ENGN 2560 Computer Vision: Homework #2
Image Stitching

Due: February 28, 2019 @ 5 p.m.

In this programming assignment you will practice feature based homography estimation to piece together a set of images to create panoramas.

An empty MATLAB script, `main.m`, is provided including few algorithmic parameter suggestions. Make use of helper methods that are called from `main.m` and make use of comments to improve code readability. Your homework submission should include your source code and a brief PDF report answering the questions with supporting visuals. Hand in your submission in a ZIP file with filename "FirstName_LastName_Homework2.zip".

Question 1

Implement an image stitching algorithm that can piece together a pair of images. Your algorithm should consist of the following steps:

- Load a pair of images known to originate from overlapping 3D scenes.
- Detect interest points/regions in each image. Ideally, the interest point/region detector should have high repeatability.
- Represent each interest point/region as a point in N dimensional space by extracting features.
- Find candidate interest point/region matches by computing distances, typically Euclidean distances, between every point in this N dimensional space. Make use of the ratio-test described in section 7.1 at `https://www.cs.ubc.ca/~lowe/papers/ijcv04.pdf` to discard outlier matches.
- Implement the RANSAC algorithm to estimate the homography matrix relating the pair of images. Specifically, take 4 candidate correspondences from the set obtained in the previous step at random and compute the resultant homography matrix. Under this homography matrix estimate, count how many of the candidate correspondences could be verified. As it is unlikely to find exact matches under the estimated homography, a candidate correspondence is considered an inlier if the norm of the displacement vector from the transformed coordinates to the matched coordinates is less than some small threshold. Iterate a predefined number of times, each time taking 4 candidate correspondences at random, and return the homography matrix estimate with the maximum number of inliers.
- Using the estimated homography matrix having the maximum number of inliers obtained in the previous step, compute the size of the composite image. To do so, one of the images could be taken as a reference. Compute the transformed coordinates of the image borders of the non-reference image. Using the transformed coordinates of the non-reference image as well as the coordinates of the reference image, create an image grid that is big enough to hold the composite and warp the non-reference image onto the reference image as follows: The pixel intensities at each grid location of the composite image needs to be determined. For each grid location, find out the coordinates of the non-reference image from which the intensity should be read. You need to make use of the function `interp2` if the coordinates fall onto a subpixel. If a given coordinate falls outside the image border an intensity value of 0 could be assumed. For grid locations where the reference and the non-reference image overlap, the maximum of the intensities or their average could be used.

Stitch the pair of images in `set1`, and `set2` using the SIFT keypoints and the SIFT features and include the visuals of the composite images. Figures 1 and 2 show the expected results of some of the steps of the above algorithm.

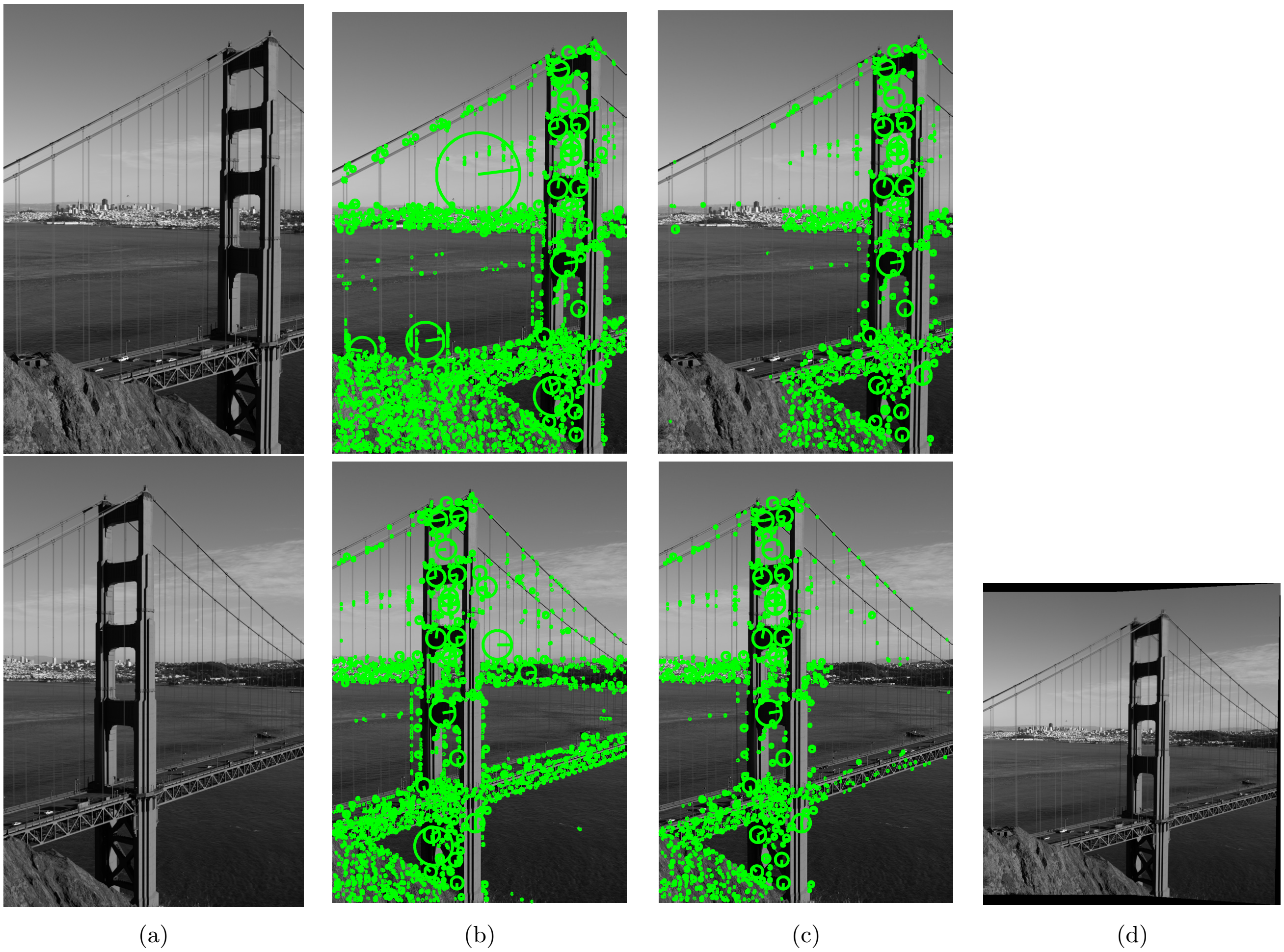


Figure 1: (a) goldengate − 02 and goldengate − 03 images (b) Detected SIFT keypoints (c) Reduced sets of keypoints after the ratio-test (d) Image stitching result

NOTES: You are not allowed to make use of the maketform, imtransform or any built-in functions alike. You need to do matrix-vector multiplication followed by interpolation to realize the homography transformation. It is highly recommended that you vectorize your code. Making excessive use of for loops can lead to a very slow implementation due to the sizes of the composite images.

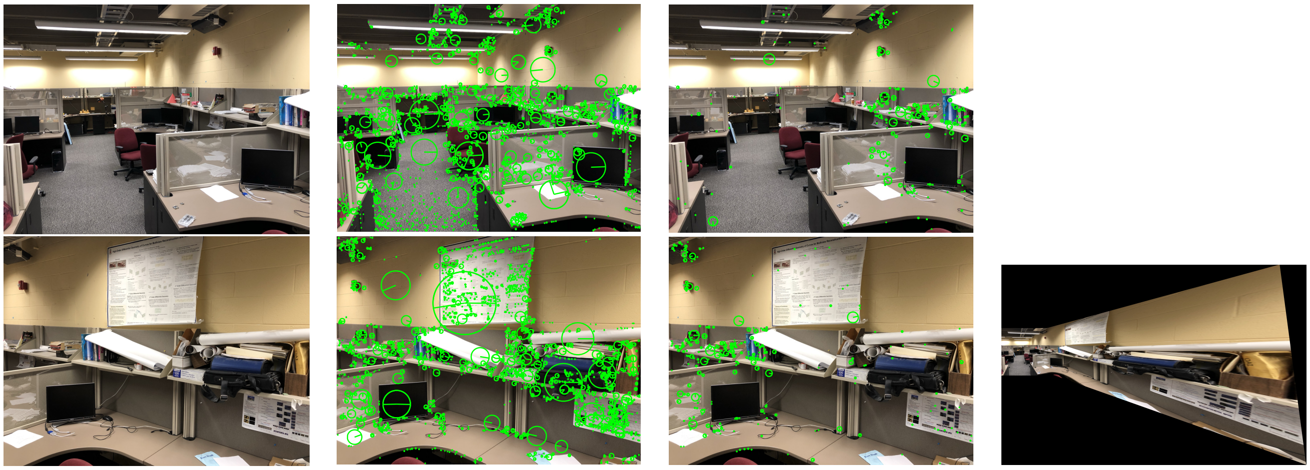
Question 2

Extend your implementation to stitch triplets of images. This can be achieved in at least two different ways. 1) Homography matrices between pairs of images could be estimated separately using the algorithm for Question 1, and a cascade of homography transforms could be used to warp images. 2) A pair of images could be stitched at a time where the composite image from the previous step is treated as the reference image for the next step.

Stitch the triplets of images in set3, and set4 using the SIFT keypoints and the SIFT features and include the visuals of the composite images. Figure 3 shows the expected results on these sets of images.

Extra credit

a) Often times, there are exposure differences between the images which result in visible seams when they are stitched. One way to eliminate these seams is to apply a blending algorithm which weighs the intensities of pixels with respect to their distances to image borders where most seams occur. Read the proposed blending algorithm



(a) (b) (c) (d)

Figure 2: (a) lems − 01 and lems − 02 images (b) Detected SIFT keypoints (c) Reduced sets of keypoints after the ratio-test (d) Image stitching result

(a) (b)

Figure 3: (a) Image stitching result of the halfdome dataset (b) Image stitching result of the hotel dataset

(a) (b)

Figure 4: (a) Image stitching result of the diamondhead dataset (b) Image stitching result of the diamondhead dataset with the blending technique discussed in the research paper

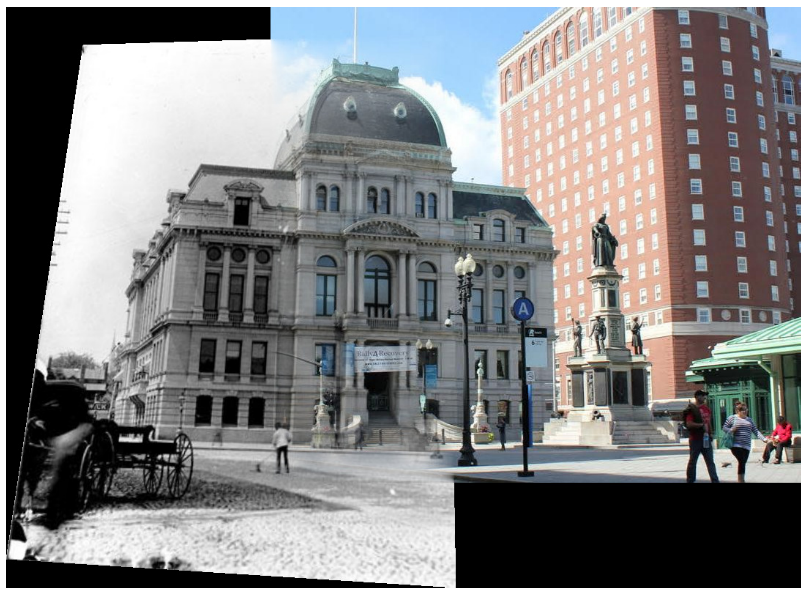


Figure 5: Historical composite of the City Hall – 1900 and 2016.

on page 138 of the following paper at `http://bigwww.epfl.ch/publications/thevenaz0701.pdf` and implement it. Apply the blending algorithm to stitch the diamondhead dataset. Figure 4 shows the differences between results with and without blending applied.

b) Providence's City Hall, located at the western end of Exchange Place, was the city's first permanent municipal building and was built in 1878. Many years later, the building remains in use as City Hall, and has not undergone much change though its surrounding has changed a lot. Using your algorithm and the blending technique introduced in part (a) try to obtain the historical composite shown in Figure 5.

c) Experiment with different interest point/region detectors as introduced in Homework 1 and extract features other than SIFT. How well they perform in the context of image stitching? Do they have any advantages over a SIFT based homography estimation?