COMP 5421 Project 1 – Report

Member(s):
Yuen Hei Chit Adrian
20029090

1. Introduction

The CVTool class are implemented to achieve feature detection by SIFT, SURF, MSER and Harris corner detection algorithms. The repair of damaged images using homography is also implemented. The computation of epipolar lines using 8 or more points is also supported. A user interface providing the visualization of the above is also developed.

2. Feature Detection

The following feature detectors and descriptors were used in this project:

I. Scale-Invariant Feature Transform (SIFT) detector and descriptor
II. Speeded-Up Robust Features (SURF) detector and descriptor
III. Maximally Stable Extremal Regions (MSER) detector and descriptor
IV. Harris Corner (HARRIS) detector and descriptor

3. Experimental Data

In the experiment to repair a damaged view of a scene using a transformed view of the same scene, 3 datasets (pairs of images) were prepared. In conjunction with the given dataset and source code, we name the first pair A and B, the second pair C and D, and the last pair E and F.

A and B are part of the given dataset and is are two views of (probably European) architecture. A is partially damaged.

C and D are part of the given dataset and is are two views of a mechanial valve on a table. C is partially damaged.

E and F are part of the given dataset and is are two views of clustered view of a robotics labatory. E is partially damaged.

The experimental data may be found under the subdirectory phase1-data/.

4. Experiment Procedure

We utilize the four methods listed in (2) to repair the damaged image using homography. With each method, keypoints are extracted from both damaged and transformed images. Feature vectors are extracted using the keypoints, and these feature vectors are matched using the nearest-neighbour approach. Then, all the matched keypoints, corresponding to matched feature vectors, are using to find the homography, i.e. the 3-by-3 perspective transform matrix , between the two views. The transformed image is transformed using the transform matrix, and the damaged region cropped and overlayed on the damaged image.

We define the error between single channels of the repaired image and the original, undamaged image as follows:

Channel Total Squared Error (TSE) =
sum (square (repaired image channel – original image channel)))

We define the error between the repaired image and the original, undamaged image as follows:

Error = mean (red channel TSE, green channel TSE, blue channel TSE)

This error is then used as a metric to compare performances between the feature detectors in the task of image repairing with homography.

5. Experiment Results and Discussion

Error	SIFT	SURF	MSER	HARRIS	Dataset Mean
A & B	103265.33	108696.33	97957.67	229374.67	134823
C & D	553725	520898	676824.67	962750.67	678549
E & F	1437975.67	623355.33	946218.33	1382370.67	1097479
Method Mean	698321	417649	573666	858164	

From the (dataset) mean error results, we see that the difficultly of repairing E by F is higher than that of C by D, and is in turn more difficult than that of A & B. Our explaination for this is that the views A and B were taken, most likely, under natural sunlight. These views are inherently more uniformly illuminated, when compared to the illumination the view C, D, E and F, which were taking in an indoor environment. Repairing E by F is the most difficult because of the cluttered scene of the lab, hence keypoints may not be well-matched, resulting in a poor homography.

From the (method) mean error results, we can ranked that method by the accuracy in completing the image repairing task. Ranking first is SURF, second is MSER, third is SIFT and the last is HARRIS.

It is not surprising that the HARRIS feature (corner) detection has the worst performance. It is not scale-invariant. Keypoint detection are not robust to forward translation (relative to camera). Camera rotations would also cause keypoints to vary under this relatively straight-forward corner detector.

SURF trumps SIFT approximates the Hessian matrix and has a lower feature space dimensionality than SIFT, resulting in less noise in the feature vectors. The author of SURF showed that SURF is more robust than SIFT on ill-formed, elongated structures. SIFT assigns an orientation to keypoints, and in cluttered scenes, this may also be a source of noise in feature point matching.

SURF also outperformed MSER based on the experimental results. MSER is inherently a region detector and is robust to affine (continuous) transformation. However, in cluttered environments or complex scenes, there are frequently small patches on the images. MSER may perform well under purer translation, however, the combination of rotation and translation changes the scene significantly. We conjecture that MSER performs worse, especially in the E & F dataset, since rotation and translation does not preserve the ordering and neighbouring of the small patches the MSER tends to capture.

6. Epipolar Lines

In terms of epipolar lines in pure forward translation, the epipole in both images should have the same coordinates.

In terms of epipolar lines in pure translation, the epipoles in theory does not exist as the line connecting camera centers does not intercept the projection planes.

In terms of epipolear lines in pure rotation, the line connecting the two camera centers have length 0 as the camera centers are the same point. Epipoles does not exist on the projective plane.

7. Future work

More analysis is needed to be done upon the epipolar line parts of the project.

Given more time, tuning of the parameters of the feature detectors and descriptors could be doneto obtain better keypoints, hence lower errors in image repairing.