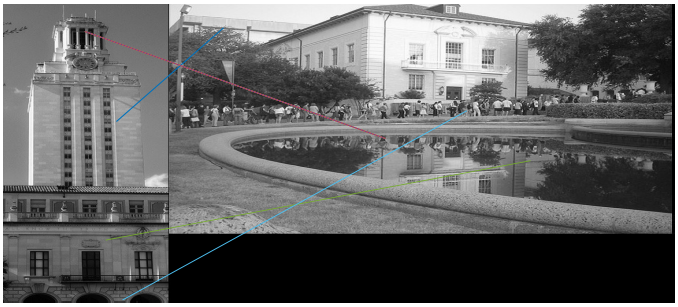


Figure 12: Detection Decision for RANSAC: scene2.jpg (detect nothing as desired).