CMSC733 Project 1
Correspondences and Mosaic

Vikas Chandrakant Raykar

May 11, 2004

1 Part A

In this part we need to find pairs of features in two perspective images such that each pair corresponds to the projection of the same scene point.

1.1 Harris Corner Detector

Interest points are obtained using the Harris corner detector [1]. The Harris corner detector first computes a corner response function. Corners are detected by finding the local maxima in the corner response function. The program returns the corners in the decreasing order of their corner response function value. I have not implemented a threshold, rather the user can ask for the number of corners (however it is possible that the number of corners returned can be less than that asked for). The two parameters that need to be tuned are the σ of the Gaussian kernel and the half-width of local-maxima neighborhood. Figure 1 shows a sample image, the detected 100 corners and the corner response function.
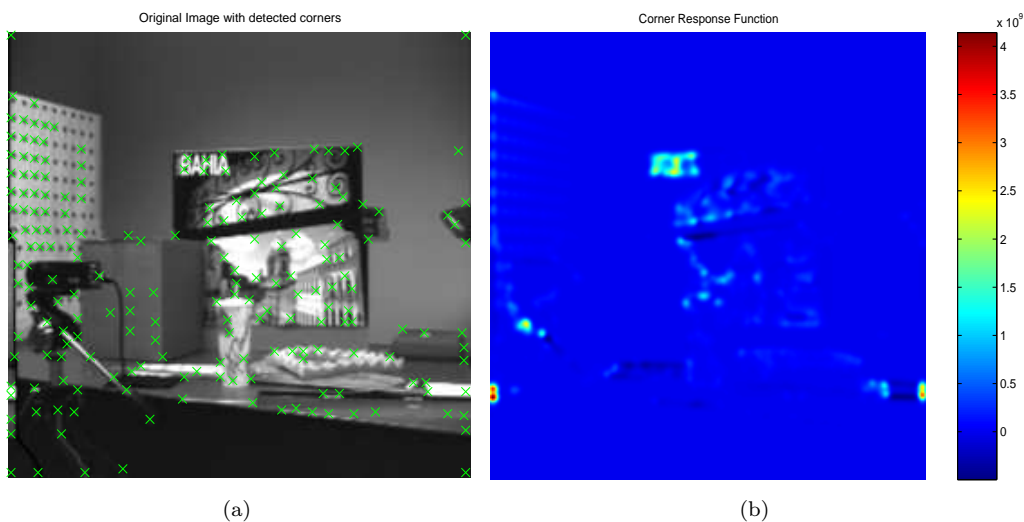

Figure 1: (a) Original image with detected corners (b) corner response function.

---

[1]C. J. Harris and M. Stephens. A combined corner and edge detector. In Proc. 4th Alvey Vision Conf., Manchester, pages 147-151, 1988.

1.2 Feature Correspondence

I have used Normalized Cross Correlation (NCC) as a similarity measure. First find the corner points in one image using the Harris corner detector. For each corner point in the first image find the corresponding point in the second image by finding the peak in the NCC function of a region around the interest point computed over a given search window. The correspondences are returned in the decreasing order of there normalized cross correlation. Figure 2 shows the two images overlaid and the corresponding matches found for the four possible pairs in the five given images. The feature correspondence is not robust enough. There are a lot of mismatches or outliers. We need to use RANSAC to find the inliers. Also note that in the regions of the checker board where there is sort of uniform texture there are a lot of mismatches.
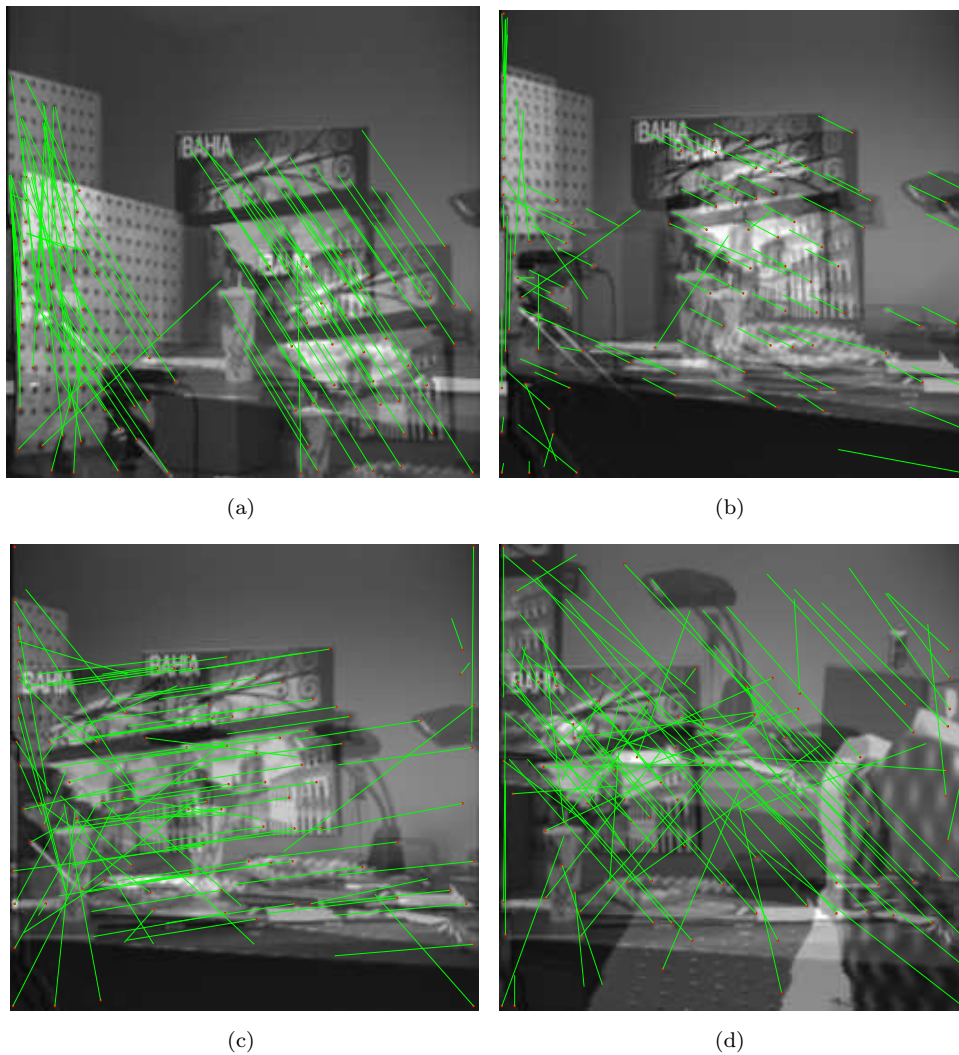

(a)   (b)

(c)   (d)

Figure 2: The two images overlaid and the corresponding matches found for the four possible pairs in the five given images.

2   Part B

Two images of the same scene taken by the same camera but from different directions are related by a homography H.

2.1   Camera Rotation $\equiv$ Homography

Consider a point x in image 1 which is the projection of the point X. Then x is related to X by the 3×4 camera projection matrix P

$$x = PX = K[I|0]X \tag{1}$$

where K is the matrix made of the intrinsic camera calibration parameters. If we consider the origin to be centered at the camera position corresponding to image 1, the rotation matrix is identity and the translation is 0.

Let $x^{'}$ be the projection of the same point in image 2. The difference between image 1 and 2 is that for image 2 the camera is rotated by a rotation matrix R.

$$x' = K[R|0]X = KR[I|0]X = KRK^{-1}K[I|0]X = KRK^{-1}x = Hx \tag{2}$$

where $H = KRK^{-1}$ a 3×3 matrix is a homography matrix.

The eigen vector of H corresponding to the real eigen value is the vanishing point of the rotation axis.

2.2   Computing Homography

I used the DLT algorithm with normalization to compute the homography. Since there were a lot of mismatches I used RANSAC along with DLT to estimate the homography [2]. RANSAC computes both the homography and a consistent set of correspondences. I uses the symmetric transfer error as the distance measure. Guided matching was also incorporated. Figure 3 shows the original correspondences and the set of inliers as computed using the RANSAC algorithm.

2.3   Mosaicing

Figure 4 shows the five images to be stitched together to form a mosaic. We choose the third image i.e. the image in the center as our reference and compute the homography H_{ri} between the reference image and all the other remaining 4 images. A point x_i in the ith image is related to the point x_r in the reference image by $x_r = H_{ri}x_i$. Once the homography is computed we rectify each of the four images to the reference frame. Since there can be fractional pixel values we need to do interpolation (linear). For each pixel position in the rectified image I find the corresponding pixel value in the original image. If the pixel position is fractional I interpolate using the image value at nearby pixels. I have used linear interpolation though other option can be selected. Figure 5 shows the final mosaic. Note that there are some errors. I think the homography can be improved my doing a further ML estimation. Also the correspondences extracted are not uniformly distributed throughout the image. A distributed sampling approach can be implemented along with the RANSAC algorithm such that samples are chosen more uniformly throughout the image.

---

[2] As given in Hartley and Zisserman

100 Original Correspondences   52 Correspondences after RANSAC

(a)

100 Original Correspondences   64 Correspondences after RANSAC

(b)

100 Original Correspondences   41 Correspondences after RANSAC

(c)

100 Original Correspondences   32 Correspondences after RANSAC
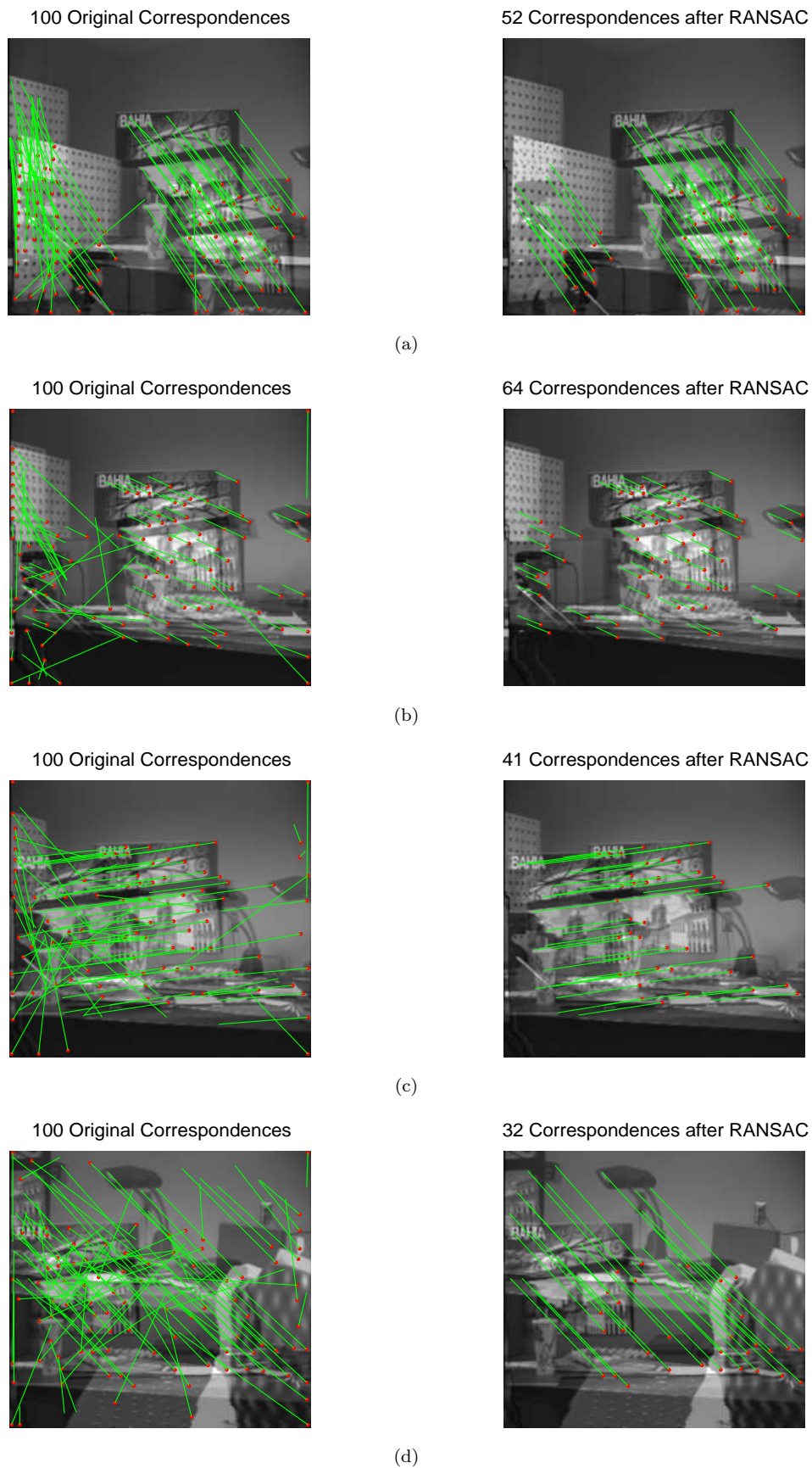
(d)

Figure 3: The two images overlaid and the corresponding matches found for the four possible pairs in the five given images. The right side shows the inliers detected by RANSAC algorithm.

Figure 4: Five image captured by the same camera with different rotations.

Figure 5: The final mosiac

3 Part C

In this part a planar scene is obscured by some wooden dowels. The task is the remove them. Since the photo is planar we can compute the homography between the reference frame and all other frames. Some correspondences on the wooden dowel may be found. But since we are using RANSAC these would be regarded as outliers and we get a perfect homography. Once we have the homography we can form a mosaic as in the previous case. Here in order to remove the wooden dowels we take the median of all the overlapping images in the mosaic. In this way since the dowels appear only in few frames at a given position by taking the median the dowel is erased. Figure 6 shows the original mosiac and the mosaic of the sequence using the median. As can be seen the banner is erased but there are slight ghost remaining. Once we have the mosiac we can get the frame back by inverse rectifying the particular frame. In order to get good results for each frame we can use that frame as the reference and repeat the procedure. Figure 7 shows the result for a frame. Figure 8 shows the result for all the frames.

(a)   (b)

Figure 6: (a) The mosaic of the video sequence (b) The mosiac using the mdeian.

(a)   (b)

Figure 7: (a)Original frame and (b) with dowels erased.

Frame 1

Frame 2

Frame 3

Frame 4

Frame 5

Frame 6

Frame 7

Frame 8

Frame 9

Frame 10

Frame 11

Frame 12

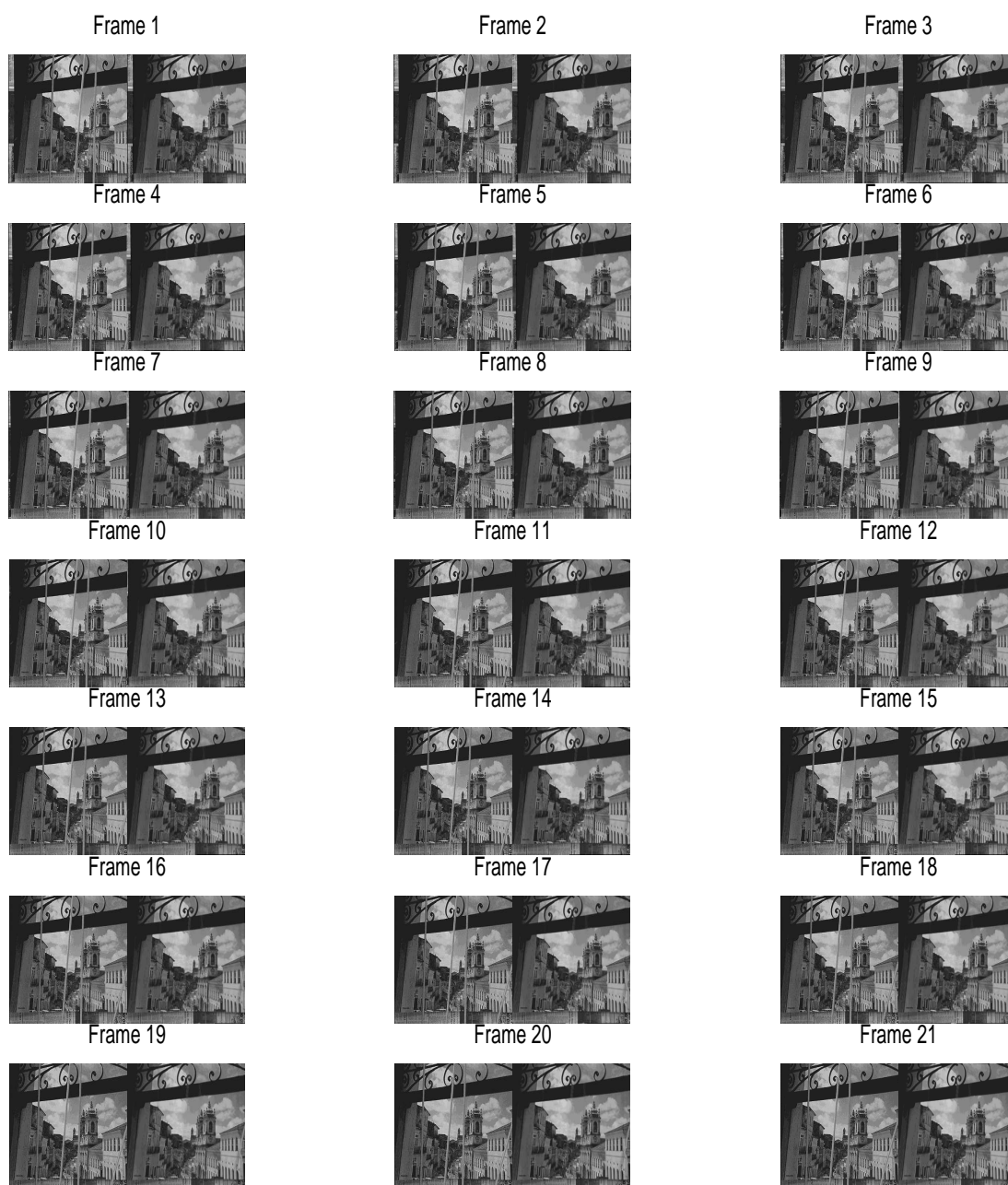Frame 13

Frame 14

Frame 15

Frame 16

Frame 17

Frame 18

Frame 19

Frame 20

Frame 21

Figure 8: Result for all the frames