2D Features

Faizudeen Kajogbola

1 UNDERSTANDING IMAGE FEATURES

1. **Pros and cons of local features (edge, corner and point) and their applications in computer vision:** Local features often give the semantic and shape information of an image. They can be used to detect geometric events such as surface orientation, depth, and color/texture discontinuities. They also present the following advantages:

   - Locality
   - Distinctiveness
   - Quantity
   - Efficieincy
   - Generalilty

   However, in searching for local features, photometric events such as changes in brightness and intensity may also be detected.

   Feature points can be used by higher-level computer vision algorithm for:

   - Object recognition
   - Image alignment
   - 3D reconstruction
   - Robot navigation
   - Indexing and database retrieval

2. **Criteria of designing a good edge detector, an example and the rough process:** A good edge detector must:

   - Minimize the probability of false positives and false negatives (i.e. must be robust to noise).
   - Localize detected edges such that they are as close as possible to the true edges.
   - Must return a single point for each true edge point by minimizing the number of local maxima around the edge point.

   Canny edge detector is an example of a good edge detector.

   Edge detection as described by Canny is his paper [**?**] can generally be divided into 4 steps:

   - Image filtering with derivative of Gaussian
   - Finding the magnitude and orientation of gradient
   - Employing hysteresis thresholding by using high threshold to start and low threshold to link edge curves
   - Performing non-maximum suppression to reduce multi-pixel wide ridges into single-pixel lines

3. **The core matehmatical ideas of feature detection:** Mathematically, we can percieve feature detection by considering the sum square difference (SSD) (or correlation) of a window with a version of itself that is shifted by u in the x-direction and v in the y-direction. We can define the error function as:

$$E(u,v) = \sum_{(x,y)\in W} \left(I(x+u,y+v) - I(x,y)\right)^2$$

with $I(x,y)$ as image pixels and W as the window.

From Taylor's series approximation, we have:

$$I(x+u,y+v) \approx I(x,y) + \frac{\partial I(x,y)}{\partial x}u + \frac{\partial I(x,y)}{\partial y}v$$

Let $I_x := \frac{\partial I(x,y)}{\partial x}$ and $I_y := \frac{\partial I(x,y)}{\partial y}$

Then we can essentially rewrite the error function as:

$$E(u,v) = \sum_{(x,y)\in W} \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} I_x^2 & I_xI_y \\ I_yI_x & I_y^2 \end{bmatrix} \begin{bmatrix} u & v \end{bmatrix}$$

The nature of a feature is dictated by the shape of the error function $E(u,v)$. Considering an SSD error function, a relatively flat error function implies a flat region; a valley implies an edge; while a sharp drop in the center of an otherwise flat error function implies a corner.

Since the above equation is a quadratic, we can effectively characterize the error $E(u,v)$ and analyze it's shape by examining the properties of the matrix:

$$H := \begin{bmatrix} I_x^2 & I_xI_y \\ I_yI_x & I_y^2 \end{bmatrix}$$

Let λ_+ and λ_- be the maximum and minimum eigenvalues of H respectively. The error function $E(u,v)$ is lower bounded by $\sum_{(x,y)\in W} \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} \lambda_- & 0 \\ 0 & \lambda_- \end{bmatrix} \begin{bmatrix} u & v \end{bmatrix}$, thus, since $E(u,v)$ experiences large changes for small shifts in the window at corners, the value of λ_- at corners must be large.

4. **Mathematical ideas of Harris corner detector:**

   The Harris corner detector reduces the computation complexity by utilizing the "Harris Operator" which is defined as:

$$f := \frac{\lambda_1\lambda_2}{\lambda_1+\lambda_2}$$

   With $\lambda_1 = \lambda_+$ and $\lambda_2 = \lambda_-$.

   This operator eliminates the need to explicitly computing the eigenvalues of H.

   The Harris conrner detector uses a measure of cornerness which is defined as $R := (\lambda_1\lambda_2) - k(\lambda_1+\lambda_2)^2$ with $k \in [0.04, 0.06]$. At corners, the value of R is large. Small R implies a flat region. At the edges, R is negative with large magnitude.

5. **Is the Harris corner detector robust with respect to intensity changes in the image? Why or why not:**

   The Harris corner detector is partially robust with respect to affine intensity changes in the image. This is because affine intensity change is a linear transformation, and the corner response is scaled linearly. However, the position of points on the new corner response with respect to the threshold may change due to the scaling.

6. **Is the Harris corner detector robust with respect to rotation? Why or why not:** The Harris corner detector is invariatnt with respect to rotation, this is beacause the eigenvalues of H do not change as a result of image rotation.

7. **The importance of invariance when describe a feature and how to achieve invariance:**

   Invariance when describing a feature is essential if we want to be able to match the same feature in a transformed version of the image. Invariance can be achieved by finding the characteristic scale of each feature and including it in the feature description.

8. **Methods for comparing two patches in an image:**

   Two patches can be compared by computing the repeatability rate which is given by:

$$repeatability = \frac{\#correspondences}{\#detected} * 100\%$$

9. **The ideas and steps of SIFT feature detection and the advantages of SIFT compared to Harris:**

   Scale Invariant Feature Transform (SIFT) involves the following steps:

   - Feature detection: Keypoints and corresponding characteristic scales on the image to be matched are defined by evaluating the extrema of the result of the difference of gaussians (DoG) function applied in scale-space.
   - Feature matching: Keypoints and corresponding characteristic scales are identified on the image to be compared; each keypoint is scaled by different factors and compared to the keypoints extracted from the image to be matched; matching points are evaluated by: $\frac{A \cap B}{A \cup B} > 60\%$

   SIFT has a significantly higher repeatability rate than the Harris detector for images scaled by a factor ≥ 1.5.

10. **The ideas of HOG as descriptors in in SIFT feature detection:**

REFERENCES