Mid Level Image Features : Shapes

Eun Yi Kim

Artificial Inteligence & Computer Vision Laboratory

I N D E X

Shapes

Approaches to Shape description

- Region based shape descriptors
- Boundary based descriptors
- Interest Operator + Descriptor

Applications

Shapes

- Shape goes one step further than color and texture.
- Color and Texture are both global attributes of an image; shape is not an image attributes
  - Shape tends to refer to a specific region of an image
  - Segmentation is still a crucial problem to be solved, so interests operator is employed for shape description
- Two-dimensional shape recognition is an important aspect of image anal ysis (image matching/retrieval)

Shape descriptors

- 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)**

Artificial Inteligence
& Computer Vision
Laboratory

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

Zero-order moment

- Why use moments?
  - Geometric moments of different orders represent spatial characteristics of the imag
e distribution

- Zero-order moment

$$A = \sum_{i=1}^{n} \sum_{j=1}^{m} B[i, j]$$

  - Total intensity of image
  - For binary image → area

Centroid

- An object's position in the image determines its spatial location.
- center of area (a centroid, center of mass) : first order moment
  - Intensity centroid
  - Geometrical center in binary image

$$\begin{cases} \bar{x} = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} jB[i,j]}{A} & : \text{ average (mean) of } j \text{ coordinates of object (1) pixels} \\ \bar{y} = \dfrac{\sum_{i=1}^{n}\sum_{j=1}^{n} iB[i,j]}{A} & : \text{ average of } i \text{ coordinates of object (1) pixels} \end{cases}$$

- A precision of tenths of a pixel is often justifiable for the centroid.
- Centroids of regions can be interesting points for analysis and matching

Second Moments

There are three second-order spatial moments of a region

- Second-order row moment

$$\mu_{rr} = \frac{1}{A} \sum_{(r,c)\in R} (r - \bar{r})^2$$

- Second-order mixed moment

$$\mu_{rc} = \frac{1}{A} \sum_{(r,c)\in R} (r - \bar{r})(c - \bar{c})$$

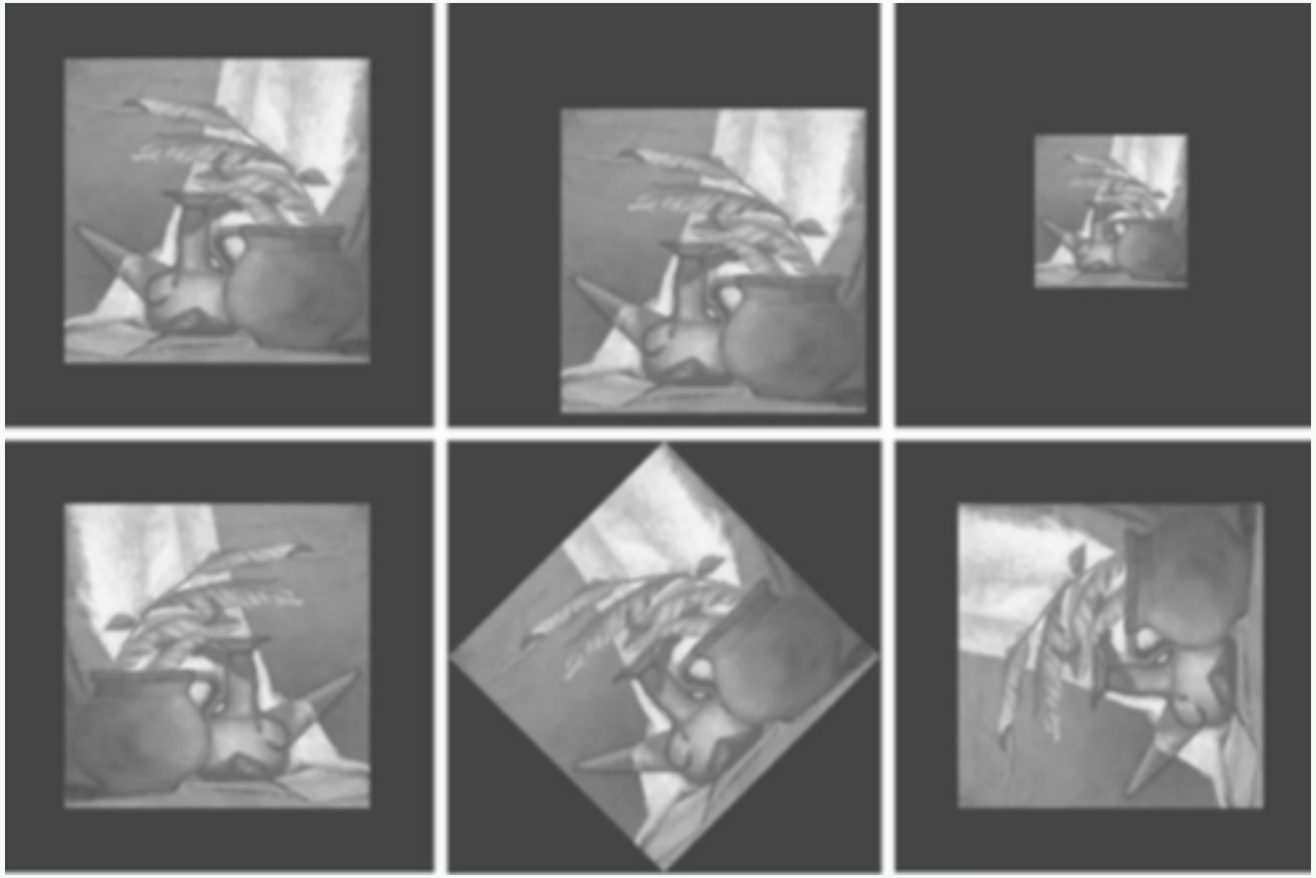
- Second-order column moment

$$\mu_{cc} = \frac{1}{A} \sum_{(r,c)\in R} (c - \bar{c})^2$$

Moment Invariants

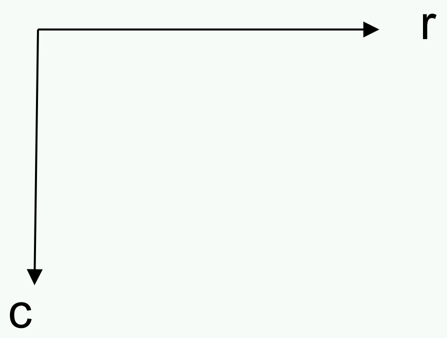
Artificial Inteligence & Computer Vision Laboratory

- Geometric transformation: translation, scale, mirroring, rotation



Contrast second moments

- For the letter 'I'
- Versus the letter 'O'
- Versus the underline '_'



Perimeter and Perimeter Length

Perimeter $P_4 = \{ (r,c) \in R \mid N_8(r,c) - R \neq \phi \}$

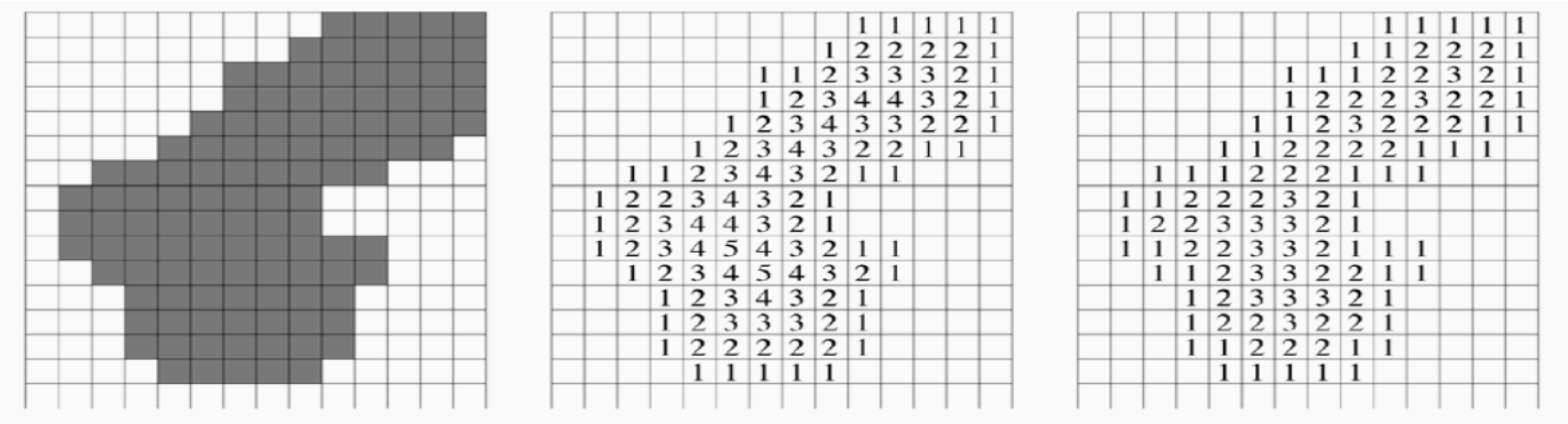
$P_8 = \{ (r,c) \in R \mid N_4(r,c) - R \neq \phi \}$

Perimeter Length

$$|P| = |\{k \mid (r_{k+1}, c_{k+1}) \in N_4(r_k, c_k)\}| + \sqrt{2}\,|\{k \mid (r_{k+1}, c_{k+1}) \in N_8(r_k, c_k) - N_4(r_k, c_k)\}|$$

Perimeter can vary significantly with object orientation

Perimeter and Perimeter Length



4-connected adjacency

8-connected adjacency

Circularity

- Common measure of circularity of a region is length of the perimeter squared divided by area

- Circularity (1): $$C_1 = \frac{|P|^2}{A}$$

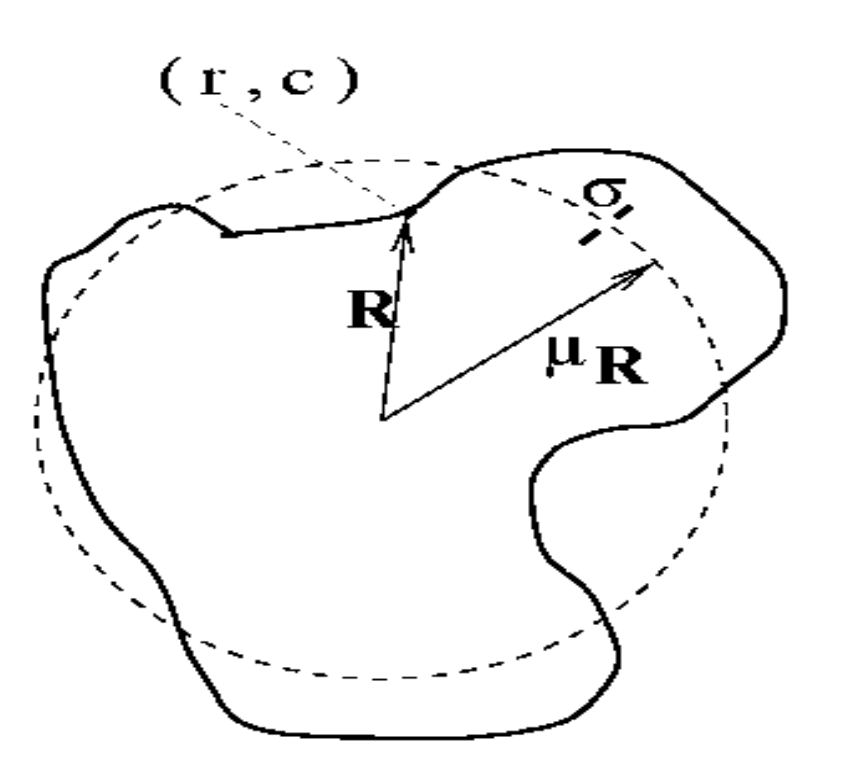
Circularity as variance of “radius”

- A second measure uses variation off of a circle
- Circularity (2): $$C_2 = \frac{\mu_R}{\sigma_R}$$

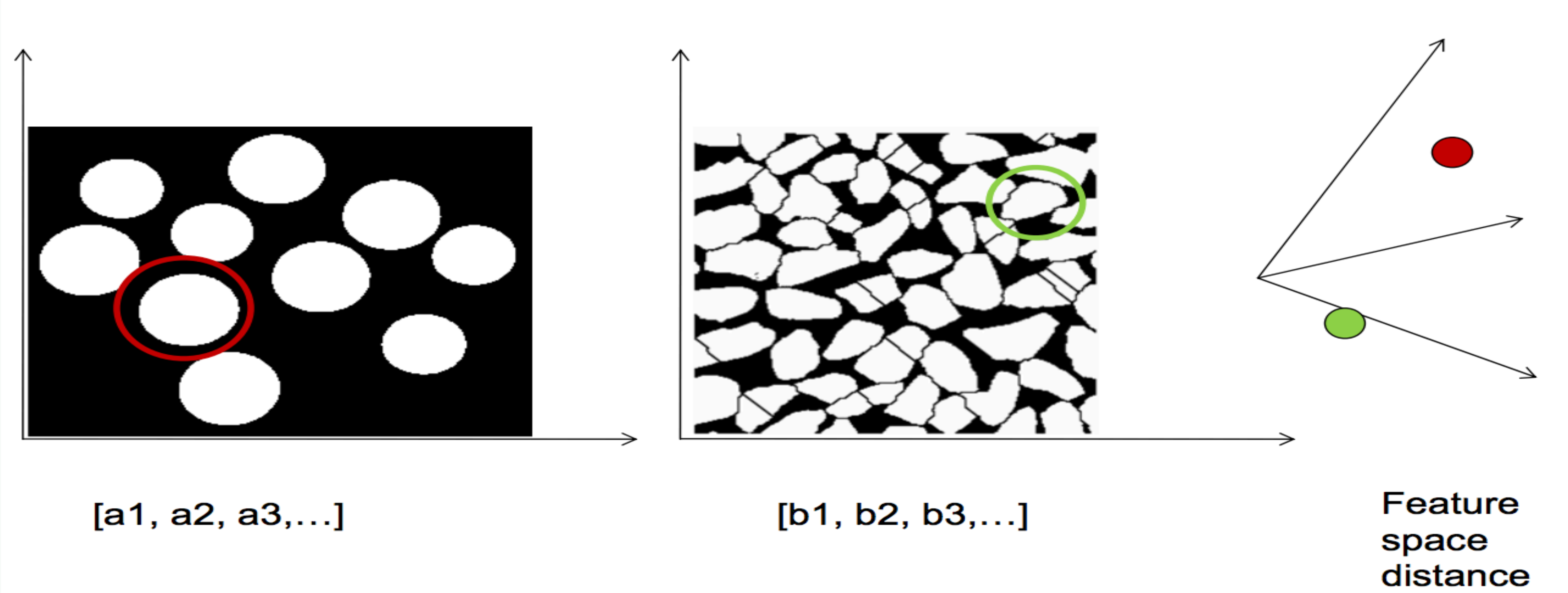
  - μ_R and σ_R^2 are the mean and variance of the distance from the centroid of the shape to the boundary pixels (r_k, c_k).

$$\mu_R = \frac{1}{K} \sum_{k=0}^{K-1} \| (r_k, c_k) - (\bar{r}, \bar{c}) \|$$

$$\sigma_R^2 = \frac{1}{K} \sum_{k=0}^{K-1} \left[\| (r_k, c_k) - (\bar{r}, \bar{c}) \| - \mu_R \right]^2$$



Invariant descriptors



Orientation (1)

- Define the orientation of an object as <u>the orientation of the axis of elongation</u>.

  ≡ axis of least <u>second order moment</u>

  variation(分散) = spread of data

  ≡ axis of least inertia

- The axis of least second moment for an object is the line which gives

$$\min_{line} \chi^2 = \min_{line} \sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}^2 B[i, j]$$

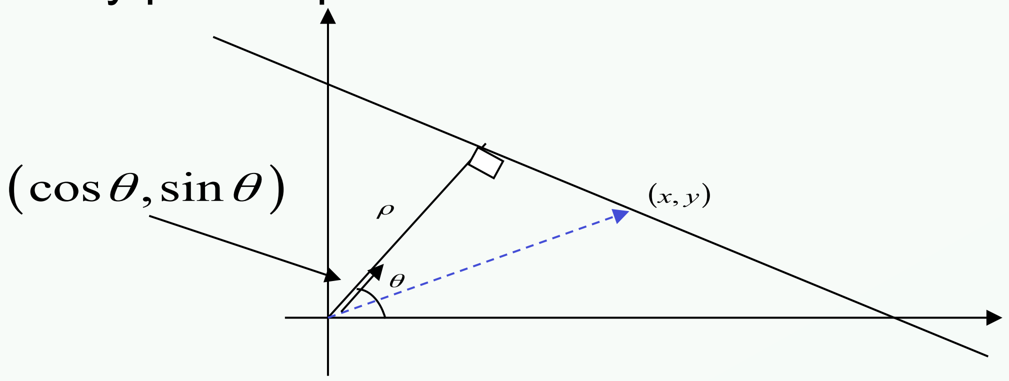
where r_{ij} the perpendicular distance from an object point $[i, j]$ to the line (axis)

Orientation (2)

Polar representation of a straight line

why polar representation instead of $y = ax + b$

cannot represent the vertical line



$$(x, y) \cdot (\cos\theta, \sin\theta) = \rho$$

$$x\cos\theta + y\sin\theta = \rho$$

projection of (x, y) onto the direction $(\cos\theta, \sin\theta)$

Then,

$$r^2 = (x\cos\theta + y\sin\theta - \rho)^2$$

$$\chi^2 = \sum_{i=1}^{n}\sum_{j=1}^{n}(x_{ij}\cos\theta + y_{ij}\sin\theta - \rho)^2 B[i, j]$$

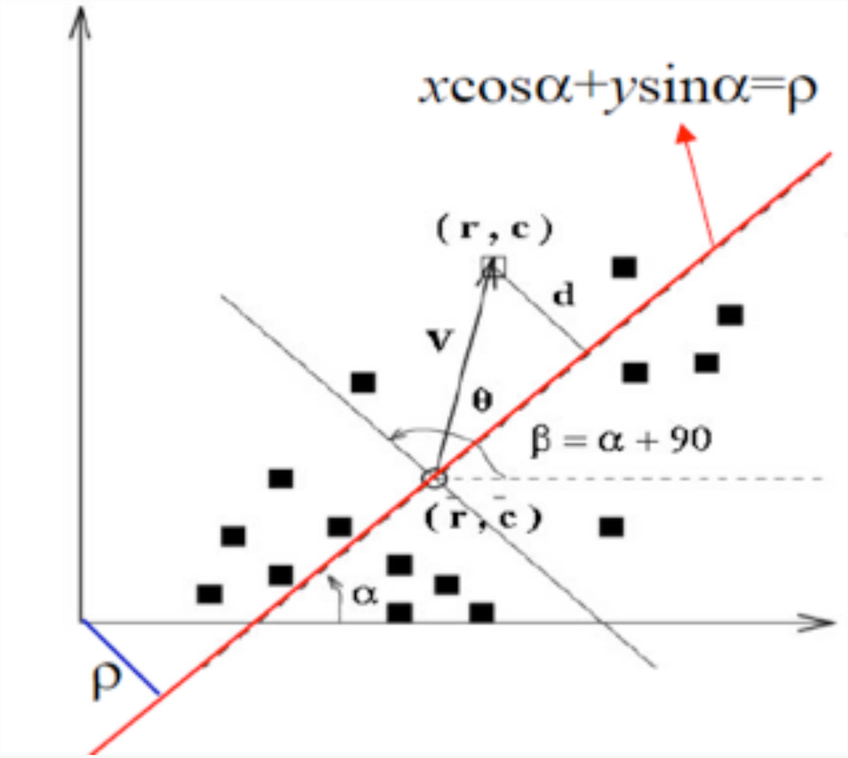
Problem: Find ρ and θ that minimizes χ^2.

Orientation (3)

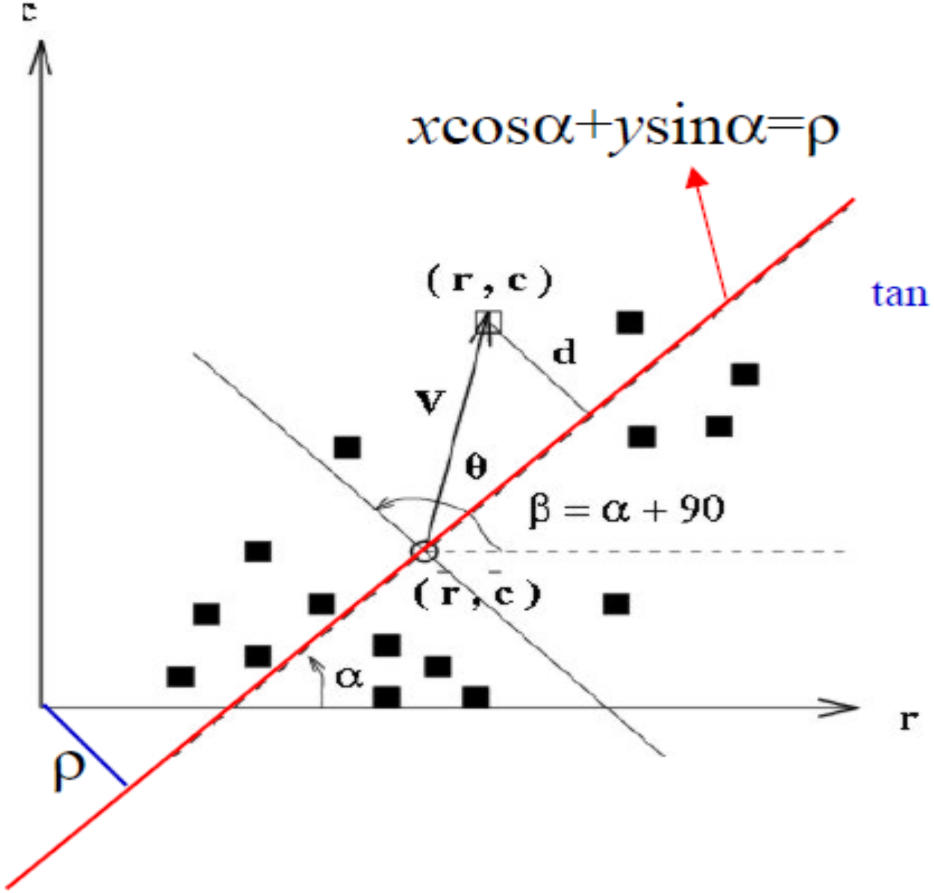
Solution: $\frac{\partial \chi^2}{\partial \rho} = 0$ and $\frac{\partial \chi^2}{\partial \theta} = 0$

- The elongation E of the object $\equiv \frac{\chi_{\max}}{\chi_{\min}}$

Orientation (4)
: Axis with Least Second Moment



Orientation (4)
: Axis with Least Second Moment



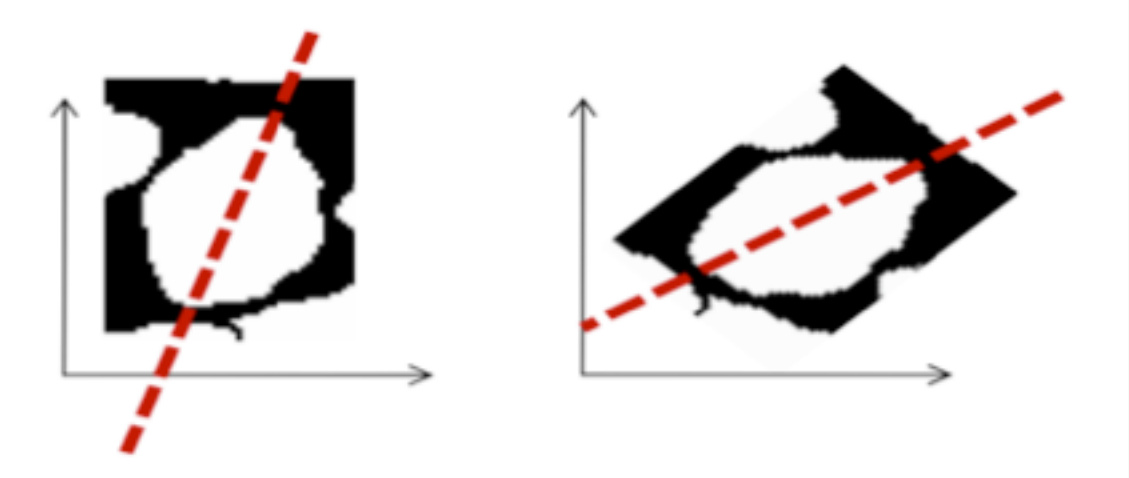
$$\tan 2\alpha = \frac{2\sum(r-\bar{r})(c-\bar{c})}{\sum(r-\bar{r})(r-\bar{r})-\sum(c-\bar{c})(c-\bar{c})}$$

$$= \frac{\frac{1}{A}2\sum(r-\bar{r})(c-\bar{c})}{\frac{1}{A}\sum(r-\bar{r})(r-\bar{r})-\frac{1}{A}\sum(c-\bar{c})(c-\bar{c})}$$

$$= \frac{2\mu_{rc}}{\mu_{rr}-\mu_{cc}}$$

Orientation (5)
: Axis with Least Second Moment

- Invariance to orientation?
  : Need a common alignment



Axis for which the squared distance to 2d object points is minimized

Basic Properties of a Region

0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0
2	2	2	2	0	0	0	0	0	1	1	1	1	1	1	0
2	2	2	2	0	0	0	0	1	1	1	1	1	1	1	1
2	2	2	2	0	0	0	0	1	1	1	1	1	1	1	1
2	2	2	2	0	0	0	0	1	1	1	1	1	1	1	1
2	2	2	2	0	0	0	0	0	1	1	1	1	1	1	0
2	2	2	2	0	0	0	0	0	0	1	1	1	1	0	0
2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0
2	2	2	2	0	0	3	3	3	0	0	0	0	0	0	0
2	2	2	2	0	0	3	3	3	0	0	0	0	0	0	0
2	2	2	2	0	0	3	3	3	0	0	0	0	0	0	0
2	2	2	2	0	0	0	0	0	0	0	0	0	0	0	0

region num.	region area	row of center	col of center	perim. length	circu-larity$_1$	circu-larity$_2$	radius mean	radius var.
1	44	6	11.5	21.2	10.2	15.4	3.33	.05
2	48	9	1.5	28	16.3	2.5	3.80	2.28
3	9	13	7	8	7.1	5.8	1.2	0.04

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

1	0
1	1

0	1
1	1

1	1
0	1

(*a*) 2×2 *external corner* patterns

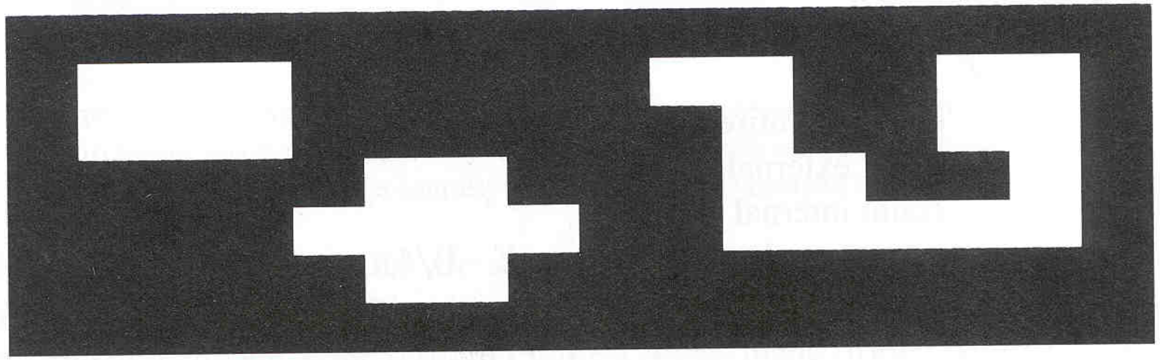
0	0
0	1

0	1
0	0

1	0
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

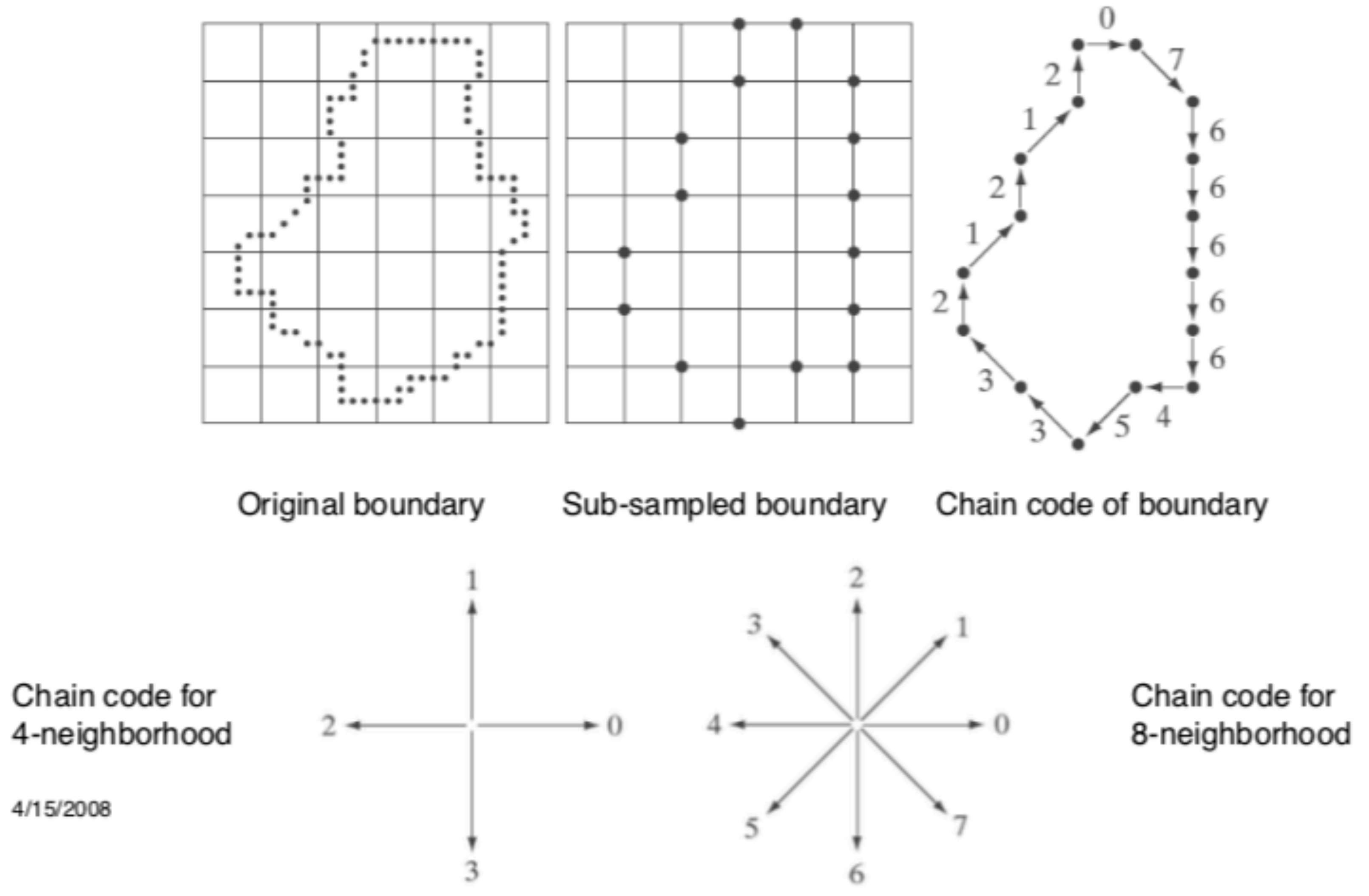
( ) External corner patterns marked with e; internal corners marked i

Boundary based Shape Descriptors

Boundary Representation

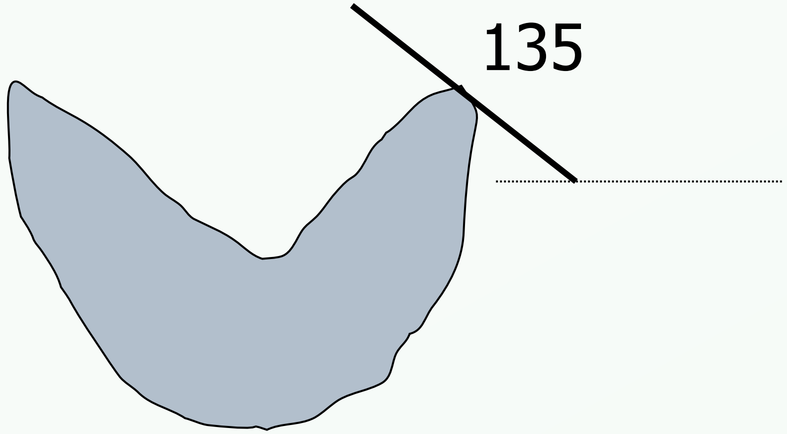
- (Freeman) Chain Code

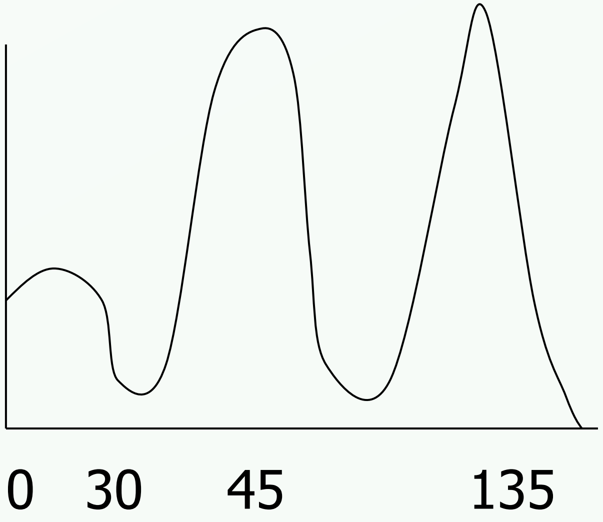
- Boundary representation
  = 0766666453321212



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

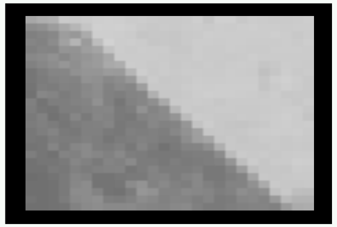
- Find “interesting” pieces of the image
  - E.g. corners, salient regions
  - Focus attention of algorithms
  - Speed up computation

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

Interest points

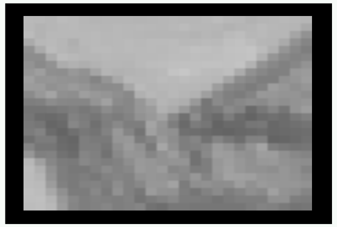
0D structure: single points

➡ not useful for matching



1D structure: lines

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



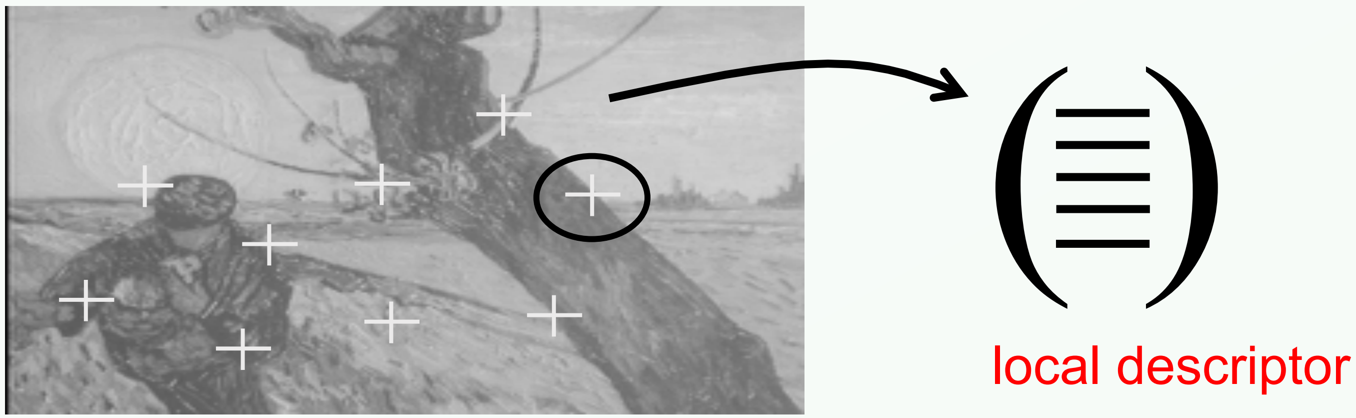
2D structure: corners

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

Interest Points have **2D** structure.

Local invariant photometric descriptors

Artificial Inteligence & Computer Vision Laboratory



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

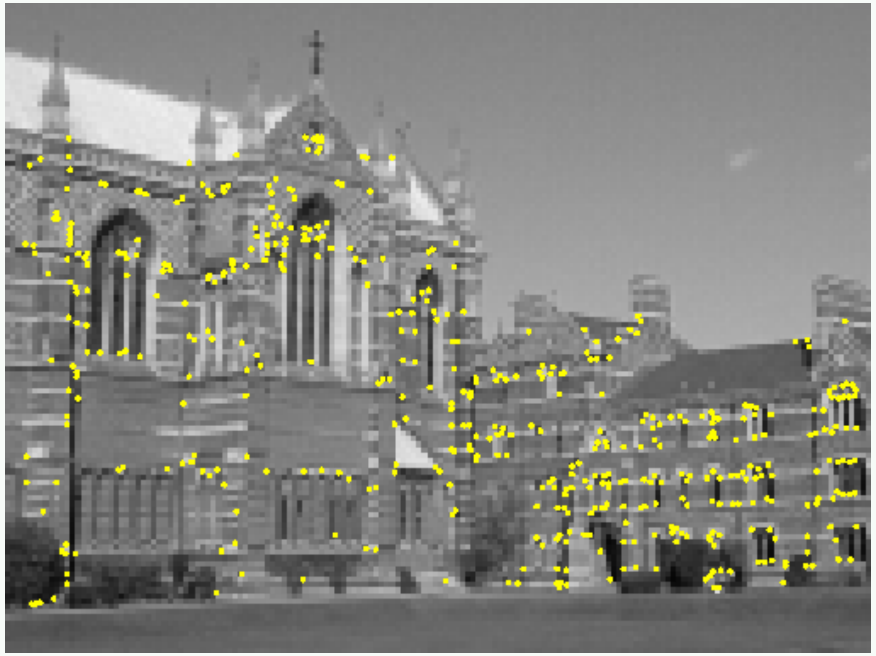
Zhang Approach

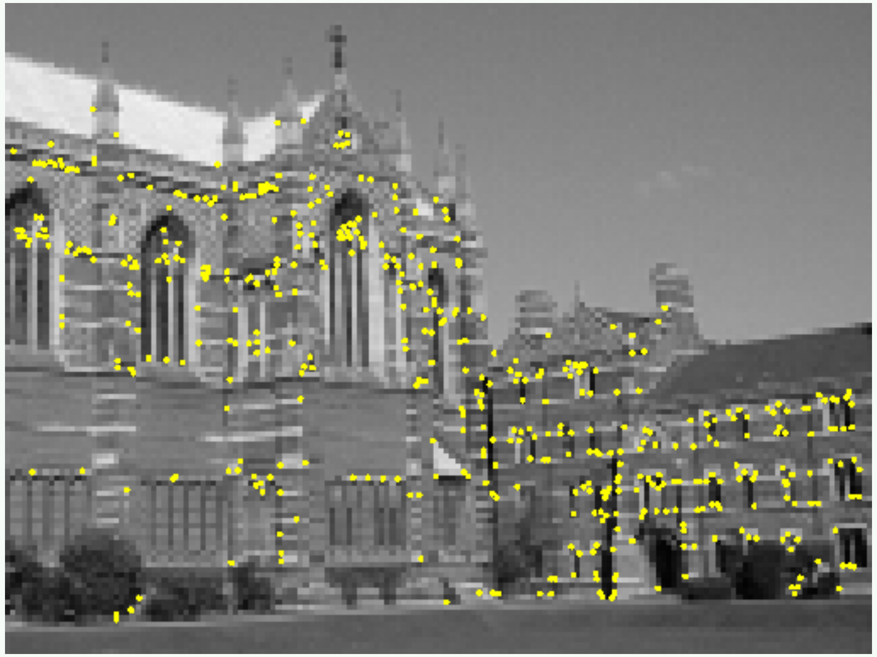
Artificial Inteligence
& Computer Vision
Laboratory

- 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.

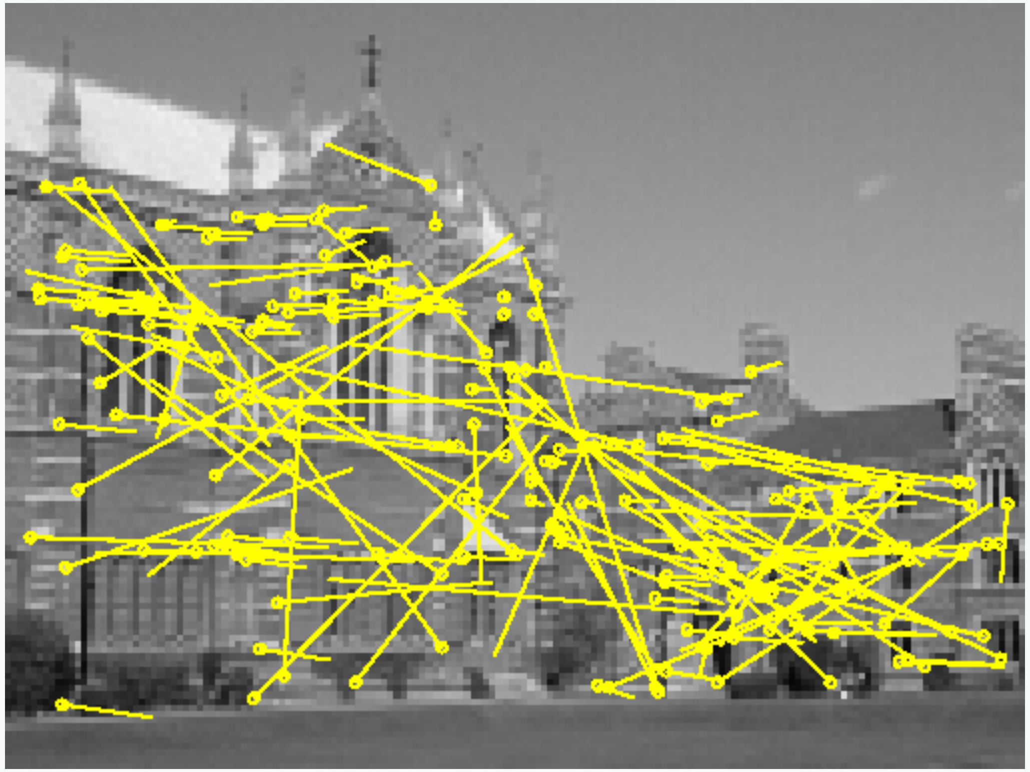
Preview: Harris detector





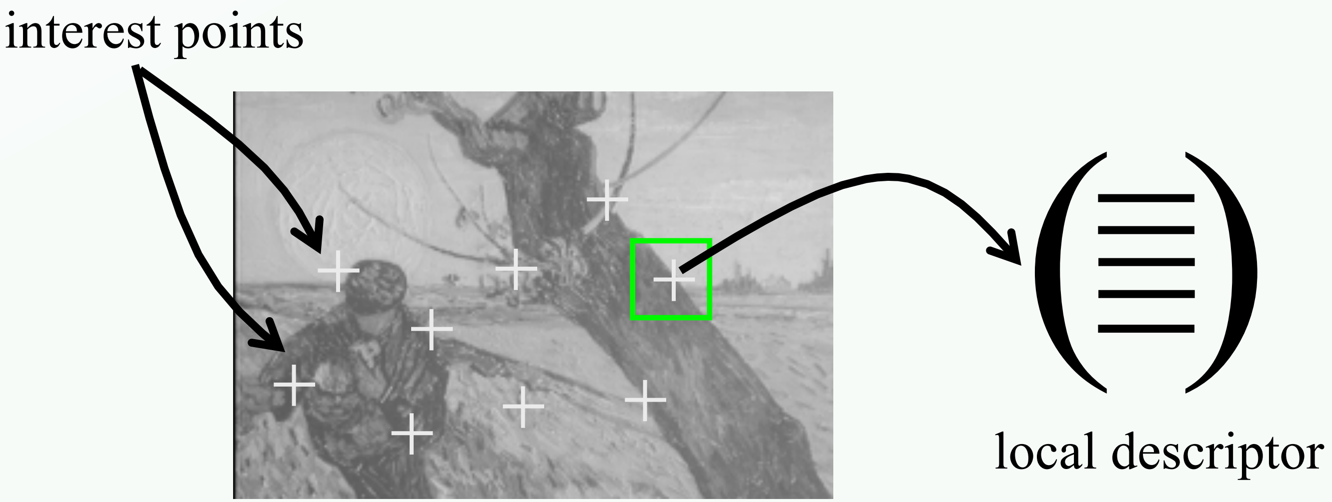
Interest points extracted with Harris (~ 500 points)

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

Artificial Inteligence & Computer Vision Laboratory

Based on the idea of auto-correlation

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 $(\Delta x, \Delta y)$

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

SSD means summed square difference

Discrete shifts can be avoided with the auto-correlation matrix

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

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

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

Artificial Inteligence & Computer Vision Laboratory

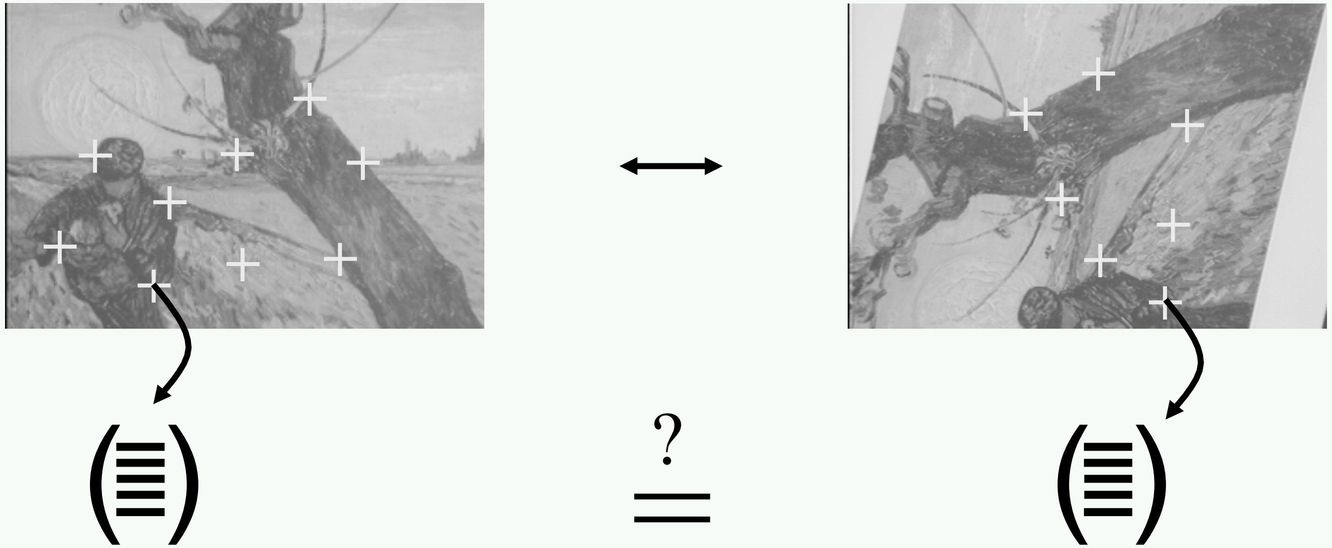
$$= \begin{pmatrix}\Delta x & \Delta y\end{pmatrix} \begin{bmatrix} \sum_{(x_k, y_k) \in W} (I_x(x_k, y_k))^2 & \sum_{(x_k, y_k) \in W} I_x(x_k, y_k) I_y(x_k, y_k) \\ \sum_{(x_k, y_k) \in W} I_x(x_k, y_k) I_y(x_k, y_k) & \sum_{(x_k, y_k) \in W} (I_y(x_k, y_k))^2 \end{bmatrix} \begin{pmatrix}\Delta x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

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

Determining correspondences



Vector comparison using a distance measure

What are some suitable distance measures?

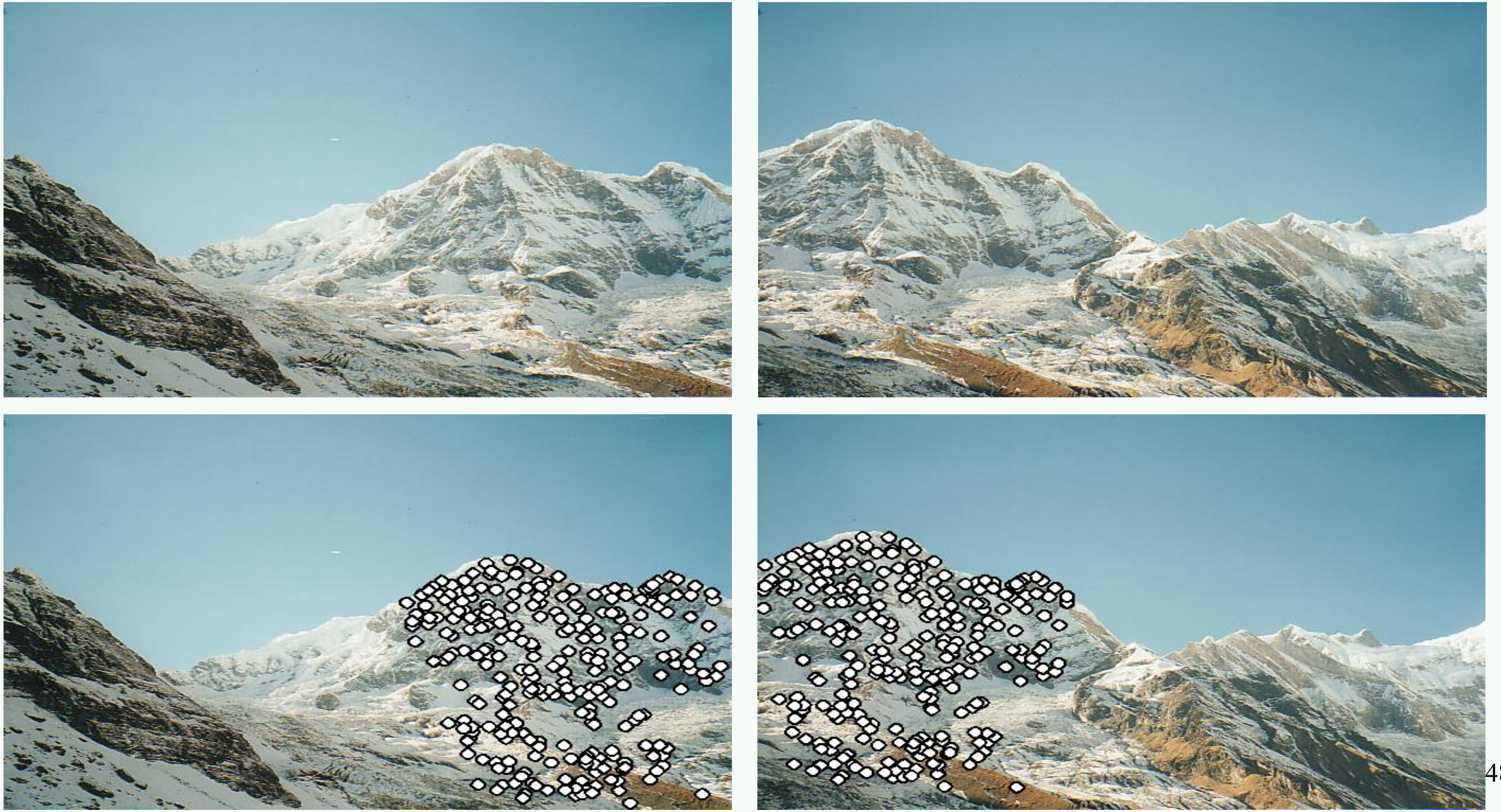
Distance Measures

- We can use the sum-square difference of the values of the pixels in a square neighborhood about the points being compared. This is the simplest measure.

$$SSD = \sum\sum(W1_{i,j} - W2_{i,j})^2$$

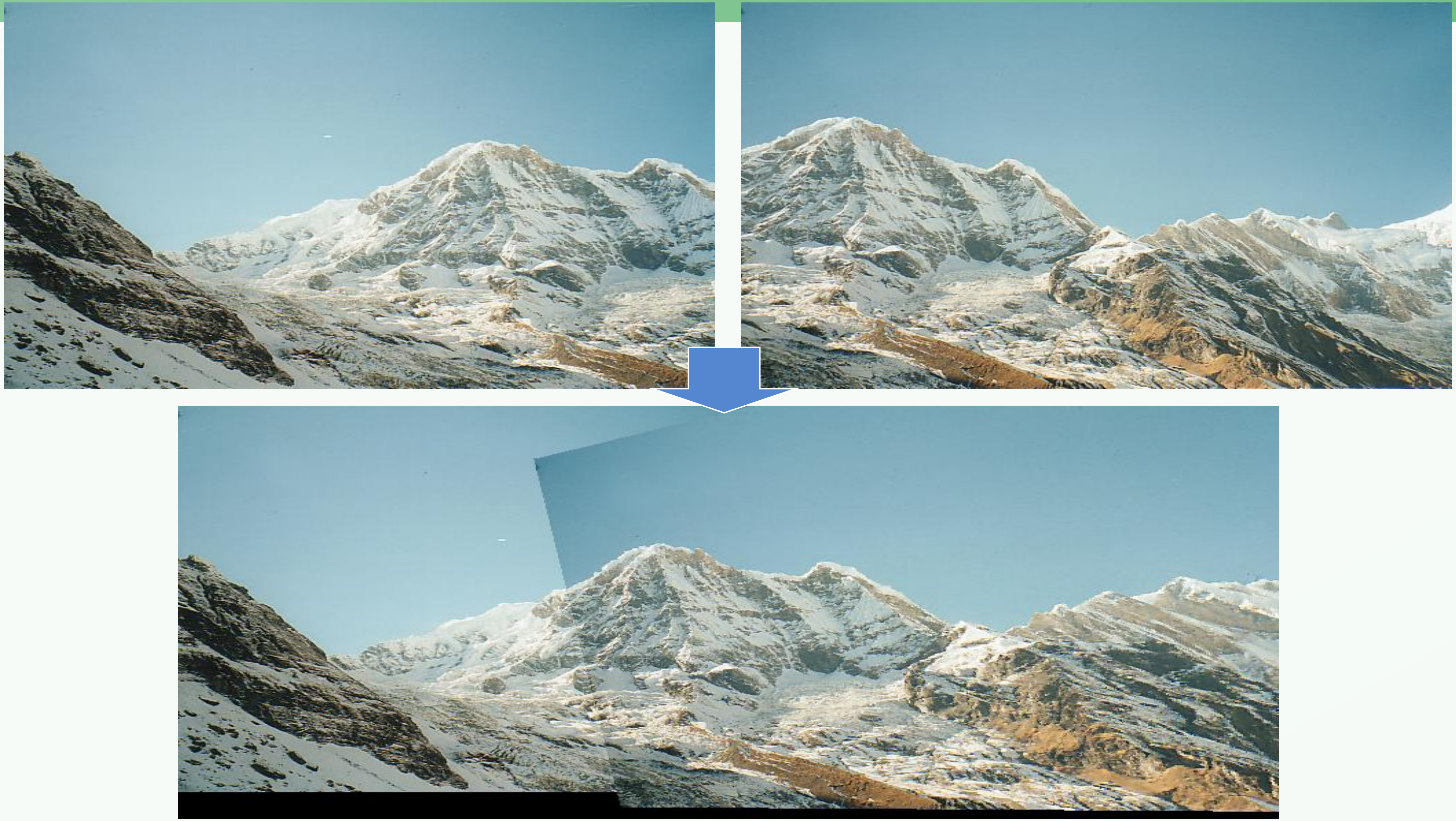
Some Matching Results from Matt Brown

Artificial Inteligence & Computer Vision Laboratory



Some Matching Results

Artificial Inteligence
& Computer Vision
Laboratory



Summary of the approach

- Basic feature matching = **Harris Corners & Correlation**
- 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

Rotation/Scale Invariance

original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	?	?	?
Is correlation invariant?	?	?	?

Rotation/Scale Invariance

original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	?	?	?
Is correlation invariant?	?	?	?

Rotation/Scale Invariance

original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	YES	?	?
Is correlation invariant?	?	?	?

Rotation/Scale Invariance

original translated rotated scaled

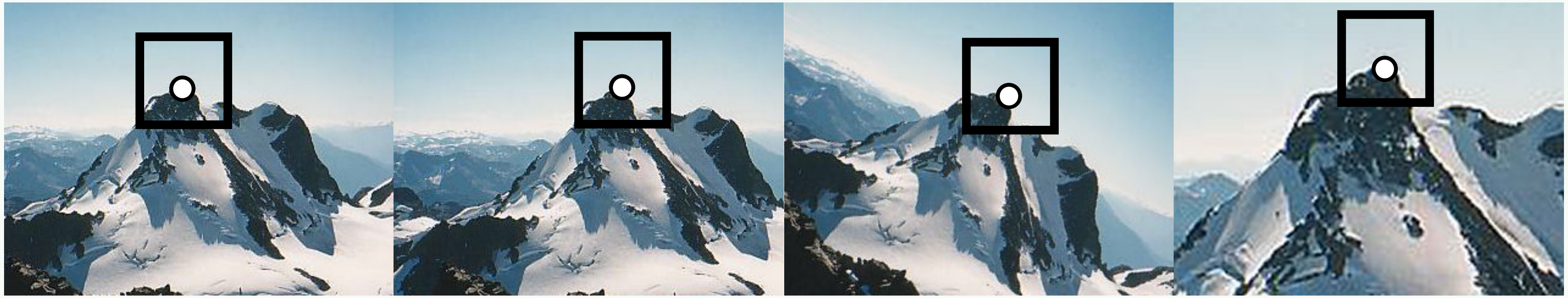
	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	?
Is correlation invariant?	?	?	?

Rotation/Scale Invariance

original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	NO
Is correlation invariant?	?	?	?

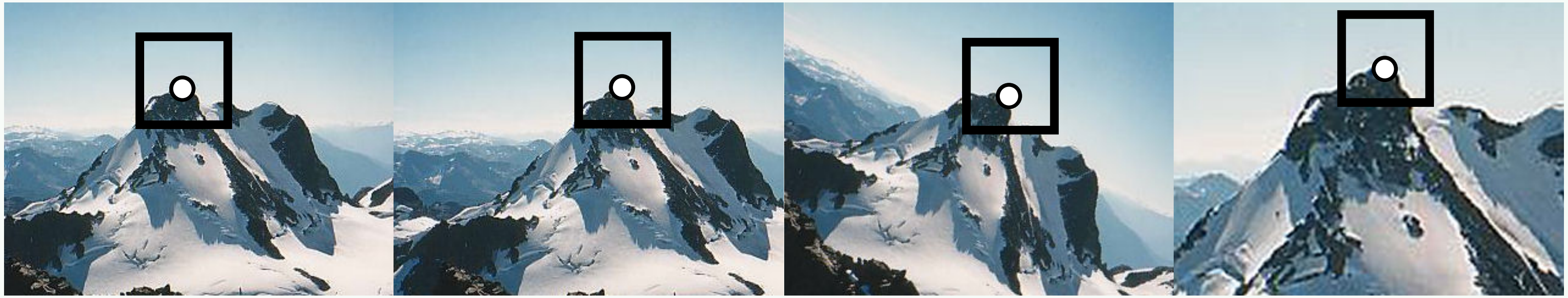
Rotation/Scale Invariance



original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	NO
Is correlation invariant?	?	?	?

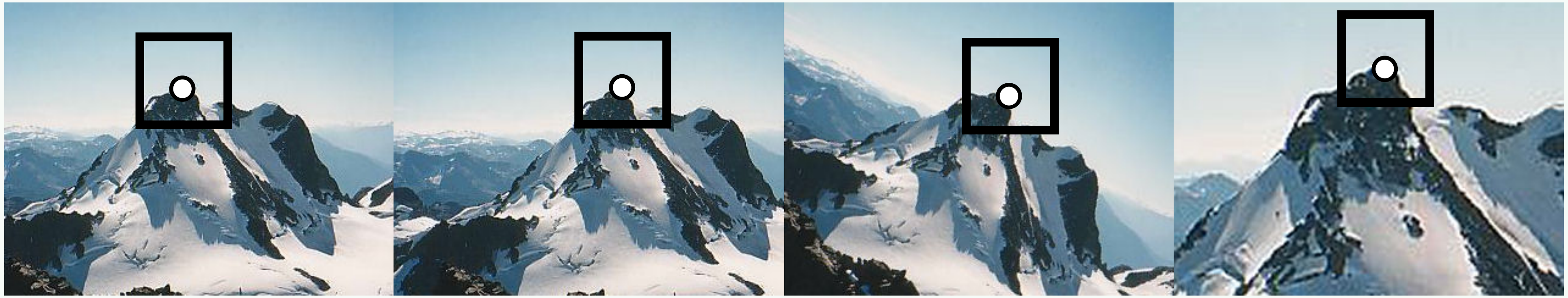
Rotation/Scale Invariance



original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	NO
Is correlation invariant?	YES	?	?

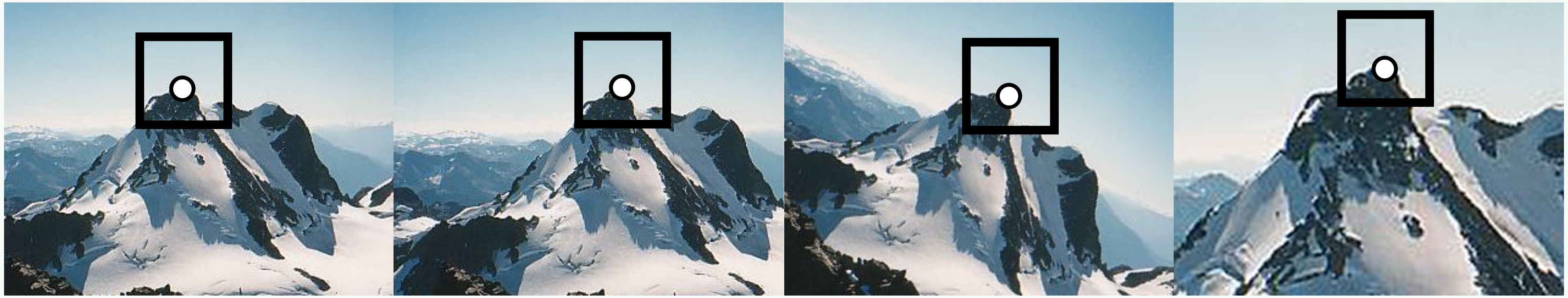
Rotation/Scale Invariance



original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	NO
Is correlation invariant?	YES	NO	?

Rotation/Scale Invariance

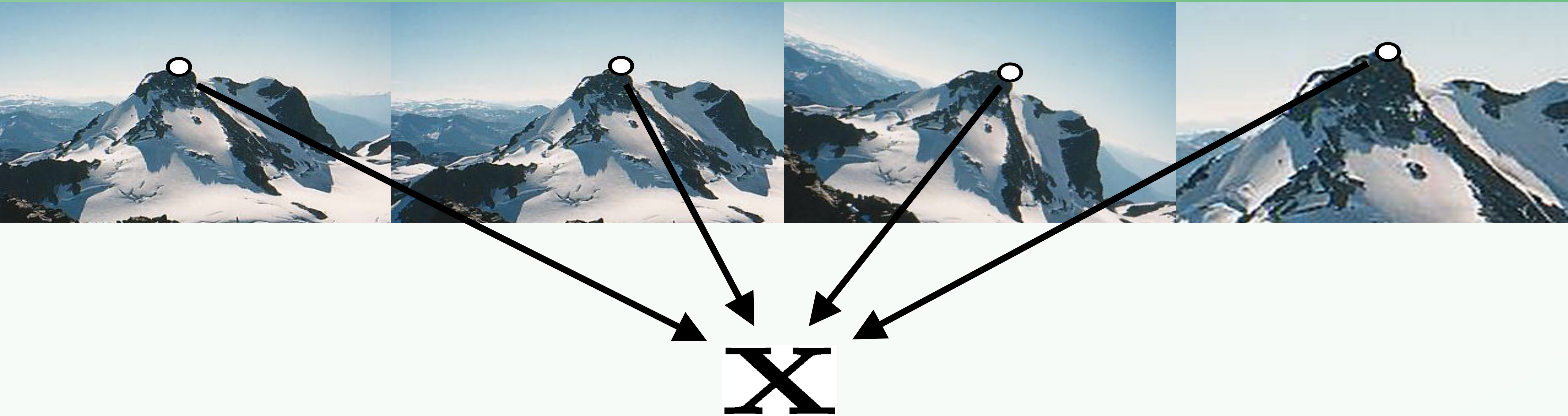


original translated rotated scaled

	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	NO
Is correlation invariant?	YES	NO	NO

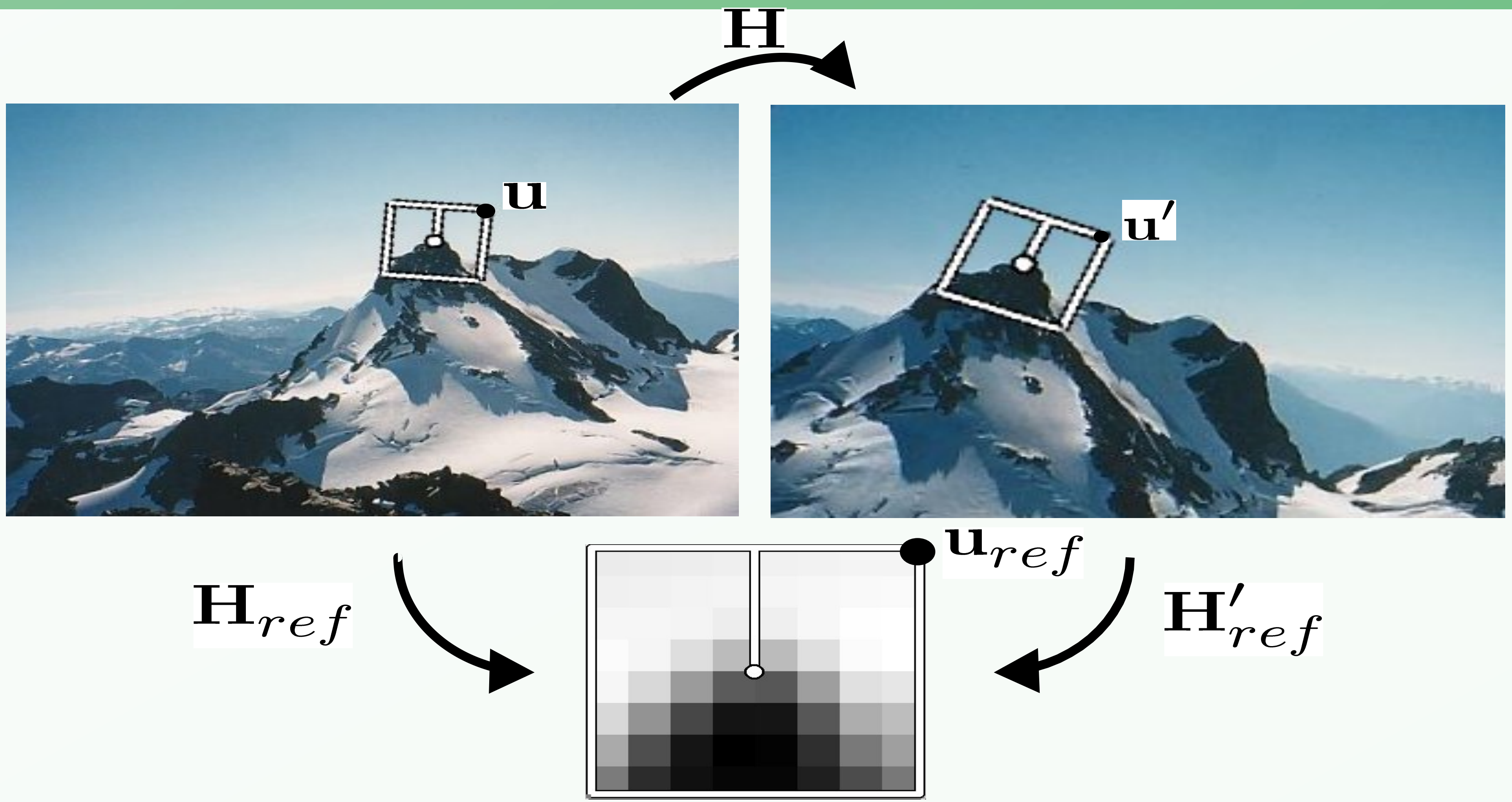
Matt Brown's Invariant Features

Artificial Inteligence & Computer Vision Laboratory



- Local image descriptors that are *invariant* (unchanged) under image t ransformations

Canonical Frames



Rotation-invariant descriptor.

Multi-Scale Oriented Patches

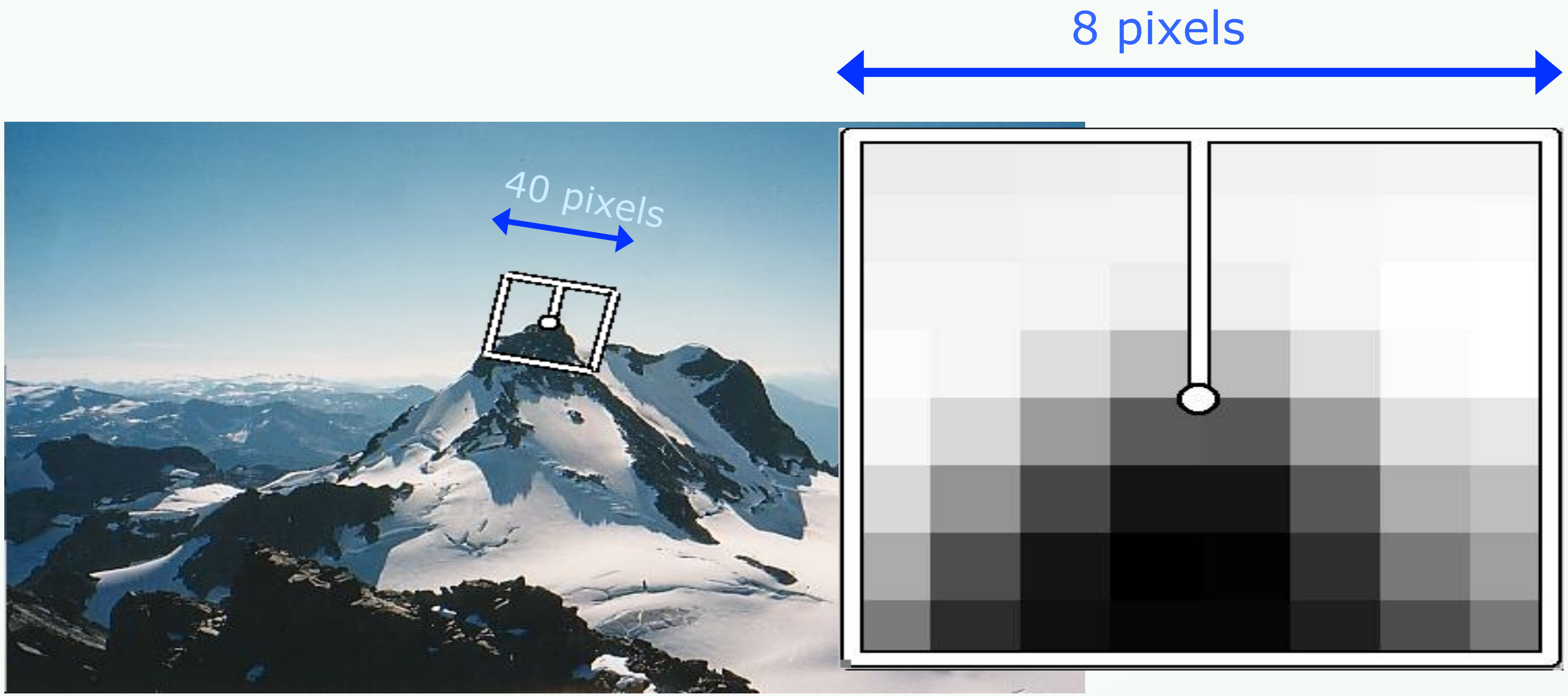
- Sample scaled, oriented patch

Multi-Scale Oriented Patches

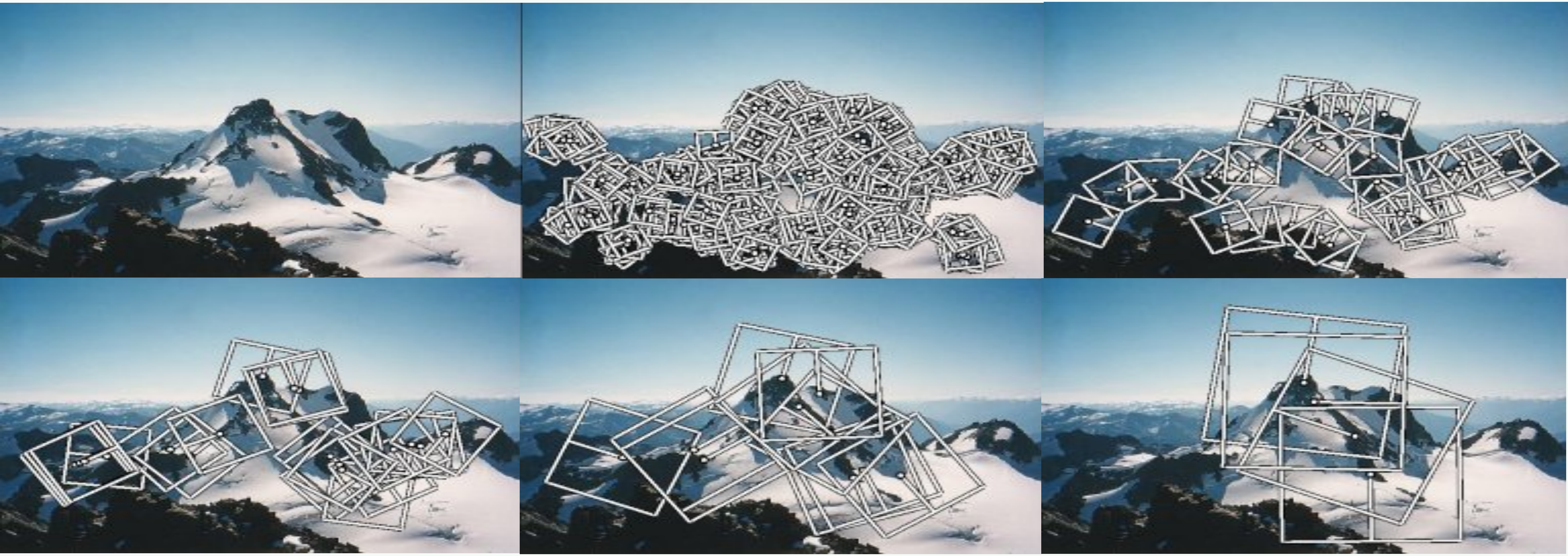
- Sample scaled, oriented patch
  - 8x8 patch, sampled at 5 x scale

Multi-Scale Oriented Patches

- Sample scaled, oriented patch
  - 8x8 patch, sampled at 5 x scale

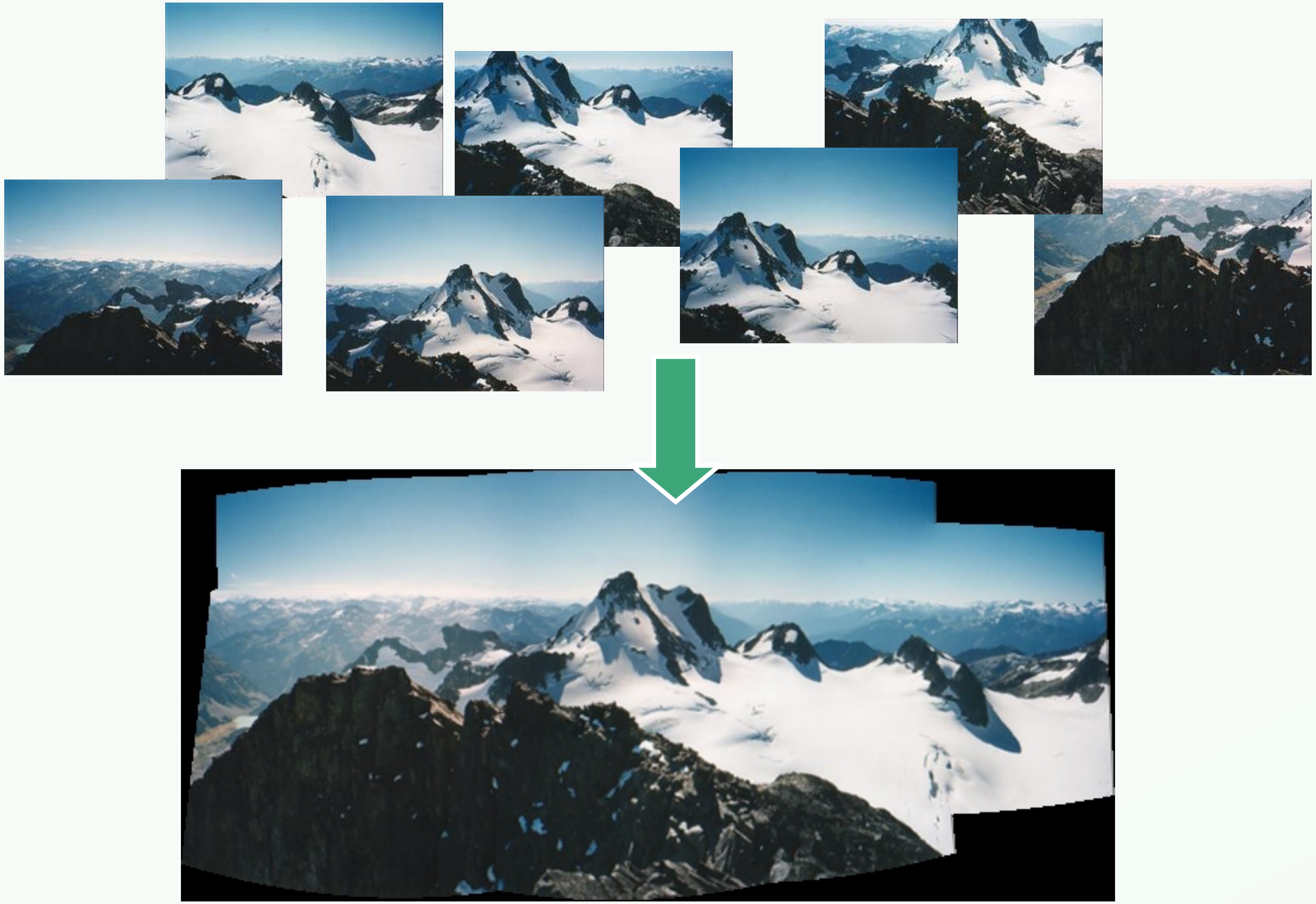


Multi-Scale Oriented Patches



- Extract oriented patches at multiple scales using dominant orientation

Application: Image Stitching



[Microsoft Digital Image Pro version 10]

Matching Interest Points: Summary

Artificial Inteligence & Computer Vision Laboratory

- Harris corners / correlation
  - Extract and match repeatable image features
  - Robust to clutter and occlusion
  - BUT not invariant to scale and rotation
- Multi-Scale Oriented Patches
  - Corners detected at multiple scales
  - Descriptors oriented using local gradient
    - Also, sample a blurred image patch
  - Invariant to scale and rotation

Leads to: **SIFT** – state of the art image features