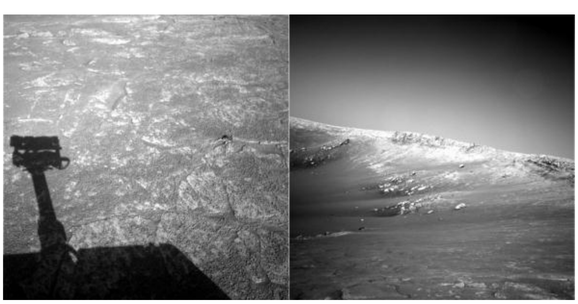
Interest points

CSE P 576
Larry Zitnick (larryz@microsoft.com)
Many slides courtesy of Steve Seitz

How can we find corresponding points?

Not always easy



NASA Mars Rover images

Answer below (look for tiny colored squares…)

NASA Mars Rover images
with SIFT feature matches
Figure by Noah Snavely

Where do humans fixate?

Called “fixation points”, and a “saccade” is the process of moving between fixation points.

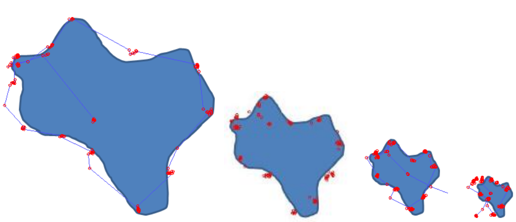
Slowly trace the outline of the above object.

Where do humans fixate?

Called “fixation points”, and a “saccade” is the process of moving between fixation points.

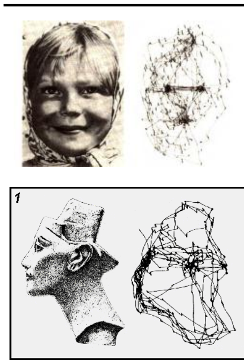
Slowly trace the outline of the above object.

Where do humans fixate?



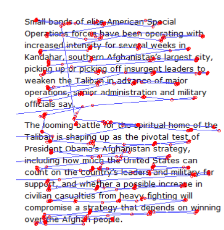
Result of one subject “me”.

Where do humans fixate?



"Eye Movements and Vision" by A. L. Yarbus; Plenum Press, New York; 1967

Top down or bottom up?



Small bands of elite American Special Operations forces have been operating with increased intensity for several weeks in Kandahar, southern Afghanistan’s largest city, picking up or picking off insurgent leaders to weaken the Taliban in advance of major operations, senior administration and military officials say.

The looming battle for the spiritual home of the Taliban is shaping up as the pivotal test of President Obama’s Afghanistan strategy, including how much the United States can count on the country’s leaders and military for support, and whether a possible increase in civilian casualties from heavy fighting will compromise a strategy that depends on winning over the Afghan people.

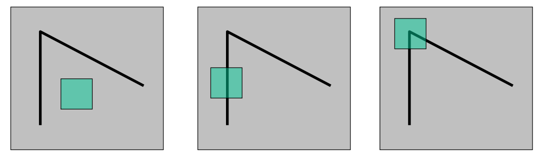
Want uniqueness

Look for image regions that are unusual
- Lead to unambiguous matches in other images

How to define “unusual”?

Local measures of uniqueness

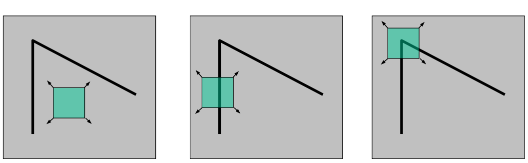
Suppose we only consider a small window of pixels
- What defines whether a feature is a good or bad candidate?



Slide adapted from Darya Frolova, Denis Simakov, Weizmann Institute.

Feature detection

Local measure of feature uniqueness
- How does the window change when you shift by a *small amount*?



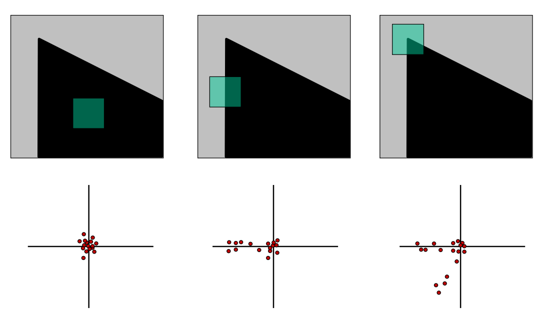
“flat” region:
no change in all directions

“edge”:
no change along the edge direction

“corner”:
significant change in all directions

Slide adapted from Darya Frolova, Denis Simakov, Weizmann Institute.

Let’s look at the gradient distributions



Principle Component Analysis

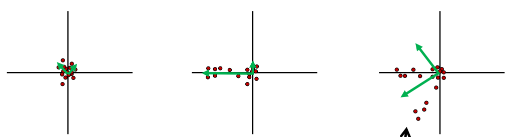
Principal component is the direction of highest variance.

Next, highest component is the direction with highest variance *orthogonal* to the previous components.

How to compute PCA components:

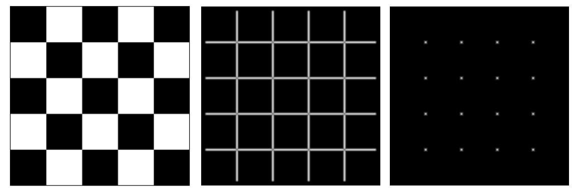
1. ~~Subtract off the mean for each data point.~~
2. Compute the covariance matrix.
3. Compute eigenvectors and eigenvalues.
4. The components are the eigenvectors ranked by the eigenvalues.

$$Hx = \lambda x$$



Both eigenvalues are large!

Simple example



I $\quad\quad$ λ_+ $\quad\quad$ λ_-

Detect peaks and threshold λ_- to find corners.

The math

To compute the eigenvalues:

1. Compute the covariance matrix.

$$H = \sum_{(u,v)} w(u,v) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \qquad I_x = \frac{\partial f}{\partial x}, I_y = \frac{\partial f}{\partial y}$$

Typically Gaussian weights

2. Compute eigenvalues.

$$H = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \lambda_{\pm} = \frac{1}{2}\left((a+d) \pm \sqrt{4bc + (a-d)^2}\right)$$

The Harris operator

λ_- is a variant of the "Harris operator" for feature detection

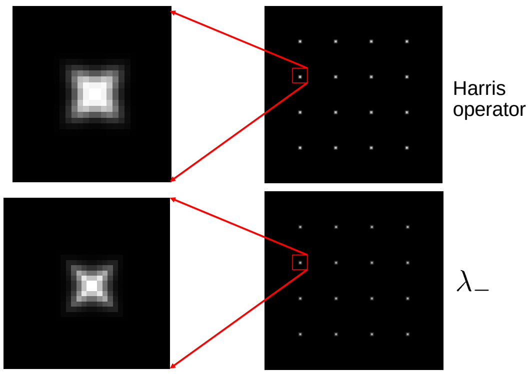
$$f = \frac{\lambda_- \lambda_+}{\lambda_- + \lambda_+} = \frac{determinant(H)}{trace(H)}$$

$$H = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \qquad \begin{aligned} det(H) &= ad - bc \\ tr(H) &= a + d \end{aligned}$$

Actually used in original paper: $f = det(H) - k(tr(H))^2$

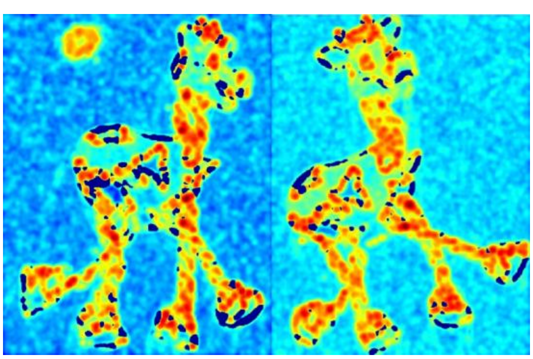
- Very similar to λ_- but less expensive (no square root)
- Called the "Harris Corner Detector" or "Harris Operator"
- Lots of other detectors, this is one of the most popular

The Harris operator

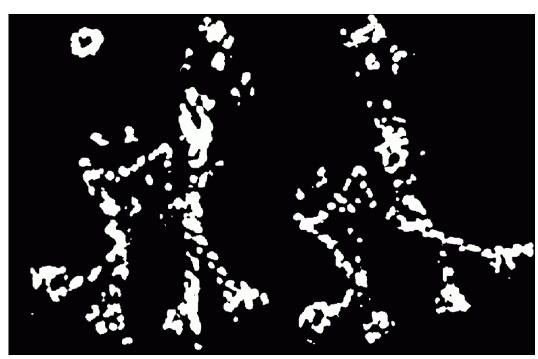


Harris detector example

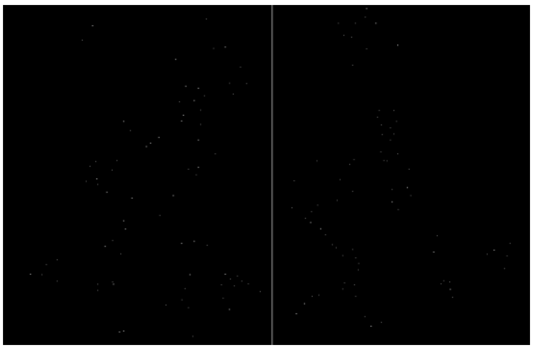
f value (red high, blue low)



Threshold (f > value)



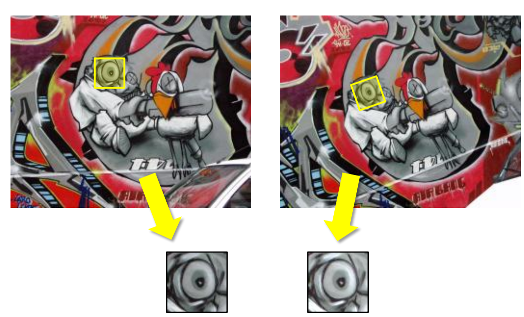
Find local maxima of f



Harris features (in red)

How can we find correspondences?

How can we find correspondences?



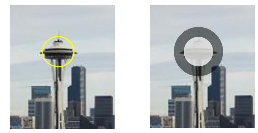
Similarity transform

Scale and rotation?

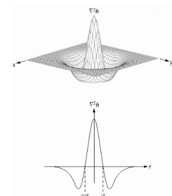
Let's look at scale first:

What is the "best" scale?

Scale



The Laplacian has a maximal response when the difference between the "inside" and "outside" of the filter is greatest.

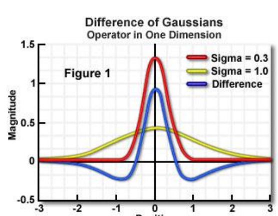


Scale

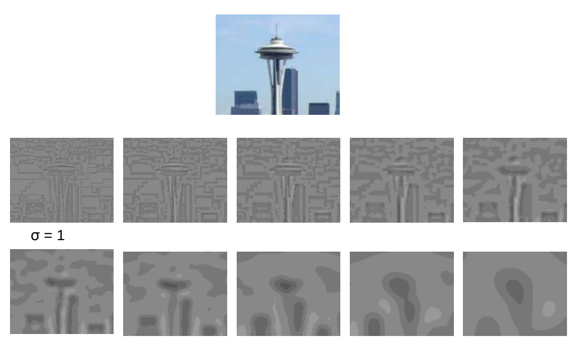
Why Gaussian?

It is invariant to scale change, i.e., $f * \mathcal{G}_\sigma * \mathcal{G}_{\sigma'} = f * \mathcal{G}_{\sigma''}$ and has several other nice properties. Lindeberg, 1994

In practice, the Laplacian is approximated using a Difference of Gaussian (DoG).



DoG example

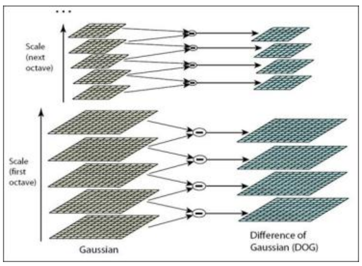


σ = 1

σ = 66

Scale

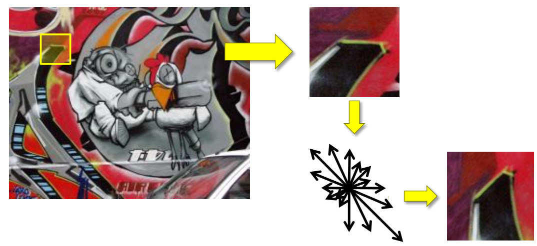
In practice the image is downsampled for larger sigmas.



Lowe, 2004.

Rotation

How to compute the rotation?



Create edge orientation histogram and find peak.

Other interest point detectors

Harris Laplace

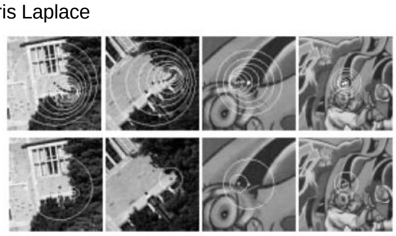
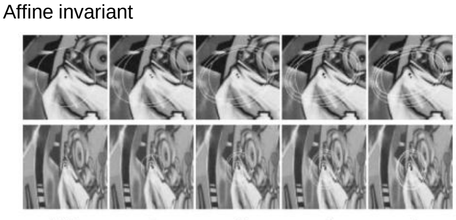


Figure 2. Scale invariant interest point detection: (Top) Initial multi-scale Harris points (selected manually) corresponding to one local structure. (Bottom) Interest points selected with the simplified Harris-Laplace approach.

Scale & Affine Invariant Interest Point Detectors
K. MIKOLAJCZYK and C. SCHMID, IJCV 2004

Other interest point detectors

Affine invariant



Scale & Affine Invariant Interest Point Detectors
K. MIKOLAJCZYK and C. SCHMID, IJCV 2004

Computationally efficient

Approximate Gaussian filters using box filters:

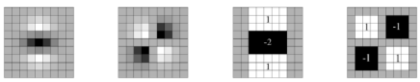


Fig. 1. Left to right: the (discretised and cropped) Gaussian second order partial derivatives in y-direction and xy-direction, and our approximations thereof using box filters. The grey regions are equal to zero.

SURF: Speeded Up Robust Features
Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, ECCV 2006

Computationally efficient

Corner detection by sampling pixels based on decision tree

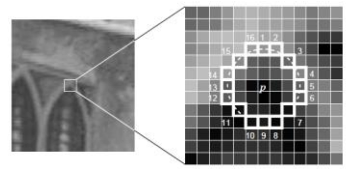


Figure 1. 12 point segment test corner detection in an image patch. The highlighted squares are the pixels used in the corner detection. The pixel at p is the centre of a candidate corner. The arc is indicated by the dashed line passes through 12 contiguous pixels which are brighter than p by more than the threshold.

Machine learning for high-speed corner detection
Edward Rosten and Tom Drummond, ECCV 2006

How well do they work in practice?

Let’s go to the videos…