CS-570 Computer Vision

Nazar Khan

Department of Computer Science
University of the Punjab

6. The Structure Tensor

Corners

- Just like edges, corners are perceptually important.
- More compact summary of an image since corners are fewer than edge pixels.
- A patch around a corner pixel is different from all other surrounding patches.

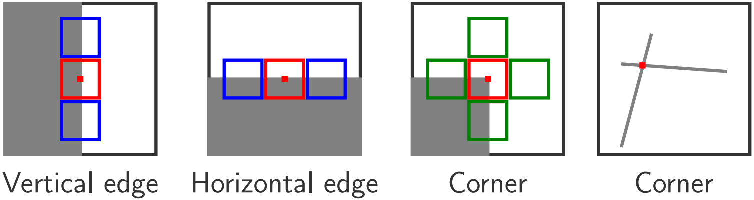


Figure: A patch containing a corner is different from all surrounding patches. Blue squares represent patches similar to the red patch. Green squares represent patches different from the red patch. Author: N. Khan (2021)

How to compare patches
Sum-squared-distance (SSD)

- For two patches P and Q of size $m \times n$ pixels, their dissimilarity can be computed using a *sum-of-squared distances*

$$SSD(P, Q) = \sum_{i=1}^{m} \sum_{j=1}^{n} (P_{ij} - Q_{ij})^2$$

- Alternatively, weighted dissimilarity can be computed as

$$SSD(P, Q) = \sum_{i=1}^{m} \sum_{j=1}^{n} w_{ij}(P_{ij} - Q_{ij})^2$$

  where weight w_{ij} determines the importance of location (i, j).
- For example, Gaussian weights give more importance to the central pixel difference.

Taylor's Approximation for 2D Functions

- Recall that Taylor's approximation for 1D functions is

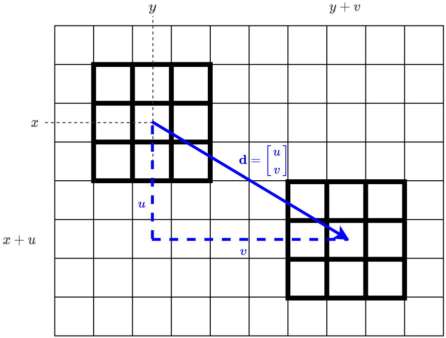
$$f(x+u) = f(x) + \frac{u}{1!}f'(x) + \frac{u^2}{2!}f''(x) + O(u^3)$$

- For 2D functions, a 2nd-order Taylor's approximation is

$$\begin{aligned} f(x+u, y+v) \approx & f(x,y) + \underbrace{\frac{u}{1!}f_x(x,y) + \frac{v}{1!}f_y(x,y)}_{\text{1st-order}} \\ & + \underbrace{\frac{u^2}{2!}f_{xx}(x,y) + \frac{v^2}{2!}f_{yy}(x,y) + \frac{2uv}{2!}f_{xy}(x,y)}_{\text{2nd-order}} \end{aligned}$$

Structure Tensor

- Let us consider patches of size 3×3 although the method works for patches of any size and shape.
- The color value of a pixel displaced from (x, y) by the direction vector $\mathbf{d} = (u, v)^T$ is $I(x + u, y + v)$.



Structure Tensor

- Weighted SSD between a patch at (x, y) and a patch displaced by the direction vector $\mathbf{d} = (u, v)^T$ is computed as

$$SSD(u, v) = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} w_{ij}(I(i+u, j+v) - I(i,j))^2$$

- Using a 1st-order Taylor's approximation

$$I(i+u, j+v) \approx I(i,j) + uI_x(i,j) + vI_y(i,j)$$

Structure Tensor

- Weighted SSD can be approximated as

$$
\begin{aligned}
SSD(u,v) &\approx \sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} w_{ij}(I(i+u,j+v)-I(i,j))^2 \\
&= \sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} w_{ij}(I(i,j)+uI_x(i,j)+vI_y(i,j)-I(i,j))^2 \\
&= \sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} w_{ij}(uI_x(i,j)+vI_y(i,j))^2 = \sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} w_{ij}(\mathbf{d}^T\nabla I_{ij})^2 \\
&= \sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} w_{ij}(\mathbf{d}^T\nabla I_{ij})(\mathbf{d}^T\nabla I_{ij})^T = \sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} w_{ij}\mathbf{d}^T\nabla I_{ij}\nabla I_{ij}^T\mathbf{d} \\
&= \mathbf{d}^T\left(\sum_{i=x-1}^{x+1}\sum_{j=y-1}^{y+1} w_{ij}\nabla I_{ij}\nabla I_{ij}^T\right)\mathbf{d} = \mathbf{d}^T A\mathbf{d}
\end{aligned}
$$

Structure Tensor

- The 2×2 matrix A is a weighted summation of the outer-products

$$\nabla I_{ij} \nabla I_{ij}^T = \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}_{ij}$$

- For Gaussian weights, A can be computed via Gaussian convolution

$$A = \begin{bmatrix} G_\rho * I_x^2 & G_\rho * I_x I_y \\ G_\rho * I_x I_y & G_\rho * I_y^2 \end{bmatrix}$$

- In this form A is known as the *structure tensor*.
- The structure tensor plays an important role in other areas of computer vision as well.

Corner Detection via Structure Tensor

- Basic idea: To find if pixel (x, y) is a corner, first find the direction in which patches become most dissimilar.
- That is, the direction $\mathbf{d} = (u, v)^T$ that maximises the SSD $\mathbf{d}^T A \mathbf{d}$ from the patch centered at (x, y).

$$\mathbf{d}^* = \arg\max_{\mathbf{d}} \mathbf{d}^T A \mathbf{d} \text{ s.t. } \|\mathbf{d}\| = 1$$

  where constraint $\|\mathbf{d}\| = 1$ ensures a non-trivial solution.
- Using the method of Lagrange multipliers, $\mathbf{d}^*$ is the eigenvector of A corresponding to the larger eigenvalue (Take-home Quiz 2).
- The SSD in the direction of any eigenvector is the corresponding eigenvalue. Prove it.

Corner Detection via Structure Tensor

- What do the eigenvalues of the structure tensor reveal about the local structure around a pixel?

$$\lambda_{\text{large}} \approx \lambda_{\text{small}} \approx 0 \implies \text{flat region}$$
$$\lambda_{\text{large}} \gg \lambda_{\text{small}} \approx 0 \implies \text{edge}$$
$$\lambda_{\text{large}} > \lambda_{\text{small}} \gg 0 \implies \text{corner}$$

- So a simple corner detection criterion could be $\lambda_{\text{small}} > \tau$.

Summary

- For 2D images, the 2×2 structure tensor is a powerful descriptor of local image regions.
- Eigenvector corresponding to larger eigenvalue represents the (local) direction of greatest rate of change in the image.
- Largest eigenvalue represents the SSD in that direction.
- Multiple uses in computer vision.