Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

16720-B Computer Vision: Homework 2

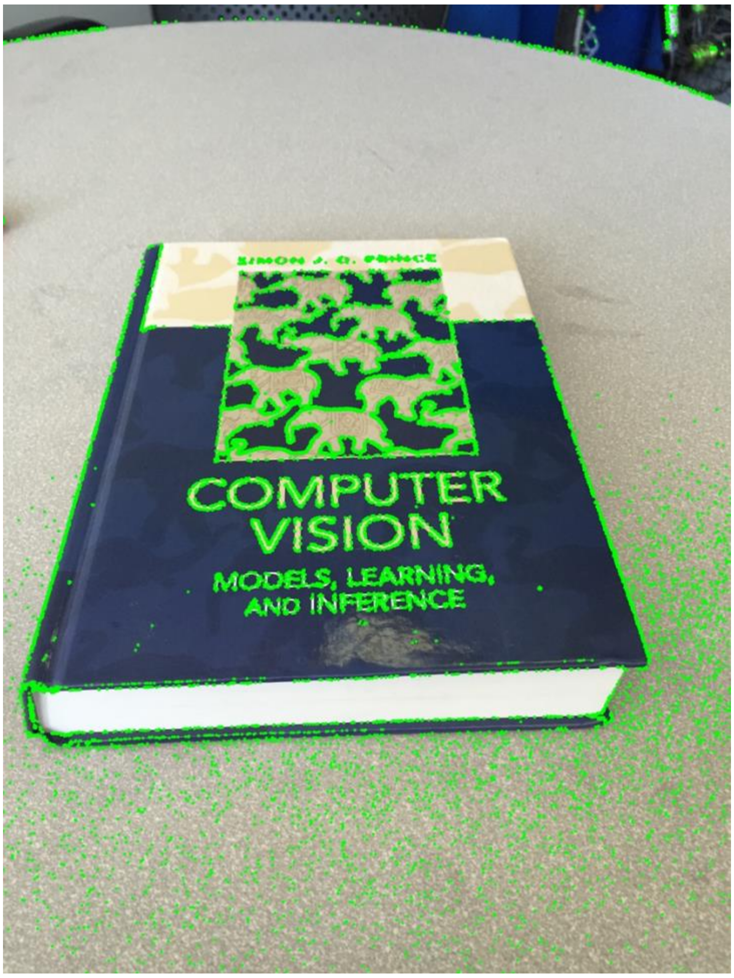
Q1.2 Image Pyramids

Figure 1: Gaussian Pyramid of Image

Figure 2: Difference of Gaussian Pyramid of Image

Q1.5 Interest Point Detection


College
Inn


COMPUTER
VISION
MODELS, LEARNING,
AND INFERENCE

Q2.4 Plot Matches

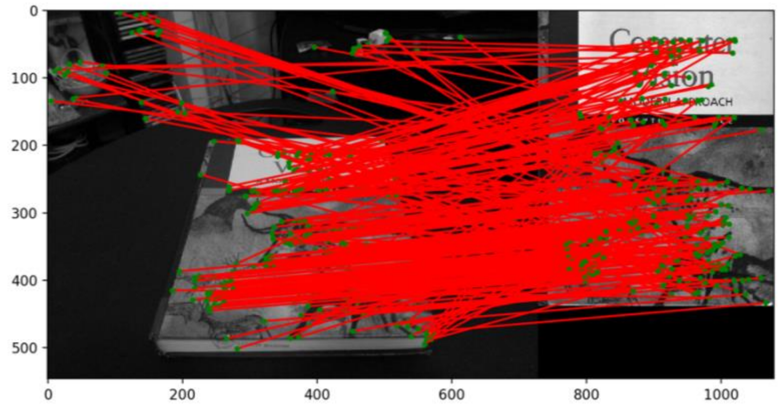


Figure 3: Book Image Matches

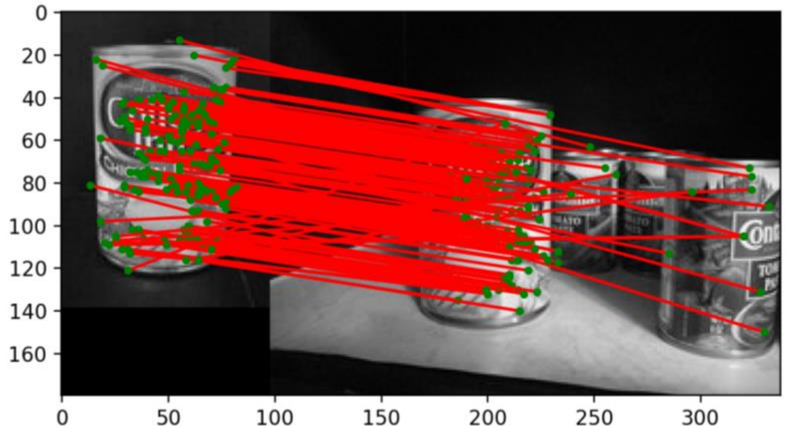


Figure 4: Chickenbroth Image Matches

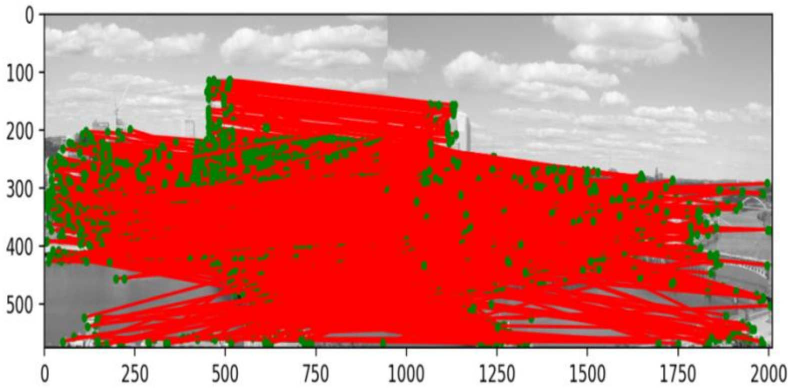


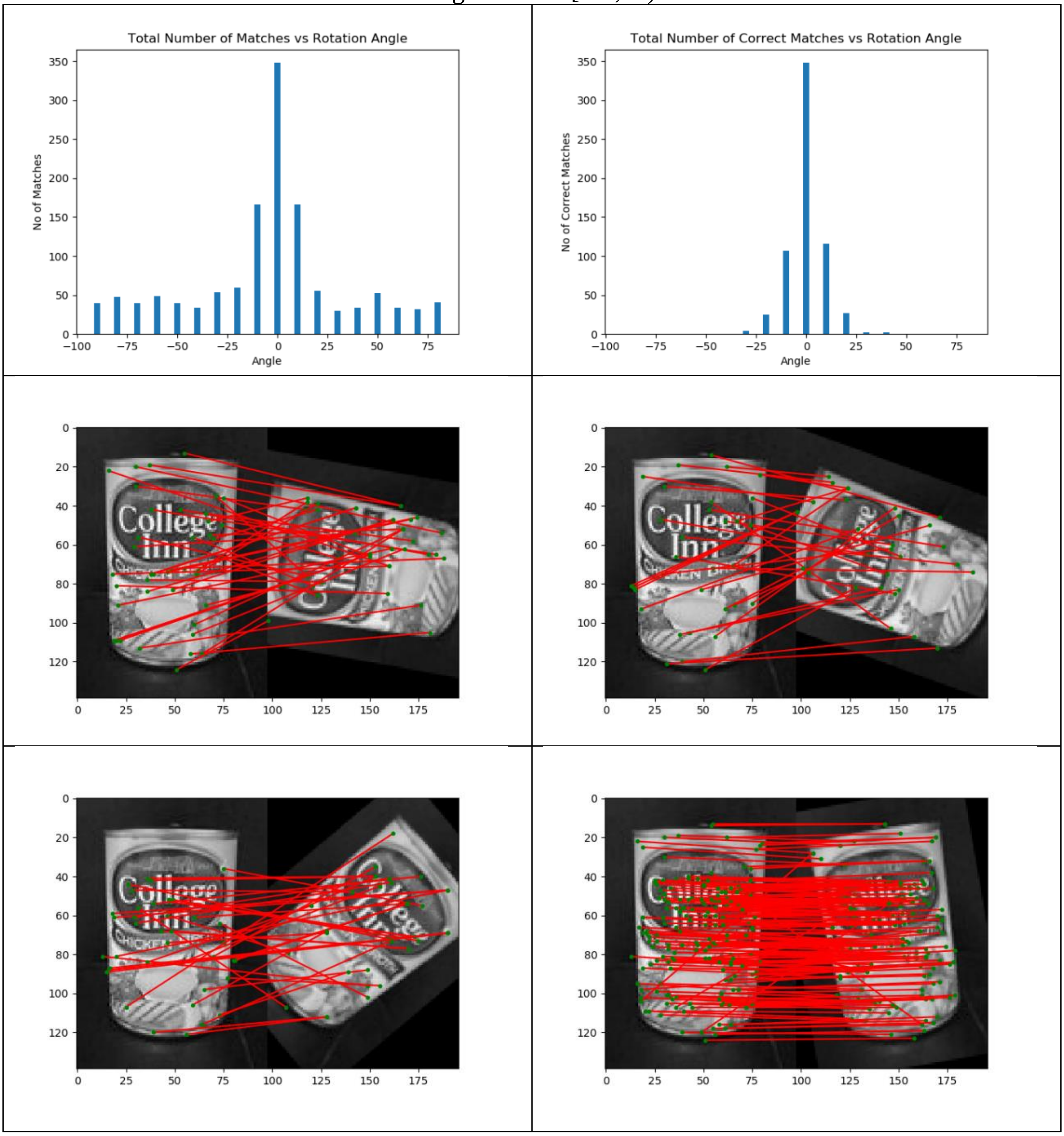
Figure 5: Incline Image Matches

It seems like BRIEF Descriptor matching works very well when there is just translation and scaling and seems to perform relatively bad when there are large rotational perturbations. For example, in the chickenbroth example we can see that some features are wrongly mapped to several other cans in the second picture.
BRIEF descriptors are local descriptors and if the images we try to compare with have very large changes or rich background, then local features of the subject in one image might get mapped to the background of another with similar local features. This will increase the number of mismatches and make it less reliable to calculate homographies in such cases.

Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

Q 2.5 (Rotation Test)

Performed the rotation at intervals of 10 degrees from [-90,90).



We can see that total number of matches reduces as the angle of rotation increases, we can also see that the number of correct matches reduces even more drastically as the angle of rotation increases.

I think this happens because the brief feature descriptor compares a fixed set of pairs of pixel positions to characterise a feature. A small thought experiment of rotating an image of a “plus” ***(+)*** sign and imagining what the BRIEF descriptor outputs can show this. It will compare different pixel intensities at different angles of rotation for essentially the same feature because it just compares values at a fixed set of locations. Hence it would fail at being rotationally invariant.

Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

Q3.1 Planar Homographies

1. We have,

$$\lambda_n \widetilde{\boldsymbol{x_n}} = \boldsymbol{H}\widetilde{u_n}$$

$$\lambda \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \phi_{13} \\ \phi_{21} & \phi_{22} & \phi_{23} \\ \phi_{31} & \phi_{32} & \phi_{33} \end{bmatrix} \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix}$$

Where,

$$\phi_{j,k} = Element\ of\ the\ Homography\ matrix\ H$$
$$\lambda = Arbitrary\ Scaling$$
$$\widetilde{x_n} = Homogeneous\ Coordinate\ in\ Camera\ View$$
$$\widetilde{u_n} = Homogeneous\ World\ Coordinates$$

If we consider homogeneous coordinates as vectors in space, then $\widetilde{x_n}$ and $H\,\widetilde{u_n}$ are in the same direction. Hence their cross-product is zero.

$$\widetilde{\boldsymbol{x_n}} \times H\widetilde{\boldsymbol{u_n}} = \boldsymbol{0}$$

Solving,

$$\lambda \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} \phi_{11}u_i + \phi_{12}v_i + \phi_{13} \\ \phi_{21}u_i + \phi_{22}v_i + \phi_{23} \\ \phi_{31}u_i + \phi_{32}v_i + \phi_{33} \end{bmatrix}$$

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \times \begin{bmatrix} \phi_{11}u_i + \phi_{12}v_i + \phi_{13} \\ \phi_{21}u_i + \phi_{22}v_i + \phi_{23} \\ \phi_{31}u_i + \phi_{32}v_i + \phi_{33} \end{bmatrix} = 0$$

$$\begin{bmatrix} y(\phi_{31}u_i + \phi_{32}v_i + \phi_{33}) - 1(\phi_{21}u_i + \phi_{22}v_i + \phi_{23}) \\ -x(\phi_{31}u_i + \phi_{32}v_i + \phi_{33}) + 1(\phi_{11}u_i + \phi_{12}v_i + \phi_{13}) \\ x(\phi_{21}u_i + \phi_{22}v_i + \phi_{23}) - y(\phi_{11}u_i + \phi_{12}v_i + \phi_{13}) \end{bmatrix} = 0$$

Converting these equations into the form of $Ah = 0$,

$$\begin{bmatrix} 0 & 0 & 0 & -u_1 & -v_1 & -1 & y_1u_1 & y_1v_1 & y_1 \\ u_1 & v_1 & 1 & 0 & 0 & 0 & -x_1u_1 & -x_1v_1 & -x_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & -u_n & -v_n & -1 & y_nu_n & y_nv_n & y_n \\ u_n & v_n & 1 & 0 & 0 & 0 & -x_nu_n & -x_nv_n & -x_n \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{12} \\ \phi_{13} \\ \phi_{21} \\ \phi_{22} \\ \phi_{23} \\ \phi_{31} \\ \phi_{32} \\ \phi_{33} \end{bmatrix}$$

Therefore we have,

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{u_1} & -\boldsymbol{v_1} & -\boldsymbol{1} & \boldsymbol{y_1u_1} & \boldsymbol{y_1v_1} & \boldsymbol{y_1} \\ \boldsymbol{u_1} & \boldsymbol{v_1} & \boldsymbol{1} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{x_1u_1} & -\boldsymbol{x_1v_1} & -\boldsymbol{x_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{u_n} & -\boldsymbol{v_n} & -\boldsymbol{1} & \boldsymbol{y_nu_n} & \boldsymbol{y_nv_n} & \boldsymbol{y_n} \\ \boldsymbol{u_n} & \boldsymbol{v_n} & \boldsymbol{1} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{x_nu_n} & -\boldsymbol{x_nv_n} & -\boldsymbol{x_n} \end{bmatrix}$$

Arcot Sridhar Kartik Sai Vishnu Kumar
karcotsr@andrew.cmu.edu

2.
There are nine elements in h. These elements of h are the elements of the Homography (H) between the camera and worldviews.

3.
There are nine elements in the homography, however we can scale the homography such that one of the elements can have a fixed value, since there is an arbitrary scale factor in our equations which we can solve for later. Hence we essentially need to solve for eight variables. Therefore H has 8 degrees of freedom and we need 4 pairs of point correspondences to solve for H.

4. (Proof inferred from Simon J.D. Prince Computer Vision Textbook Appendix C7.2, C5)
The problem,

$$Ah = 0 \text{ } such \text{ } that \text{ } |h| = 1$$

Can be rephrased as,

$$arg\min_{h}[Ah] \text{ } such \text{ } that \text{ } |h| = 1$$

- This is similar to the Principal Direction problem.
- We are trying to find the direction which will be mapped to the minor axis of the ellipsoid resulting from the linear mapping A.

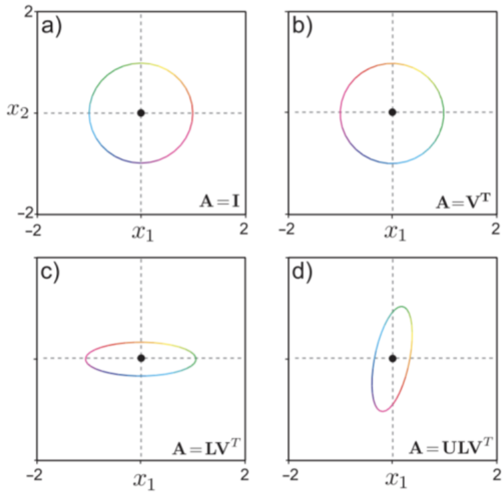


Figure C.3 Cumulative effect of SVD components for matrix $\mathbf{A}_3$. a) Original object. b) Applying matrix $\mathbf{V}^T$ rotates and reflects the object around the origin. c) Subsequently applying $\mathbf{L}$ causes stretching/compression along the coordinate axes. d) Finally, applying matrix $\mathbf{U}$ rotates and reflects this distorted structure.

Figure 6: Illustration from Simon J.D Prince Appendix C.5

- By performing Singular Value Decomposition of A-

$$SVD(A) = ULV^T$$

  1. Here U and V are orthogonal matrices which are essentially rotations
  2. L is the diagonal vector with eigen values which is the scaling of the new basis vectors we get from V.
  3. SVD can be looked at as a Rotation by V, Scaling by L and another rotation by U.

- Therefore from the above propositions and figure above we can say that the minor axis we are looking for is the row in V_T which is being scaled by the smallest eigenvalue in the diagonal matrix L.
- Ideally, if there is no noise in the matches we have from the image to world coordinates, then the minimisation problem will find a h such that Ah=0.

Q6.1

Figure 7: Warped but clipped incline_R.png

Figure 8: Clipped Panorama

Q6.2

Figure 9: Warped and non-clipped incline_R.png

Figure 10: Panorama generated with no clipping

Q6.3

Figure 10: Panorama generated with no clipping

Q 7.2

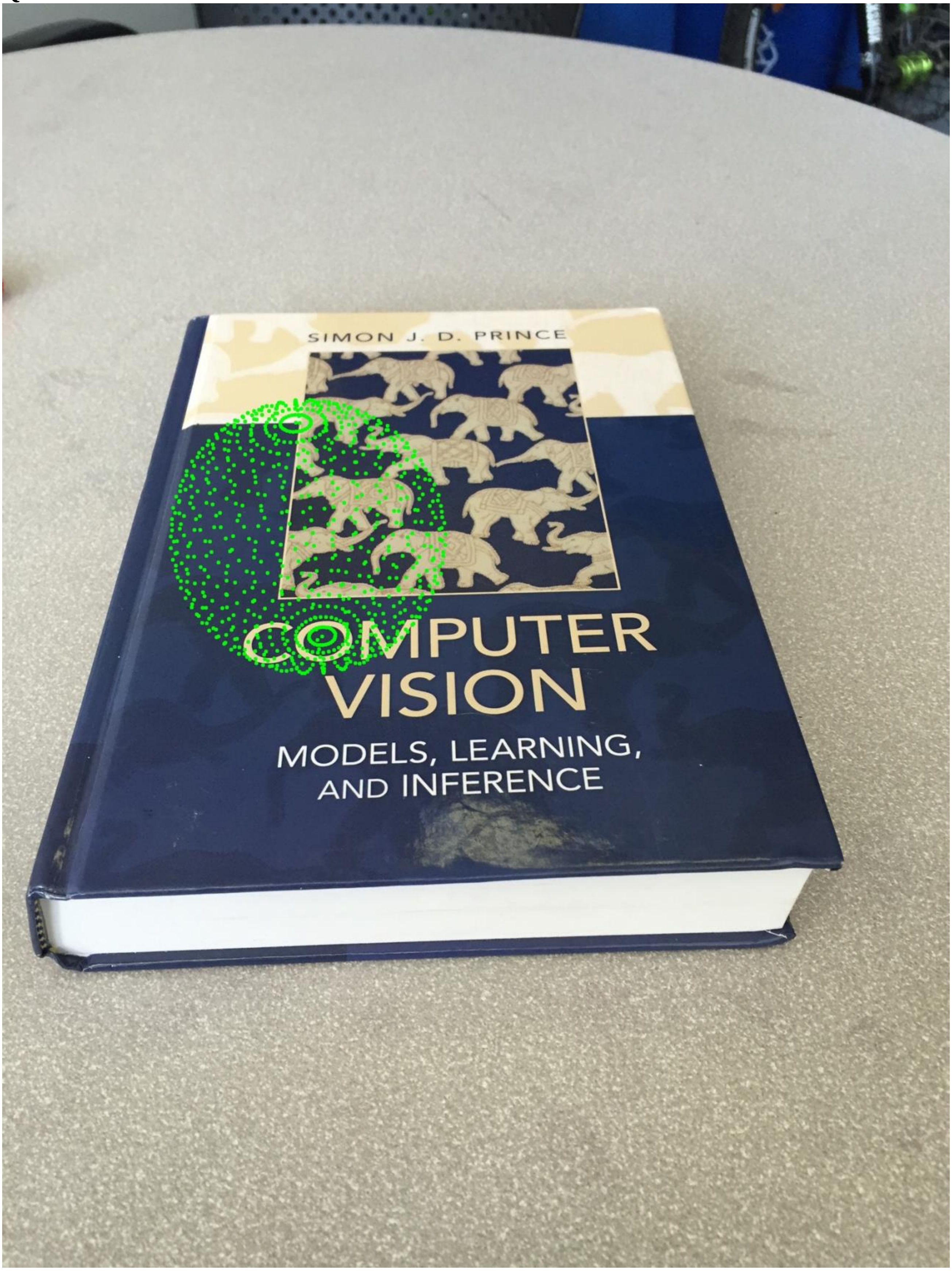


Figure 11: Ball projected on Book