Mid Level Image Features : Shapes

Eun Yi Kim

Artificial Inteligence & Computer Vision Laboratory

Shape descriptors

Artificial Inteligence & Computer Vision Laboratory

- There are three approaches to defining shapes

1. **Shape represented by its region descriptors – Simple !!**
2. **Shape represented by its Boundary**
3. **Shape represented by its interests points (corners)**

Region based Shape Descriptors

Geometric and Shape Properties

- area
- centroid
- perimeter
- perimeter length
- circularity, elongation
- mean and standard deviation of radial distance
- second order moments (row, column, mixed)
- bounding box
- extremal axis length from bounding box
- lengths and orientations of axes of best-fit ellipse

Often want features independent of position, orientation, scale

Topological Region Descriptors

- Topological properties: properties of image preserved under rubber-sheet distortions
  - # holes in the image
  - # connected components

H=2, C=1

H=0, C=3

H=1, C=1

H=2, C=1

Topological Region Descriptors : Hole Counting

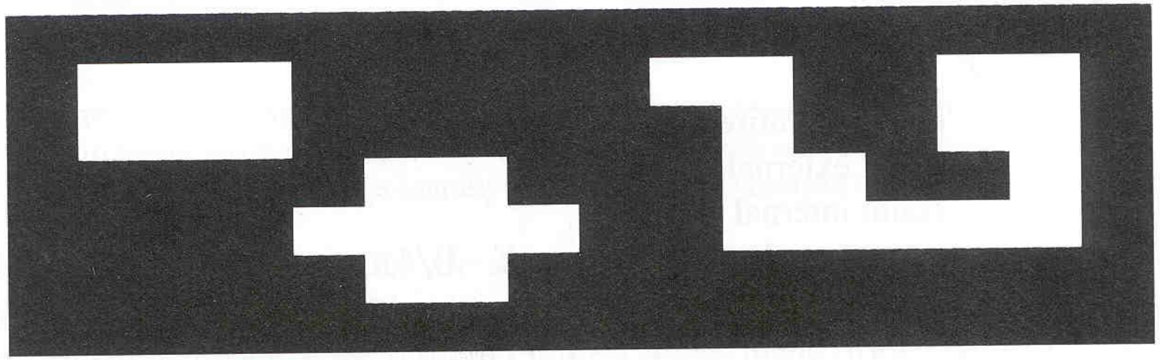
- “external corner” has 3(1)s and 1(0)
- “internal corner” has 3(0)s and 1(1)
- Holes computed from only these patterns!

1	1	1	0	0	1	1	1
1	0	1	1	1	1	0	1

(*a*) 2 × 2 *external corner* patterns

0	0	0	1	1	0	0	0
0	1	0	0	0	0	1	0

(*b*) 2 × 2 *internal corner* patterns



(*c*) Three bright holes in dark background

Topological Region Descriptors : Hole Counting Algorithm

Input a binary image and output the number of holes it contains.

M is a binary image of **R** rows of **C** columns.
1 represents material through which light has not passed;
0 represents absence of material indicated by light passing.
Each region of 0s must be 4-connected and all image border pixels must be 1s.
E is the count of *external corners* (3 ones and 1 zero)
I is the count of *internal corners* (3 zeros and 1 one)

```
integer procedure Count_Holes(M)
{
    examine entire image, 2 rows at a time;
    count external corners E;
    count internal corners I;
    return(number_of_holes = (E - I)/4);
}
```

Topological Region Descriptors : Hole Counting Example

Artificial Inteligence & Computer Vision Laboratory

(E-I)/4

	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	e	i
0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		
1	1	0	0	0	1	1	1	1	1	0	0	1	1	0	0	1		
2	1	0	0	0	1	1	1	1	1	1	0	1	1	0	0	1		
3	1	1	1	1	1	0	0	1	1	1	0	0	1	1	0	1		
4	1	1	1	1	0	0	0	0	1	1	0	0	0	0	0	1		
5	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1		
6	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1		

(*d*) Binary input image 7 rows high and 16 columns wide

	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	e	i
0	e			e					e		e		e		e		6	0
1									e	i							1	1
2	e			e	e		e				i	e	e	i			6	2
3				e	i		i	e				i		i			2	4
4				e	i		i	e		e					e		4	2
5					e		e										2	0
6																	0	0

(e) External corner patterns marked with e; internal corners marked i

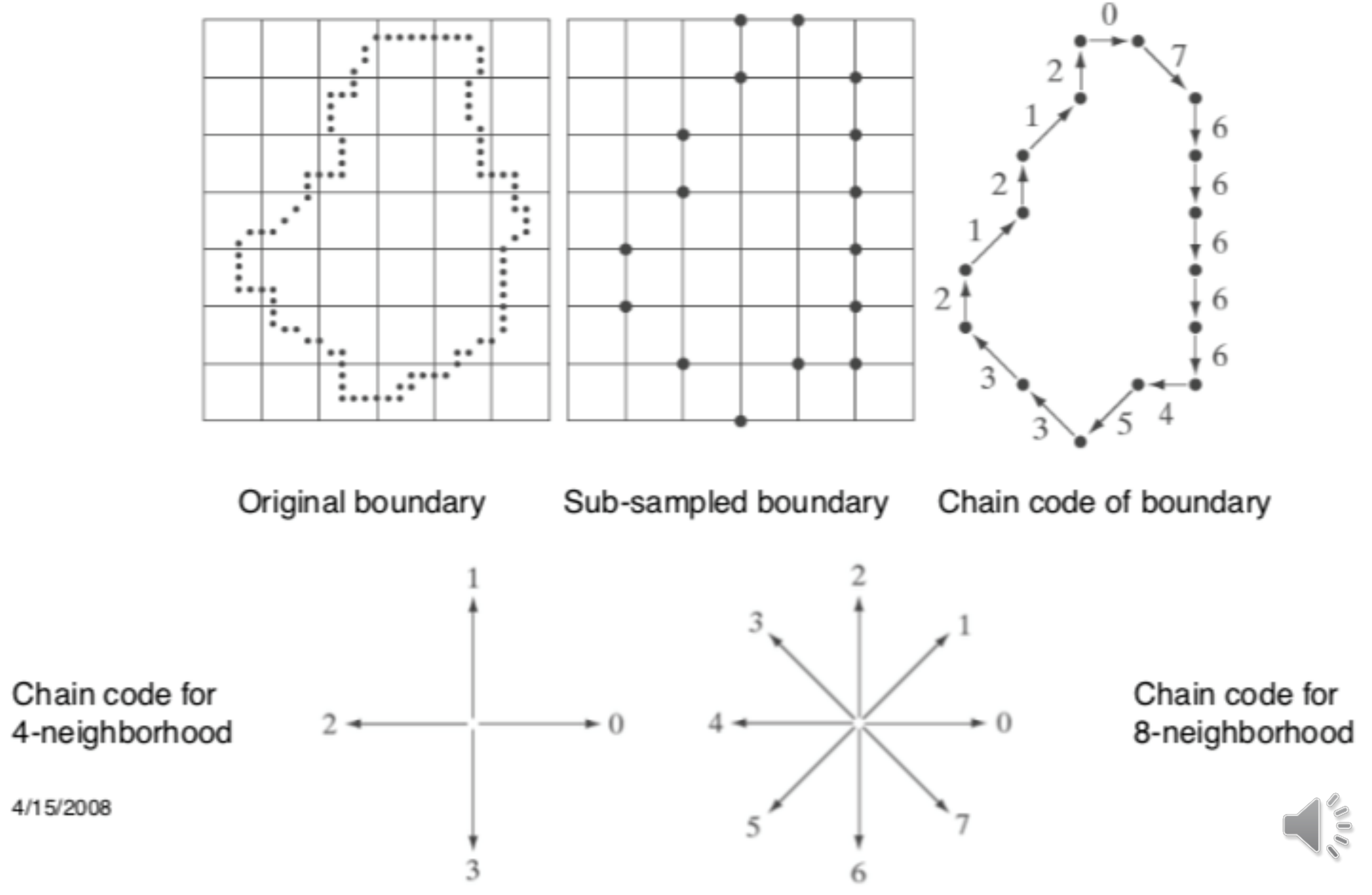
Boundary based Shape Descriptors

Boundary Representation

Artificial Inteligence & Computer Vision Laboratory

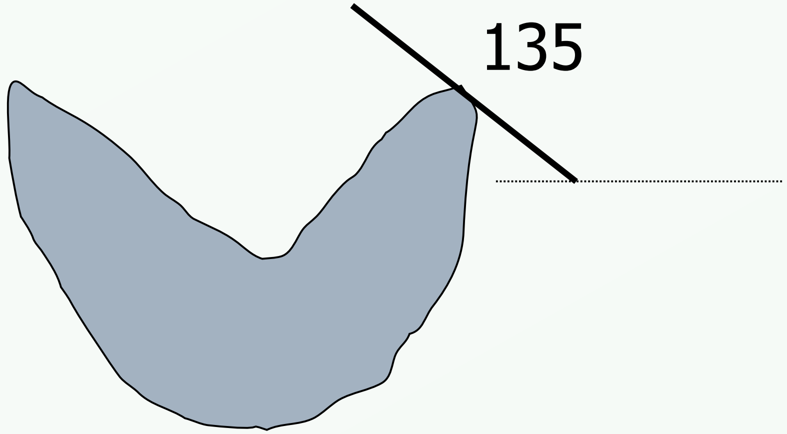
- (Freeman) Chain Code

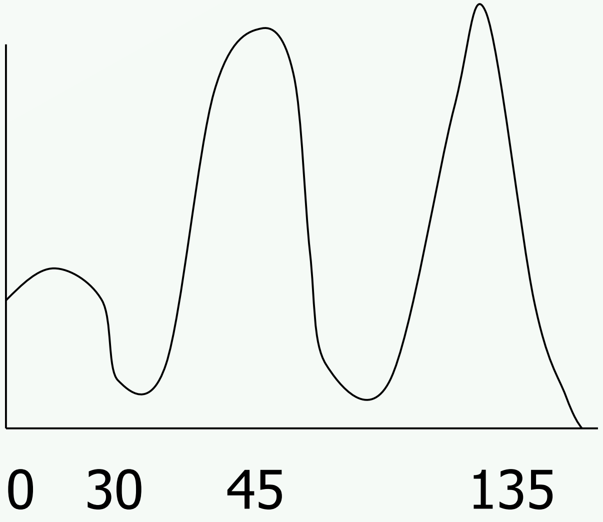
- Boundary representation = 076666453321212



Boundary Representation

- Tangent-Angle histograms





Is this feature invariant to starting point?
Is it invariant to size, translation, rotation?

Interest Operator + Descriptor

- Harris operator
- Multi-scaled operator
- SIFT (scale invariant feature transform)
- HOG (histogram of oriented gradient)

Introduction to Interest Operators

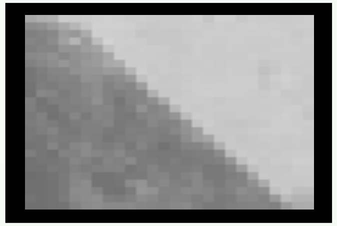
- Find “interesting” pieces of the image
  - E.g. corners, salient regions
  - Focus attention of algorithms
  - Speed up computation

- Many possible uses in matching/recognition
  - Search
  - Object recognition
  - Image alignment & stitching
  - Stereo
  - Tracking
  - …

Interest points

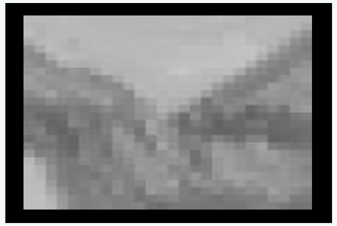
0D structure: single points

➡ not useful for matching



1D structure: lines

➡ edge, can be localised in 1D, subject to the aper ture problem



2D structure: corners

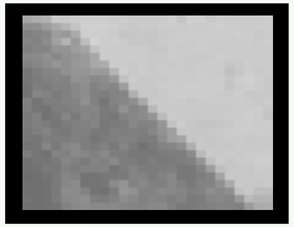
➡ corner, or **interest point**, can be localised in 2D, good for matching

Interest Points have **2D** structure.

Interest points

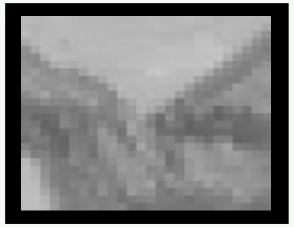
0D structure: single points

➡ not useful for matching



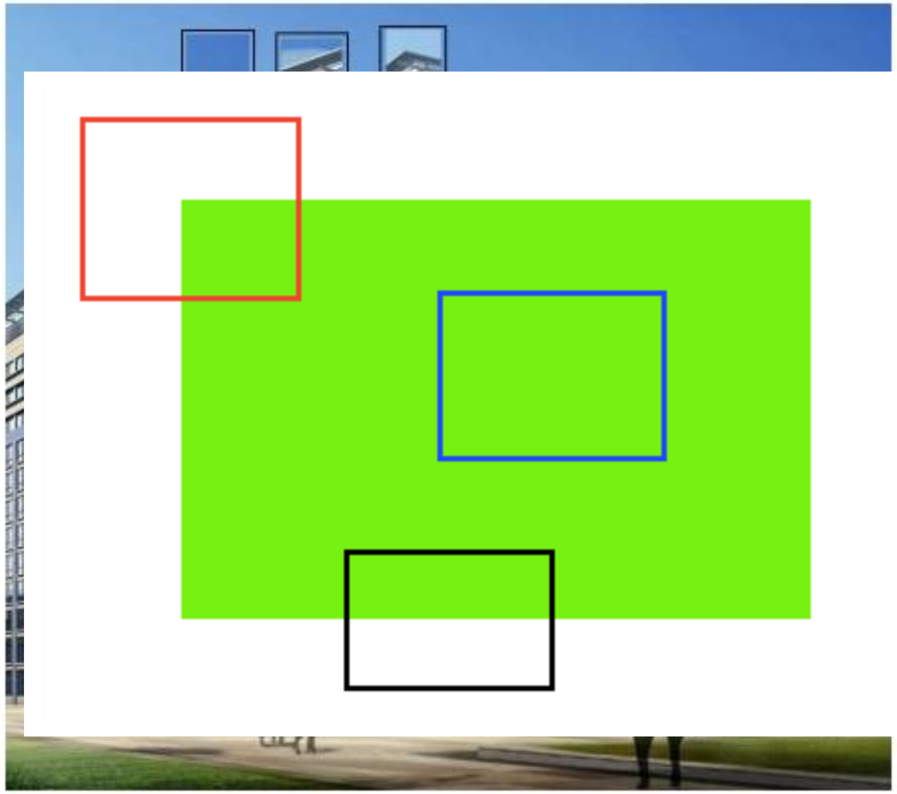
1D structure: lines

➡ edge, can be localised in 1D, subject to th
e aperture problem



2D structure: corners

➡ corner, or **interest point**, can be localized in 2D, good for matching



Interest Points have **2D** structure.

Local invariant photometric descriptors

Artificial Inteligence & Computer Vision Laboratory

local descriptor

- *Local* : robust to occlusion/clutter + no segmentation
- *Photometric* : (use pixel values) distinctive descriptions
- *Invariant* : to image transformations + illumination changes

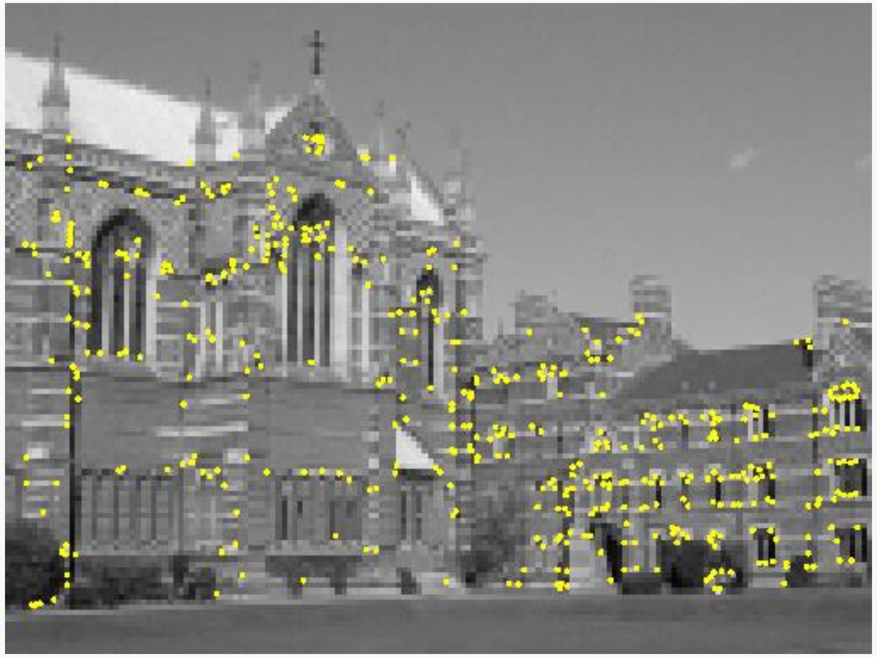
Zhang Approach

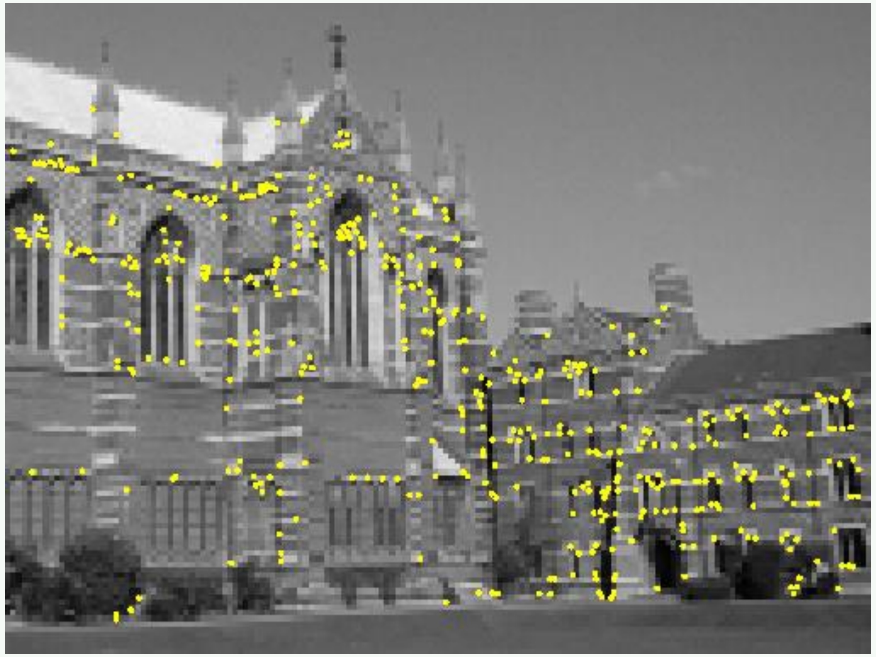
- Extraction of interest points with the Harris detector
- Comparison of points with cross-correlation
- Verification with the fundamental matrix

The fundamental matrix maps points from the first image to corresponding p oints in the second matrix using a homography, that is determined through th e solution of a set of equations that usually minimizes a least square error.

Preview: Harris detector

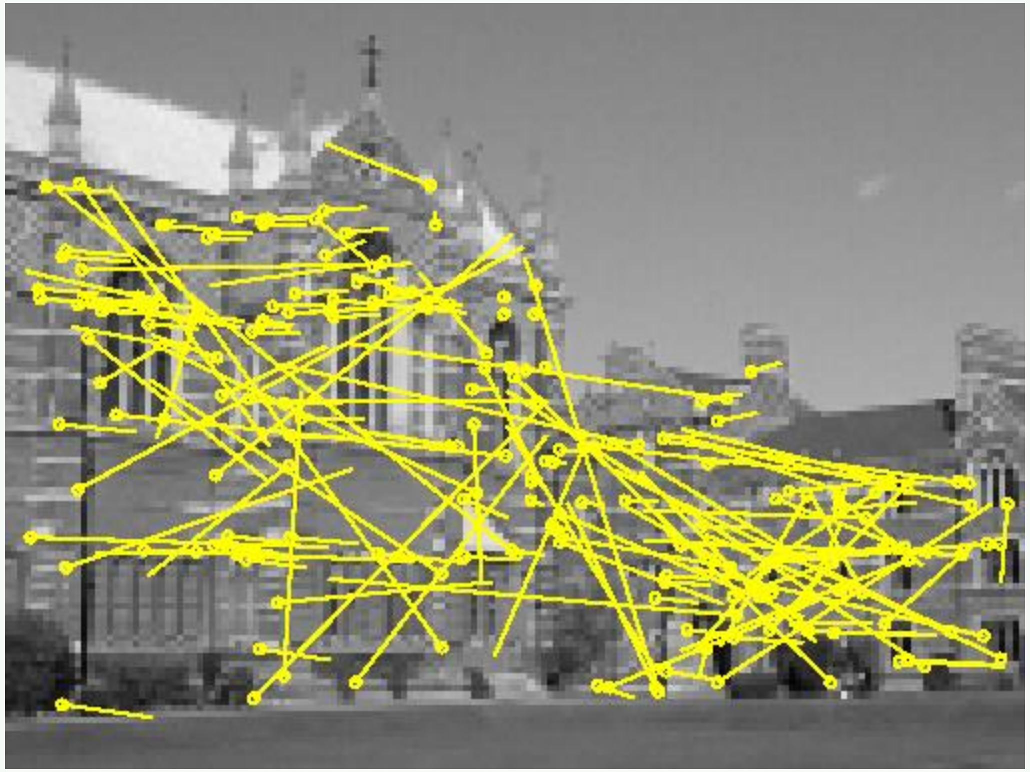
Artificial Inteligence
& Computer Vision
Laboratory





Interest points extracted with Harris (~ 500 points)

Cross-correlation matching



Initial matches – motion vectors (188 pairs)

General Interest Detector/Descriptor Approach

1) Extraction of interest points

2) Computation of local descriptors

3) Determining correspondences

4) Selection of similar images

1. Harris detector

Based on the idea of auto-correlation

$$\rho(dr,dc) = \frac{\sum_i \sum_j I[i,j]I[i+dr,j+dc]}{\sum_i \sum_j I^2[i,j]} = \frac{I[i,j]\circ I_d[i,j]}{I[i,j]\circ I[i,j]}$$

Important difference in all directions => interest point

Background
: Moravec Corner Detector

Artificial Inteligence & Computer Vision Laboratory

- 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 fn (SSD) for a point (x, y) and a shift $(\Delta x, \Delta y)$

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

SSD means summed square difference

Discrete shifts can be avoided with the auto-correlation matrix

with $$I(x_k + \Delta x, y_k + \Delta y) = I(x_k, y_k) + \overbrace{(I_x(x_k, y_k) \quad I_y(x_k, y_k))}^{\text{what is this?}} \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix}$$

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

Harris detector

Rewrite as inner (dot) product

$$f(x,y) = \sum_{(x_k,y_k)\epsilon W} \left(\begin{bmatrix} I_x(x_k,y_k) & I_y(x_k,y_k) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \right)^2$$

$$= \sum_{(x_k,y_k)\epsilon W} \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} \begin{bmatrix} I_x(x_k,y_k) \\ I_y(x_k,y_k) \end{bmatrix} \begin{bmatrix} I_x(x_k,y_k) & I_y(x_k,y_k) \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

The center portion is a 2x2 matrix

Have we seen this matrix before?

$$= \sum_W \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

$$= \begin{bmatrix} \Delta x & \Delta y \end{bmatrix} \sum_W \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$$

Harris detector

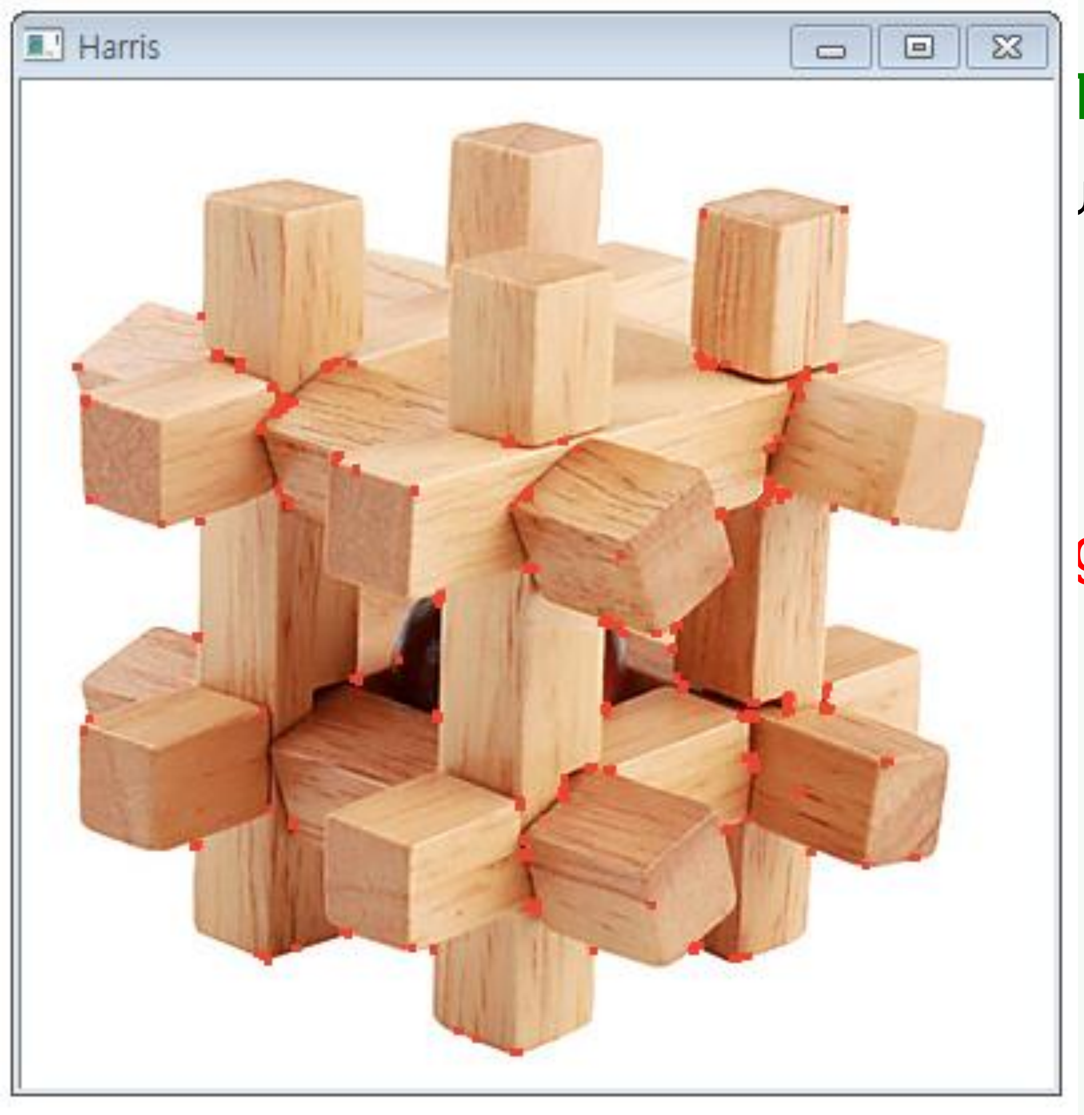
$$= \begin{pmatrix} \Delta x & \Delta y \end{pmatrix} \begin{bmatrix} \sum_{(x_k, y_k) \in W} (I_x(x_k, y_k))^2 & \sum_{(x_k, y_k) \in W} I_x(x_k, y_k) I_y(x_k, y_k) \\ \sum_{(x_k, y_k) \in W} I_x(x_k, y_k) I_y(x_k, y_k) & \sum_{(x_k, y_k) \in W} (I_y(x_k, y_k))^2 \end{bmatrix} \begin{pmatrix} \Delta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

Auto-correlation matrix M

Harris detection

- Auto-correlation matrix
  - captures the structure of the local neighborhood
  - measure based on **eigenvalues** of M
    - 2 strong eigenvalues => interest point
    - 1 strong eigenvalue => contour
    - 0 eigenvalue => uniform region
- Interest point detection
  - threshold on the eigenvalues
  - local maximum for localization

Some Details from the Harris Paper

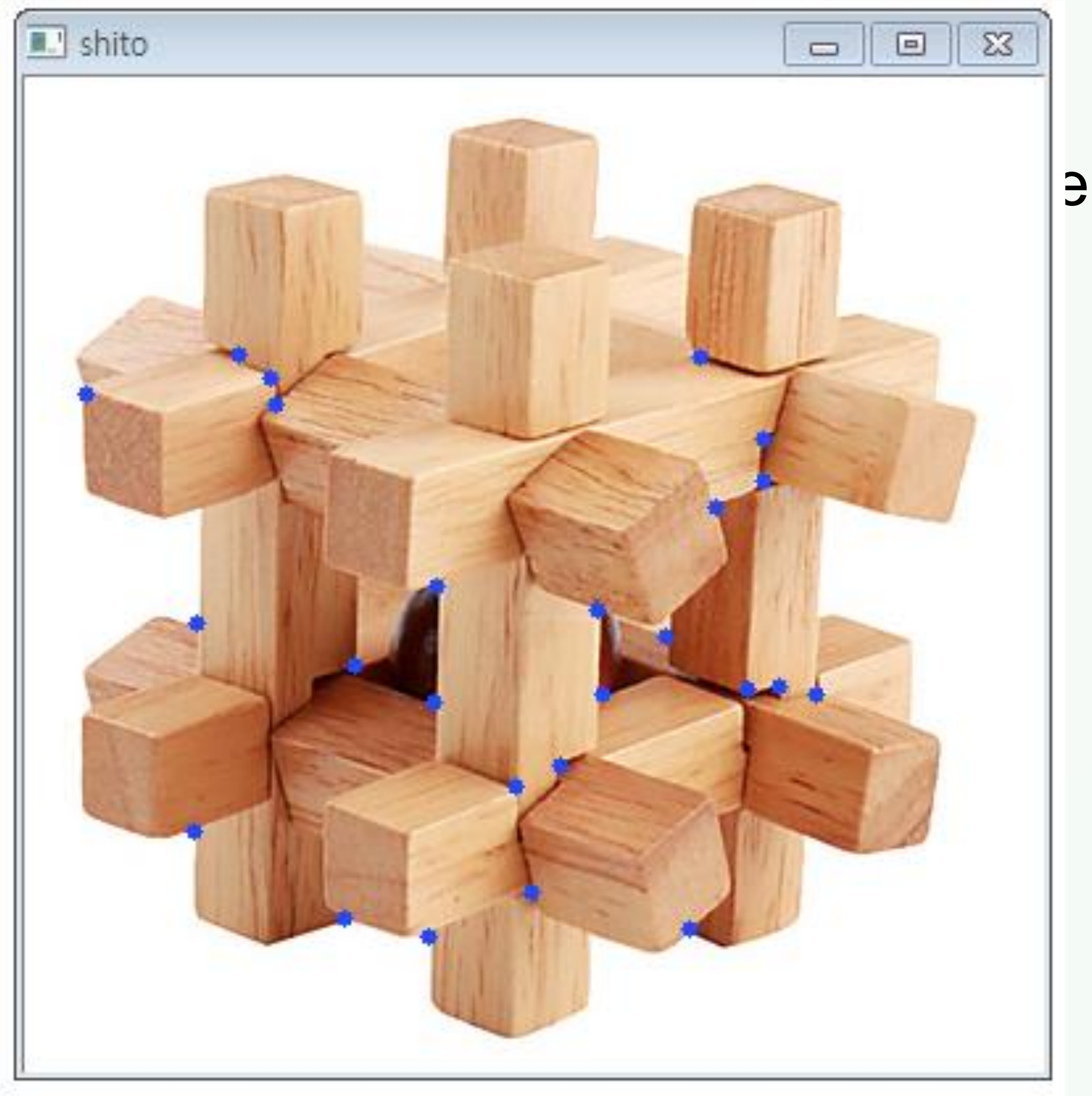


Tr(M

ues.

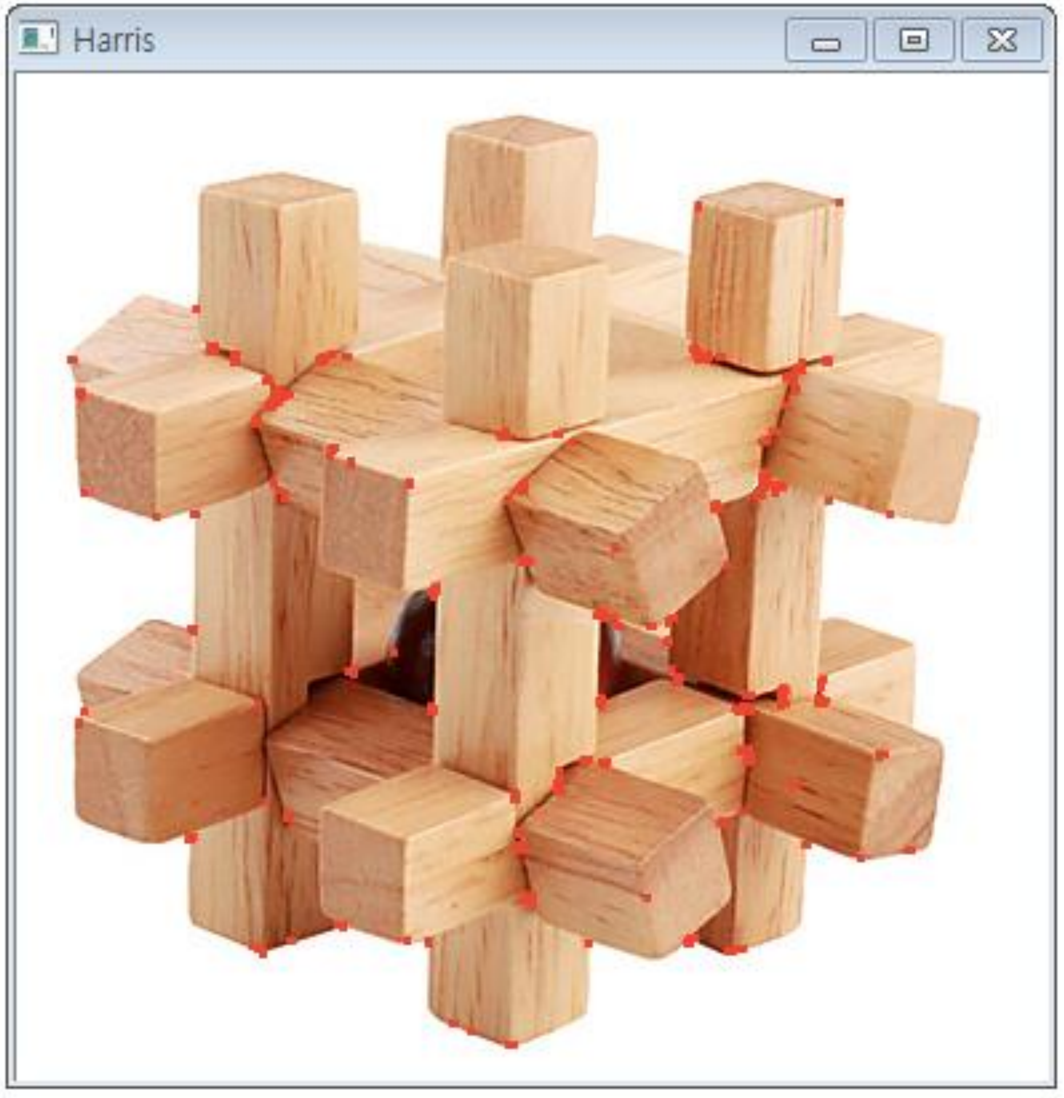
ges

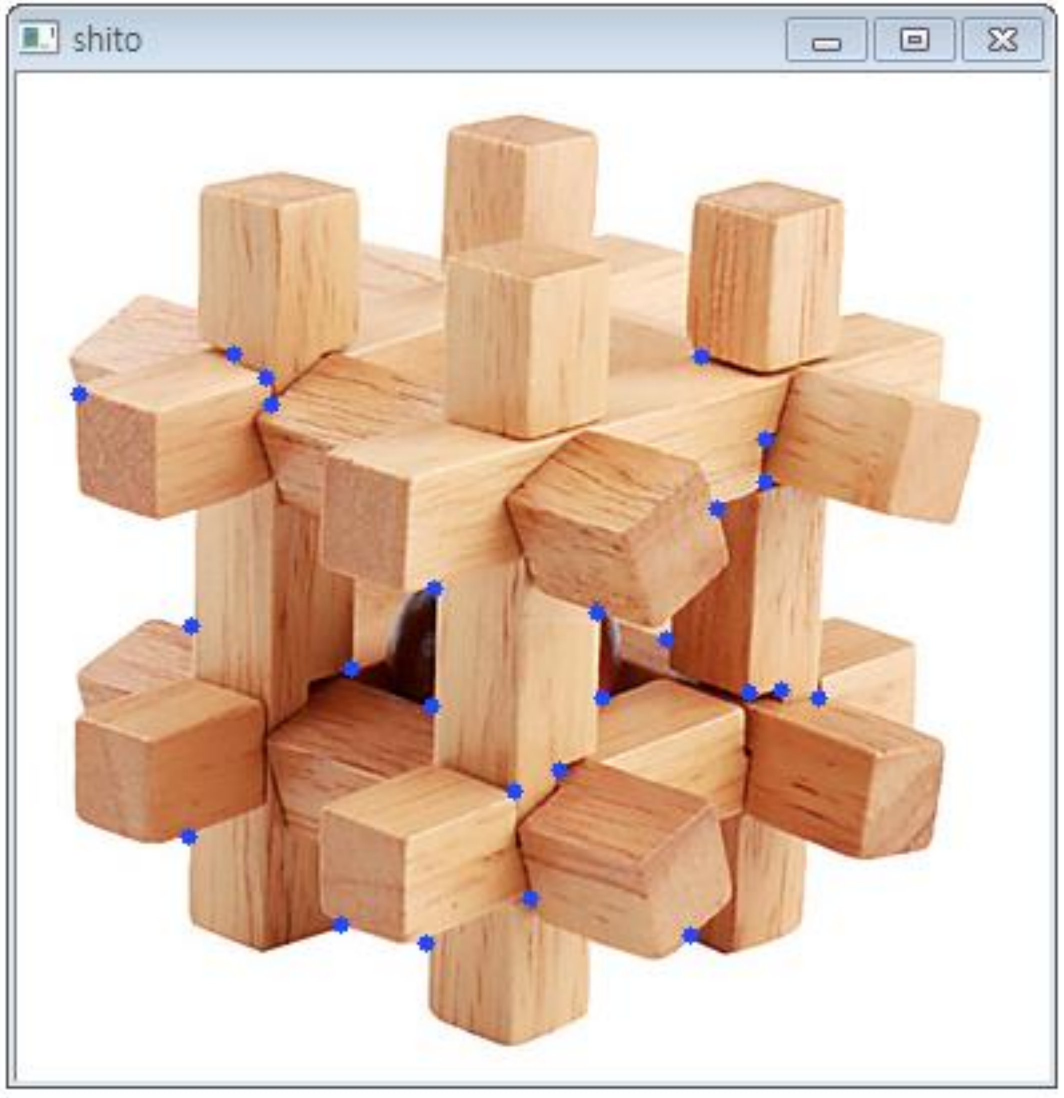
= m



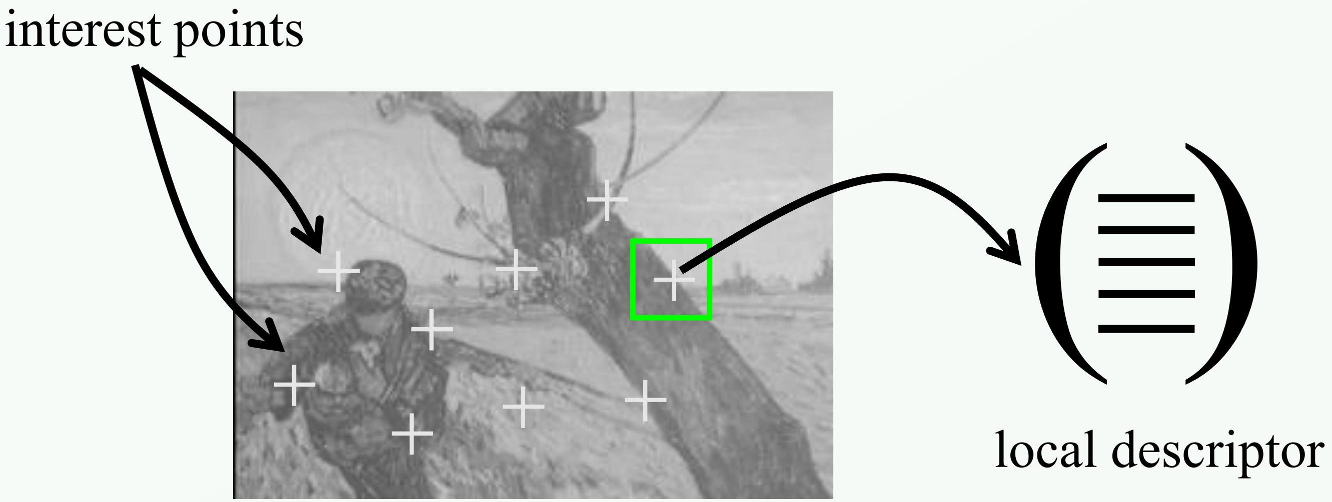
e

Some Details from the Harris Paper





General Interest Detector/Descriptor Approach



1) Extraction of interest points

2) Computation of local descriptors

3) Determining correspondences

4) Selection of similar images