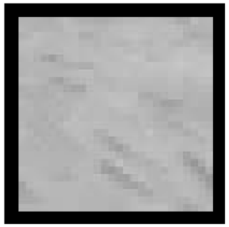
Interest Operators

All lectures are from posted research papers.

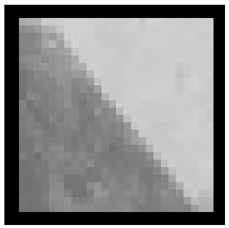
- **Harris Corner Detector: the first and most basic interest operator**
- **SIFT interest point detector and region descriptor**
- **Kadir Entropy Detector and its use in object recognition**

Interest points



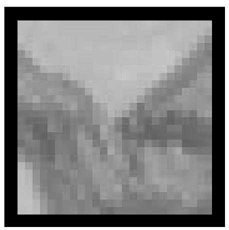
0D structure

➡ not useful for matching



1D structure

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



2D structure

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

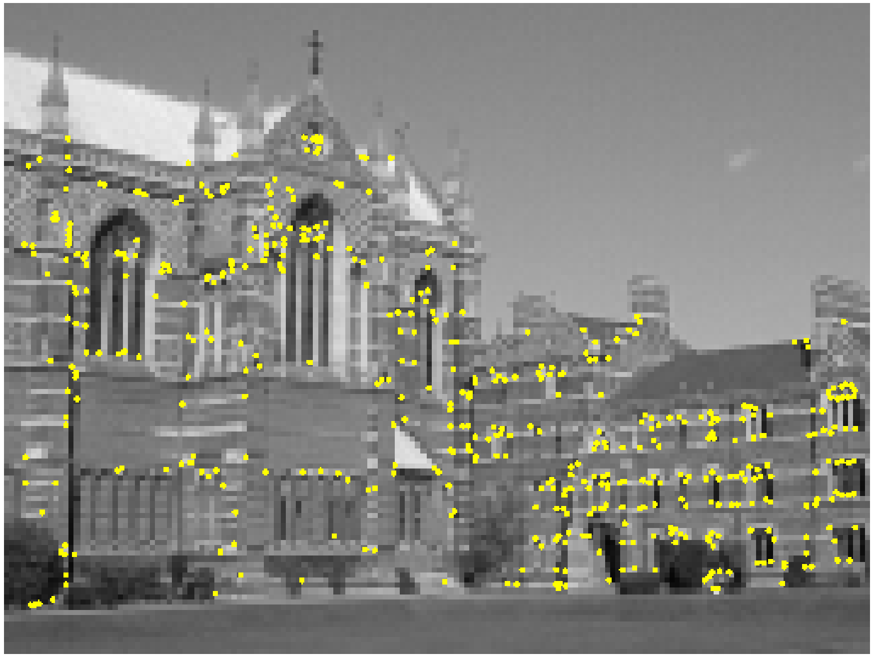
Interest Points have **2D** structure.

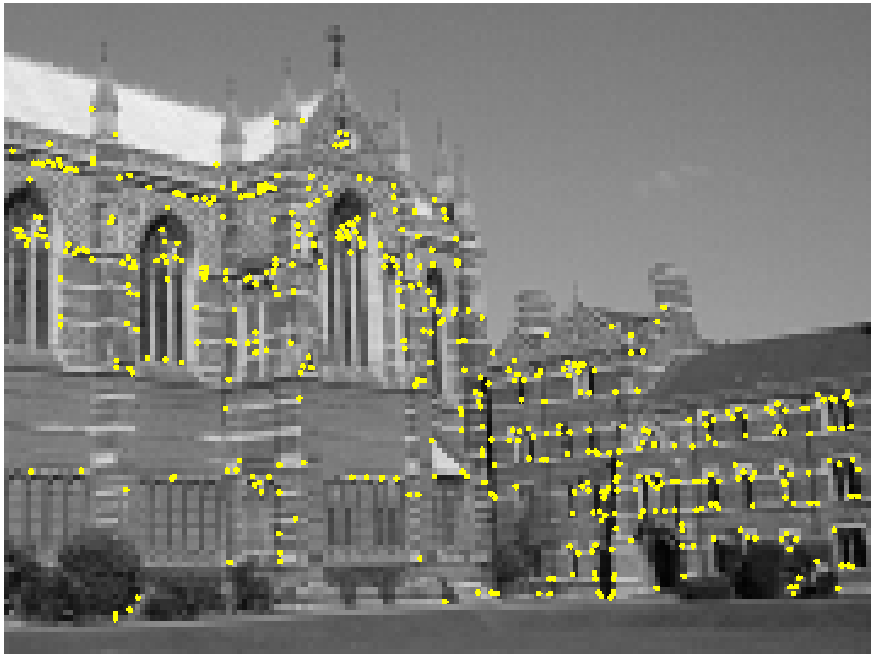
Simple Approach

- 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 points in the second matrix using a homography that is determined through the solution of a set of equations that usually minimizes a least square error. CH 12-13

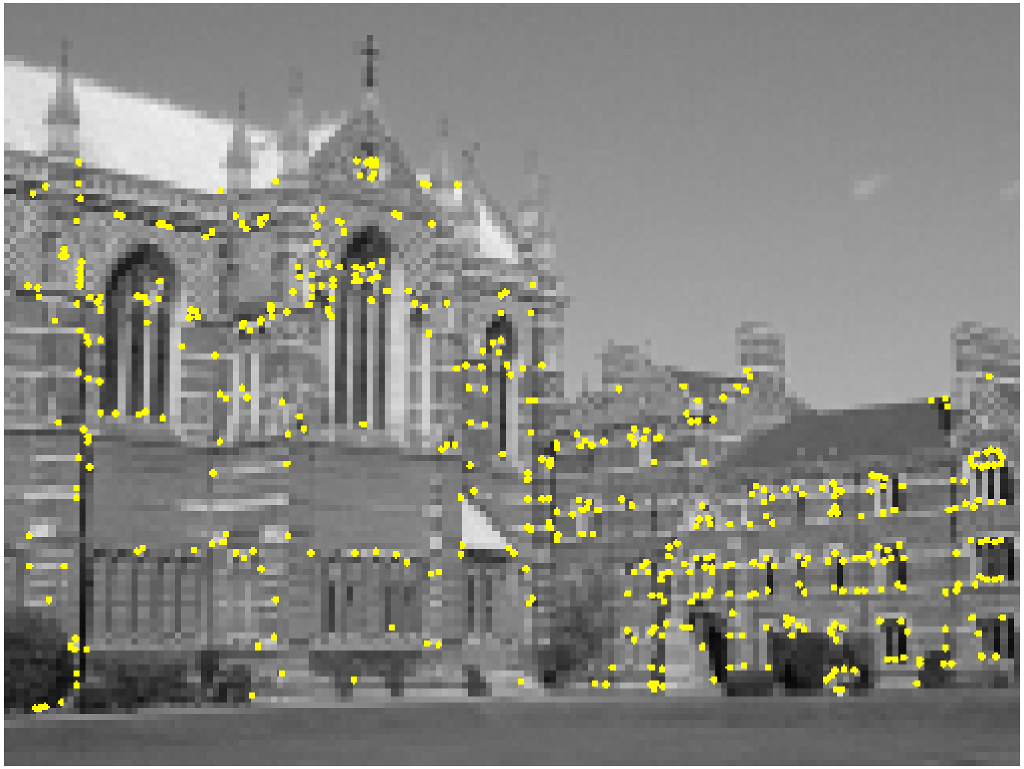
Preview: Harris detector



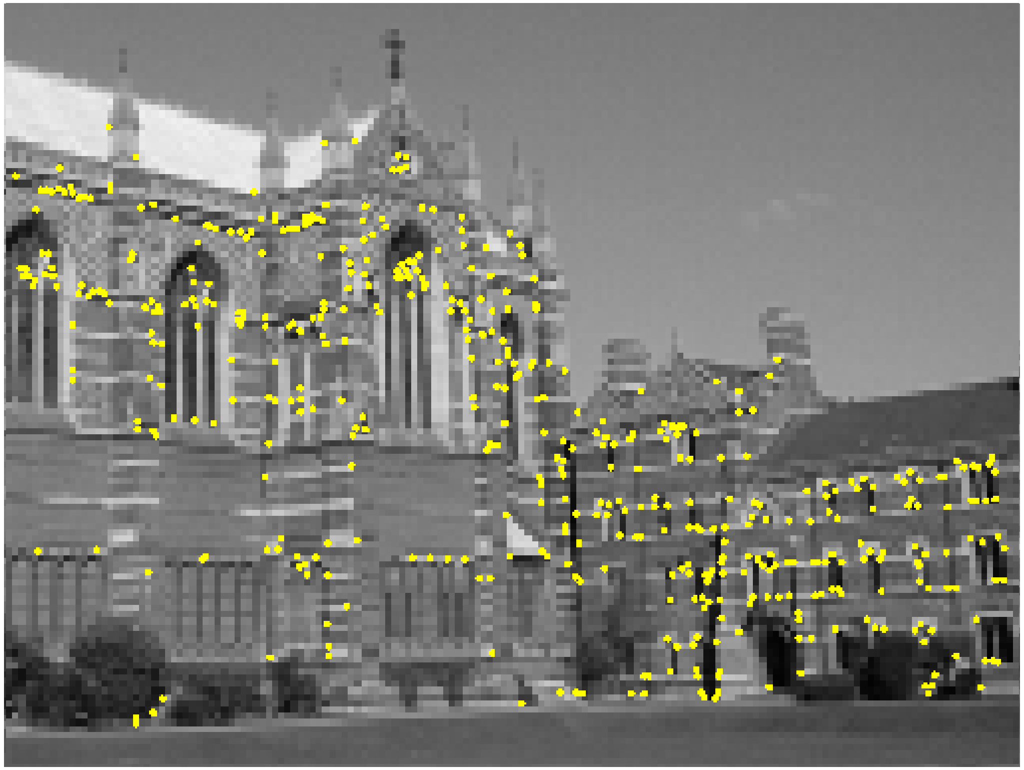


Interest points extracted with Harris (~ 500 points)

Harris detector

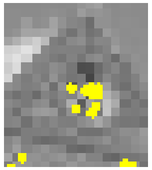


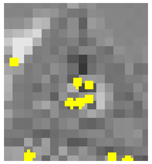
Harris detector

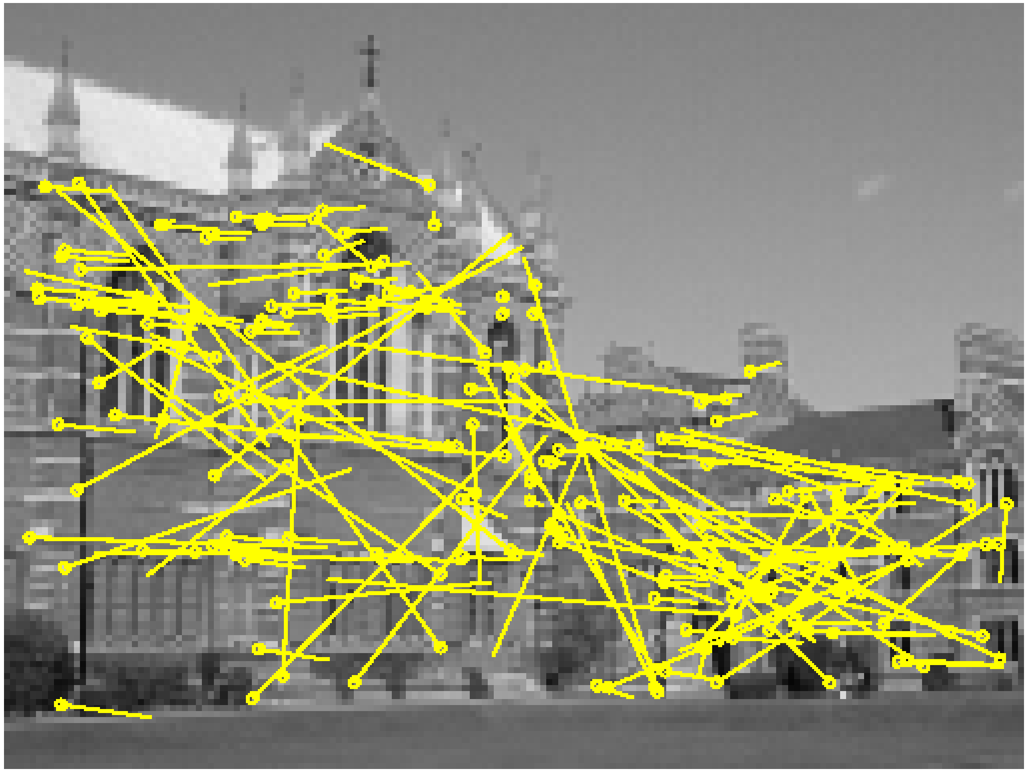


Cross-correlation matching

Match two points based on how similar their neighbor-hoods are.







Initial matches – motion vectors (188 pairs)

Problem with Matching

- You get a lot of false matches.
- There is really only one (3D) transformation between the two views.
- So all the matches should be consistent.
- Ransac is an algorithm that chooses consistent matches and throws the outliers out.

Homography

In the field of computer vision, any two images of the same planar surface in space are related by a homography (assuming a pinhole camera model).

--

Mathematical definition

Homogeneous coordinates are used, because matrix multiplication cannot directly perform

Given:

$$p_a = \begin{bmatrix} x_a \\ y_a \\ 1 \end{bmatrix}, p'_b = \begin{bmatrix} w'x_b \\ w'y_b \\ w' \end{bmatrix}, \mathbf{H}_{ab} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

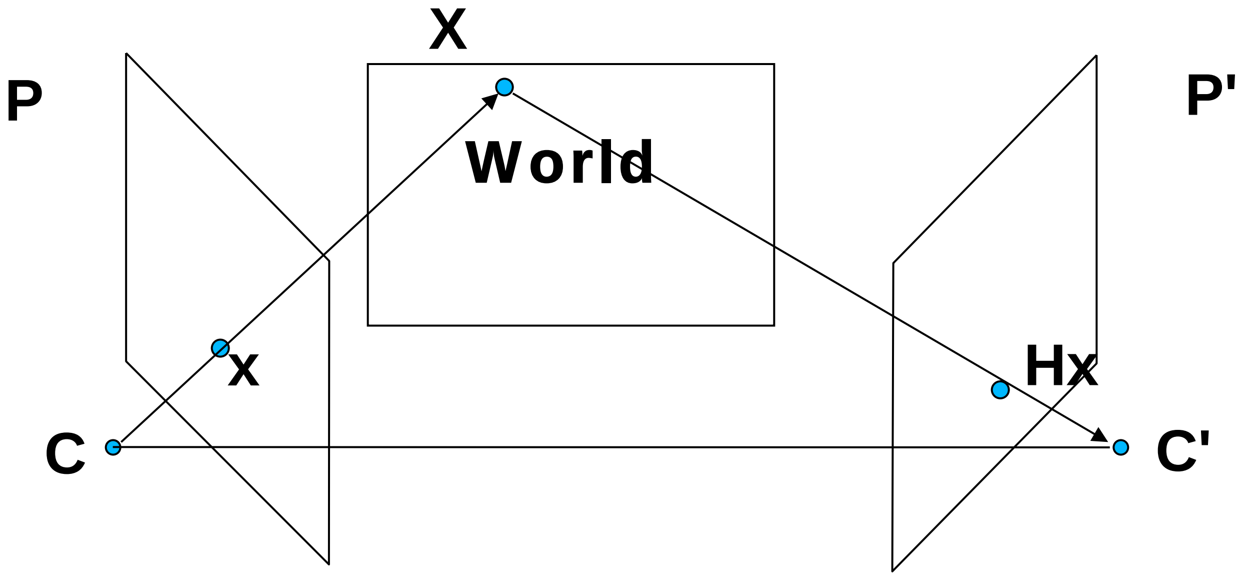
Then:

$$p'_b = \mathbf{H}_{ab} p_a$$

From Wikipedia

$$
\begin{pmatrix}
8.7976964e\text{-}01 & 3.1245438e\text{-}01 & -3.9430589e{+}01 \\
-1.8389418e\text{-}01 & 9.3847198e\text{-}01 & 1.5315784e{+}02 \\
1.9641425e\text{-}04 & -1.6015275e\text{-}05 & 1.0000000e{+}00
\end{pmatrix}
$$

Plane Transfer Homography



- Because we assume the world is a plane, x and transferred points x’ are related by a homography.
- If world plane coordinate is p, then
- x=Ap and x’=A’p.
- $x' = A'A^{-1}x$.

RANSAC for Fundamental Matrix

Step 1. Extract features

Step 2. Compute a set of potential matches

Step 3. do

Step 3.1 select minimal sample (i.e. 7 matches)

Step 3.2 compute solution(s) for F

(generate hypothesis)

Step 3.3 determine inliers (verify hypothesis)

until a large enough set of the matches become inliers

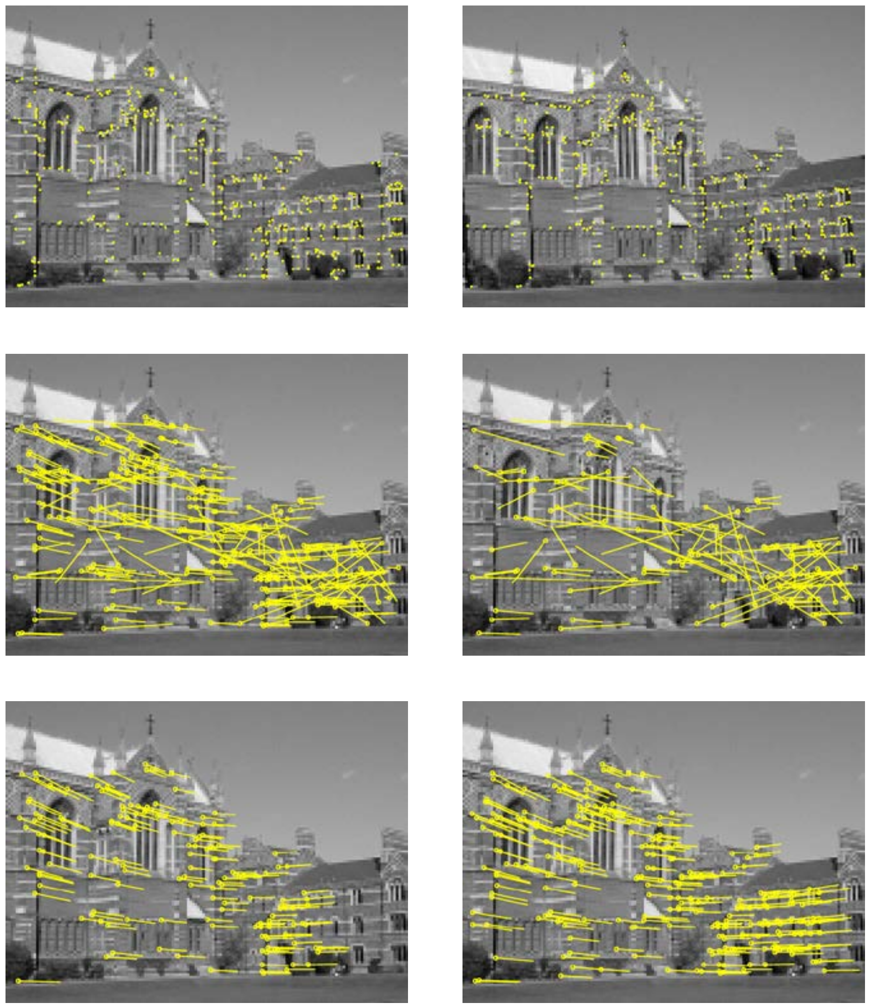
Step 4. Compute F based on all inliers

Step 5. Look for additional matches

Step 6. Refine F based on all correct matches

Example: robust computation

from H&Z



Interest points (500/image)
(640x480)

Putative correspondences (268)
(Best match,SSD<20,±320)

Outliers (117)
(*t*=1.25 pixel; 43 iterations)

Inliers (151)

Final inliers (262)

Corner detection

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 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

what is this?

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

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

Harris detector

Rewrite as inner (dot) product

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

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

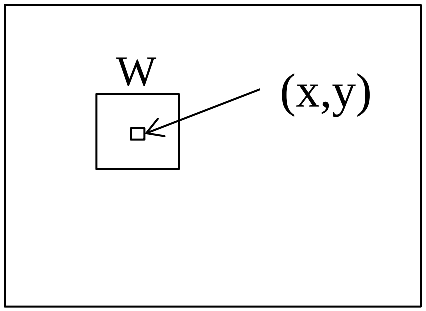
The center portion is a 2x2 matrix

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

Harris Corner Detector

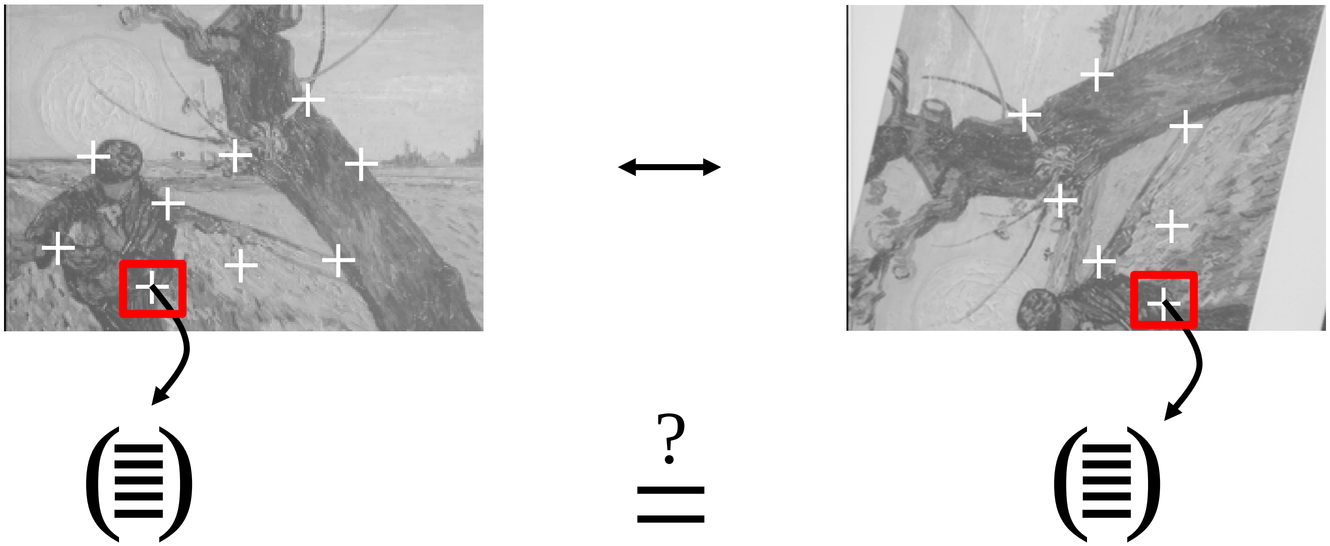
- Corner strength $R = \text{Det}(M) - k\,\text{Tr}(M)^2$
- Let α and β be the two eigenvalues
- $\text{Tr}(M) = \alpha + \beta$
- $\text{Det}(M) = \alpha\beta$

- R is positive for corners, negative for edges, and small for flat regions
- Selects corner pixels that are 8-way local maxima

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \qquad \begin{aligned} \det(\mathbf{A}) &= a_{11}a_{22} - a_{12}a_{21} \\ \text{tr}(\mathbf{A}) &= a_{11} + a_{22} \end{aligned}$$

$R = \text{Det}(M) - k\,\text{Tr}(M)^2$ is the Harris Corner Detector.

Now we need a descriptor

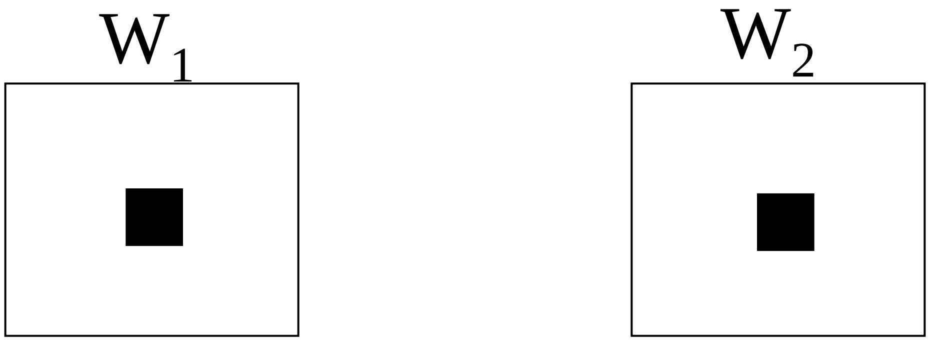


Vector comparison using a distance measure

How do we compare the two regions?

Distance Measures

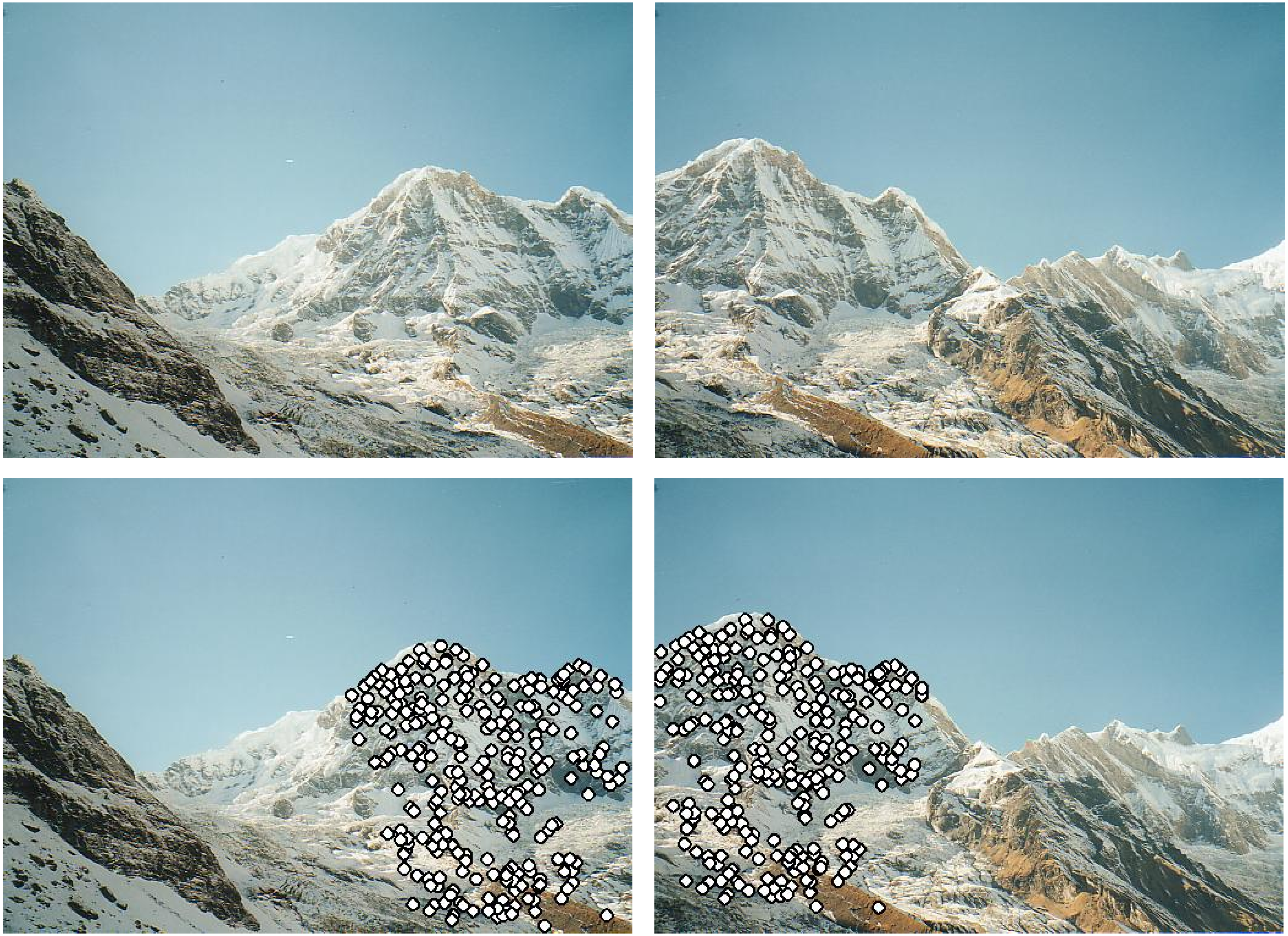
- We can use the sum-square difference of the values of the pixels in a square neighborhood about the points being compared.



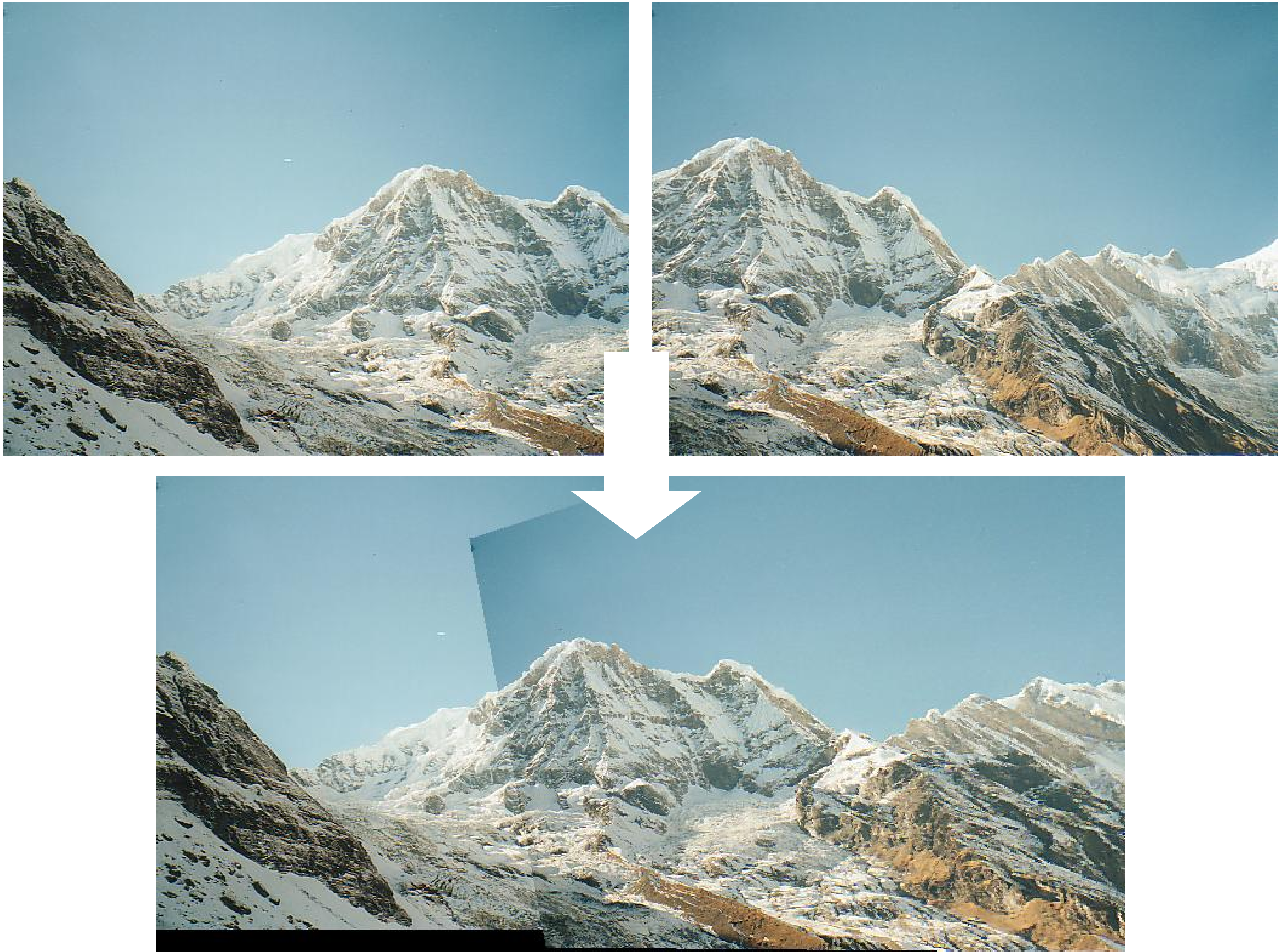
$$SSD = \sum\sum(W_1(i,j) - (W_2(i,j))^2$$

Works when the motion is mainly a translation.

Some Matching Results from Matt Brown



Some Matching Results



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

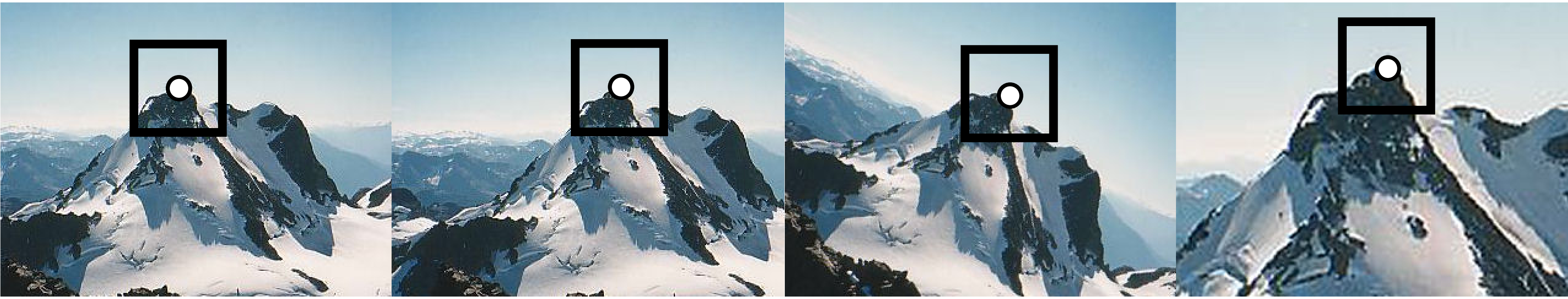
	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	?
Is correlation invariant?	?	?	?

Rotation/Scale Invariance

original translated rotated scaled

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

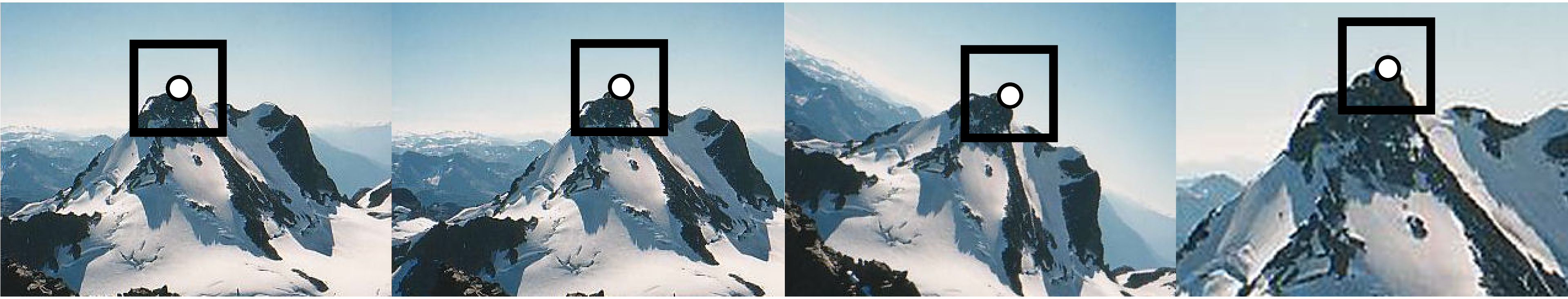
Rotation/Scale Invariance



original translated rotated scaled

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

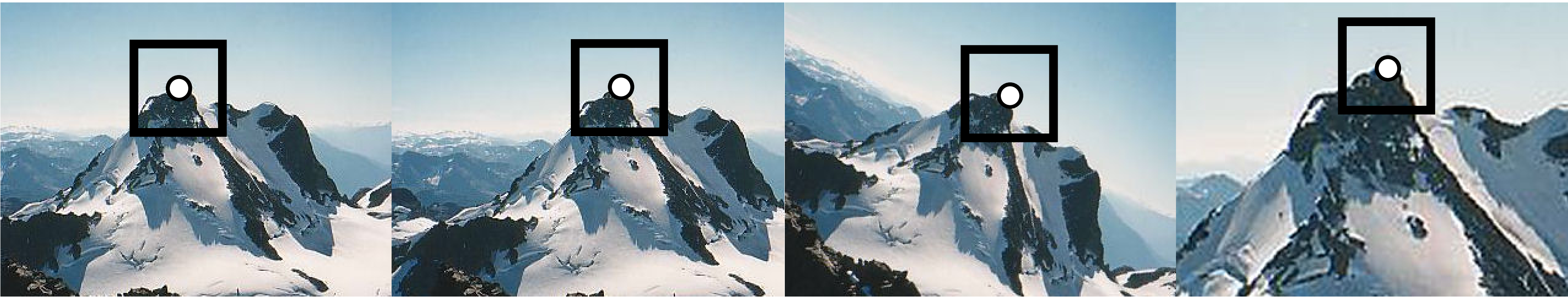
Rotation/Scale Invariance



original translated rotated scaled

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

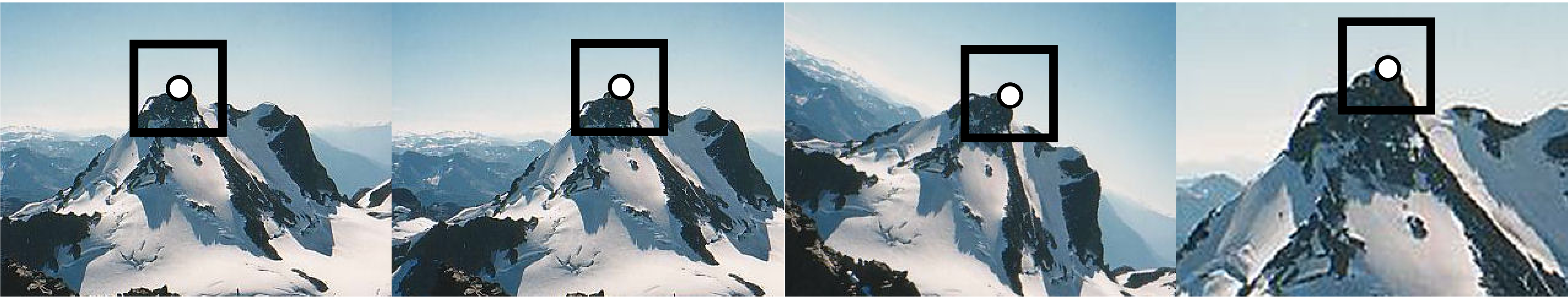
Rotation/Scale Invariance



original translated rotated scaled

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

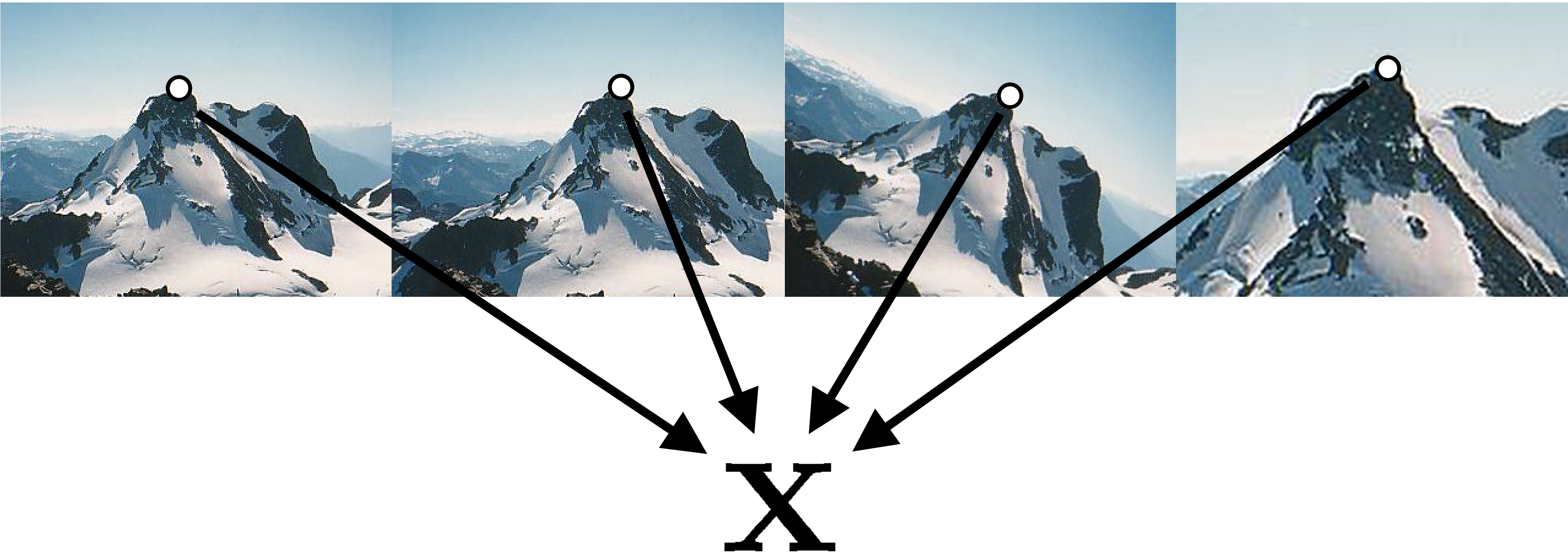
Rotation/Scale Invariance



original translated rotated scaled

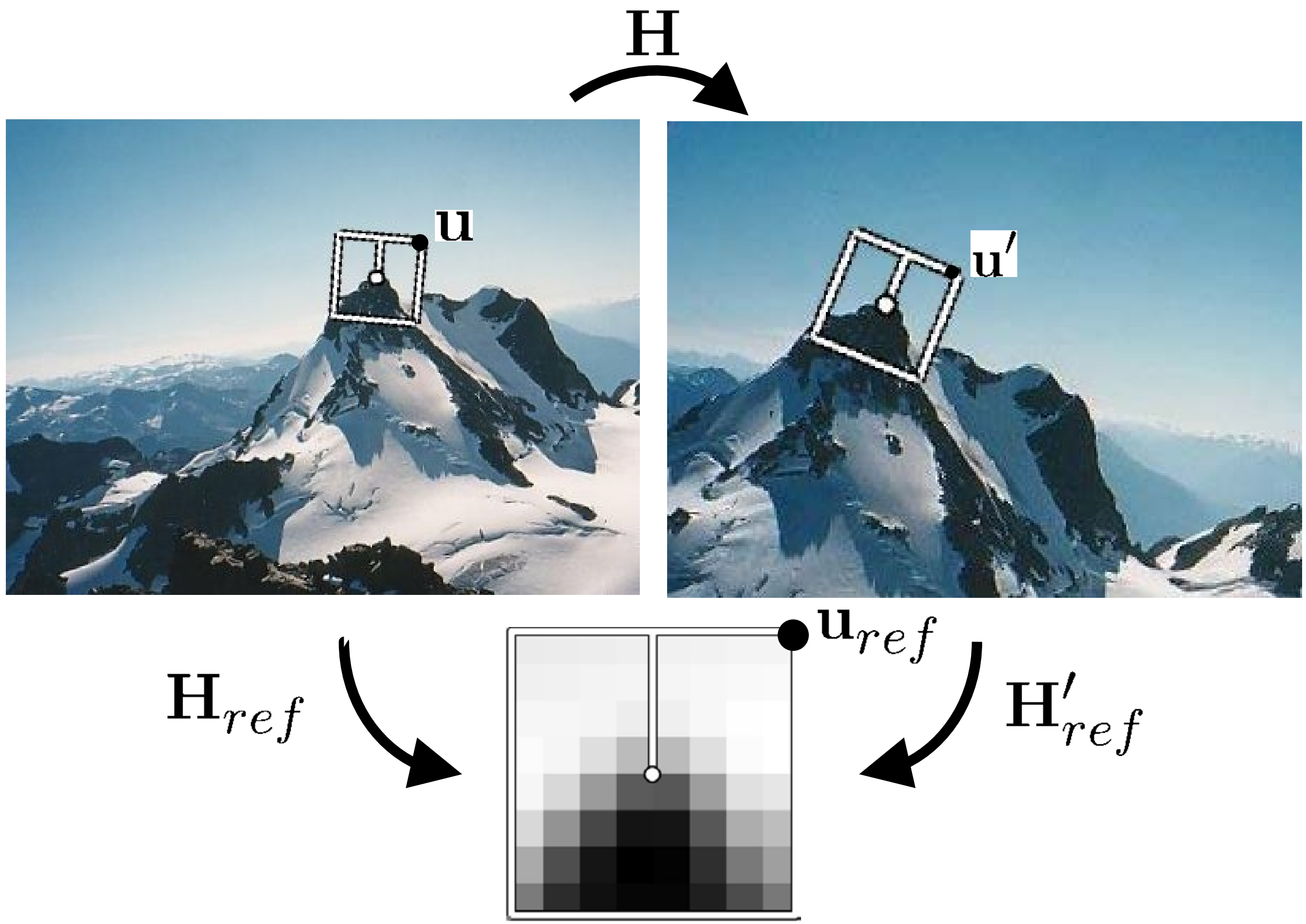
	Translation	Rotation	Scale
Is Harris invariant?	YES	YES	NO
Is correlation invariant?	YES	NO	NO

Matt Brown’s Invariant Features

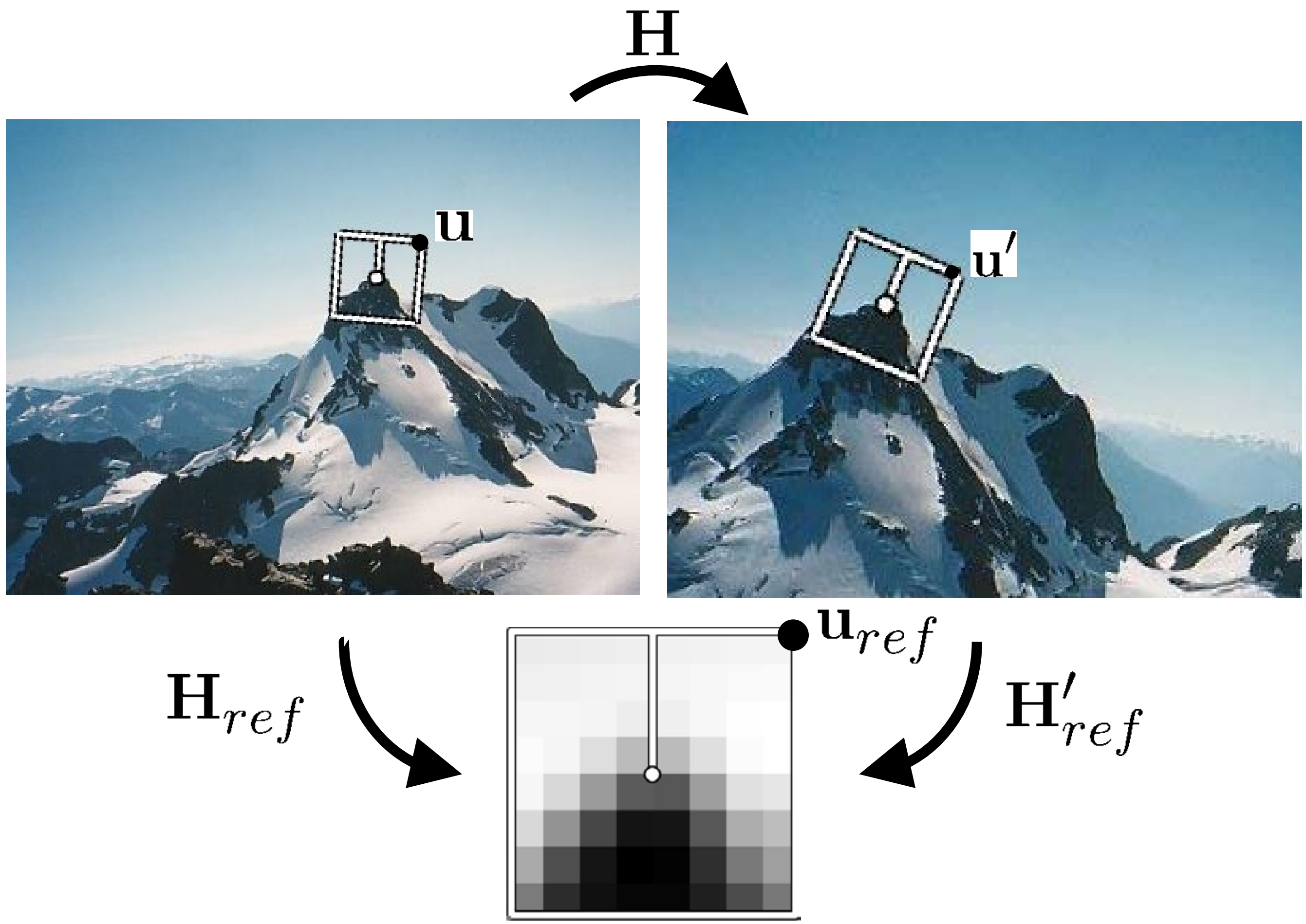


- Local image descriptors that are *invariant* (unchanged) under image transformations

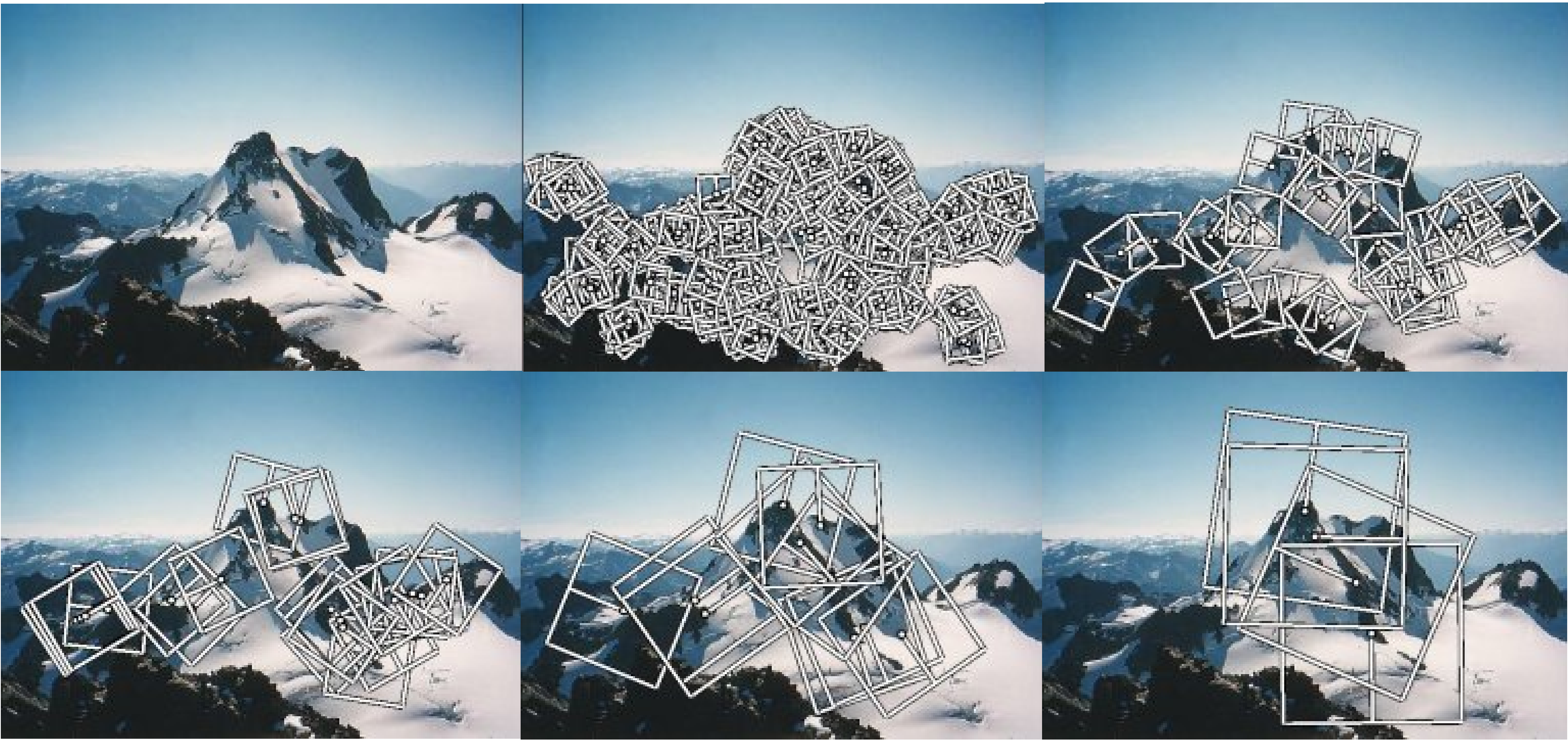
Canonical Frames



Canonical Frames



Multi-Scale Oriented Patches



- Extract oriented patches at multiple scales using dominant orientation

Multi-Scale Oriented Patches

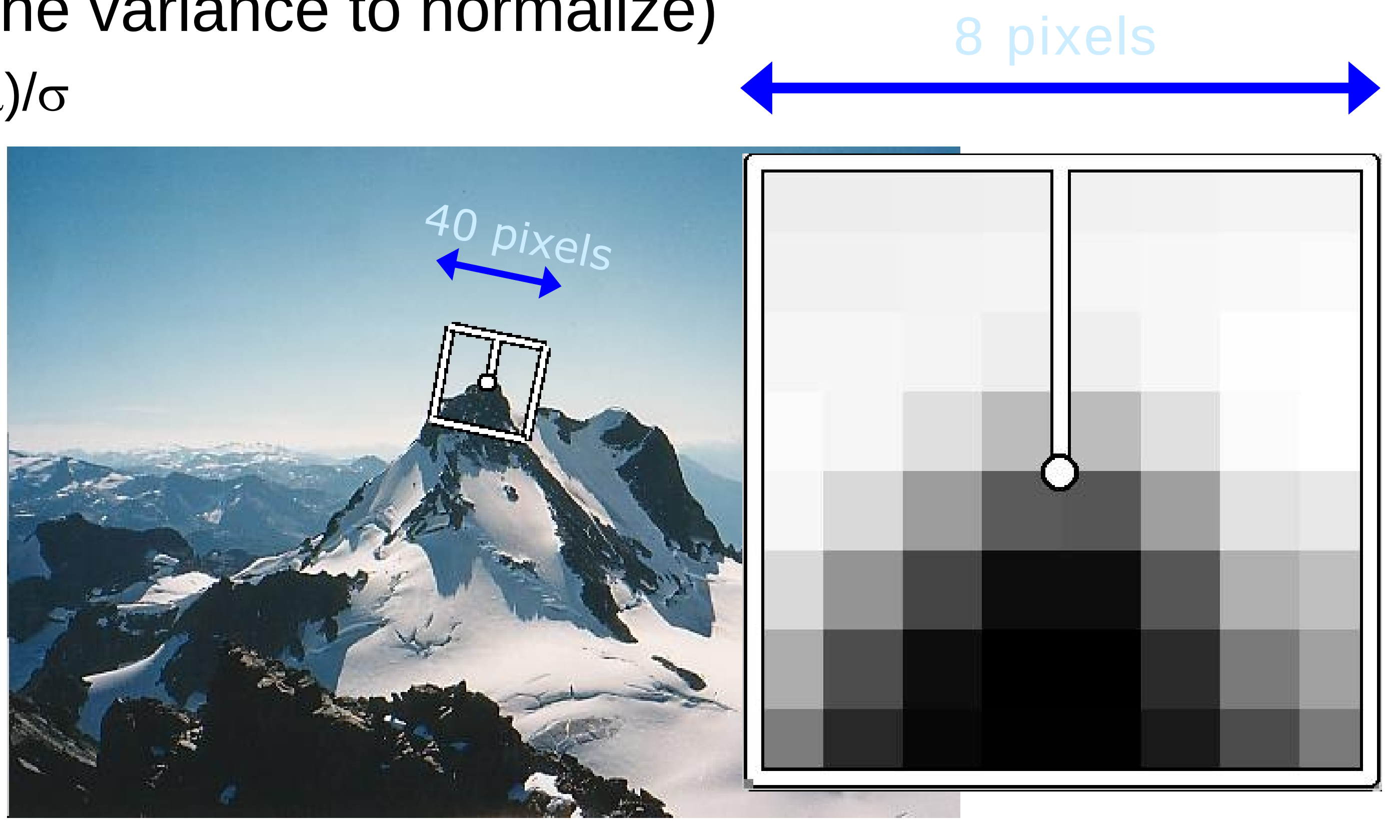
- Sample scaled, oriented patch

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
- Bias/gain normalized (subtract the mean of a patch and divide by the variance to normalize)
  - $I' = (I - \mu)/\sigma$



Matching Interest Points: Summary

- 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