

## **SIFT ALGORITHM & FEATURE MATCHING USING HOMOGRAPHY**

Scale Invariant Feature Transform produces reasonably invariant 128 dimensional vectors with detection carried out by:

- Scale Space extrema detection using Laplacian of Gaussians by difference in the sigma values.
- Key point localisation by different octaves and DoG that contribute with considerations of 26 neighbours in obtaining the key points among the levels and  $k$  is set to have levels in the octave.
- Orientation assignment is done by calculation of the histograms of the gradients. The 4X4 blocks have 8 gradient directions which assign the orientation to the key point and create the 128 dimensional key point descriptor.

This process of obtaining the key points and the descriptors is carried out on both the test and the target image. Hence there needs to be a way in which the key points need to be matched by Brute Force Matcher. The BF matcher is used with L2 norm distance used to calculate the minimum distances between the feature vectors of the descriptors between the target and the test images.

According to Lowe's ratio, only the good matches are considered and all the other matches are omitted. These matches are then used by the RANSAC algorithm to only account for the inliers while calculation of the homography.

Hence the homography matrix shows the transformation between both the target and the test image objects. Inliers are important for the robustness of this process. This enables the process of image stitching by obtaining key points and matching by finding the homographic transforms.

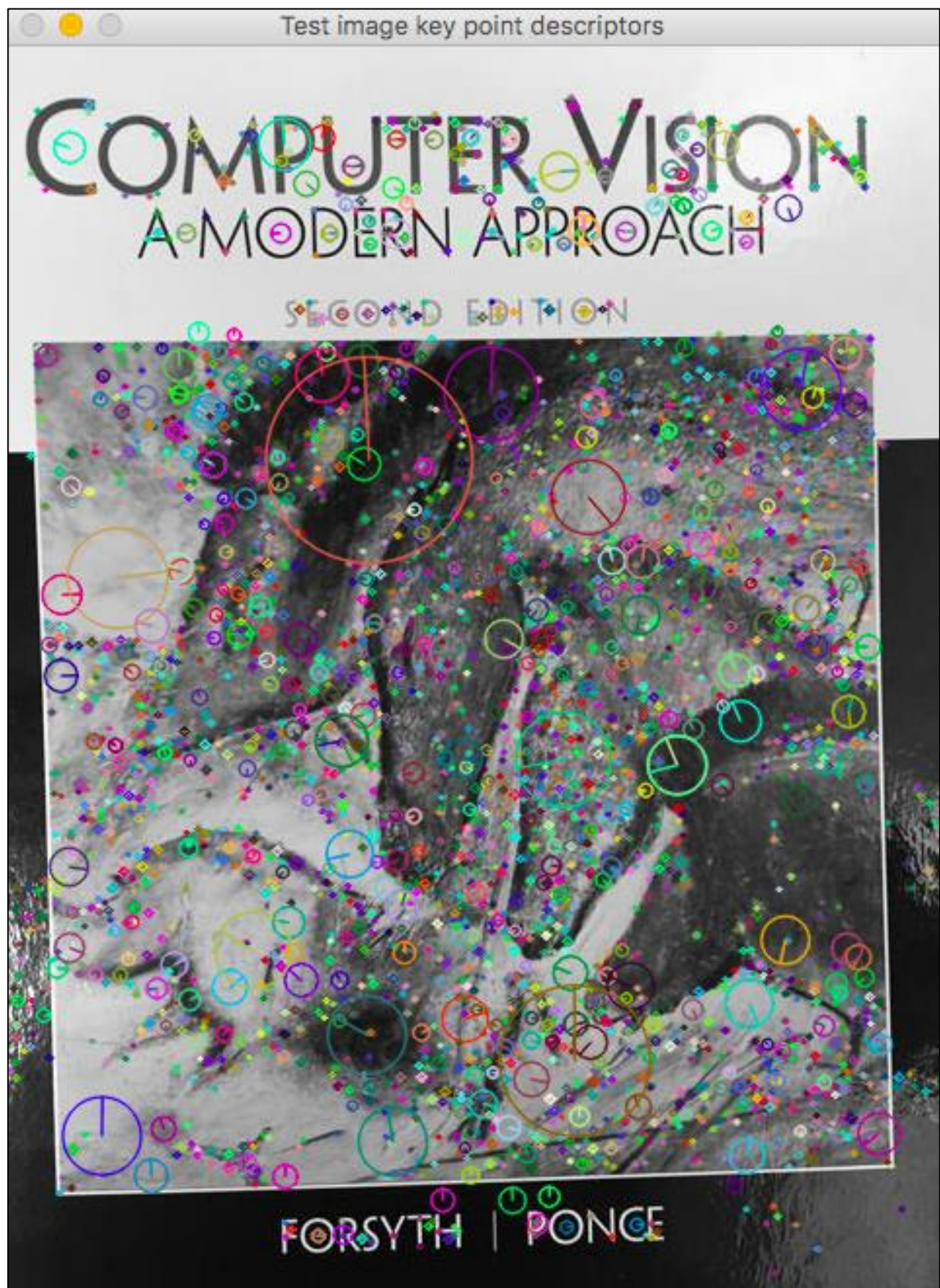
## RESULTS

Test Image	Target Image	Total good matches Lowe's ratio	Amount consistent with Homography	Percentage of inliers (in %)
Image1	Image3	420	369	87.86
Image1	Image4	240	183	76.25
Image1	Image5	562	525	93.42
Image2	Image3	194	165	85.05
Image2	Image4	210	185	88.09
Image2	Image5	48	35	72.91

## INTERMEDIATE RESULTS

CASE 1 with Image 1 as the test image and Image 3,4,5 as targets

- Test image: Image1
- Target image: Image3

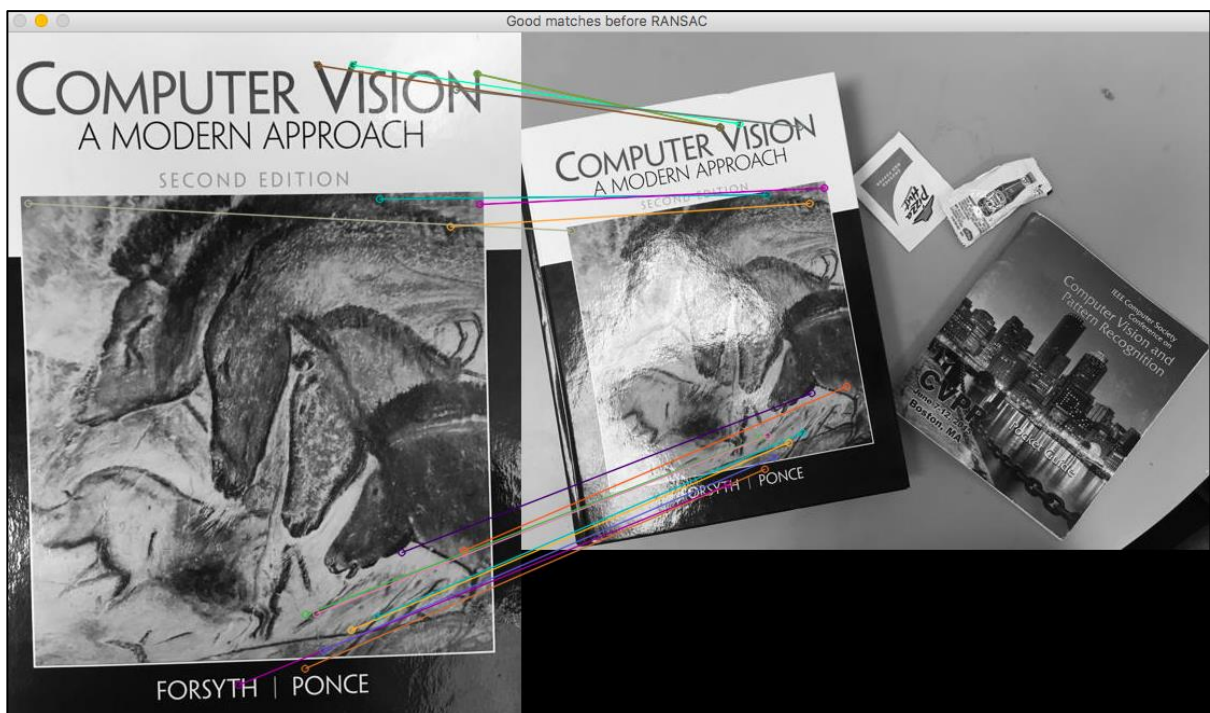


Test Image 1





Target Image 3



Match between top 20 descriptors with minimum L2 distance  
according Lowe's ratio before RANSAC



### Match between Inliers only after RANSAC

```
The total number of good matches found are: 420
The total number of consistent value inliers with the homography are: 369
The homography matrix found is as follows:
[[ 5.41381183e-01  7.32981440e-02  2.22433696e+01]
 [ -1.02250794e-01  5.34014772e-01  9.96393481e+01]
 [ -5.05601737e-05 -5.40824826e-05  1.00000000e+00]]
```

Results and the Homography matrix obtained.

- Test image: Image1
- Target image: Image4



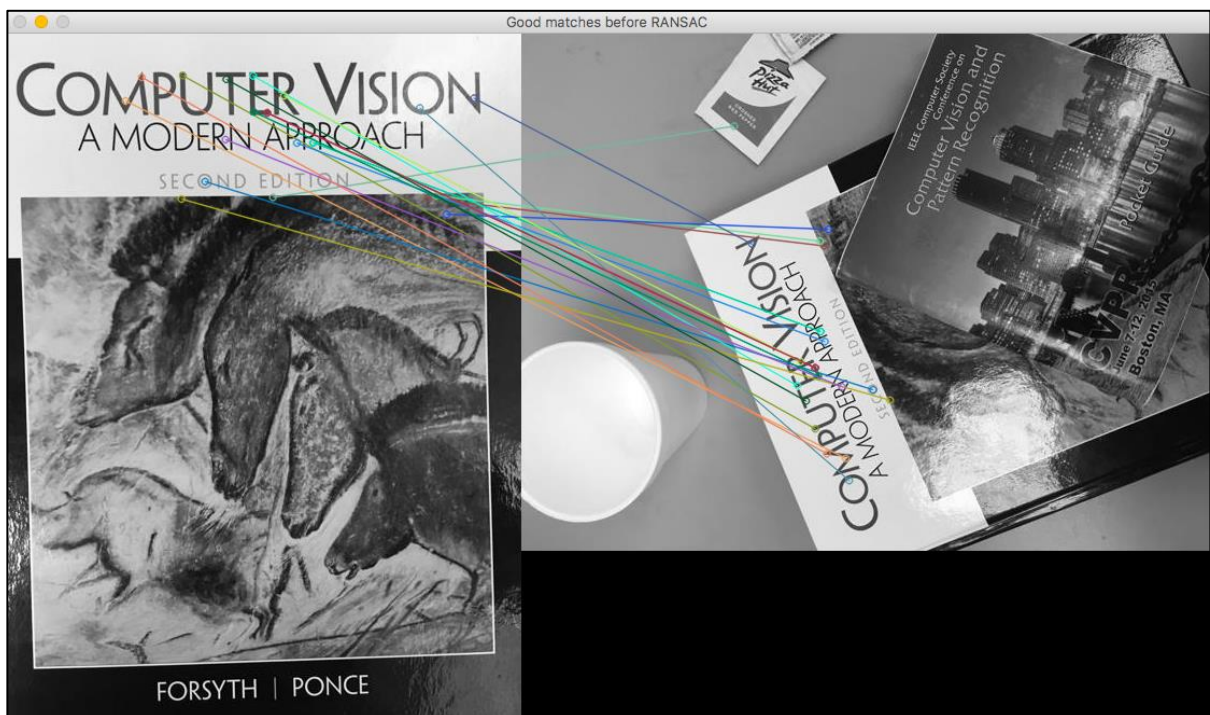


Test Image 1





Target Image 4



Match between top 20 descriptors with minimum L2 distance  
according Lowe's ratio before RANSAC



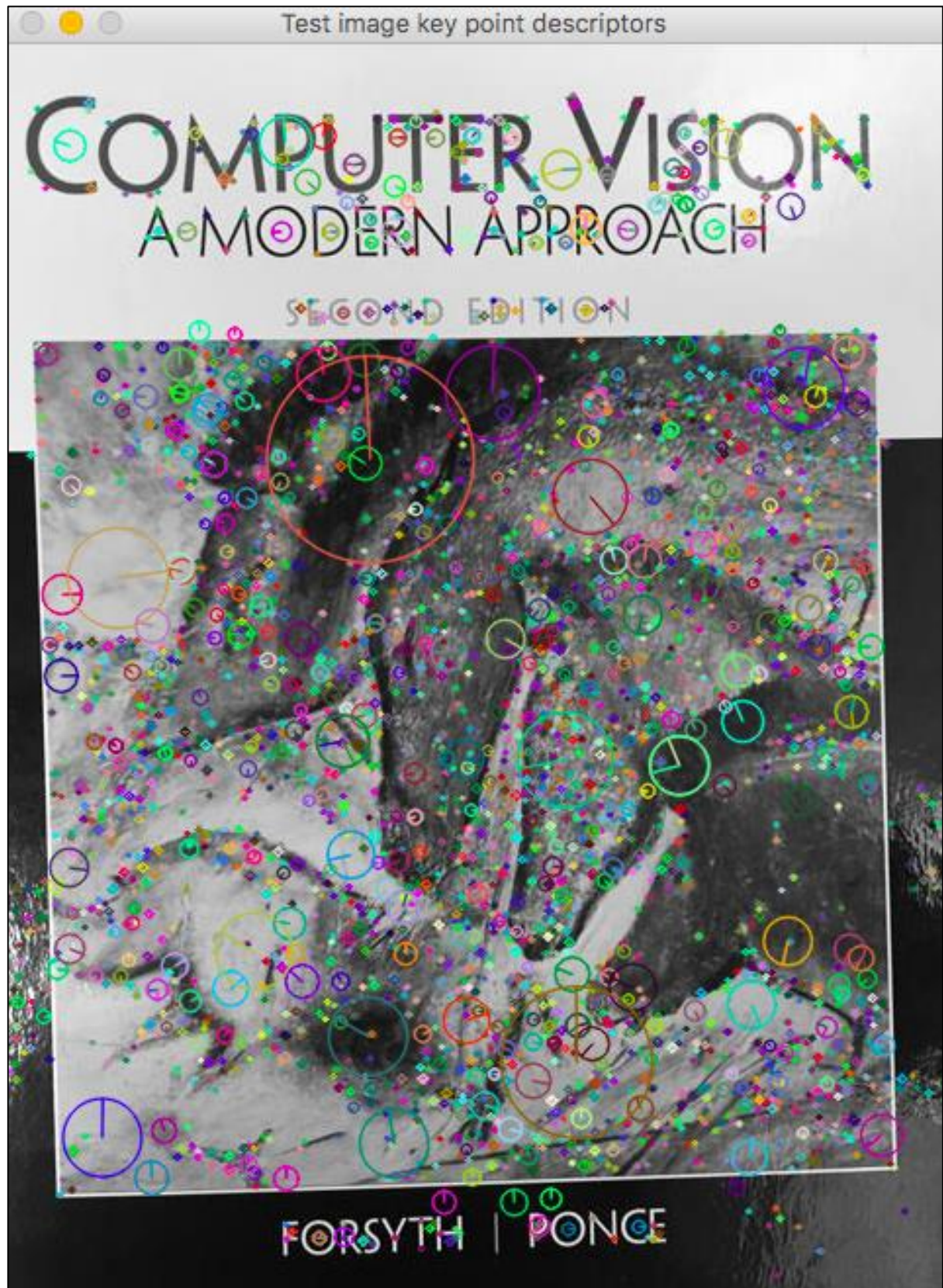
### Match between Inliers only after RANSAC

```
The total number of good matches found are: 240  
The total number of consistent value inliers with the homography are: 183  
The homography matrix found is as follows:  
[[ -2.54290026e-01  5.32202346e-01  2.93174854e+02]  
 [ -6.05687884e-01 -3.08570248e-01  4.76691655e+02]  
 [  2.51801196e-05 -2.06339495e-04  1.00000000e+00]]
```

Results and the Homography matrix obtained.

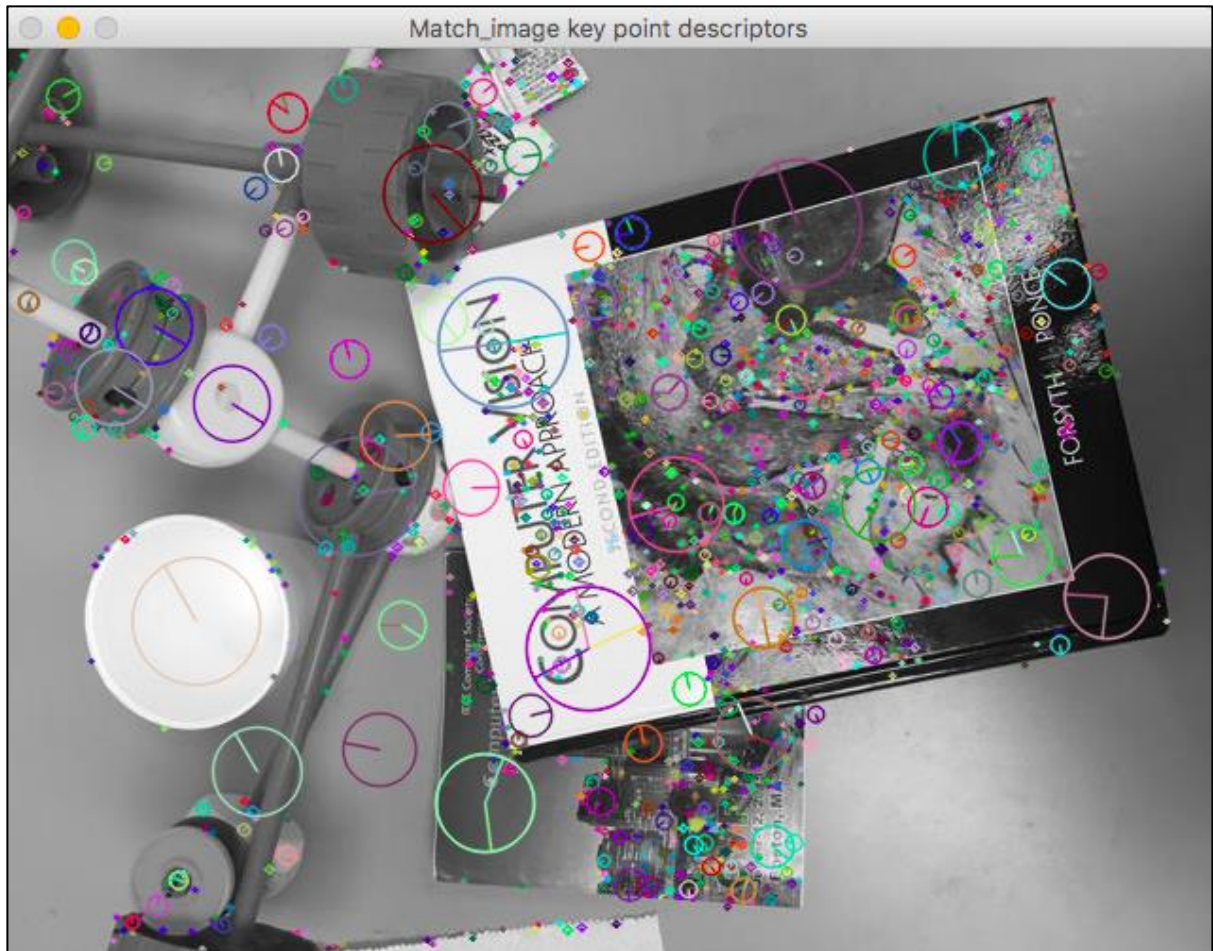
- Test image: Image1
- Target image: Image5





Test Image 1



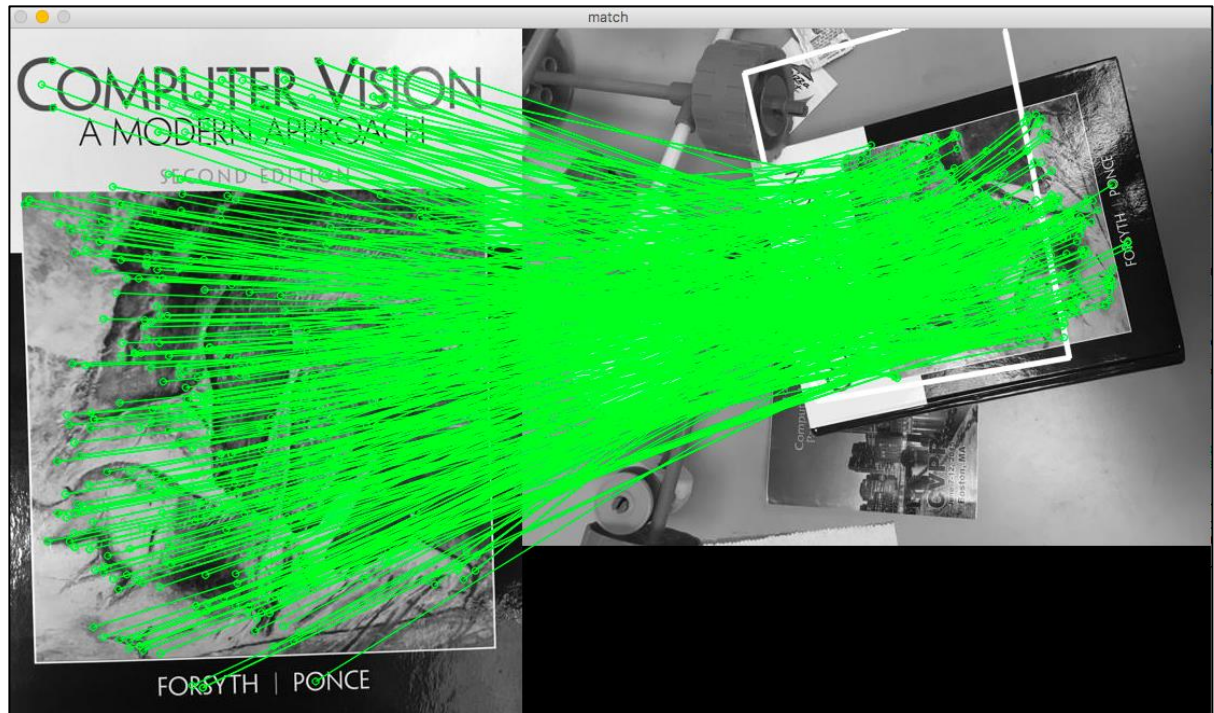


Target Image 5



Match between top 20 descriptors with minimum L2 distance  
according Lowe's ratio before RANSAC





Match between Inliers only after RANSAC

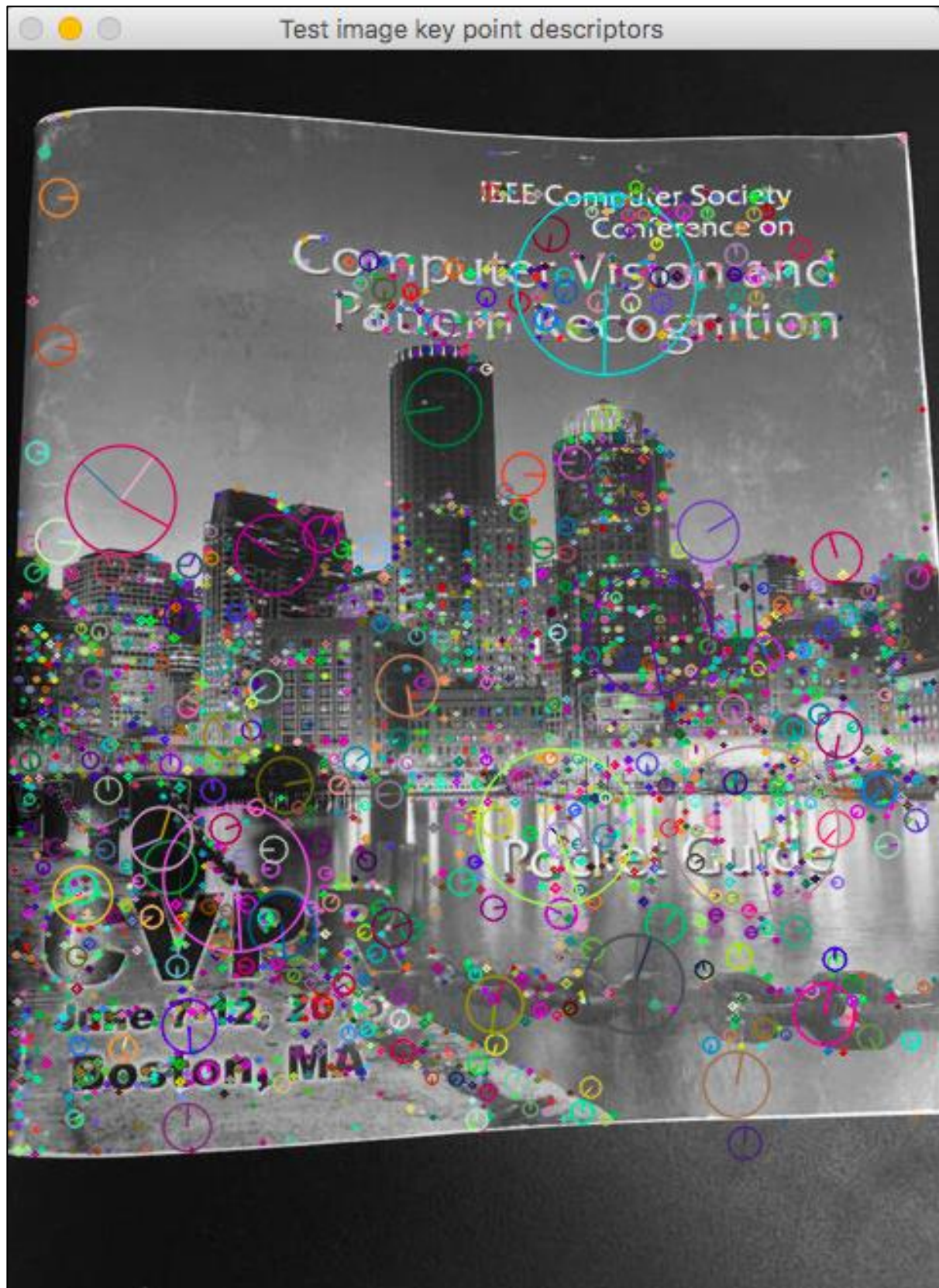
```
The total number of good matches found are: 562
The total number of consistent value inliers with the homography are: 525
The homography matrix found is as follows:
[[ -1.18028827e-01  4.07428330e-01  2.74530892e+02]
 [ -4.77027906e-01 -1.32868172e-01  3.45322623e+02]
 [ -5.26740430e-05 -1.60259647e-04  1.00000000e+00]]
```

Results and the Homography matrix obtained.

---

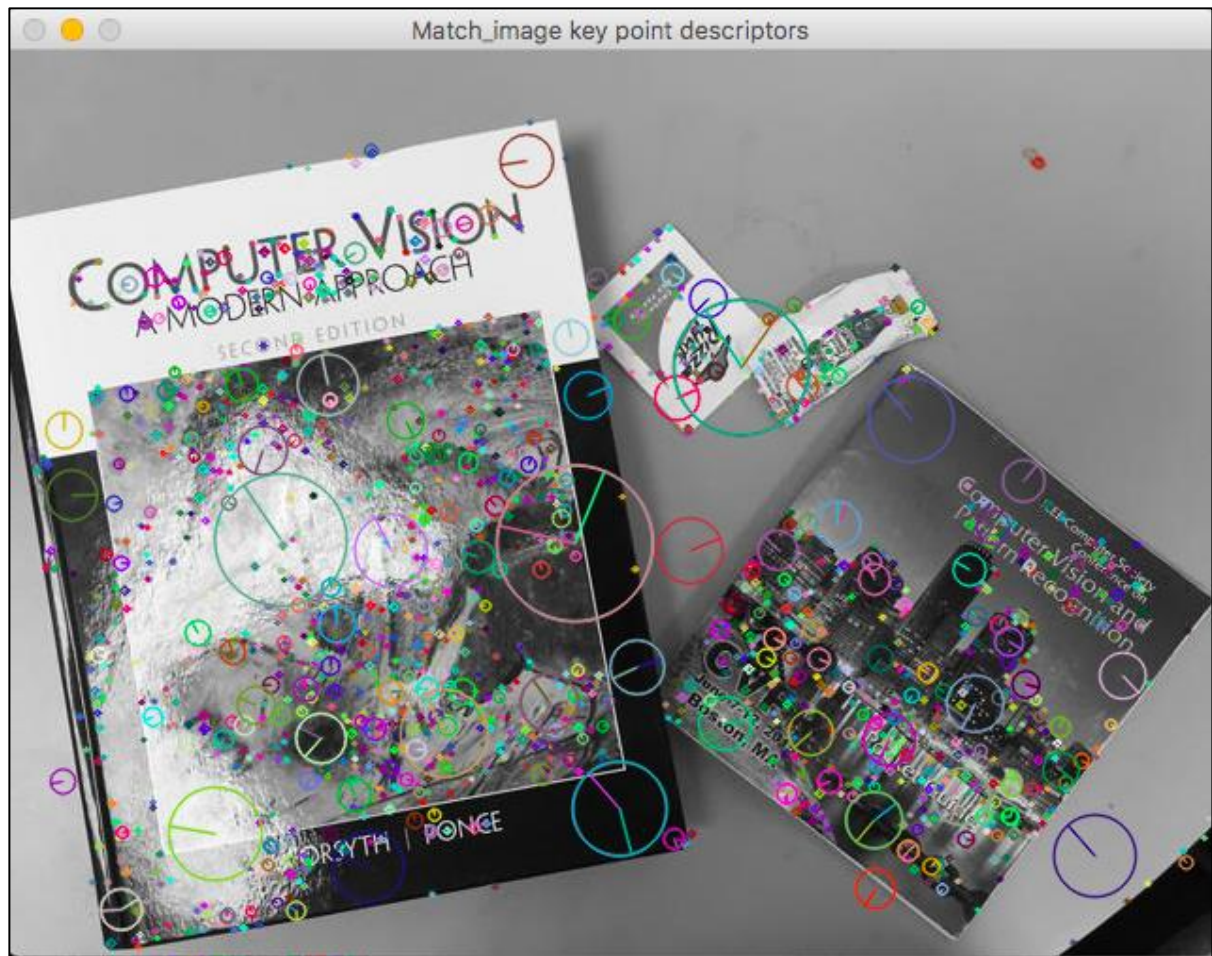
CASE 2 with Image 2 as the test image and Image 3,4,5 as targets

- Test image: Image2
- Target image: Image3

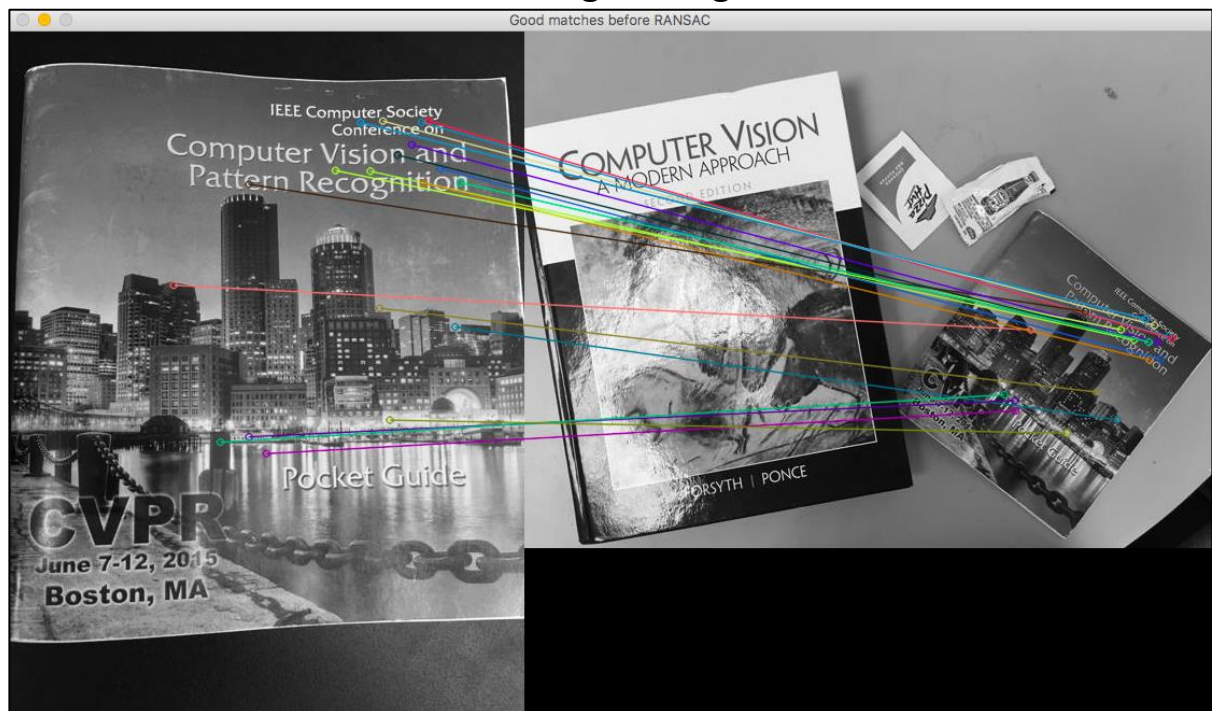


Test Image 2





Target Image 3



Match between top 20 descriptors with minimum L2 distance according Lowe's ratio before RANSAC



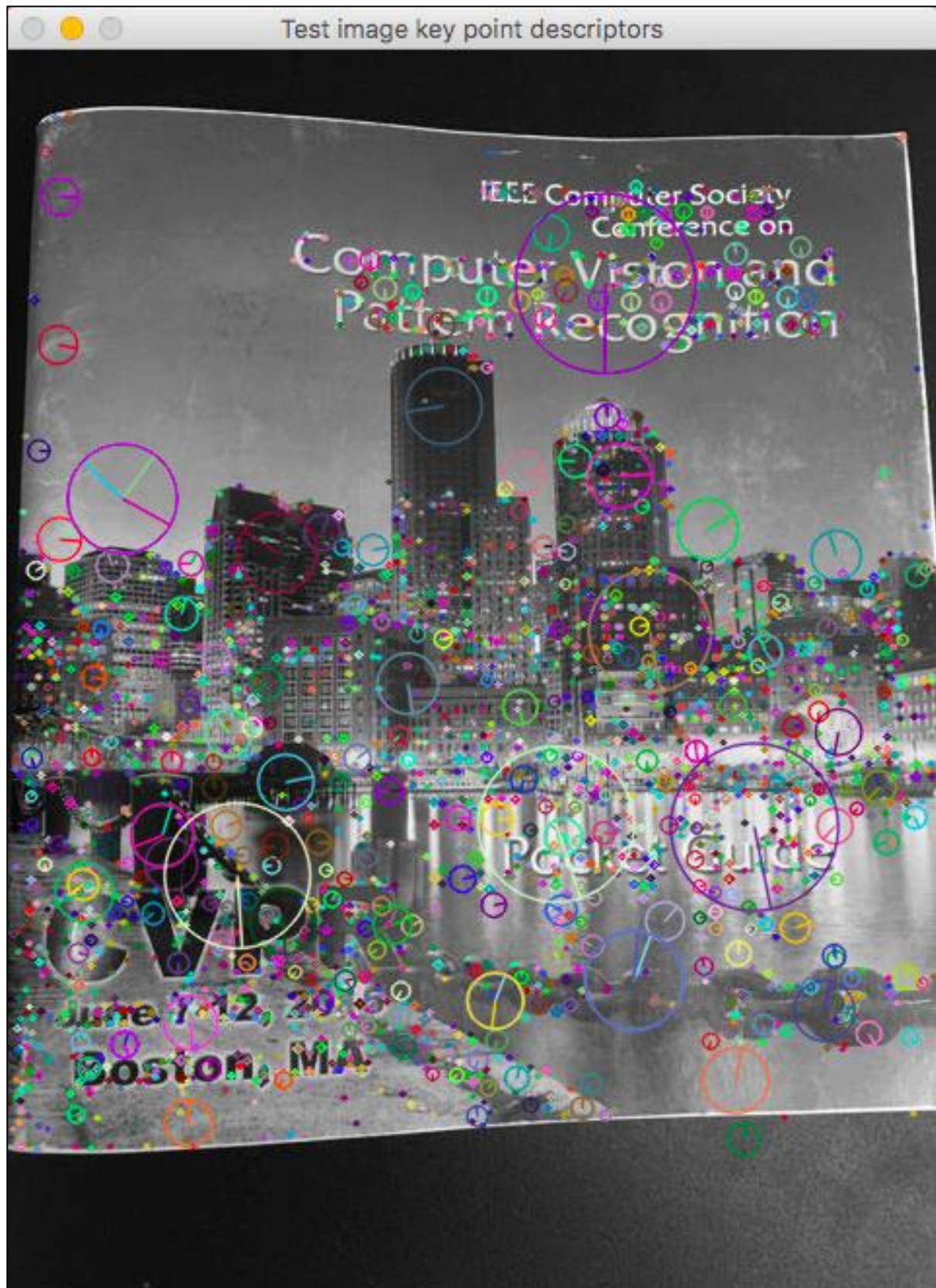
### Match between Inliers only after RANSAC

```
The total number of good matches found are: 194
The total number of consistent value inliers with the homography are: 165
The homography matrix found is as follows:
[[ 2.55982148e-01 -1.63149091e-01  4.82765148e+02]
 [ 2.11429953e-01  4.31831990e-01  1.50200415e+02]
 [-1.89712756e-04  2.46512055e-04  1.00000000e+00]]
```

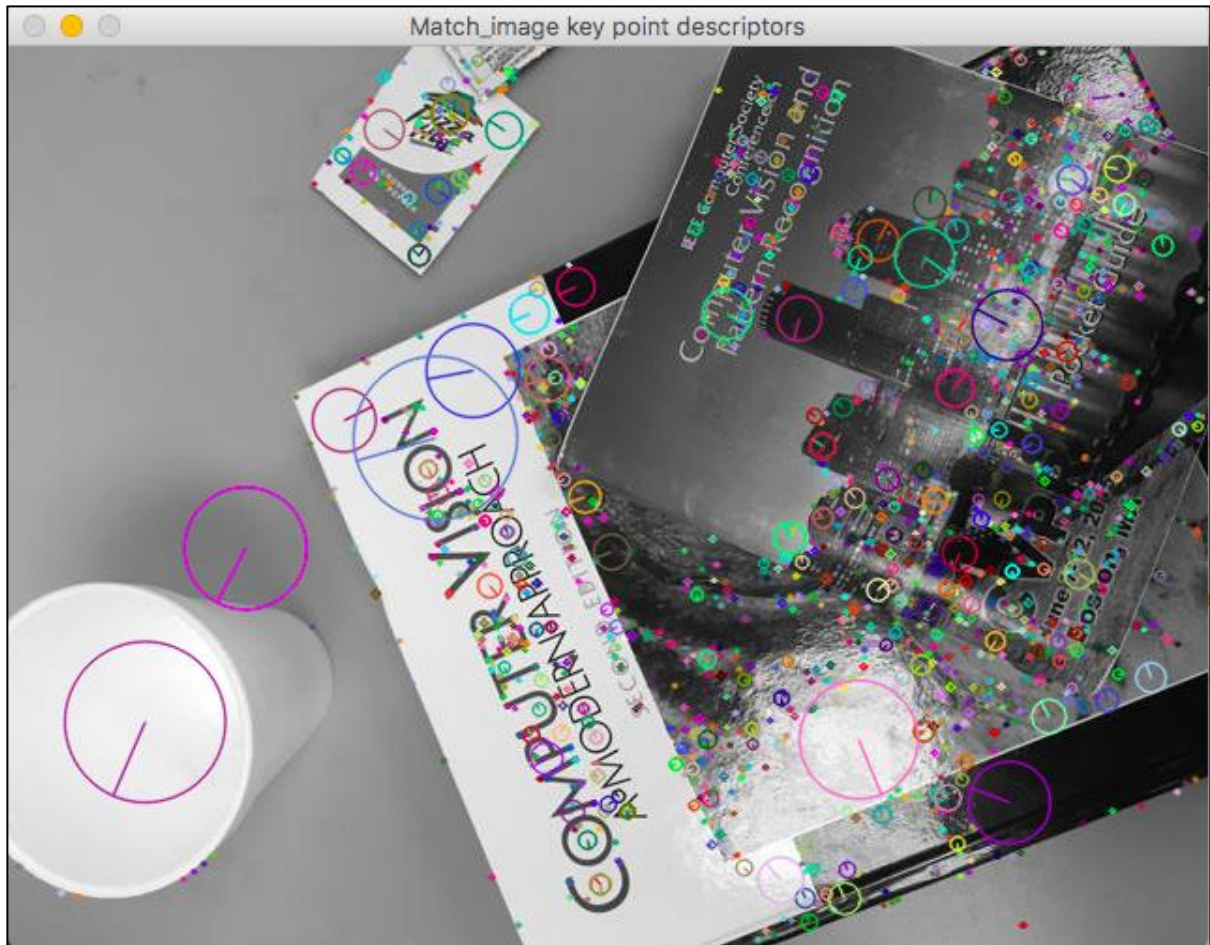
Results and the Homography matrix obtained.

- Test image: Image2
- Target image: Image4





Test Image 2

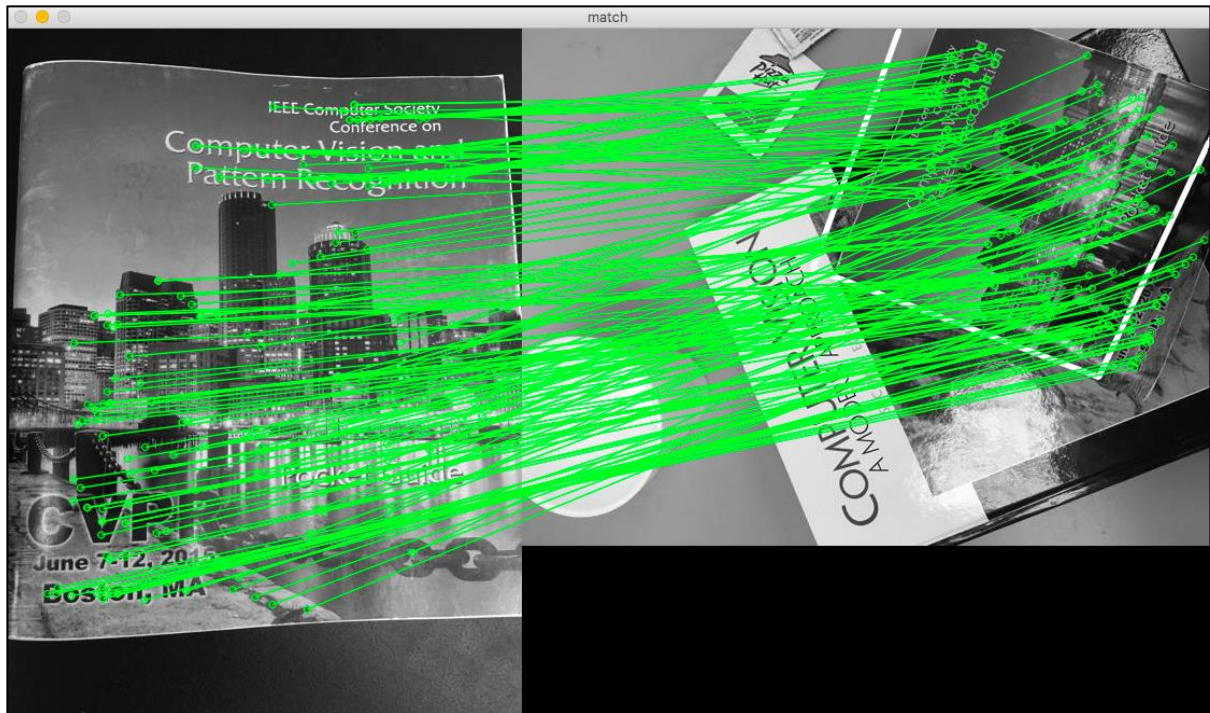


Target Image 4



Match between top 20 descriptors with minimum L2 distance according Lowe's ratio before RANSAC



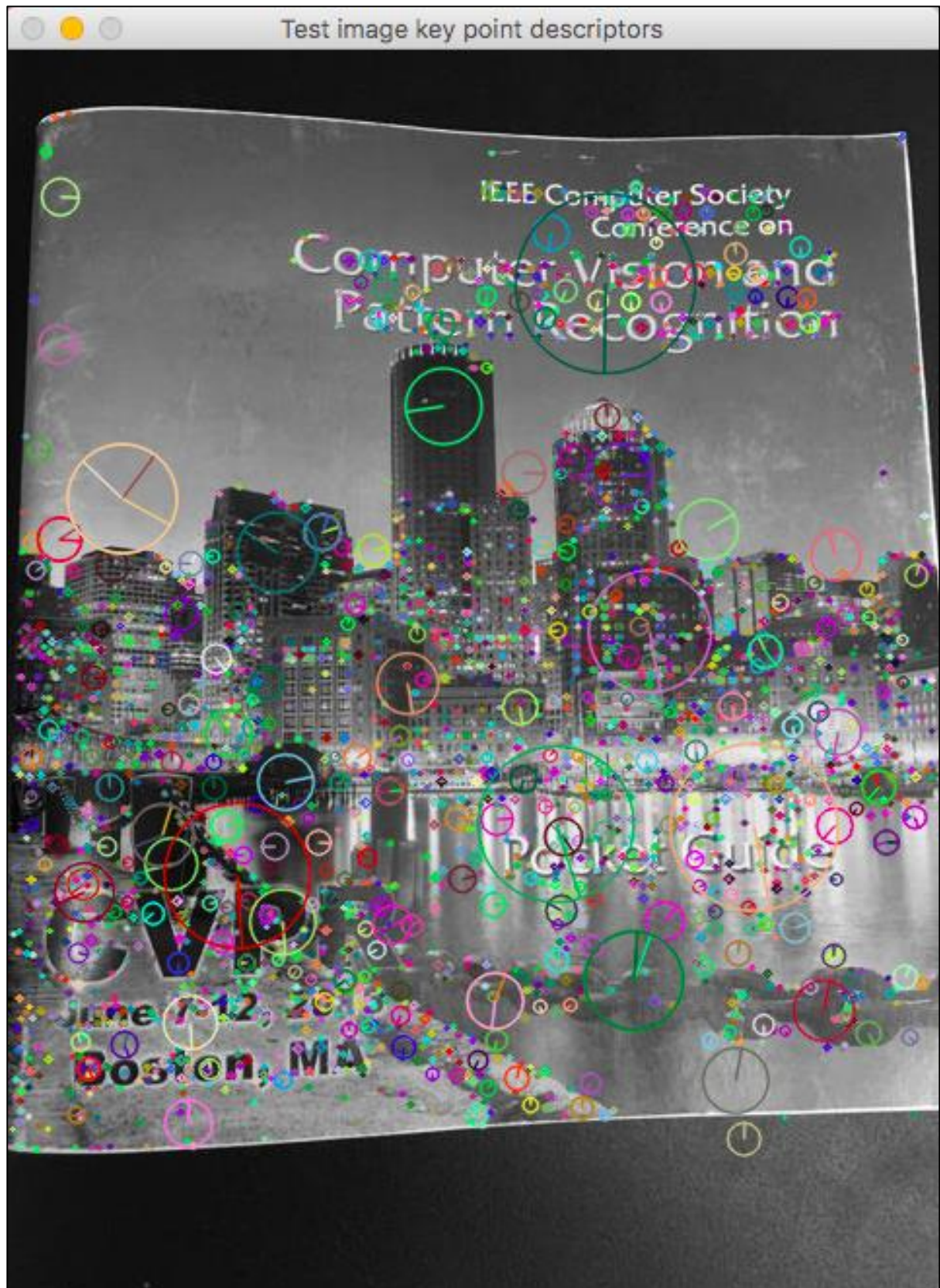


### Match between Inliers only after RANSAC

```
The total number of good matches found are: 210  
The total number of consistent value inliers with the homography are: 185  
The homography matrix found is as follows:  
[[ 1.39689732e-01  6.05250061e-01  2.70750538e+02]  
 [ -5.61297616e-01  2.40851508e-01  2.22543128e+02]  
 [ -1.78467339e-04  9.41572804e-05  1.00000000e+00]]
```

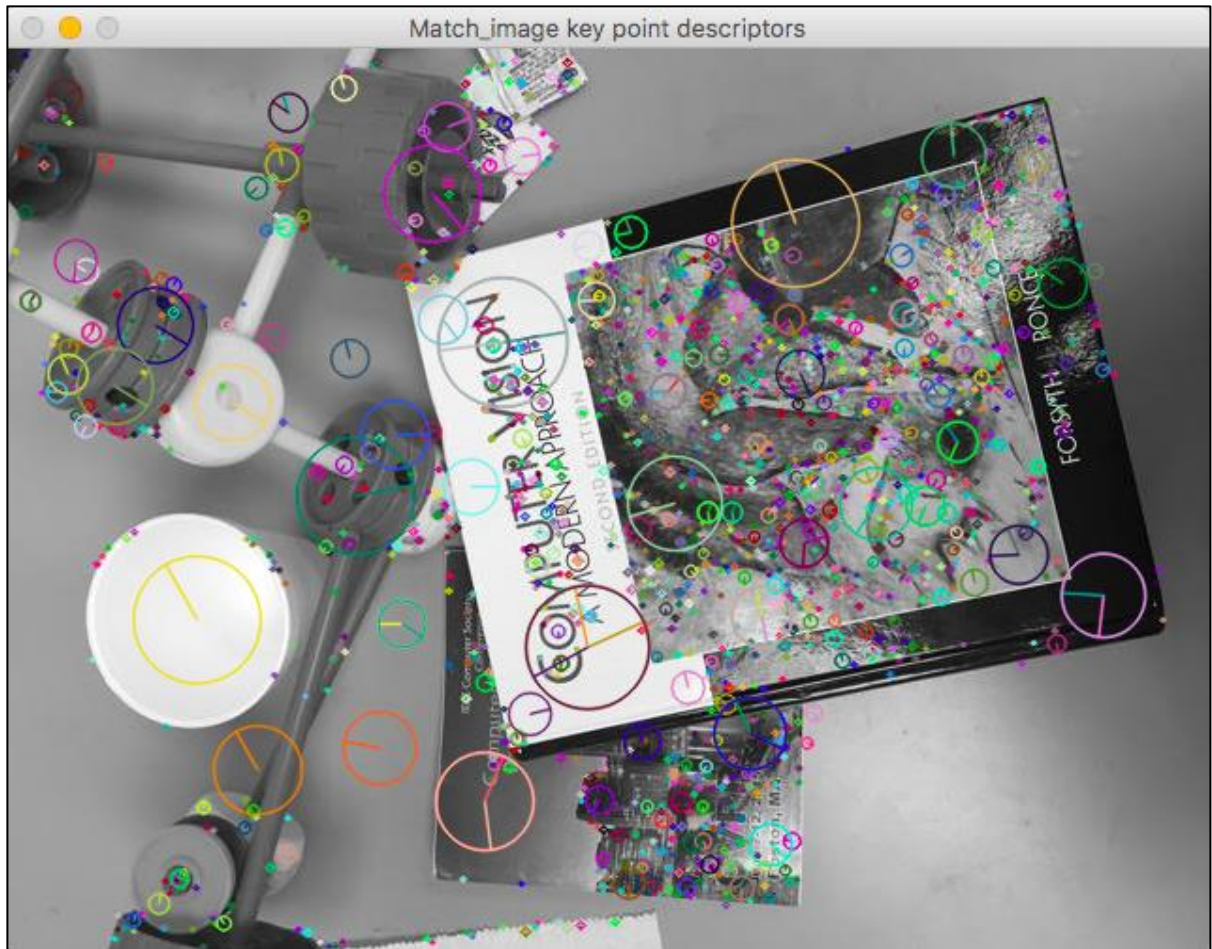
Results and the Homography matrix obtained.

- 
- Test image: Image2
  - Target image: Image5

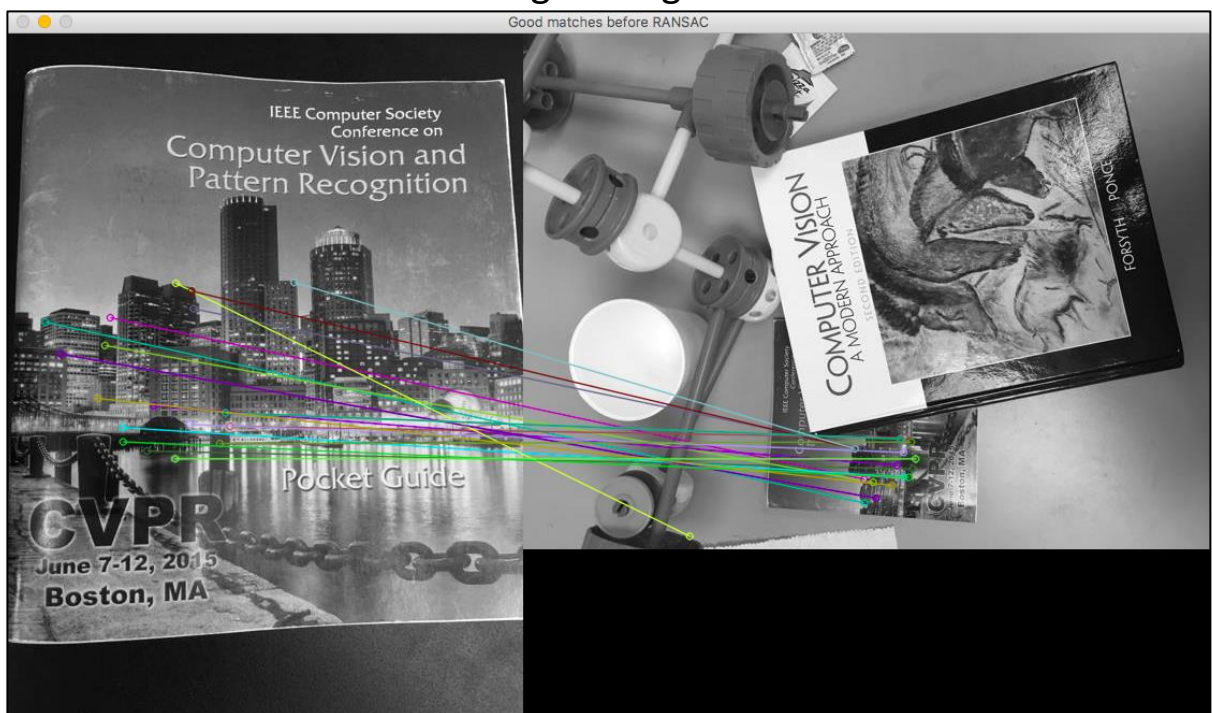


Test Image 2





Target Image 5



Match between top 20 descriptors with minimum L2 distance  
according Lowe's ratio before RANSAC



Match between Inliers only after RANSAC

```
The total number of good matches found are: 48
The total number of consistent value inliers with the homography are: 35
The homography matrix found is as follows:
[[ -5.06183515e-02  4.21519330e-01  2.16788477e+02]
 [ -4.54676565e-01  8.17136672e-02  4.44576300e+02]
 [ -1.97598700e-04  1.48423377e-04  1.00000000e+00]]
```

Results and the Homography matrix obtained.

## ANALYSIS BASED ON THE RESULTS

SIFT features are utilised to obtain the key points and the descriptors in each of the cash with the matching between the descriptors done with the aid of the brute force matcher with the good match criterion decided by D.Lowe ratio.

These descriptors are considered and the Homography matrix is computed with consideration of RANSAC algorithm for robust visual navigation between the 128 dimensional vector.

While carrying out this process only the inliers yield the best matches after the RANSAC algorithm is applied which increases the accuracy. The number of Inliers to the total good matches explains the total number of matches actually in the image that are matched between the test and the target image.



This percentage is higher in the case that the test image and the target image have large number of similar descriptors with minimum L2 distance between the feature vectors.

Consider the CASE 1: Image 1 is matched with Images 3, 4 and 5, highest inlier percentage is found in the case of Image 5 since it has rotated, scaled and translated version of itself with complete visibility and occludes the other image. Higher scores are also expected since there is complete visibility of the key points and the descriptors with similar intensity and variations are present. Similar illumination along with low image noise and small change in viewpoint also assist it.

The least amount of matches and inliers ratio is present in Image 4 since the other book is occluding the original test image to be matched. Since large number of key points to be matched with the test image are missing along with the clipped and cropped portions, it has the least matches as well as the least inliers due to miss out of the key points and the corresponding descriptors. Large amount of rotation is an hindrance too.

Image 1 and Image 3 have better accuracy than Image 4 since there are more key points that are visible as well as there is no occlusion of the other book. It is lesser than the case of Image 5 since there are many key points and descriptors similar to the test image with complete visibility which reduces the number of inliers and hence has a lower percentage of inliers to good matches that have been obtained. Low angle of rotation along with similar viewpoints assist this process.

Consider the CASE 2: Image 2 is matched with Images 3, 4 and 5, highest inlier percentage is found in the case of Image 4 since it has rotated, scaled and translated version of itself with complete visibility and occludes the other image. Higher scores are also

expected since there is complete visibility of the key points and the descriptors with similar intensity and variations are present.

The least amount of matches and inliers ratio is present in Image 5 since the other book is occluding the original test image to be matched. Since large number of key points to be matched with the test image are missing along with the clipped and cropped portions, it has the least matches as well as the least inliers due to miss out of the key points and the corresponding descriptors.

Image 1 and Image 3 have better accuracy than Image 5 since there are more key points that are visible as well as there is no occlusion of the other book. It is lesser than the case of Image 3 since there are many key points and descriptors similar to the test image with complete visibility which reduces the number of inliers and hence has a lower percentage of inliers to good matches that have been obtained.

In general, it is important to note that the SIFT algorithm works for grey scale images only and hence the illumination and intensity matter while calculation of the key points. Computationally faster results can be obtained by using SURF, ORB for feature detection.

Bf matcher was utilised for obtained matches with L2 norm criterion which could have been easily replaced by FLANN matcher for quicker computation as well as robust results.

RANSAC is a visual navigation technique which can also be substituted by Iterative closest point and similar algorithms for higher inlier percentage in the good matches.

Best results with the configuration set up in the problem statement can be obtained in the cases with good illumination, low rotation, small change in view point along with large number of similar key points availability along with low occluded regions.



