5.3.9 Line detection by local pre-processing operators

Special local operators:

- line finding
- line thinning
- line filling
- interest points
- nonlinear operators, mathematical morphology techniques, coming next.

- edge detector is a general tool not depending on image content
- detected edges are to some degree robust – do not depend much on small changes in illumination, viewpoint change, etc.

- richer features can be reliably detected
- special detectors can outperform simple edge detectors in some classes of applications
  - Line detectors
    used to detect linear objects such as dimension lines in engineering drawings or railways or roads in satellite images
  - corner detectors, interest point detectors
    used mainly to register two or more images one to the other

Line finding operators

- find very thin curves in the image
- assume that curves do not bend sharply
- line may be detected by local convolution
- four such 3×3 kernels detect lines rotated modulo 45°.

$$h_1 = \begin{bmatrix} -1 & -1 & -1 \\ 2 & 2 & 2 \\ -1 & -1 & -1 \end{bmatrix} , \quad h_2 = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} , \quad h_3 = \begin{bmatrix} -1 & 2 & -1 \\ -1 & 2 & -1 \\ -1 & 2 & -1 \end{bmatrix} , \quad \ldots$$

$$(5.70)$$

- 5×5 masks also common
- line detectors may produce more lines than needed
- non-linear constraints may reduce this

5.3.10 Detection of corners (interest points)

- finding pairs of corresponding points in two similar images
- all possible pairs of pixels should be examined to solve this **correspondence problem**, — very computationally expensive
- process is simplified if correspondence of **interest points** is used
- interest point should have some typical local property
- **corners** are very good interest points
- **Corner detectors**
  - input – gray-level image
  - output – values proportional to likelihood that pixel is a corner
- **Interest points** — by thresholding
- corners serve better than lines in correspondence problems
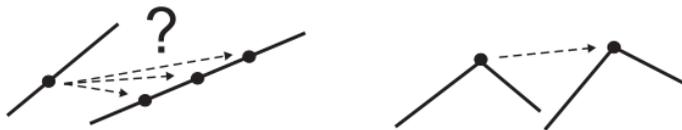- due to the **aperture problem**

Figure 5.34: Ambiguity of lines for matching and unambiguity of corners.

- Edge detectors not stable at corners
- gradient at the tip of the corner is ambiguous
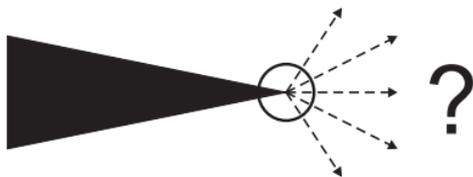- this observation is used in corner detectors

Figure 5.35: Ambiguity of edge detector at the corner tip.

- **corner** can be defined as a pixel, a small neighborhood of which has two dominant and different edge directions
- this definition is not precise – isolated point of local intensity maximum or minimum, line endings, or an abrupt change in the curvature of a curve gives a response similar to a corner
- yet – such detectors are named corner detectors and are widely used
- if "true" corners have to be detected then some additional constraints must be used

- corner detectors are usually not robust

- may be overcome by manual expert supervision or large redundancies to prevent individual errors from dominating the task = many more corners are detected in two or more images than are needed for estimating a transformation sought between these images

- **Moravec detector** [Moravec 77] – simplest

- maximal in pixels with high contrast

- = corners and sharp edges

$$\mathrm{MO}(i,j) = \frac{1}{8} \sum_{k=i-1}^{i+1} \sum_{l=j-1}^{j+1} \left| f(k,l) - f(i,j) \right|. \tag{5.71}$$

- computationally more expensive corner operators give better results – possibly based on facet models

- image f approximated in neighborhood of pixel (i, j) by cubic polynomial with coefficients c_k:

$$f(i, j) = c_1 + c_2\, x + c_3\, y + c_4\, x^2 + c_5\, x\, y + c_6\, y^2 + c_7\, x^3 + c_8\, x^2\, y + c_9\, x\, y^2 + c_{10}\, y^3 \tag{5.72}$$

- Zuniga–Haralick operator ZH

$$\mathrm{ZH}(i, j) = \frac{-2\,(c_2^2\, c_6 - c_2\, c_3\, c_5 + c_3^2\, c_4)}{\left(c_2^2 + c_3^2\right)^{3/2}}\,. \tag{5.73}$$

- Kitchen-Rosenfeld KR operator is similar

$$\mathrm{KR}(i, j) = \frac{-2\,(c_2^2\, c_6 - c_2\, c_3\, c_5 + c_3^2\, c_4)}{c_2^2 + c_3^2}$$

- ZH operator did outperform the KR corner detector in test images [Haralick and Shapi

- **Harris corner detector** [Harris and Stephen 88] improved upon Moravec's
- added differential of the corner score (sum of square differences)
- image patch $W \in f$ shifted by $\Delta x, \Delta y$
- sum of square differences S between values of the image f given by the patch W and its shifted variant by $\Delta x, \Delta y$ given by:

$$S_W(\Delta x, \Delta y) = \sum_{x_i \in W} \sum_{y_i \in W} \left(f(x_i, y_i) - f(x_i - \Delta x, y_i - \Delta y) \right)^2. \tag{5.74}$$

- corner point not suffering from aperture problem

- $\Rightarrow$ high response of $S_W(\Delta x, \Delta y)$ for all Δx, Δy

- if the shifted image patch is approximated by the first-order Taylor expansion

$$f(x_i - \Delta x, y_i - \Delta y) \approx f(x_i, y_i) + \left[\frac{\partial f(x_i, y_i)}{\partial x}, \frac{\partial f(x_i, y_i)}{\partial y} \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} , \qquad (5.75)$$

- then minimum of $S_W(\Delta x, \Delta y)$ can be obtained analytically

- substitute this approximation given by equation (5.75) into equation (5.74)

$$
\begin{aligned}
S(x,y) &= \sum_{x_i \in W} \sum_{y_i \in W} \left(f(x_i, y_i) - f(x_i, y_i) - \left[\frac{\partial f(x_i, y_i)}{\partial x}, \frac{\partial f(x_i, y_i)}{\partial y} \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \\
&= \sum_{x_i \in W} \sum_{y_i \in W} \left(- \left[\frac{\partial f(x_i, y_i)}{\partial x}, \frac{\partial f(x_i, y_i)}{\partial y} \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \\
&= \sum_{x_i \in W} \sum_{y_i \in W} \left(\left[\frac{\partial f(x_i, y_i)}{\partial x}, \frac{\partial f(x_i, y_i)}{\partial y} \right] \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2 \\
&\quad \text{having in mind that } \mathbf{u}^2 = \mathbf{u}^\top \mathbf{u} \\
&= \sum_{x_i \in W} \sum_{y_i \in W} [\Delta x, \Delta y] \left(\begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \right) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \\
&= [\Delta x, \Delta y] \left(\sum_{x_i \in W} \sum_{y_i \in W} \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix} \right) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \\
&= [\Delta x, \Delta y] \, A_W(x, y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} ,
\end{aligned}
$$

where the Harris matrix $A_W(x, y)$ is the second derivative of S around the point $(x, y) = (0, 0)$

- A is

$$A(x,y) = \begin{bmatrix} \sum\limits_{x_i \in W} \sum\limits_{y_i \in W} \frac{\partial^2 f(x_i,y_i)}{\partial x^2} & \sum\limits_{x_i \in W} \sum\limits_{y_i \in W} \frac{\partial f(x_i,y_i)}{\partial x} \frac{\partial f(x_i,y_i)}{\partial y} \\ \sum\limits_{x_i \in W} \sum\limits_{y_i \in W} \frac{\partial f(x_i,y_i)}{\partial x} \frac{\partial f(x_i,y_i)}{\partial y} & \sum\limits_{x_i \in W} \sum\limits_{y_i \in W} \frac{\partial^2 f(x_i,y_i)}{\partial y^2} \end{bmatrix} . \quad (5.76)$$

- usually isotropic window is used, such as a Gaussian
- $\Rightarrow$ response will also be isotropic

- local structure matrix A represents the neighborhood
- Harris matrix A is symmetric and positive semi-definite
- its main modes of variation correspond to partial derivatives in orthogonal directions and are reflected in eigenvalues λ_1, λ_2 of matrix A
- modes of variations can be found using PCA

- three distinct cases can appear:

    1. Both eigenvalues are small. This means that image f is flat in the examined pixel. There are no edges or corners in this location.

    2. One eigenvalue is small and the second one large. The local neighborhood is ridge-shaped. Significant change of image f occurs if a small movement is made perpendicularly to the ridge.

    3. Both eigenvalues are rather large. A small shift in any direction causes significant change of image f. A corner is found.
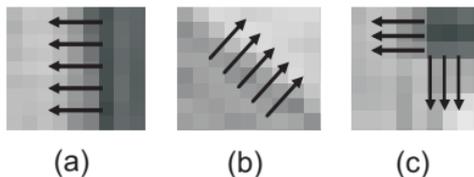

(a)   (b)   (c)

Figure 5.36: Illustration of the decision within Harris corner detector according to eigenvalues of the local structure matrix. (a), (b) Ridge detected, no corner at this position. (c) Corner detected.

- exact eigenvalue computation can be avoided by calculating the response function
  $R(A) = \det(A) - \kappa \operatorname{trace}^2(A)$

- κ from 0.04 to 0.15 reported as appropriate

Figure 5.37: Example of Harris corners in the image. *Courtesy of Martin Urban, Czech Technical University in Prague, who used such images for 3D reconstruction. A color version of this figure may be seen in the color inset—Plate 1.*

Algorithm 5.5: Harris corner detector

1. Filter the image with a Gaussian.

2. Estimate intensity gradient in two perpendicular directions for each pixel, $\frac{\partial f(x,y)}{\partial x}$, $\frac{\partial f(x,y)}{\partial y}$. This is performed by twice using a 1D convolution with the kernel approximating the derivative.

3. For each pixel and a given neighborhood window:
   - Calculate the local structure matrix A.
   - Evaluate the response function $R(A)$.

4. Choose the best candidates for corners by selecting a threshold on the response function $R(A)$ and perform non-maximal suppression.

- Harris corner detector very popular
- advantages = insensitivity to
    - 2D shift and rotation
    - small illumination variations
    - small viewpoint change
    - + low computational requirements
- disadvantages = not invariant to
    - larger scale change
    - viewpoint changes
    - significant changes in contrast

Many more corner-like detectors exist[Mikolajczyk and Schmid 04; Mikolajczyk et al. 05].

5.3.11   Detection of maximally stable extremal regions

- Harris detector ... invariant to rotation and translation
- but — if image is rescaled or transformed projectively, output of Harris detector changes rapidly

Figure 5.38: Example of MSERs detected in the image. Regions with the red border are results of the algorithm on the increasingly ordered list of intensities. Regions with the green borders come from the list with decreasing ordering. *Courtesy of Jiri Matas, Czech Technical University in Prague, who used such images for 3D reconstruction. A color version of this figure may be seen in the color inset—Plate 2.*

Maximally Stable Extremal Regions (MSERs) [Matas et al. 02]

- can be repeatedly detected after translations and rotations
- AND after similarity and affine transforms
- MSER detection process:
  - Imagine all possible thresholdings of an input gray-level image I
  - pixels below threshold ... 'black'
  - above or equal ... 'white'
  - imagine a movie of thresholded images I_t, with frame t corresponding to threshold t
  - $\Rightarrow$ we first see a white image
  - black spots corresponding to local intensity minima will appear and grow
  - regions corresponding to two local minima will merge
  - the last image will be black
- union of all connected components of all frames of the movie is identical to the set of all maximal regions
- minimal regions could be obtained by inverting the intensity of I and running the same process
- local binarization is stable over a large range of thresholds in certain regions

- such regions are of interest since they have the following properties:
    - Invariance to monotonic transformation. $M \colon \mathcal{S} \to \mathcal{S}$ of image intensities.
    - The set of extremal regions is unchanged after transformation M,
      $I(p) < I(q) \to M\big(I(p)\big) = I'(p) < I'(q) = M\big(I(q)\big)$
      since M does not affect adjacency (and thus contiguity). The intensity ordering is preserved.
    - Invariance to adjacency preserving (continuous) transformation. $T \colon \mathcal{D} \to \mathcal{D}$ on the image domain.
    - Stability, since only extremal regions whose support is virtually unchanged over a range of thresholds are selected.
    - Multi-scale detection. Since no smoothing is involved, both very fine and very large structure is detected.
    - The set of all extremal regions can be enumerated in $\mathcal{O}(n \log \log n)$, i.e., almost in linear time for 8 bit images.

Algorithm 5.6: Enumeration of Extremal Regions.

Input: Image I.

Output: List of nested extremal regions.

1. For all pixels sorted by intensity:
    - Place pixel in the image.
    - Update the connected component structure.
    - Update the area for the effected connected component.

2. For all connected components:
    - Local minima of the rate of change of the connected component area define stable thresholds.

As pixels ordered by intensity are placed in the image (either in decreasing or increasing order), the list of connected components and their areas is maintained using the efficient union-find algorithm [Sedgewick 98].

- process produces data structure holding the area of each connected component as a function of a threshold

- a merge of two components is viewed as the end of existence of the smaller component and the insertion of all pixels of the smaller component into the larger one

- intensity levels that are local minima of the rate of change of the area function are selected as thresholds

- output ... each MSER represented by a local intensity minimum (or maximum) and a threshold

- Algorithm 5.6 and efficient watershed algorithm [Vincent and Soille 91] are essentially identical

- but ... structure of outputs is different

- in watershed computation, focus is on thresholds where regions merge and watershed basins touch
(such thresholds are highly unstable – after a merge, the region area changes abruptly)

- in MSER detection, range of thresholds is sought that leaves the watershed basin effectively unchanged

- MSER has shown the highest repeatability of affine-invariant detectors [Mikolajczyk et Frauendorfer and Bischof 05]