CSE/ECE 344/544 Quiz 5 Solutions – Computer Vision – Winter 2023 Apr. 27, 2023

Time: 30 minutes Max Marks: 14 (UG/PG)

Instructions:
- Do not plagiarize. Do not assist your classmates in plagiarism.
- Show your full solution for the questions to get full credit.
- Attempt all questions that you can.
- In the unlikely case a question is not clear, discuss it with an invigilator. Please ensure that you clearly include any assumptions you make, even after clarification from the invigilator.

1. (2+1+1+3=7 points) Answer the following about Harris corner detectors.

   a) Is the gradient of the image required for detecting the Harris corners in the image? If yes, explain how are gradients computed, else, explain how Harris corners are detected in the image directly.

   b) Justify or refute. The Harris corner detection relies on the approximate SSD (sum of squared differences) error computed over small local displacements of a patch/window.

   c) Justify or refute. The SSD error is computed explicitly for each patch and the maximum is chosen based on the eigenvalues of the error matrix.

   d) Write the matrix for which the eigenvalues are computed. Explain each component and provide the rationale (and one of the formulae) of using the eigenvalues for identifying Harris corners.

Solution:

a) Yes, the gradient of the image is required for detecting the Harris corners in an image. The gradients can be computed by devising a filter that performs smoothing and differenciation, e.g., discrete approximations of a derivative of a Gaussian like the Sobel operator. The image is convolved with such a filter for differentiating it wrt x and wrt y. The two images combined give the gradient of intensity at each pixel.

b) Correct. The Harris corner detector maximizes the (first order, i.e., up to first derivative in the Taylor series) approximation of the SSD error computed over a patch or window for small displacements.

c) Incorrect. The SSD error is never computed explicitly, instead the quadratic form of the approximate error is parametrized by the gradient image and used for identifying the corners by analysing eigenvalues of the parametrized matrix.

d) The slide screenshot in Fig. 1 answers this question. Here, $\mathbf{I}_x$ and $\mathbf{I}_y$ are the partial derivatives of

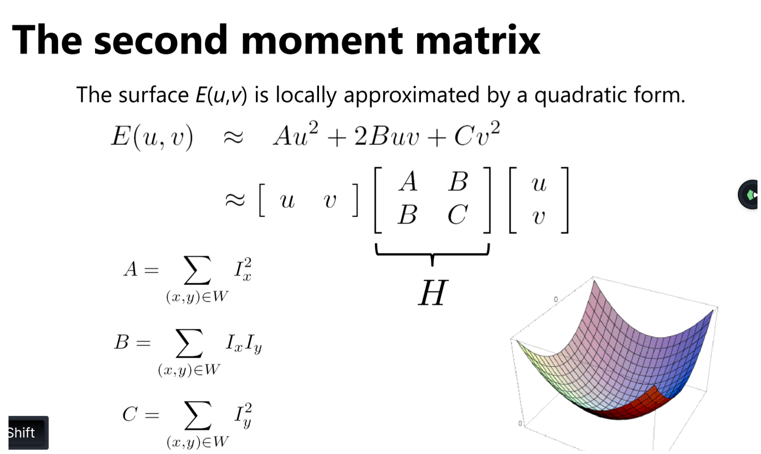

Figure 1: Quadratic form of the error matrix.

the Image intensity with respect to x and y respectively.

2. (2+2=4 points) Answer the following about Random Sample Consensus (RANSAC).

   a) Given a set of 3D points, provide the steps for using RANSAC for fitting a plane. You may assume that the acceptable error threshold for plane fitting is known to be $\tau = 3$cm.

   b) Let there be two images (from two different viewpoints) of a scene that comprises of two planes. One of the planes has many chessboard patterns (the one we used for calibration) stuck on it, while the other has very little texture (i.e., only a few corners). Write the steps to estimate a planar Homography between the two images. Which plane do you think is likely to be detected first by RANSAC? Explain why.

   Solution:

   a)

   1. Input the data set of the 3D points $[\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$, the residual threshold τ and the residual function (usually $\widehat{n}^\top \mathbf{x} + d$ for planes, where $\widehat{n}$ is the normal, and the d is the offset.). Also input (or calculate based on estimatie of percentage of expected inliers) the maximum number of iterations to be done.
   2. Repeat:
      (a) Randomly sample (with replacement) 3 points. Reject them if they are collinear. Else, generate a model hypothesis by fitting a plane to the three points (estimating the normal and the offset $(\widehat{n}, d)$.
      (b) Compute the residual for each point with respect to the model hypothesis.
      (c) Retain the number of *inlier* points for which the residuals are lower than τ.
   3. Pick the model hypothesis that has the maximum number of inliers (highest consensus) and perform least squares fit using all selected inliers to obtain the plane parameters.

   b)

   1. Input the data set of the potential (or putative) 2D point correspondences (matches) $[\mathbf{x}_1^{(1)}, \mathbf{x}_2^{(1)}, \ldots, \mathbf{x}_N^{(1)}]$ and $[\mathbf{x}_1^{(2)}, \mathbf{x}_2^{(2)}, \ldots, \mathbf{x}_N^{(2)}]$. Input the residual threshold τ and the residual function (symmetric transfer error, algebraic error.). Also input (or calculate based on estimatie of percentage of expected inliers) the maximum number of iterations to be done.
   2. Repeat:
      (a) Randomly sample (with replacement) 4 point correspondences. Generate a model hypothesis by fitting a plane to the three points (estimating the homography $\mathbf{H}$.
      (b) Compute the residual for each point with respect to the model hypothesis.
      (c) Retain the number of *inlier* points for which the residuals are lower than τ.
   3. Pick the model hypothesis that has the maximum number of inliers (highest consensus) and perform a least squares fit using all selected inliers to obtain the Homography.

   The plane that has the chessboard patterns stuck on it is more likely to be detected first with RANSAC as it has more number of corners points and therefore will allow RANSAC converge quickly.

3. (1+2=3 points) Answer the following about Scale Invariant Feature Transform (SIFT).

   a) Why are SIFT feature descriptors said to be rotation invariant?

   b) List the advantages of using histograms of gradient orientations as a choice of feature descriptors.

   Solution:

   a) SIFT features are said to be rotation invariant because they use blob / corner detection that only relies on the local gradient magnitudes (via eigenvalues, similar to Harris corners).

b) 1. Histograms approximate the distribution of gradient directions in a local neighborhood of the keypoints, which makes the description more robust to illumination and scale changes.
2. The histogram based descriptors can be compared with other similar descriptors using efficient distance measures like EUclidean, Chi-squared, etc.
3. Histogram of gradient orientations are invariant to rotations as well, so long as the bins are organized with respect to a domainant rotation.