Interest Operators

• Find “interesting” pieces of the image

• Multiple possible uses
  – image matching
    • stereo pairs
    • tracking in videos
    • creating panoramas
  – object recognition
Goal: Local invariant photometric descriptors

- Local: robust to occlusion/clutter + no segmentation
- Photometric: distinctive
- Invariant: to image transformations + illumination changes
History - Matching

Matching based on correlation alone
Matching based on line segments

⇒ Not very discriminating (why?)

⇒ Solution: matching with interest points & correlation

[ A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry,
Z. Zhang, R. Deriche, O. Faugeras and Q. Luong,
Artificial Intelligence 1995 ]
Approach

• Extraction of interest points with the Harris detector

• Comparison of points with cross-correlation

• Verification with the fundamental matrix (later in the course)
Harris detector

Interest points extracted with Harris (~ 500 points)
Cross-correlation matching

Initial matches (188 pairs)
Global constraints

Robust estimation of the fundamental matrix

99 inliers

89 outliers
Summary of the approach

• Very good results in the presence of occlusion and clutter
  – local information
  – discriminant greyvalue information
  – robust estimation of the global relation between images
  – for limited view point changes

• Solution for more general view point changes
  – wide baseline matching (different viewpoint, scale and rotation)
  – local invariant descriptors based on greyvalue information
History - Recognition

Problems: occlusion, clutter, image transformations, distinctiveness

⇒ Solution: recognition with local photometric invariants

[ Local greyvalue invariants for image retrieval, 
  C. Schmid and R. Mohr, 
  PAMI 1997 ]
1) Extraction of interest points (characteristic locations)
2) Computation of local descriptors
3) Determining correspondences
4) Selection of similar images
Interest points

Geometric features
- repeatable under transformations

2D characteristics of the signal
- high informational content

Comparison of different detectors [Schmid98]  ➔  Harris detector
Harris detector

Based on the idea of auto-correlation

Important difference in all directions => interest point
Background: Moravec Corner Detector

- take a window $w$ in the image
- shift it in four directions $(1,0), (0,1), (1,1), (-1,1)$
- compute a difference for each
- compute the min difference at each pixel
- local maxima in the min image are the corners

$$E(x,y) = \sum_{u,v \text{ in } w} w(u,v) |I(x+u,y+v) - I(u,v)|^2$$
Shortcomings of Moravec Operator

• Only tries 4 shifts. We’d like to consider “all” shifts.

• Uses a discrete rectangular window. We’d like to use a smooth circular (or later elliptical) window.

• Uses a simple min function. We’d like to characterize variation with respect to direction.

Result: Harris Operator
Harris detector

Auto-correlation function for a point \((x, y)\) and a shift \((\Delta x, \Delta y)\)

\[
f(x, y) = \sum_{(x_k, y_k) \in W} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2
\]

Discrete shifts can be avoided with the auto-correlation matrix

with \( I(x_k + \Delta x, y_k + \Delta y) = I(x_k, y_k) + (I_x(x_k, y_k) \quad I_y(x_k, y_k)) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \)

\[
f(x, y) = \sum_{(x_k, y_k) \in W} \left( \begin{pmatrix} I_x(x_k, y_k) & I_y(x_k, y_k) \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \right)^2
\]
Harris Math Manipulation

\[ f(x, y) = \sum_{(x_k, y_k) \in W} (I(x_k, y_k) - I(x_k + \Delta x, y_k + \Delta y))^2 \]

\[ = \sum_{(x_k, y_k) \in W} (I(x_k, y_k) - (I(x_k, y_k) + (I_x(x_k, y_k) I_y(x_k, y_k))(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}))^2 \]

\[ = \sum_{(x_k, y_k) \in W} (I(x_k, y_k) - (I_x(x_k, y_k) I_y(x_k, y_k)))(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix})^2 \]

\[ = \sum_{(x_k, y_k) \in W} (I_x(x_k, y_k) I_y(x_k, y_k))(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix})^2 \]

\[ = \sum_{(x_k, y_k) \in W} (I_x(x_k, y_k) I_y(x_k, y_k))(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix})(I_x(x_k, y_k) I_y(x_k, y_k))(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}) \]

\[ = \sum_w (\Delta x \Delta y)(I_x I_y)(I_x I_y)(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix})(I_x I_y)(I_x I_y)(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}) \]

\[ = \sum_w (\Delta x \Delta y)(I_x I_x I_y I_y)(I_x I_y I_y I_x)(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}) \]

\[ = (\Delta x \Delta y)(\begin{pmatrix} \sum_w I_x^2 & \sum_w I_x I_y \\ \sum_w I_x I_y & \sum_w I_y^2 \end{pmatrix})(\begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}) \]
Harris detector

\[
\begin{bmatrix}
\sum (I_x(x_k, y_k))^2 & \sum I_x(x_k, y_k)I_y(x_k, y_k) \\
\sum I_x(x_k, y_k)I_y(x_k, y_k) & \sum (I_y(x_k, y_k))^2
\end{bmatrix}
\begin{bmatrix}
\Delta x \\
\Delta y
\end{bmatrix}
\]

Auto-correlation matrix M
Harris detection

- **Auto-correlation matrix**
  - captures the structure of the local neighborhood
  - measure based on eigenvalues of $M$ which form a rotationally invariant descriptor.
    - 2 strong eigenvalues $\Rightarrow$ interest point
    - 1 strong eigenvalue $\Rightarrow$ contour
    - 0 eigenvalue $\Rightarrow$ uniform region

- **Interest point detection**
  - threshold on the eigenvalues
  - local maximum for localization
Some Details from the Harris Paper

- Let $\alpha$ and $\beta$ be the two eigenvalues
- $\text{Tr}(M) = \alpha + \beta$
- $\text{Det}(M) = \alpha \beta$
- Response $R = \text{Det}(M) - k \text{Tr}(M)$
- $R$ is positive for corners, - for edges, and small for flat regions
- Select corner pixels that are 8-way local maxima

Trace and determinant are easy to compute.
Determining correspondences

Vector comparison using a distance measure

What are some suitable distance measures?
Some Matching Results
Summary of the approach

• Very good results in the presence of occlusion and clutter
  – local information
  – discriminant greyvalue information
  – invariance to image rotation and illumination

• Not invariance to scale and affine changes

• Solution for more general view point changes
  – local invariant descriptors to scale and rotation
  – extraction of invariant points and regions