

---

# Image Features

CSE 576, Spring 2005

# About me

---

- Ph. D., Carnegie Mellon, 1988
- Researcher, Cambridge Research Lab at DEC, 1990-1995
- Senior Researcher, Interactive Visual Media Group, Microsoft, 1995-
- Research interests:
  - computer vision (stereo, motion),  
computer graphics (image-based rendering),  
data-parallel programming



# Today's lecture

---

- What is computer vision?
  - Scale-space and pyramids
  - What are good features?
  - Feature detection
  - Feature descriptors
  - (Next lecture: feature matching)
- 
- Project 1 description and demo [Ian Simon]

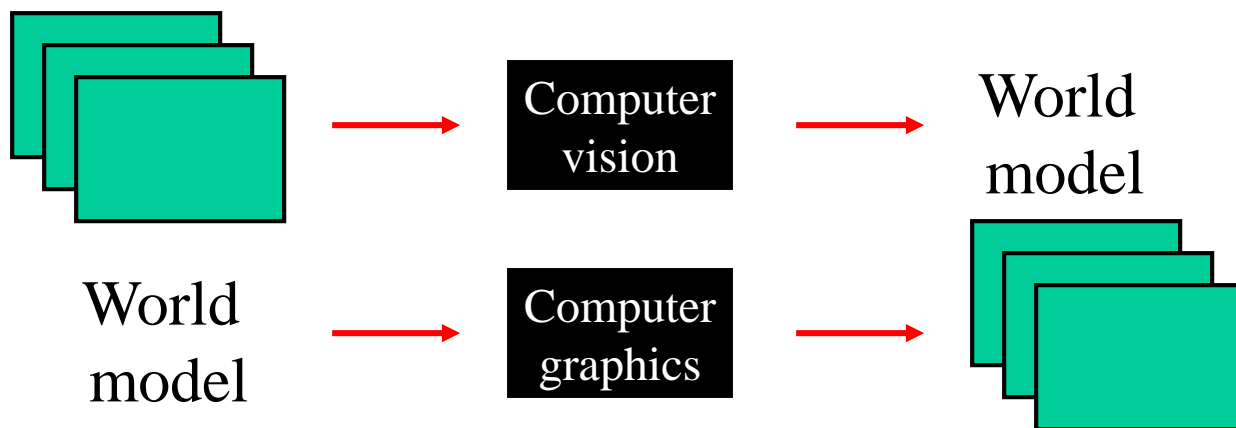
---

# What is Computer Vision?

# What is Computer Vision?

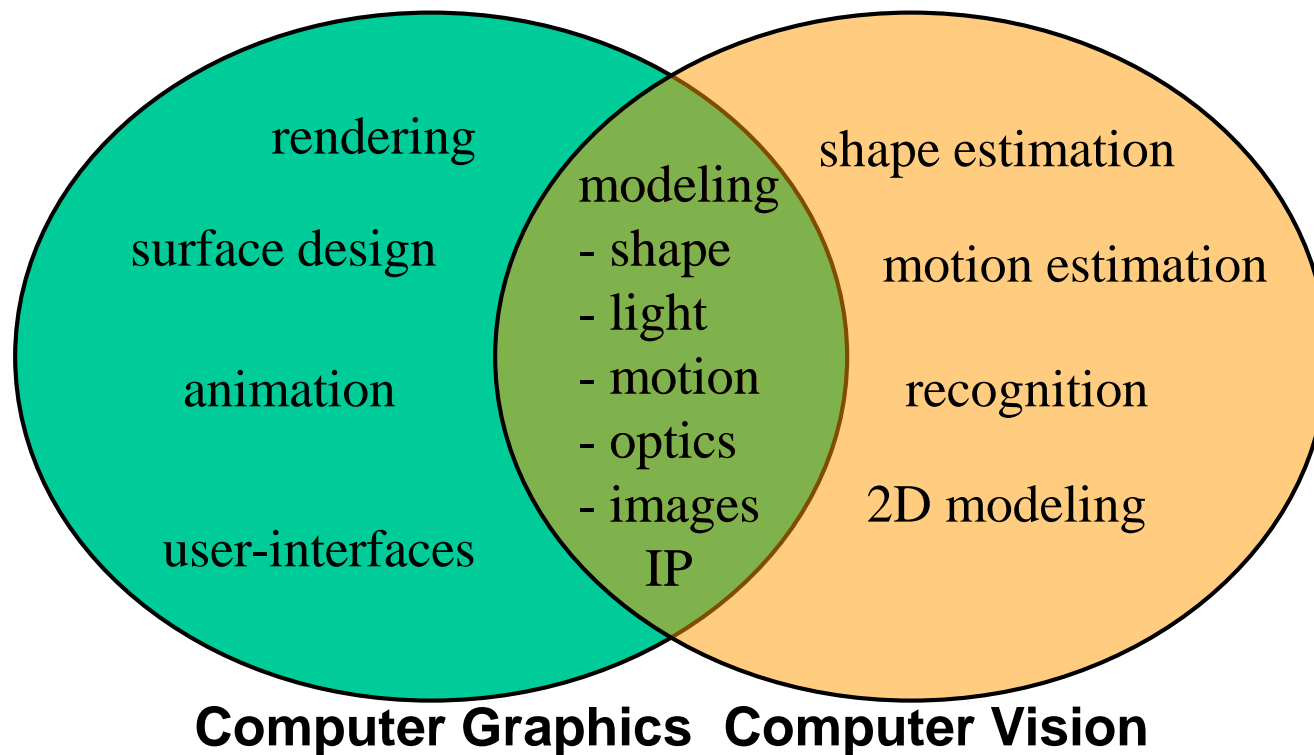
---

- Image Understanding (AI, behavior)
- A sensor modality for robotics
- Computer emulation of human vision
- Inverse of Computer Graphics



# Intersection of Vision and Graphics

---



# Computer Vision [Trucco&Verri'98]

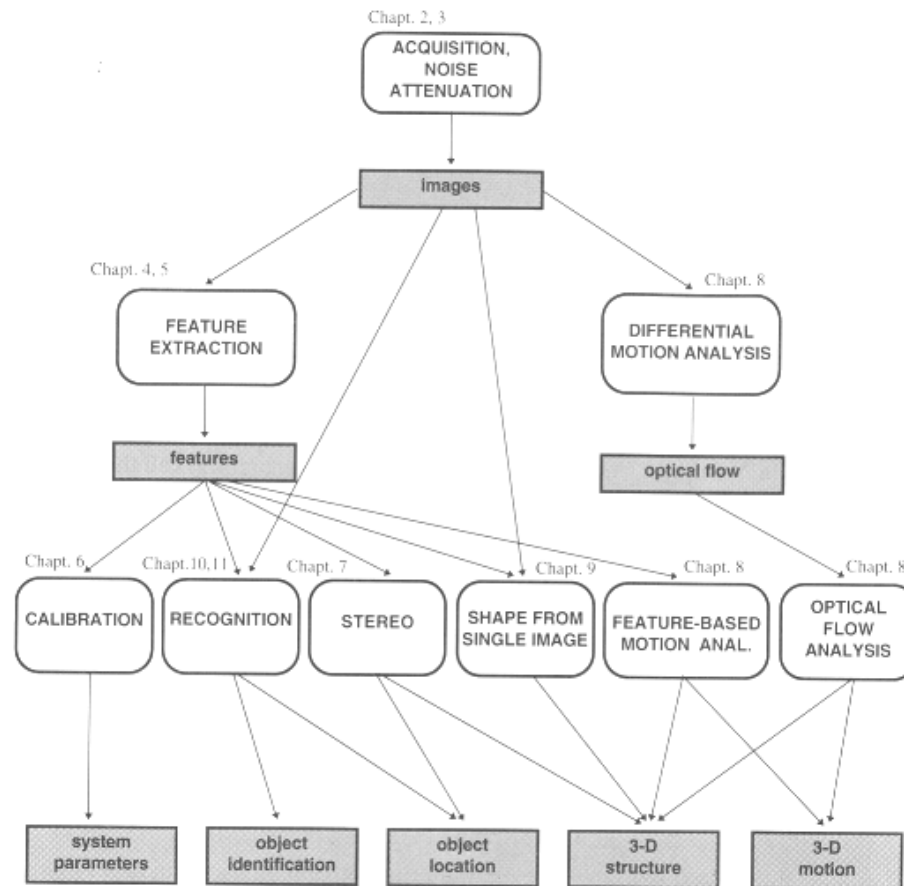
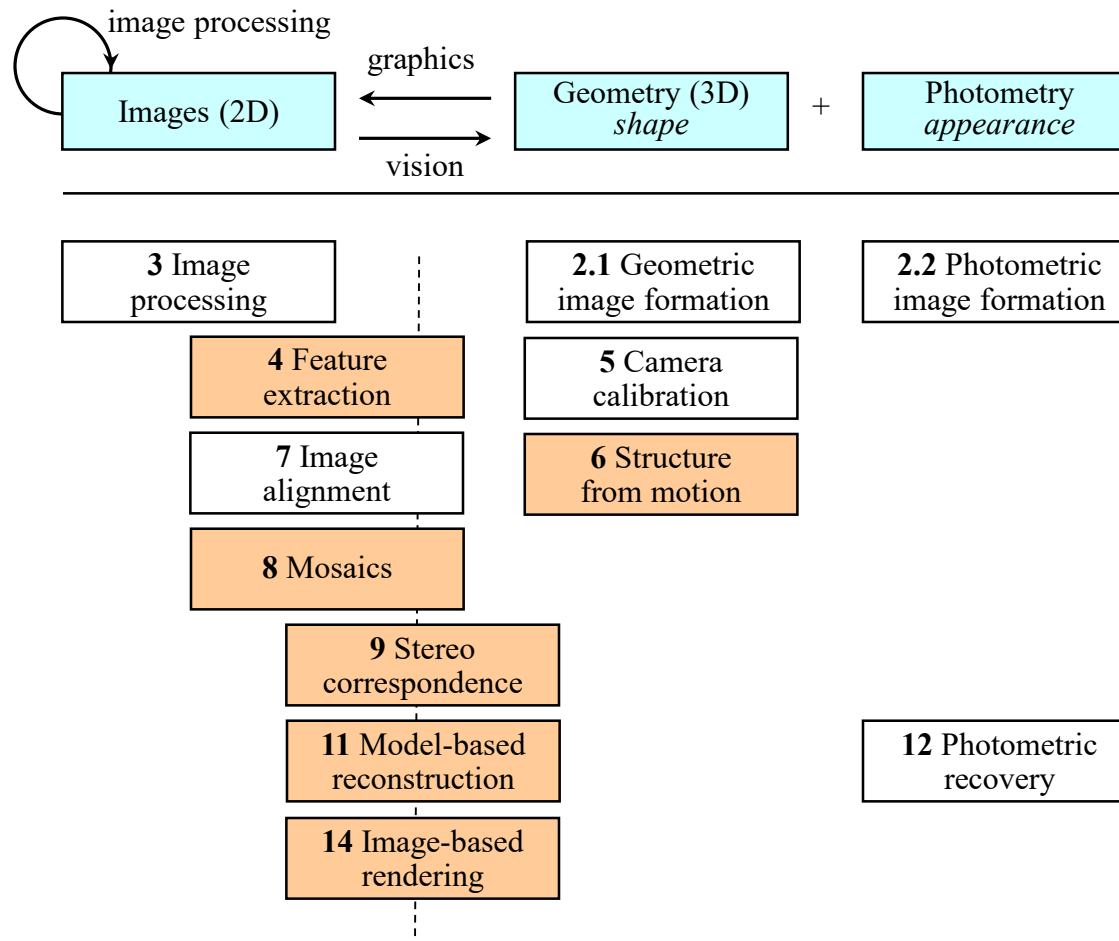


Figure 1.7 The book at a glance: method classes (white boxes), results (grey boxes), their interdependence, and where to find the various topics in this book.

# Image-Based Modeling





# Applications

---

- Geometric reconstruction: modeling, forensics, special effects (ILM, RealVis, 2D3)
- Image and video editing (Avid, Adobe)
- Webcasting and Indexing Digital Video (Virage)
- Scientific / medical applications (GE)

# Applications

---

- Tracking and surveillance (Sarnoff)
- Fingerprint recognition (Digital Persona)
- Biometrics / iris scans (Iridian Technologies)
- Vehicle safety (MobilEye)
- Drowning people (VisionIQ Inc)
- Optical motion capture (Vicon)

# Image Morphing

---



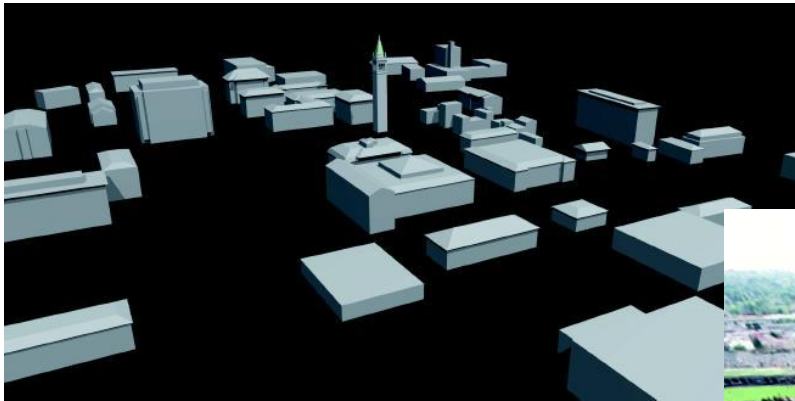
# Panoramic Mosaics

---



# 3D Shape Reconstruction

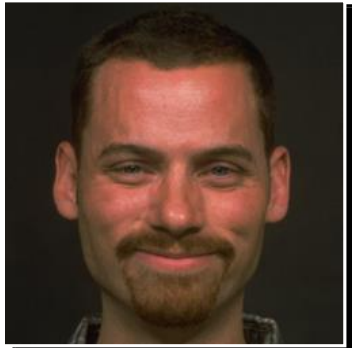
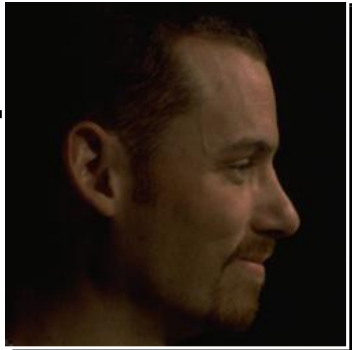
---



Debevec, Taylor, and Malik, SIGGRAPH 1996

# Face Modeling

---

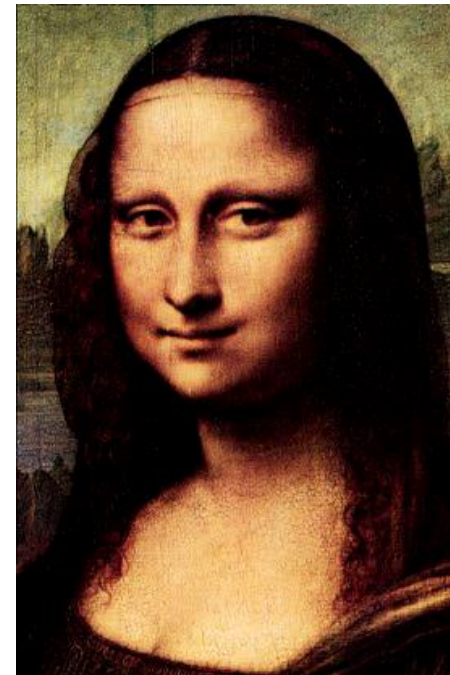
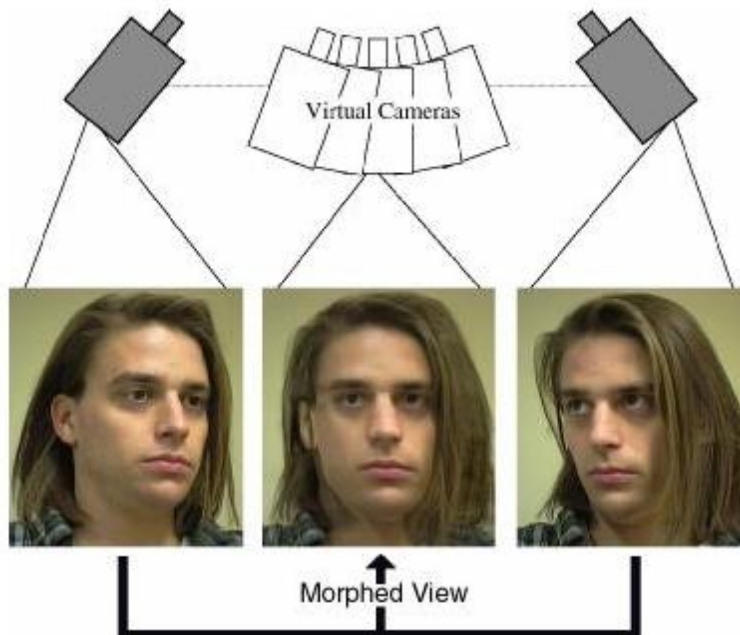




# View Morphing

---

Morph between pair of images using epipolar geometry [Seitz & Dyer, SIGGRAPH'96]



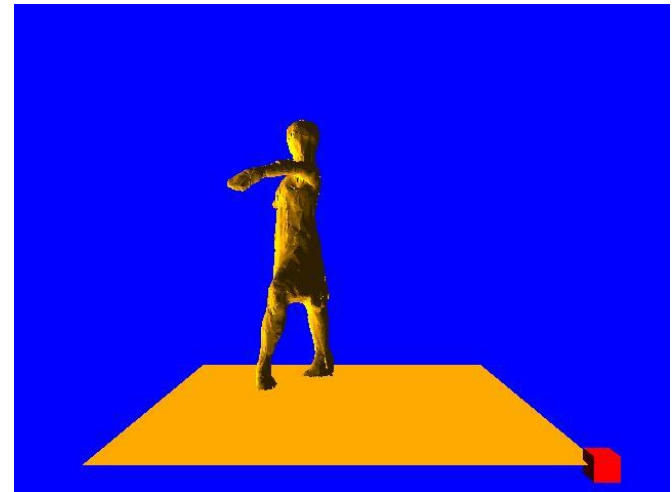
# Virtualized Reality™

---

Takeo Kanade, CMU

- collect video from 50+ stream

reconstruct 3D model sequences



<http://www.cs.cmu.edu/afs/cs/project/VirtualizedR/www/VirtualizedR.html>

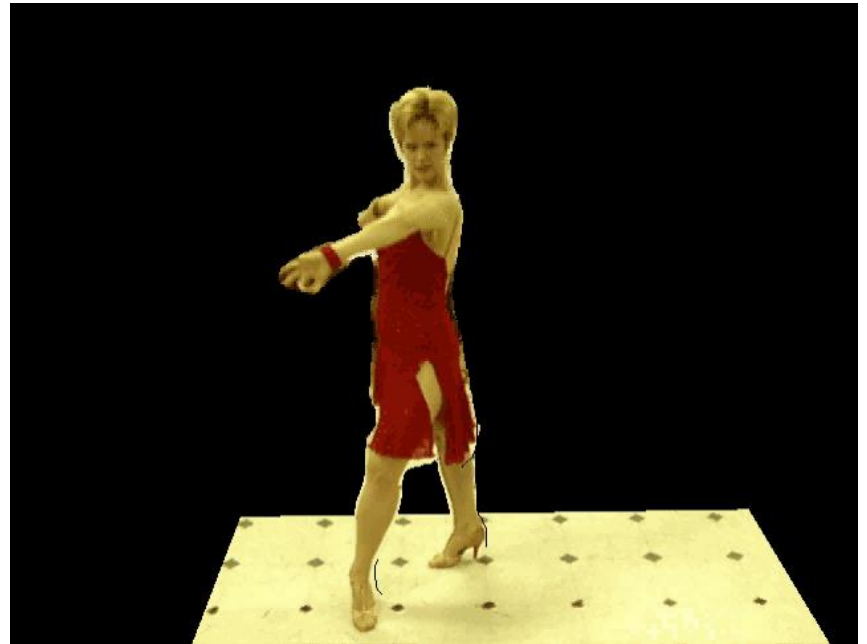


# Virtualized Reality™

---

Takeo Kanade, CMU

- generate new video



- steerable version used for SuperBowl XXV  
“eye vision” system

# Edge detection and editing

---



Elder, J. H. and R. M. Goldberg. "Image Editing in the Contour Domain,"  
Proc. IEEE: Computer Vision and Pattern Recognition, pp. 374-381, June, 1998.

# Image Enhancement

---

High dynamic range photography

[Debevec *et al.*'97; Mitsunaga & Nayar'99]

- combine several different exposures together

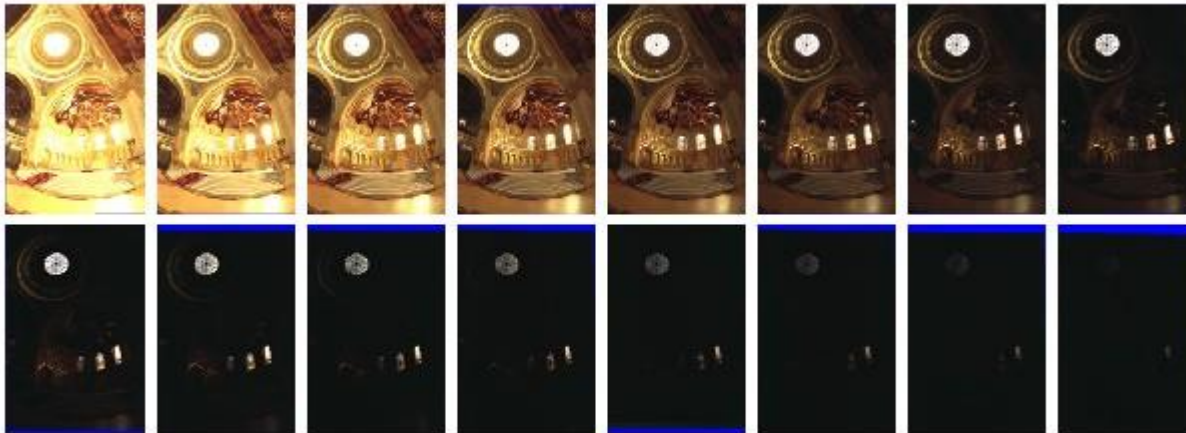


Figure 6: Sixteen photographs of a church taken at 1-stop increments from 30 sec to  $\infty$  sec. The sun is directly behind the rightmost stained glass window, making it especially bright. The blue borders seen in some of the image margins are induced by the image registration process.



# Today's lecture

---

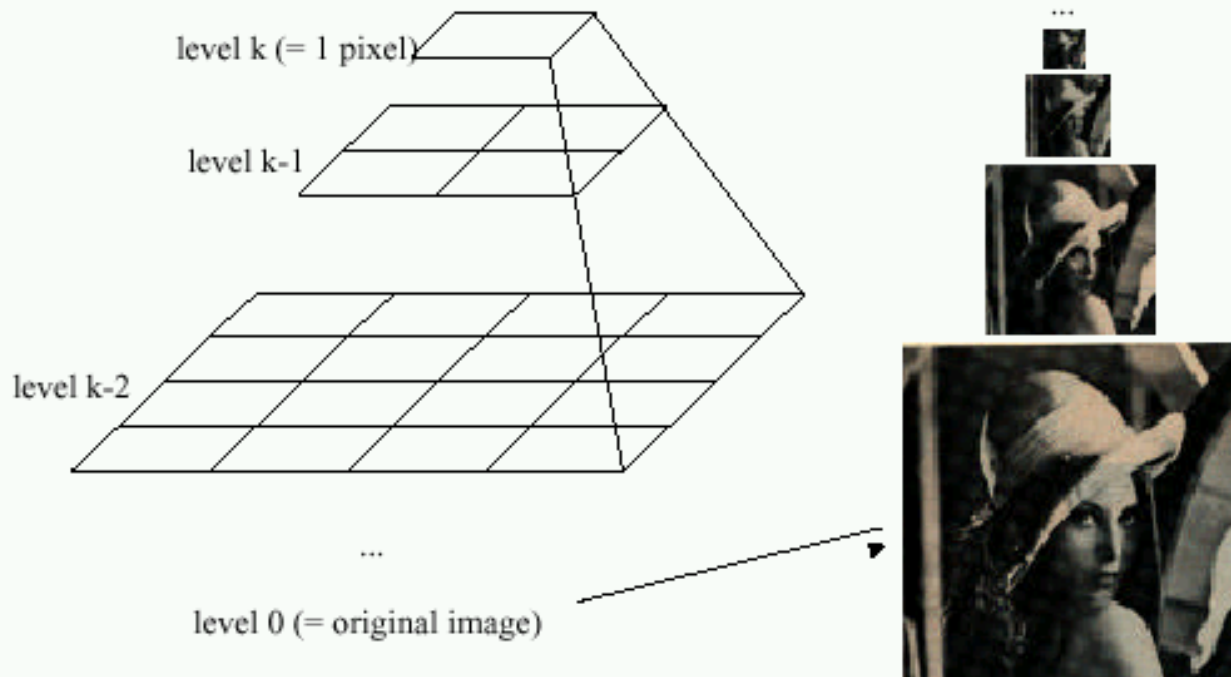
- What is computer vision?
  - **Scale-space and pyramids**
  - What are good features?
  - Feature detection
  - Feature descriptors
  - (Next lecture: feature matching)
- 
- Project 1 description and demo [Ian Simon]

---

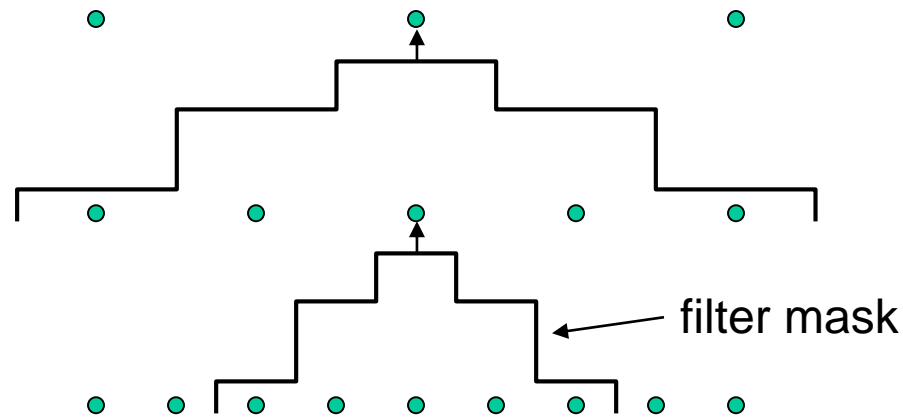
# Image Pyramids

# Image Pyramids

Idea: Represent  $N \times N$  image as a “pyramid” of  $1 \times 1, 2 \times 2, 4 \times 4, \dots, 2^k \times 2^k$  images (assuming  $N = 2^k$ )



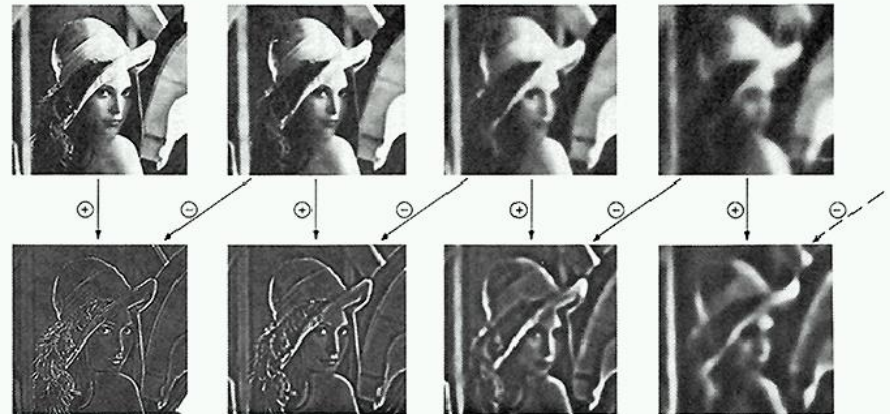
# Pyramid Creation



“Gaussian” Pyramid

“Laplacian” Pyramid

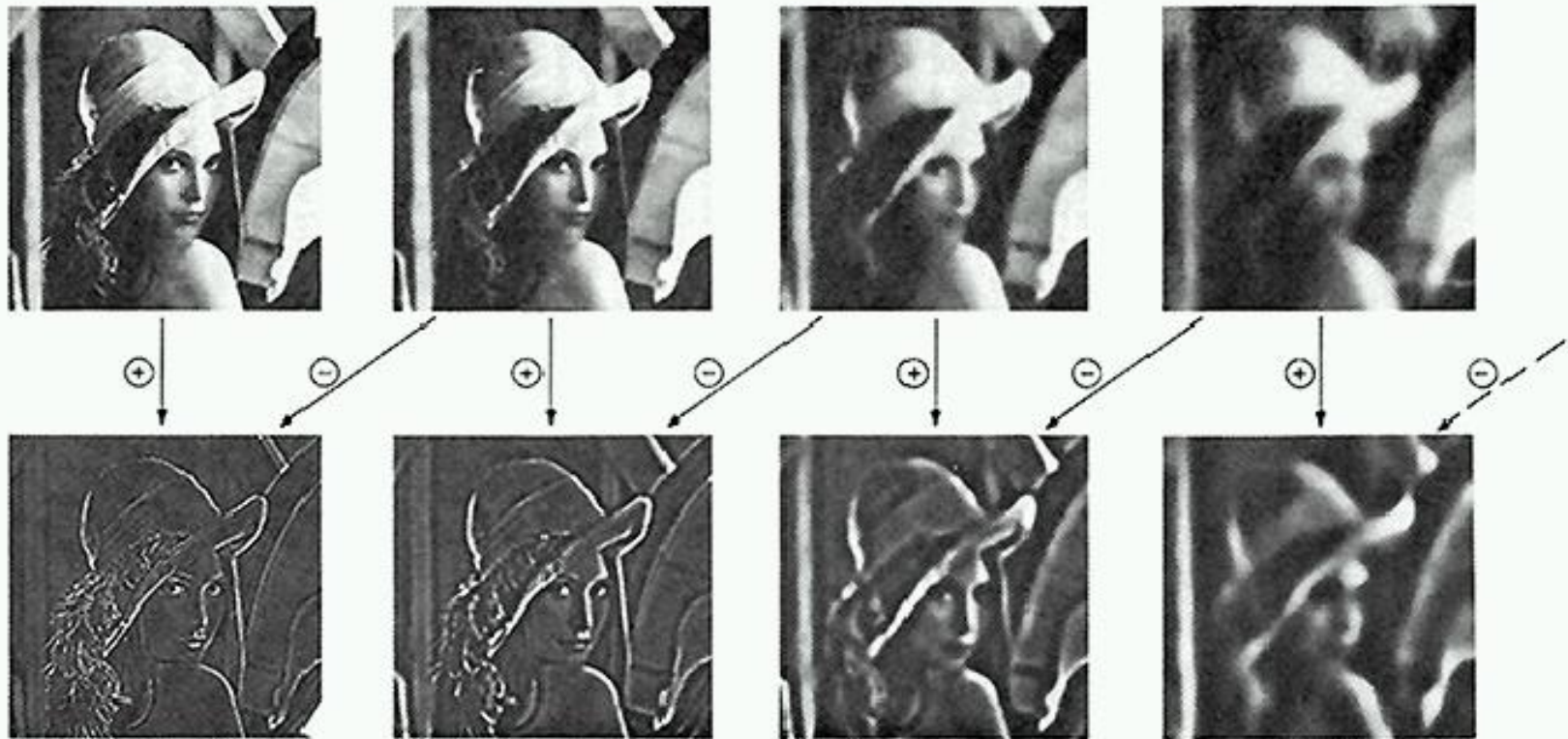
- Created from Gaussian pyramid by subtraction  
 $L_l = G_l - \text{expand}(G_{l+1})$





# Octaves in the Spatial Domain

## Lowpass Images



## Bandpass Images



# Pyramids

---

## Advantages of pyramids

- Faster than Fourier transform
- Avoids “ringing” artifacts

## Many applications

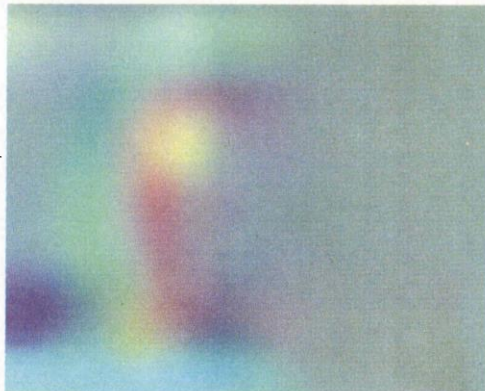
- small images faster to process
- good for multiresolution processing
- compression
- progressive transmission

Known as “MIP-maps” in graphics community

## Precursor to wavelets

- Wavelets also have these advantages

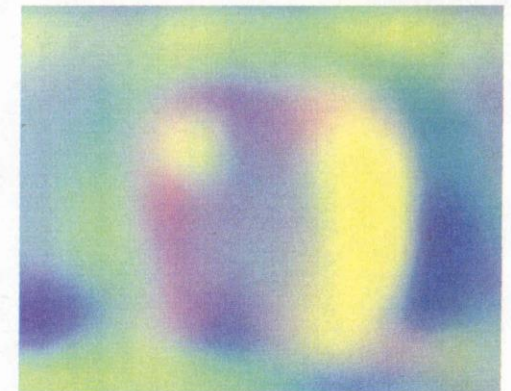
Laplacian  
level  
4



(c)

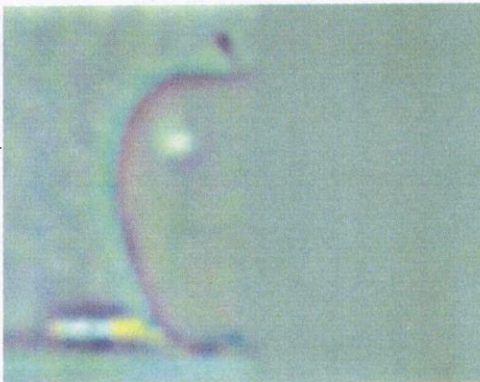


(g)

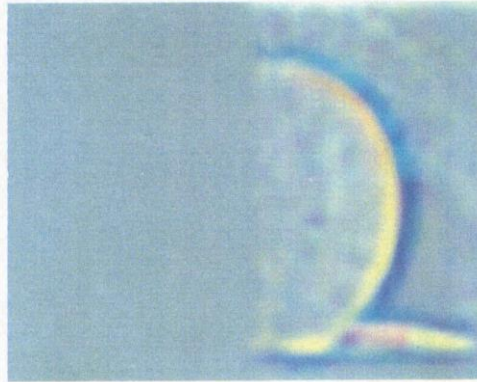


(k)

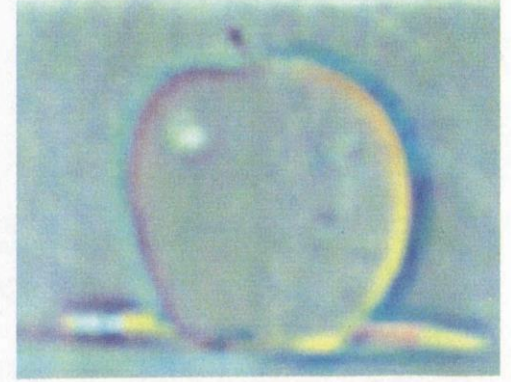
Laplacian  
level  
2



(b)

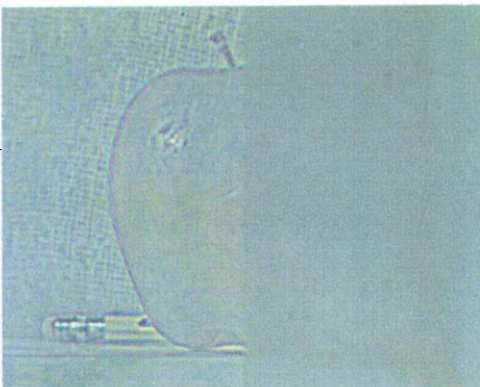


(f)

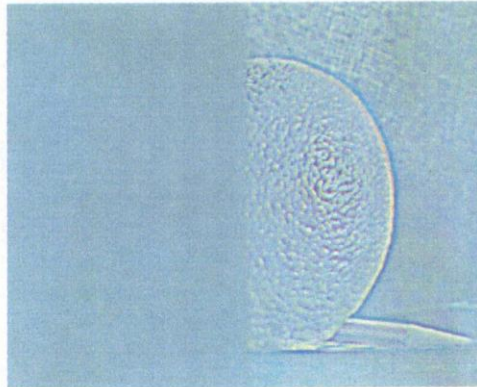


(j)

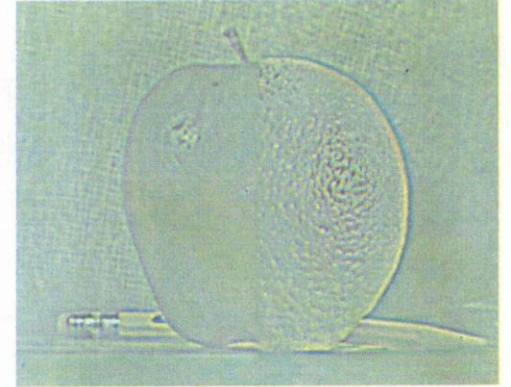
Laplacian  
level  
0



(a)



(e)



(i)

3/31/2000

left pyramid

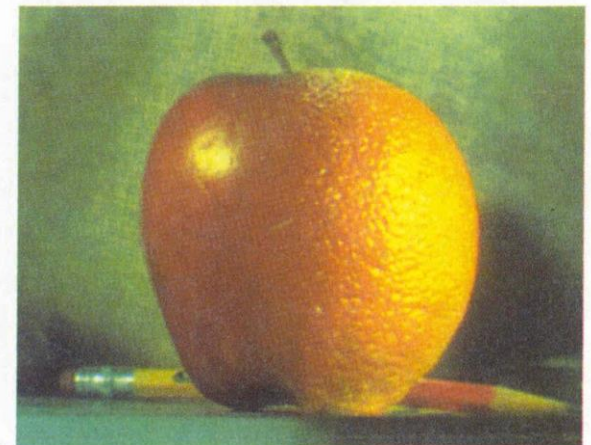
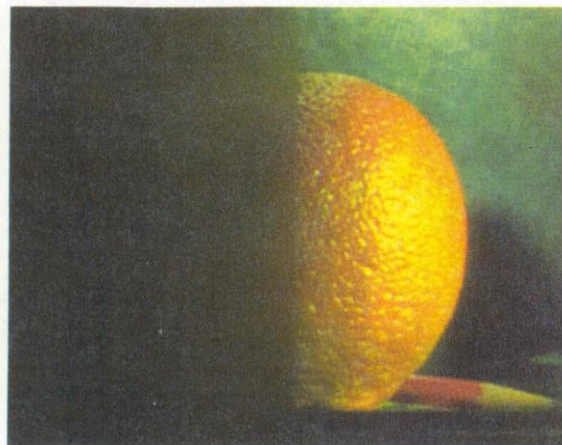
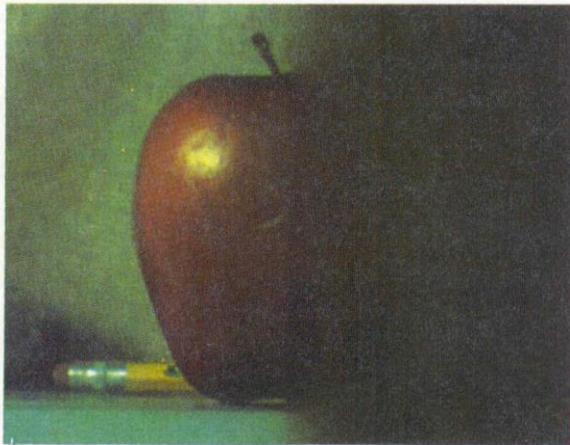
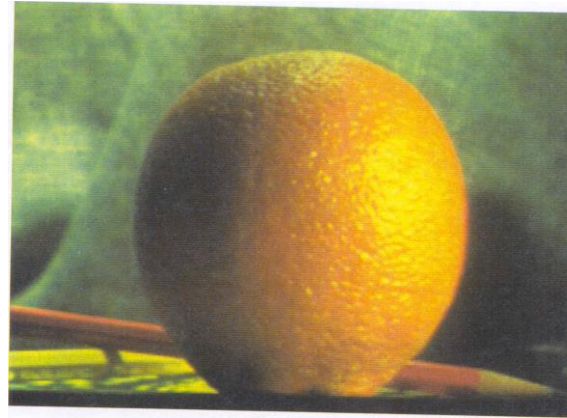
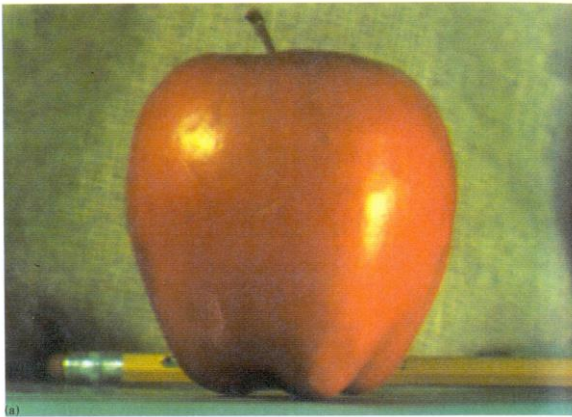
CSE 570: COMPUTER VISION  
right pyramid

blended pyramid



# Pyramid Blending

---





**original**



**smoothed (5x5 Gaussian)**

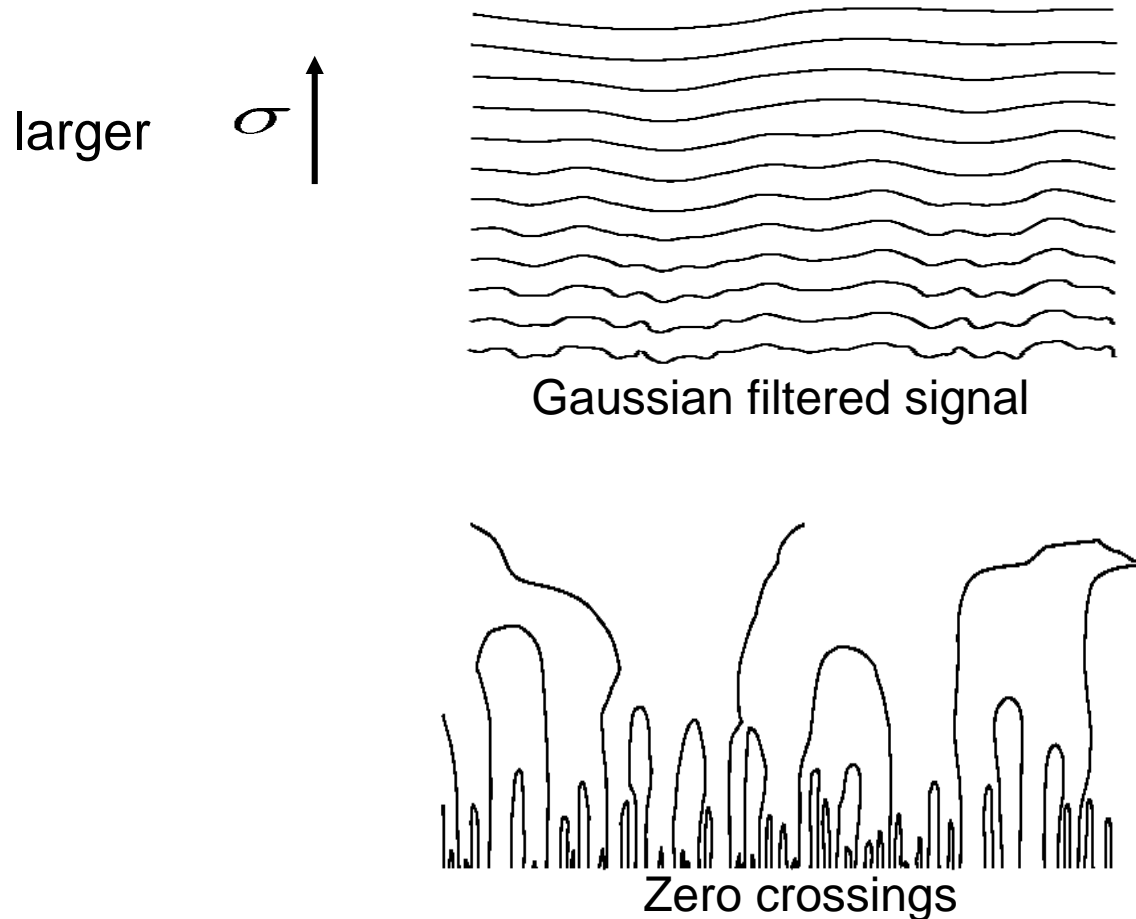


**smoothed – original**  
(scaled by 4, offset +128)

Why does  
this work?

# Scale space (Witkin 83)

---

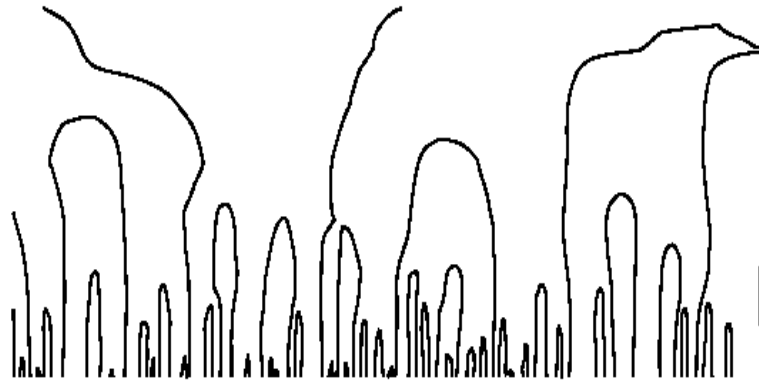


# Scale space: insights

---

As the scale is increased

- edge position can change
- edges can disappear
- new edges are not created



# Today's lecture

---

- What is computer vision?
  - Scale-space and pyramids
  - What are good features?
  - Feature detection
  - Feature descriptors
  - (Next lecture: feature matching)
- 
- Project 1 description and demo [Ian Simon]

---

These slides adapted from:

# Matching with Invariant Features

Darya Frolova, Denis Simakov  
The Weizmann Institute of Science  
March 2004



---

and...

# Real-time Object Recognition using Invariant Local Image Features

David Lowe

Computer Science Department

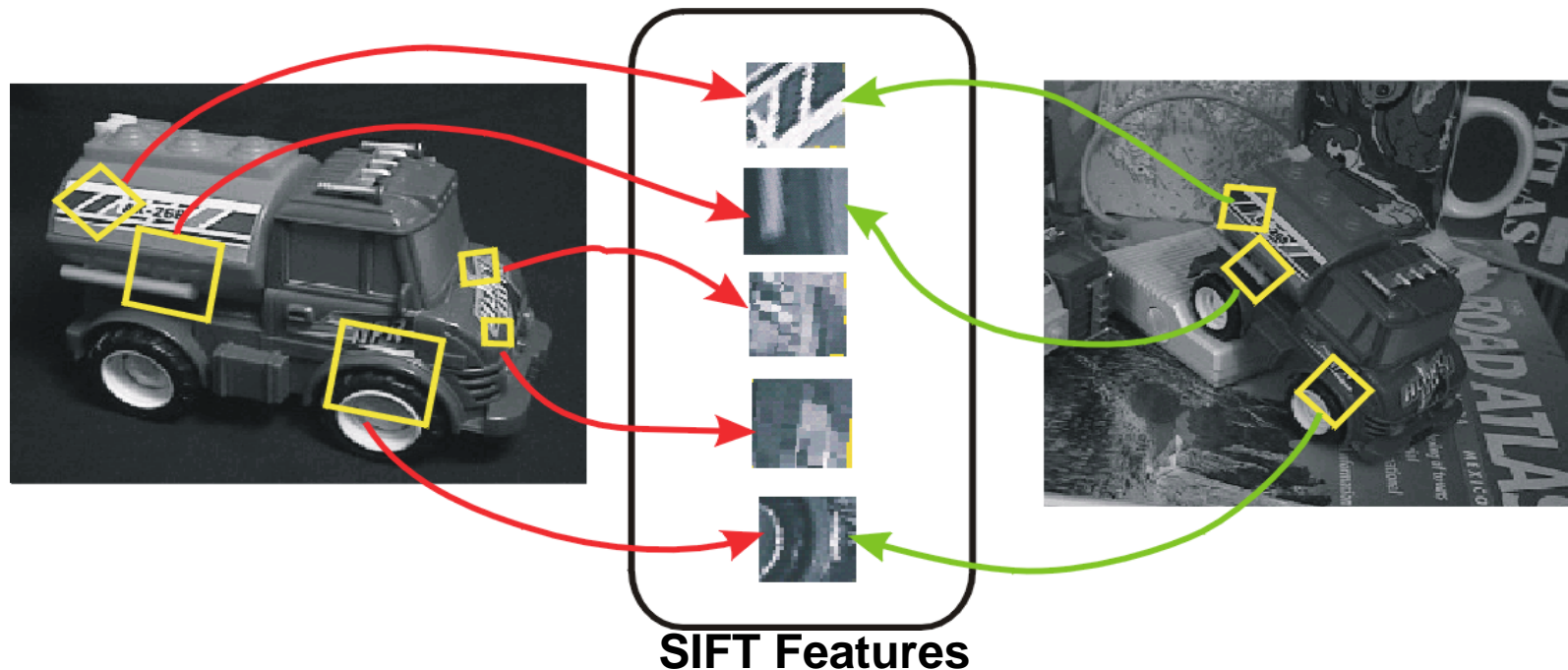
University of British Columbia

NIPS 2003 Tutorial

# Invariant Local Features

---

Image content is transformed into local feature coordinates that are invariant to translation, rotation, scale, and other imaging parameters



# Advantages of local features

---

- Locality:** features are local, so robust to occlusion and clutter (no prior segmentation)
- Distinctiveness:** individual features can be matched to a large database of objects
- Quantity:** many features can be generated for even small objects
- Efficiency:** close to real-time performance
- Extensibility:** can easily be extended to wide range of differing feature types, with each adding robustness

# More motivation...

---

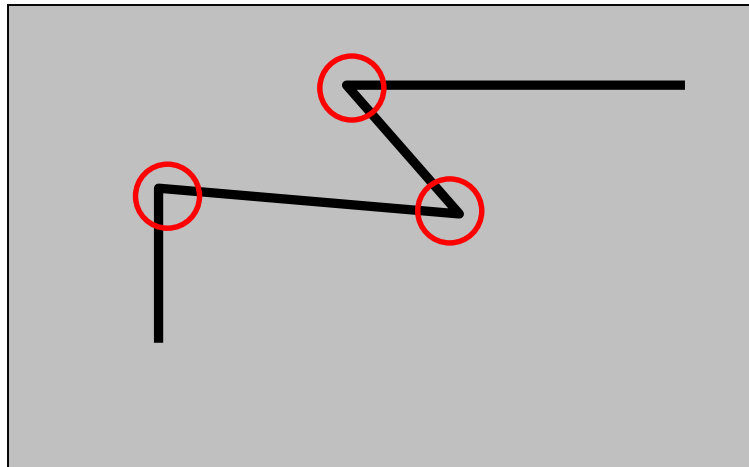
Feature points are used also for:

- Image alignment (homography, fundamental matrix)
- 3D reconstruction
- Motion tracking
- Object recognition
- Indexing and database retrieval
- Robot navigation
- ... other

# Harris corner detector

---

C.Harris, M.Stephens. “A Combined Corner and Edge Detector”. 1988

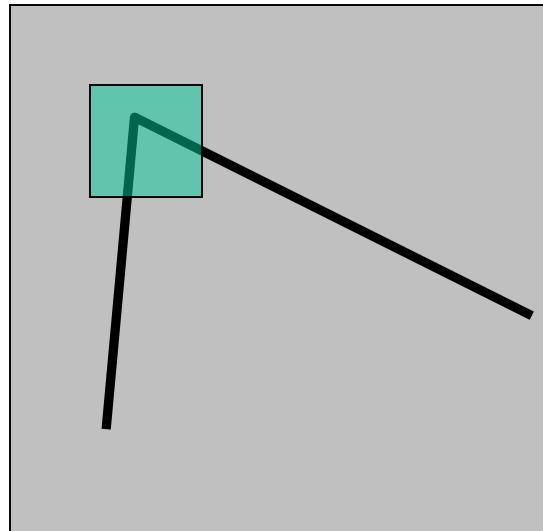


# The Basic Idea

---

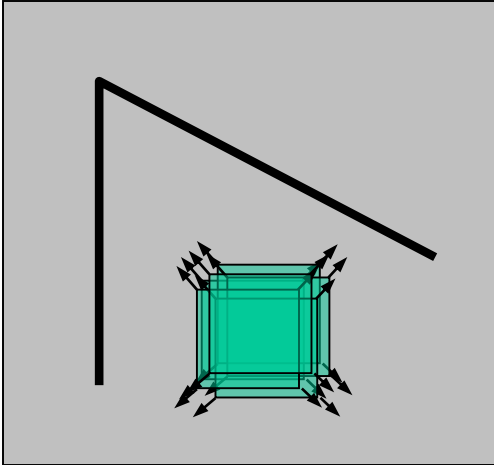
We should easily recognize the point by looking through a small window

Shifting a window in *any direction* should give a *large change* in intensity

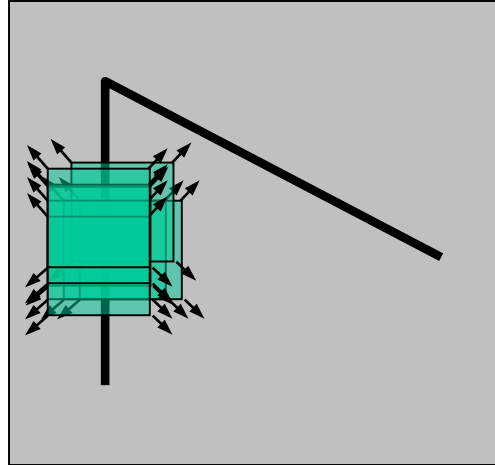


# Harris Detector: Basic Idea

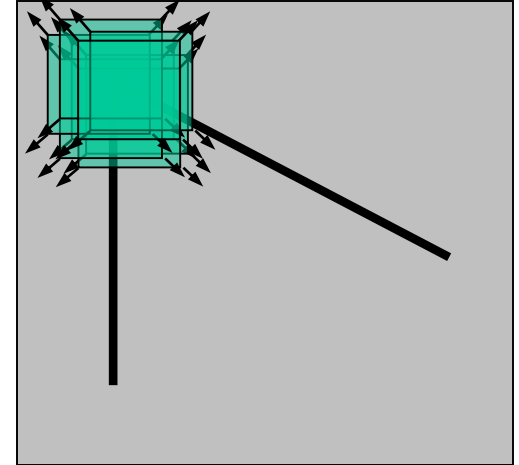
---



“flat” region:  
no change in  
all directions



“edge”:  
no change along  
the edge direction



“corner”:  
significant change  
in all directions

# Harris Detector: Mathematics

---

Change of intensity for the shift  $[u, v]$ :

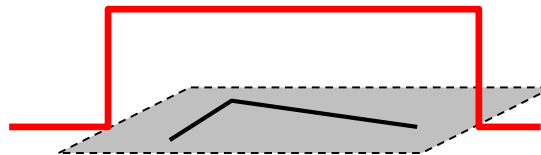
$$E(u, v) = \sum_{x, y} w(x, y) [I(x + u, y + v) - I(x, y)]^2$$

Window  
function

Shifted  
intensity

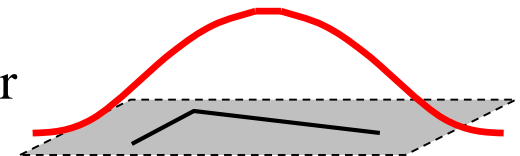
Intensity

Window function  $w(x, y) =$



1 in window, 0 outside

or



Gaussian



# Harris Detector: Mathematics

---

For small shifts  $[u, v]$  we have a *bilinear* approximation:

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

where  $M$  is a  $2 \times 2$  matrix computed from image derivatives:

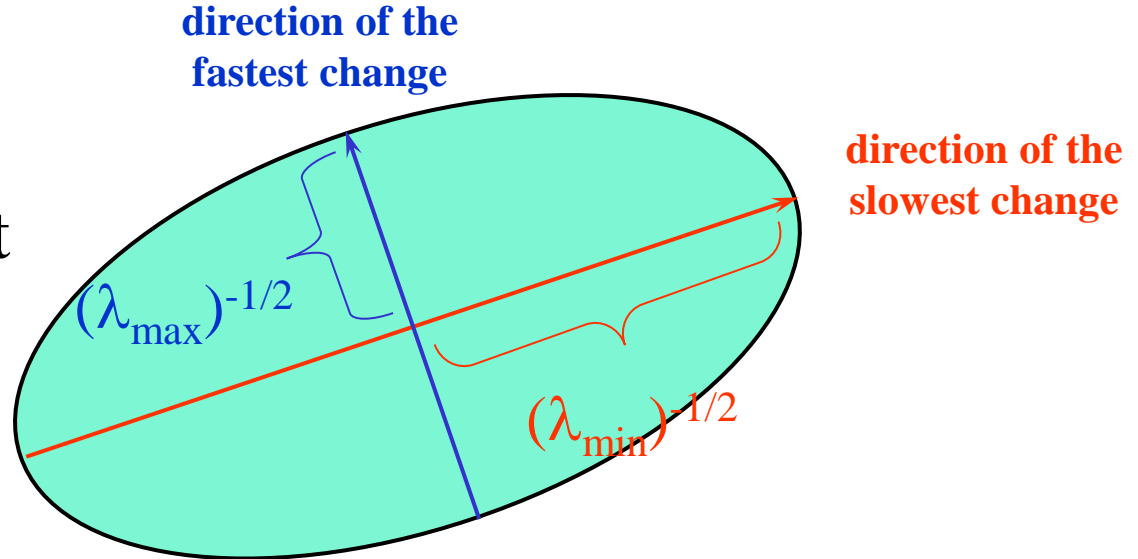
$$M = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

# Harris Detector: Mathematics

Intensity change in shifting window: eigenvalue analysis

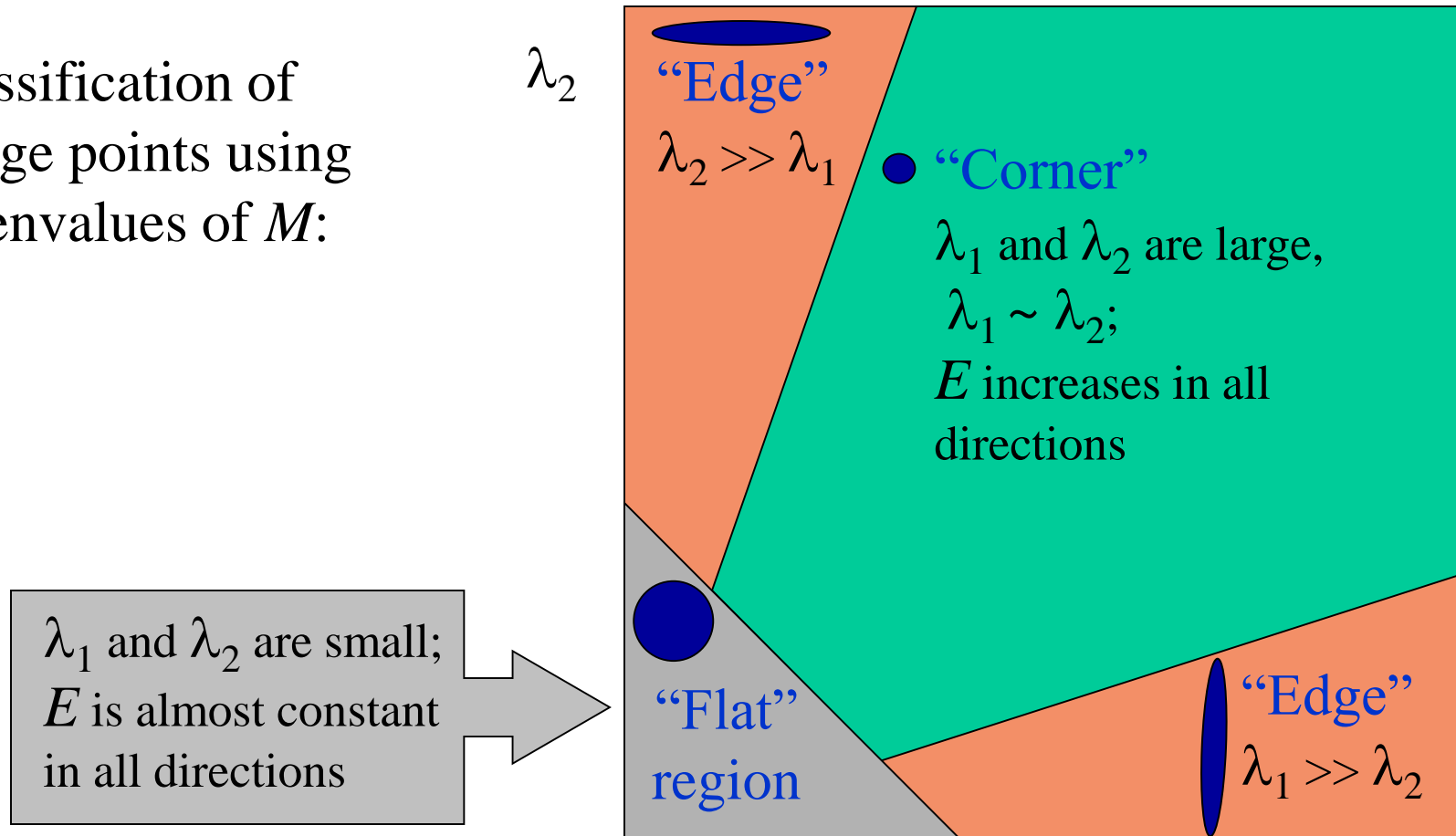
$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad \lambda_1, \lambda_2 - \text{eigenvalues of } M$$

Ellipse  $E(u, v) = \text{const}$



# Harris Detector: Mathematics

Classification of  
image points using  
eigenvalues of  $M$ :



# Harris Detector: Mathematics

---

Measure of corner response:

$$R = \det M - k (\text{trace } M)^2$$

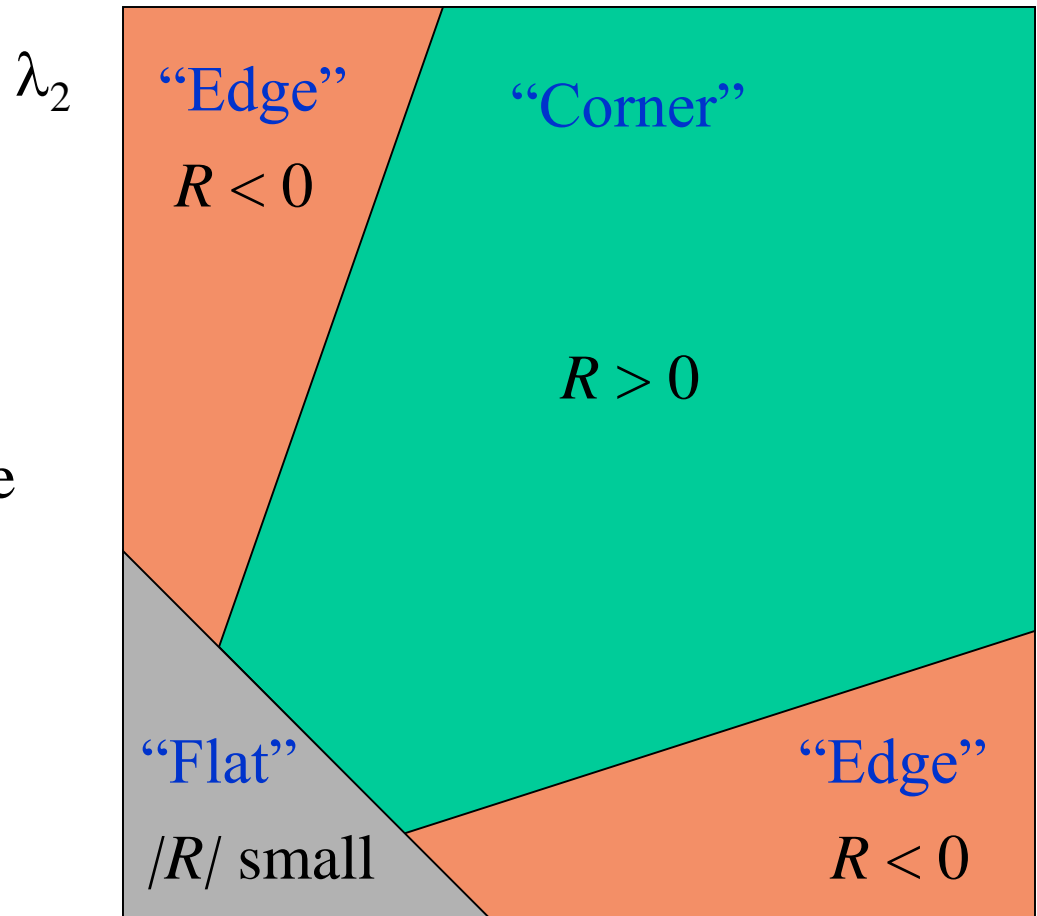
$$\det M = \lambda_1 \lambda_2$$

$$\text{trace } M = \lambda_1 + \lambda_2$$

( $k$  – empirical constant,  $k = 0.04$ - $0.06$ )

# Harris Detector: Mathematics

- $R$  depends only on eigenvalues of  $M$
- $R$  is large for a **corner**
- $R$  is negative with large magnitude for an **edge**
- $|R|$  is small for a **flat** region



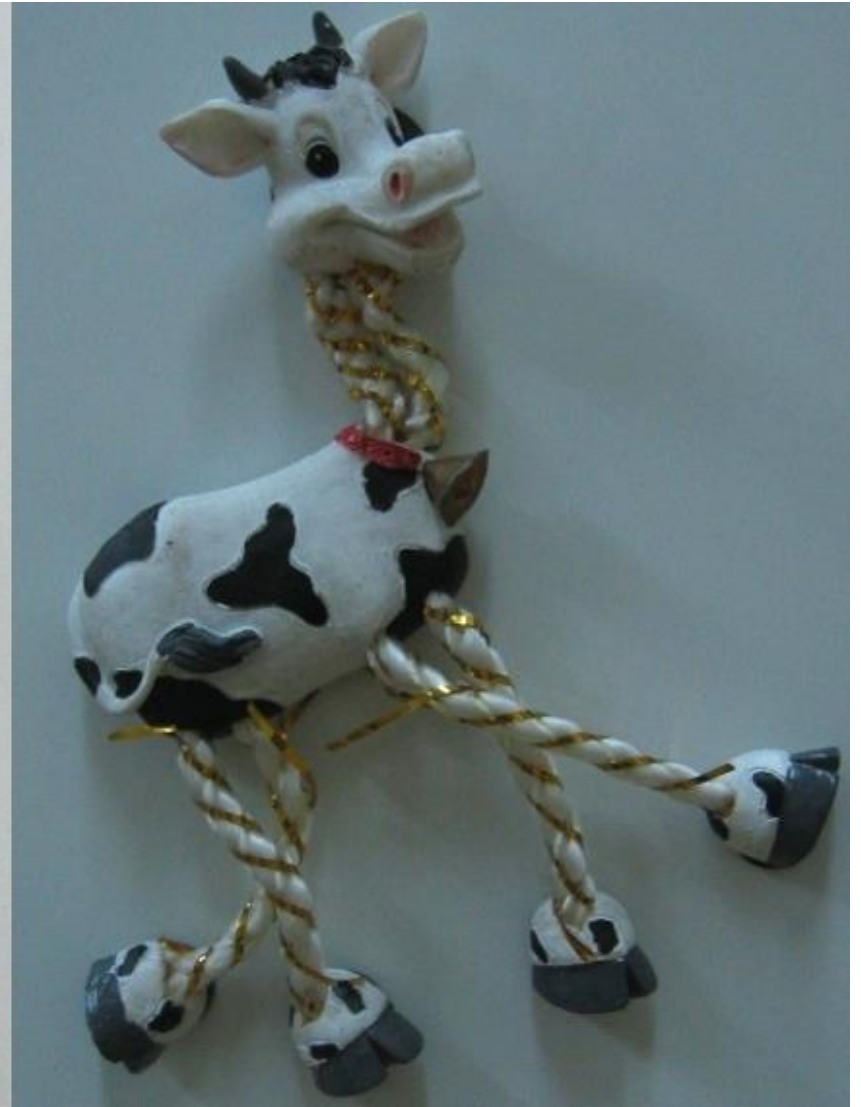
# Harris Detector

---

## The Algorithm:

- Find points with large corner response function  $R$  ( $R > \text{threshold}$ )
- Take the points of local maxima of  $R$

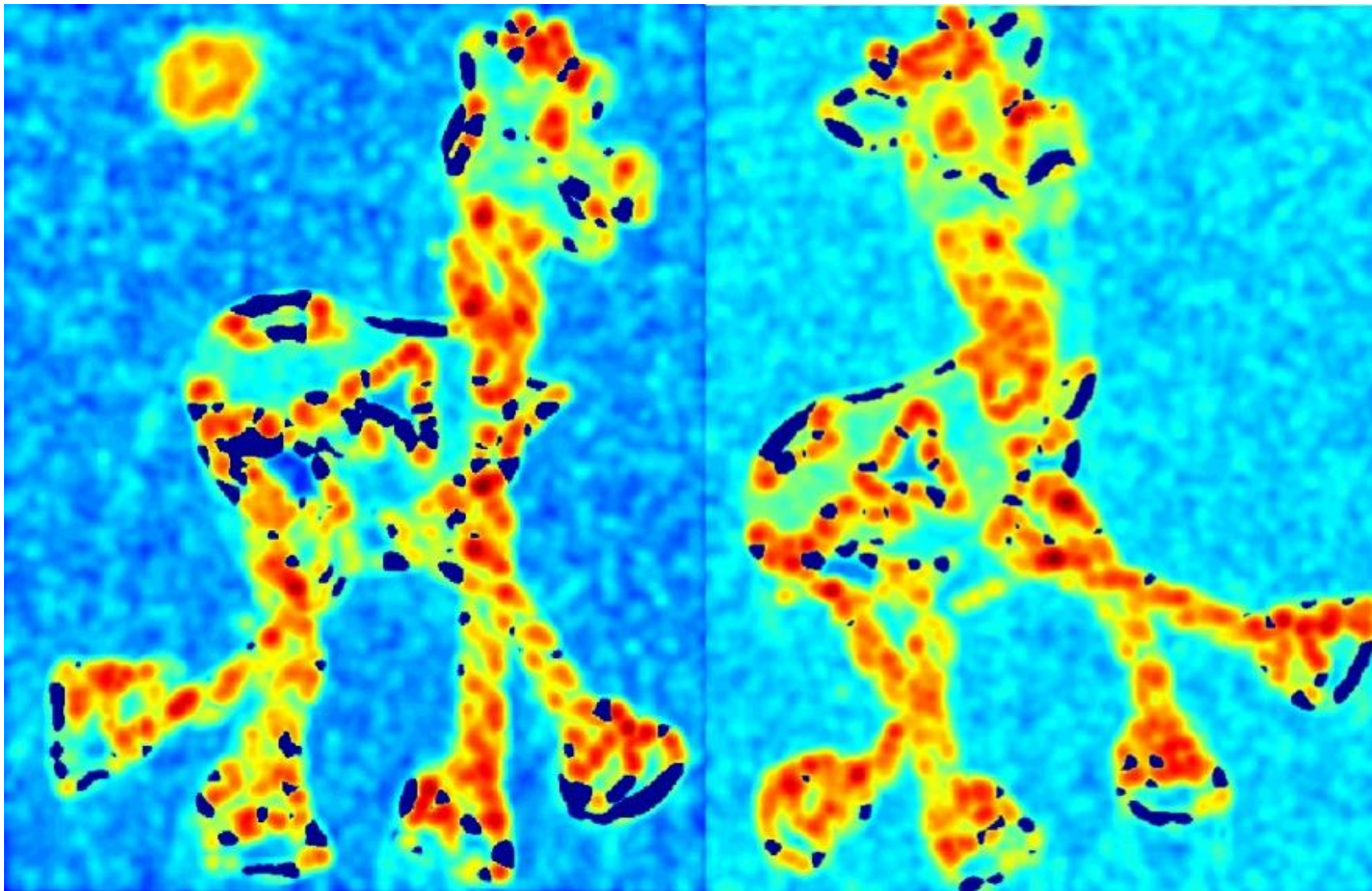
# Harris Detector: Workflow





# Harris Detector: Workflow

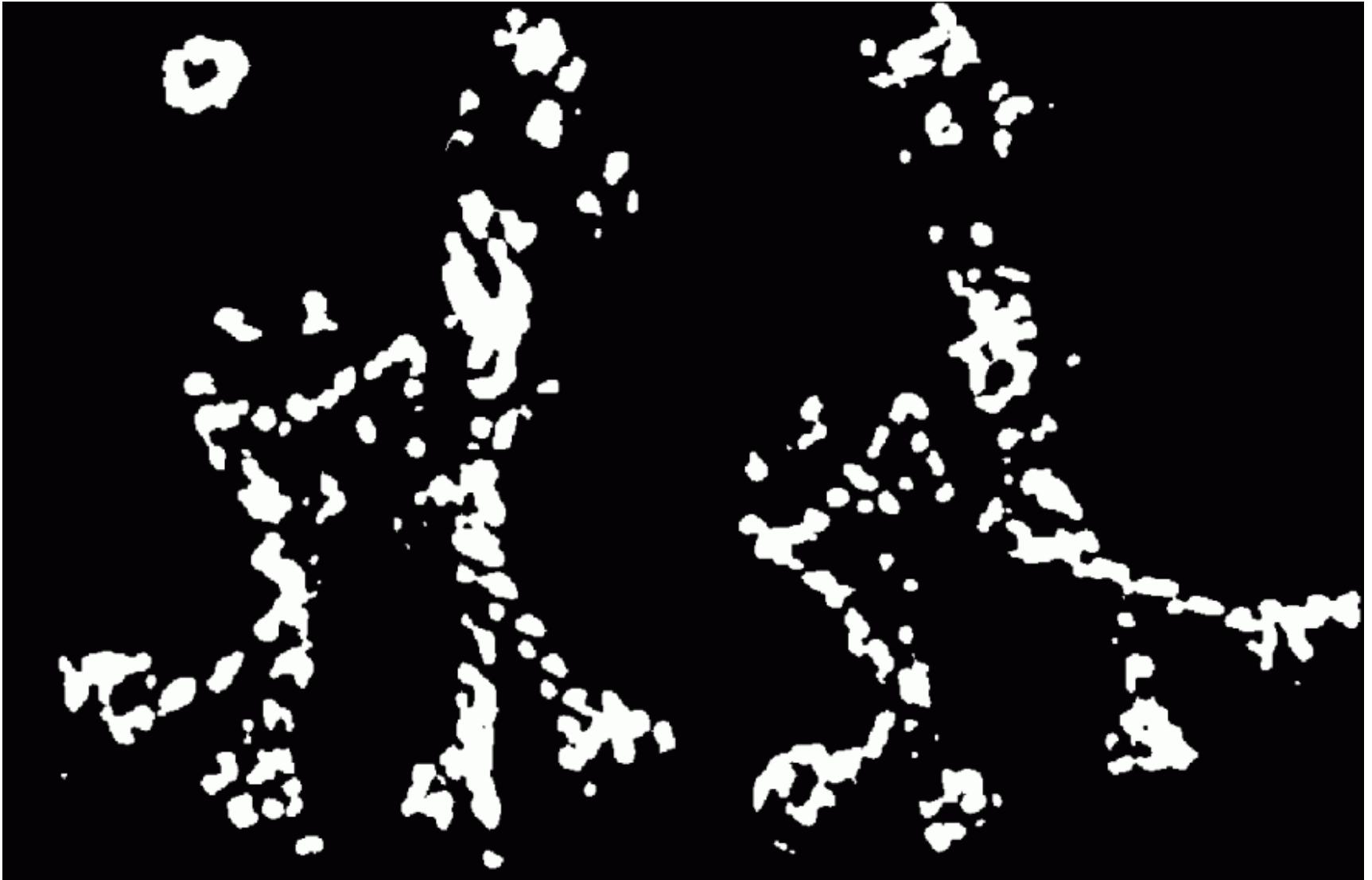
Compute corner response  $R$





# Harris Detector: Workflow

Find points with large corner response:  $R > \text{threshold}$



# Harris Detector: Workflow

Take only the points of local maxima of  $R$



# Harris Detector: Workflow



# Harris Detector: Summary

---

Average intensity change in direction  $[u, v]$  can be expressed as a bilinear form:

$$E(u, v) \cong [u, v] \ M \ \begin{bmatrix} u \\ v \end{bmatrix}$$

Describe a point in terms of eigenvalues of  $M$ :  
*measure of corner response*

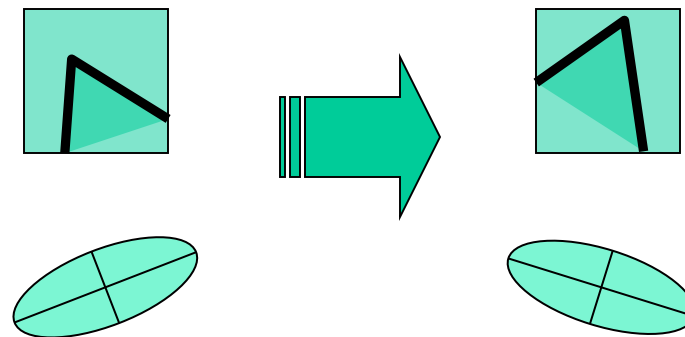
$$R = \lambda_1 \lambda_2 - k (\lambda_1 + \lambda_2)^2$$

A good (corner) point should have a *large intensity change in all directions*, i.e.  $R$  should be large positive

# Harris Detector: Some Properties

---

## Rotation invariance



Ellipse rotates but its shape (i.e. eigenvalues) remains the same

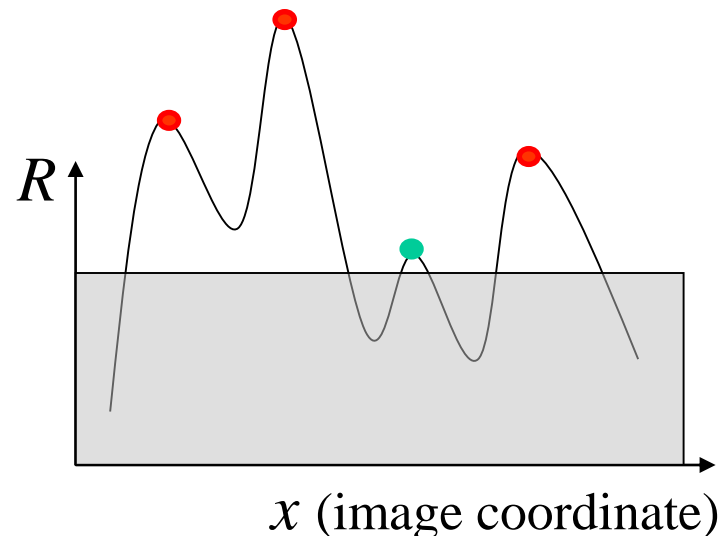
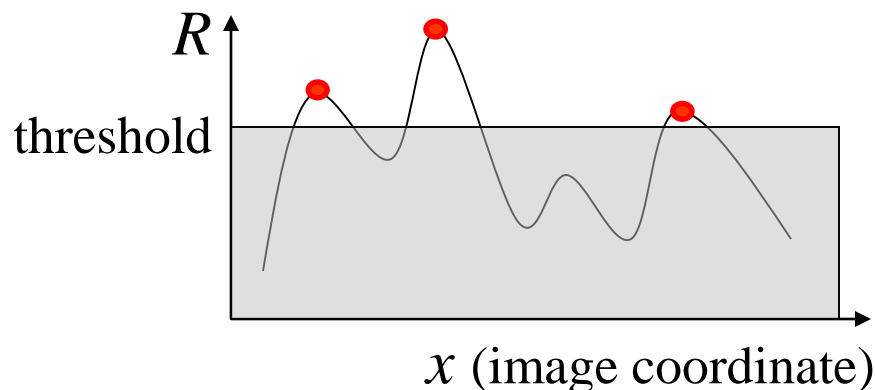
*Corner response  $R$  is invariant to image rotation*

# Harris Detector: Some Properties

---

## Partial invariance to *affine intensity* change

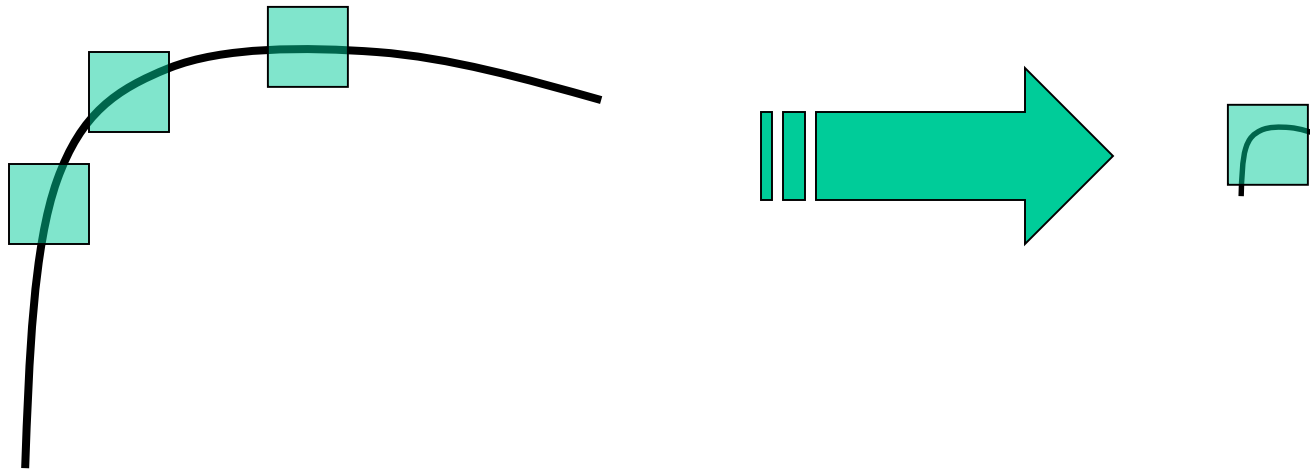
- ✓ Only derivatives are used  $\Rightarrow$  invariance to intensity shift  $I \rightarrow I + b$
- ✓ Intensity scale:  $I \rightarrow a I$



# Harris Detector: Some Properties

---

But: non-invariant to *image scale*!



All points will be  
classified as **edges**

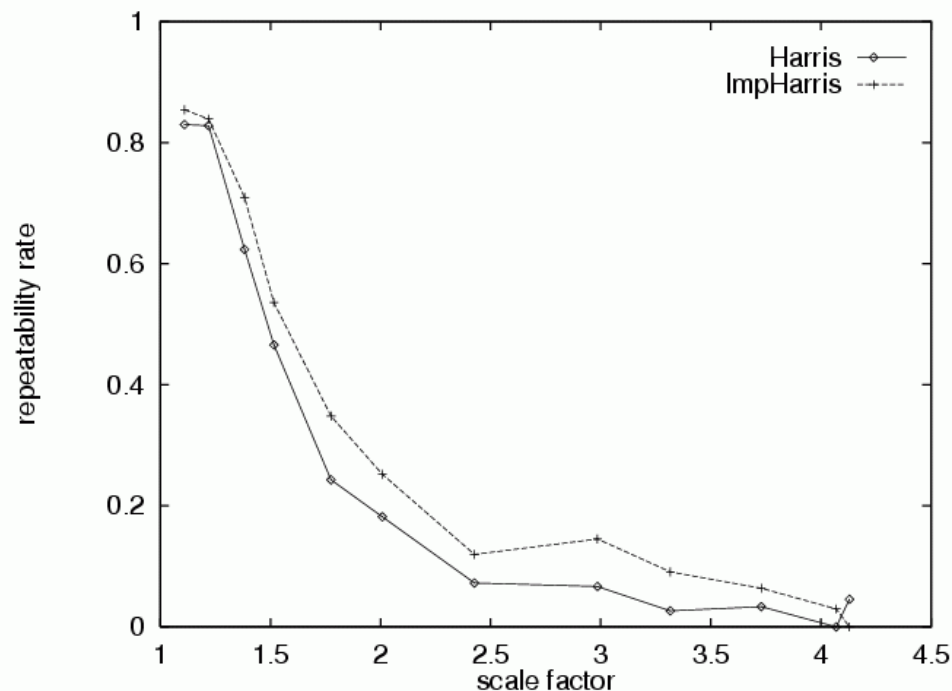
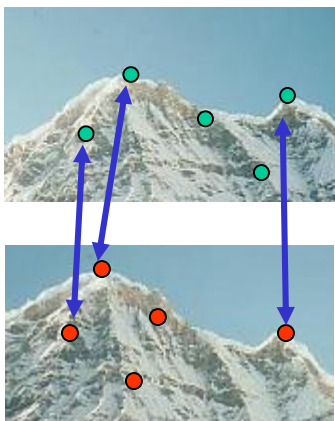
**Corner !**

# Harris Detector: Some Properties

## Quality of Harris detector for different scale changes

Repeatability rate:

$$\frac{\# \text{ correspondences}}{\# \text{ possible correspondences}}$$



C.Schmid et.al. “Evaluation of Interest Point Detectors”. IJCV 2000

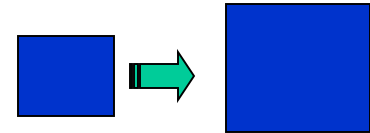
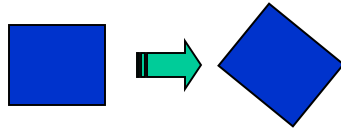


# Models of Image Change

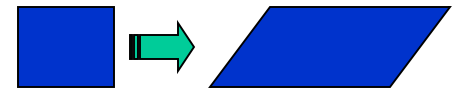
---

## Geometry

- Rotation
- Similarity (rotation + uniform scale)

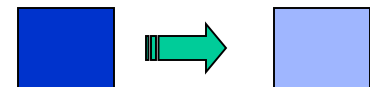


- Affine (scale dependent on direction)  
valid for: orthographic camera, locally planar object



## Photometry

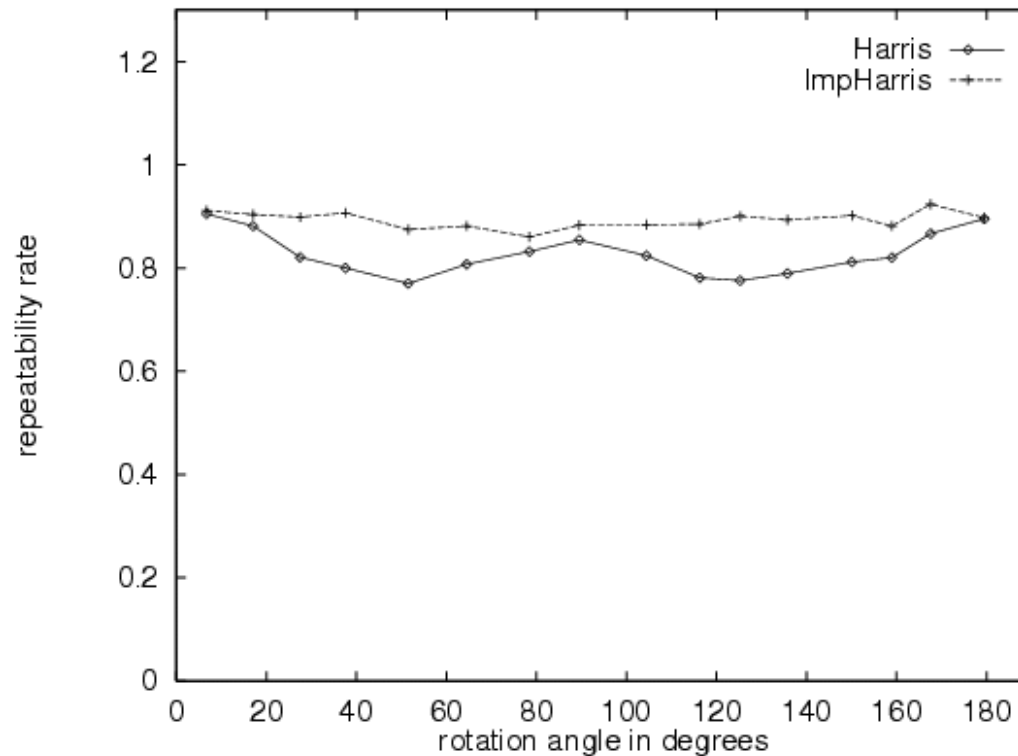
- Affine intensity change ( $I \rightarrow a I + b$ )



# Rotation Invariant Detection

---

## Harris Corner Detector



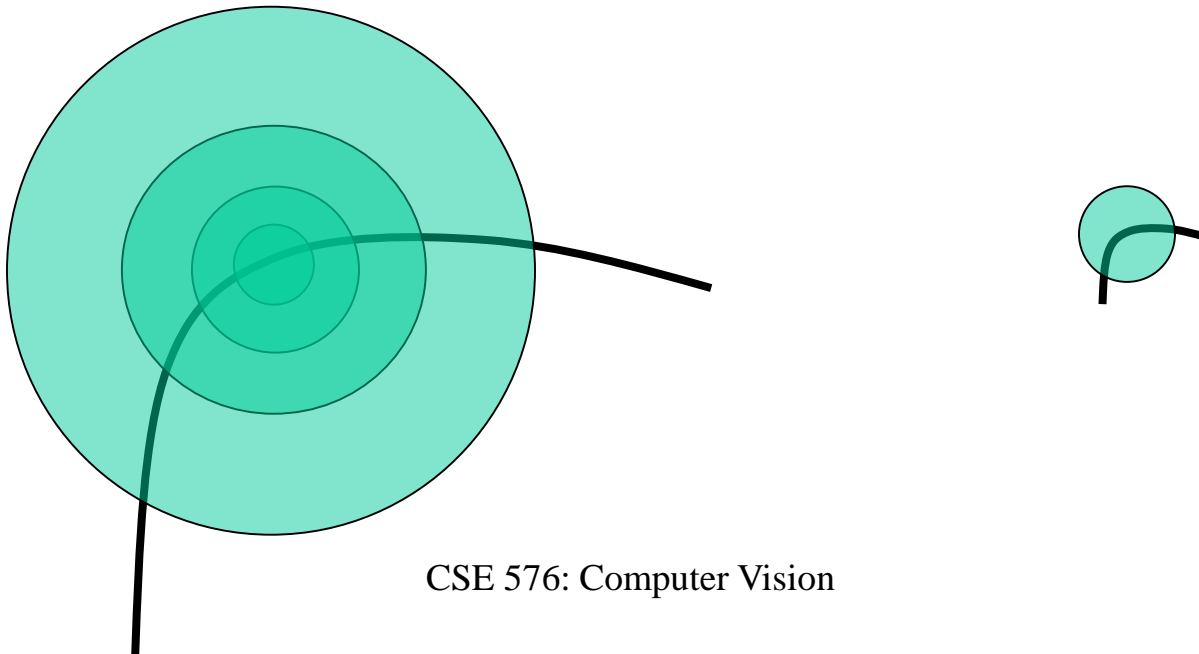
C.Schmid et.al. “Evaluation of Interest Point Detectors”. IJCV 2000

# Scale Invariant Detection

---

Consider regions (e.g. circles) of different sizes around a point  
a point

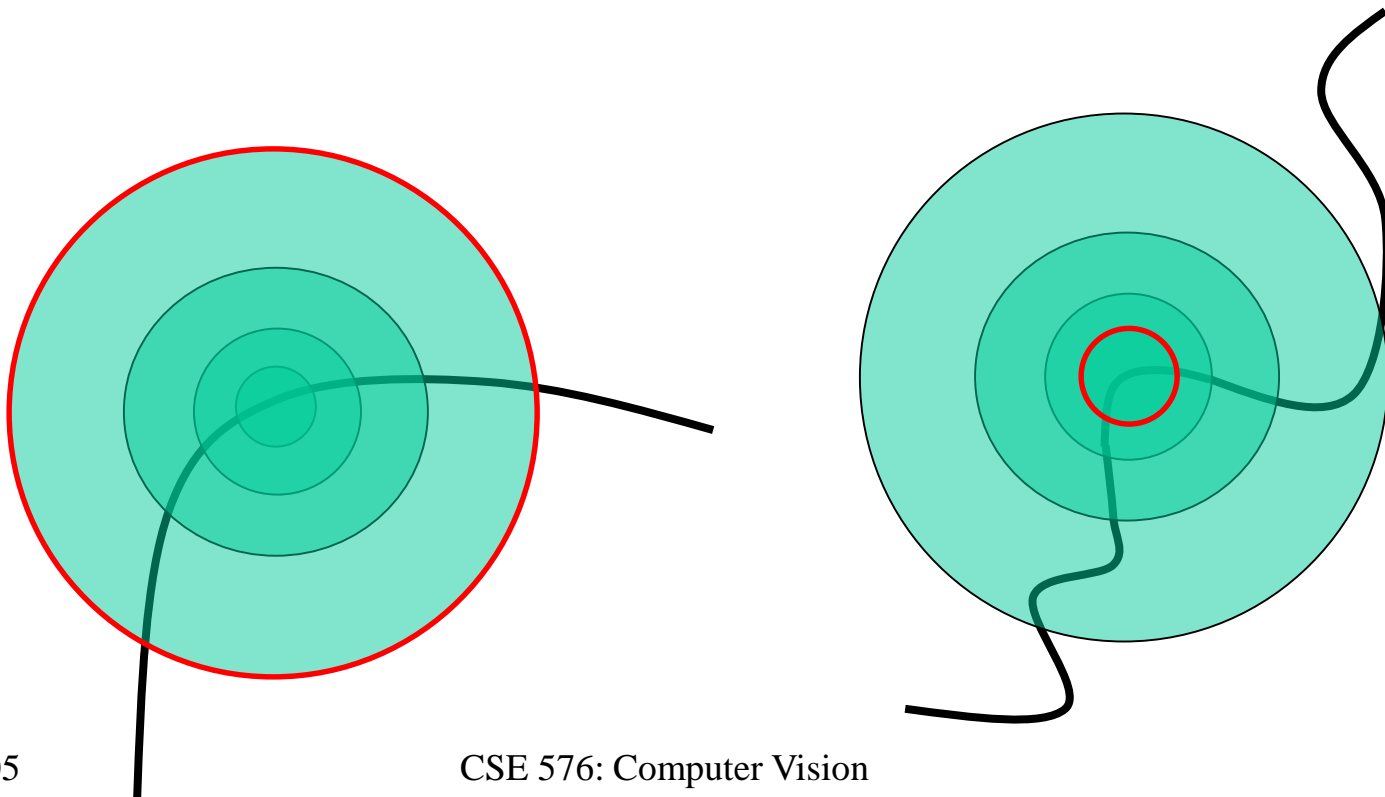
Regions of corresponding sizes will look the same in  
both images



# Scale Invariant Detection

---

The problem: how do we choose corresponding circles *independently* in each image?



# Scale invariance

---

**Requires a method to repeatably select points in location and scale:**

The only reasonable scale-space kernel is a Gaussian (Koenderink, 1984; Lindeberg, 1994)

An efficient choice is to detect peaks in the difference of Gaussian pyramid (Burt & Adelson, 1983; Crowley & Parker, 1984 – but examining more scales)

Difference-of-Gaussian with constant ratio of scales is a close approximation to Lindeberg's scale-normalized Laplacian (can be shown from the heat diffusion equation)

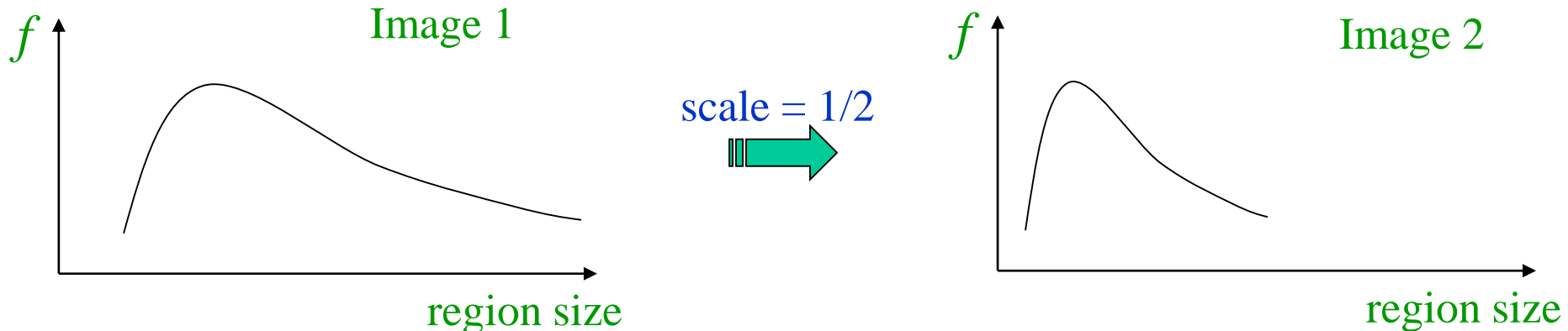
# Scale Invariant Detection

Solution:

- Design a function on the region (circle), which is “scale invariant” (the same for corresponding regions, even if they are at different scales)

Example: average intensity. For corresponding regions (even of different sizes) it will be the same.

- For a point in one image, we can consider it as a function of region size (circle radius)



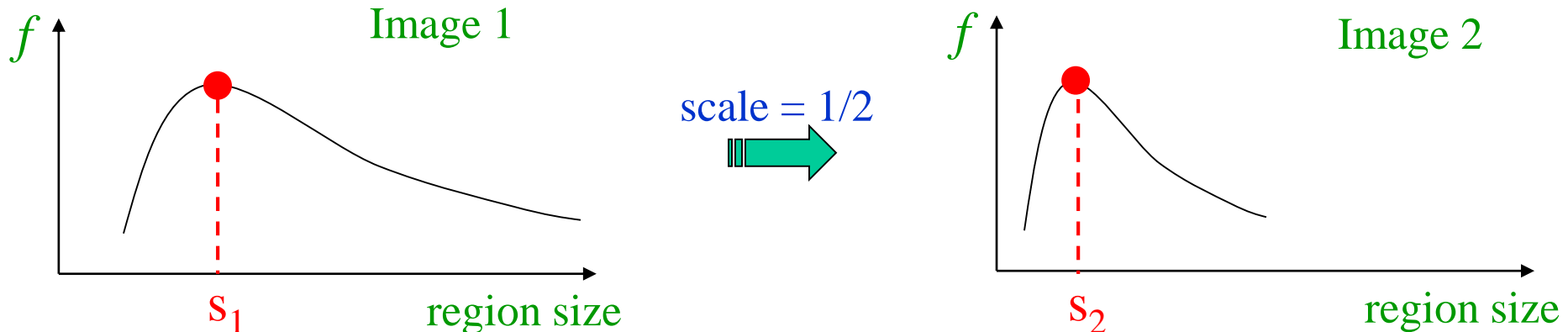
# Scale Invariant Detection

Common approach:

Take a local maximum of this function

Observation: region size, for which the maximum is achieved, should be *invariant* to image scale.

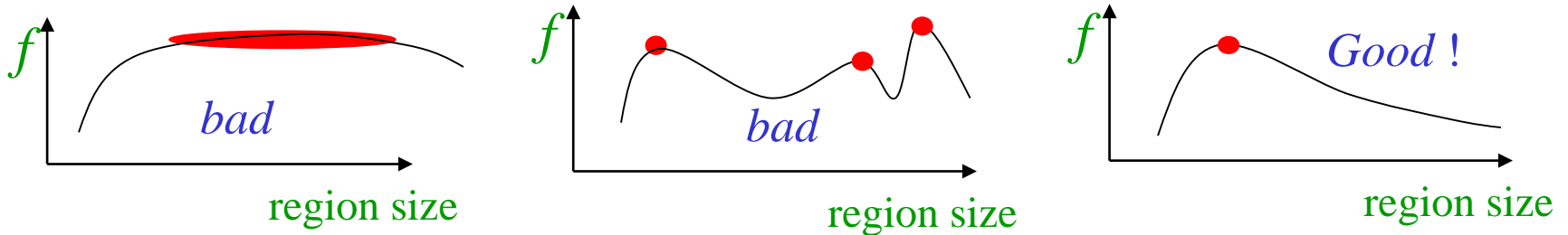
Important: this scale invariant region size is found in each image **independently**!



# Scale Invariant Detection

---

A “good” function for scale detection:  
has one stable sharp peak



- For usual images: a good function would be a one which responds to contrast (sharp local intensity change)



# Scale Invariant Detection

Functions for determining scale

$$f = \text{Kernel} * \text{Image}$$

Kernels:

$$L = \sigma^2 \left( G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma) \right)$$

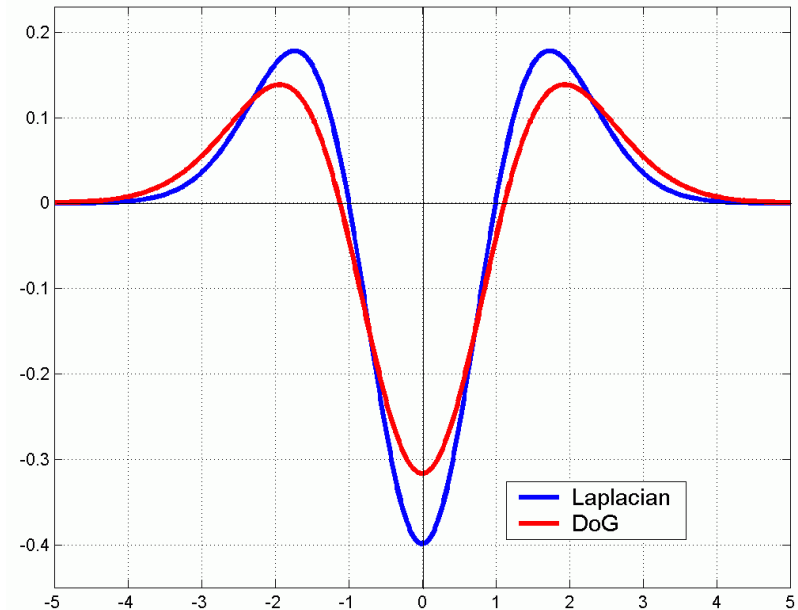
(Laplacian)

$$\text{DoG} = G(x, y, k\sigma) - G(x, y, \sigma)$$

(Difference of Gaussians)

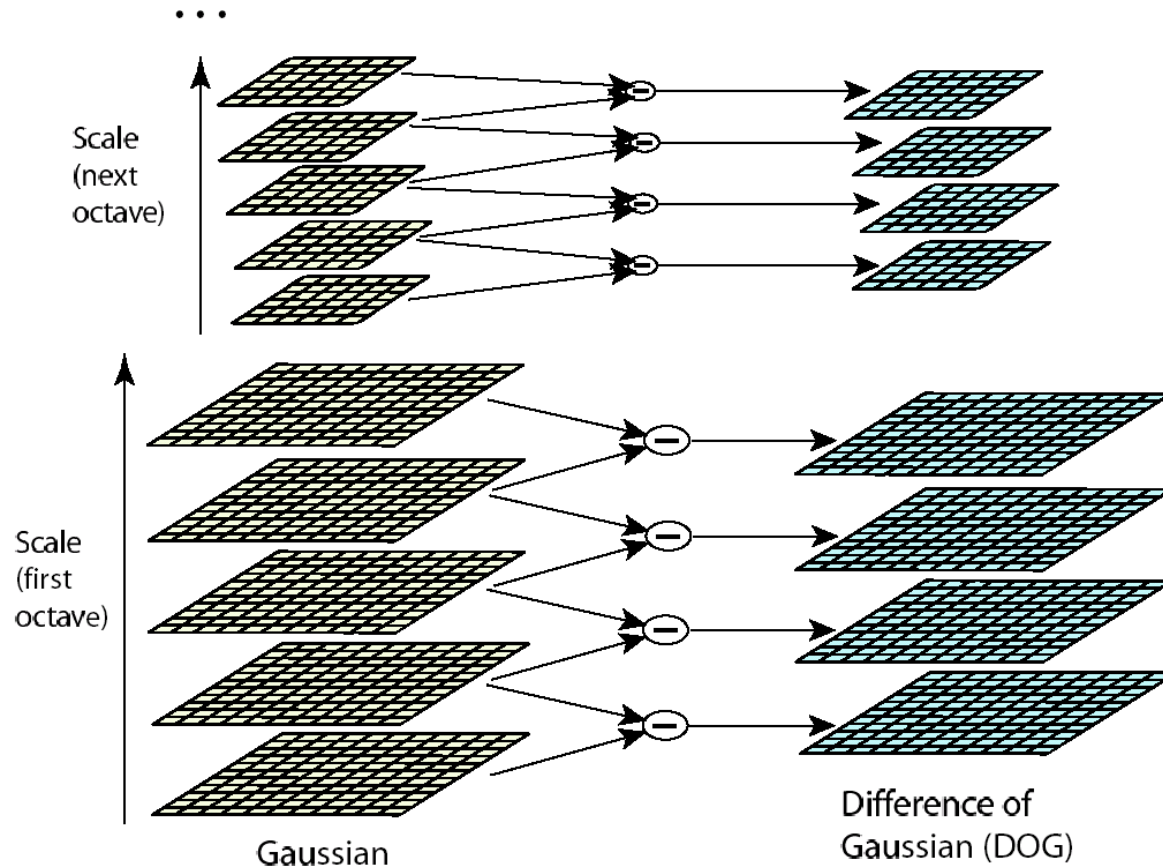
where Gaussian

$$G(x, y, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2 + y^2}{2\sigma^2}}$$



Note: both kernels are invariant to  
*scale* and *rotation*

# Scale space: one octave at a time



# Key point localization

---

Detect maxima and minima of difference-of-Gaussian in scale space

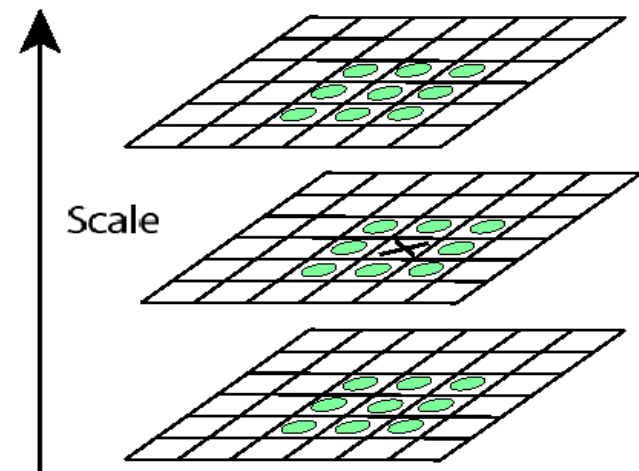
Fit a quadratic to surrounding values for sub-pixel and sub-scale interpolation (Brown & Lowe, 2002)

Taylor expansion around point:

$$D(\mathbf{x}) = D + \frac{\partial D^T}{\partial \mathbf{x}} \mathbf{x} + \frac{1}{2} \mathbf{x}^T \frac{\partial^2 D}{\partial \mathbf{x}^2} \mathbf{x}$$

Offset of extremum (use finite differences for derivatives):

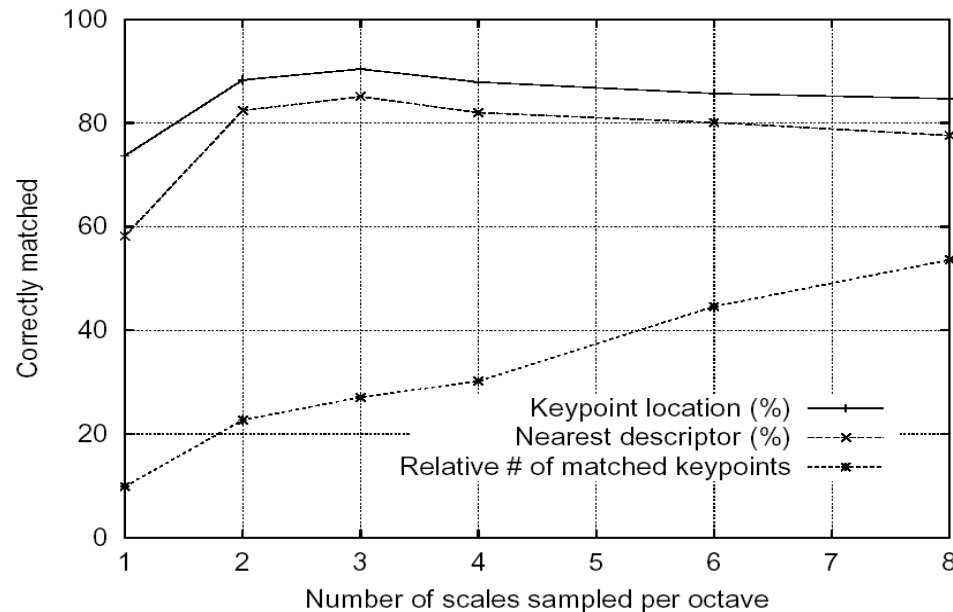
$$\hat{\mathbf{x}} = -\frac{\partial^2 D^{-1}}{\partial \mathbf{x}^2} \frac{\partial D}{\partial \mathbf{x}}$$



# Sampling frequency for scale

---

More points are found as sampling frequency increases, but accuracy of matching decreases after 3 scales/octave



# Eliminating unstable keypoints

---

Discard points with DOG value below threshold (low contrast)

However, points along edges may have high contrast in one direction but low in another

Compute principal curvatures from eigenvalues of 2x2 Hessian matrix, and limit ratio (Harris approach):

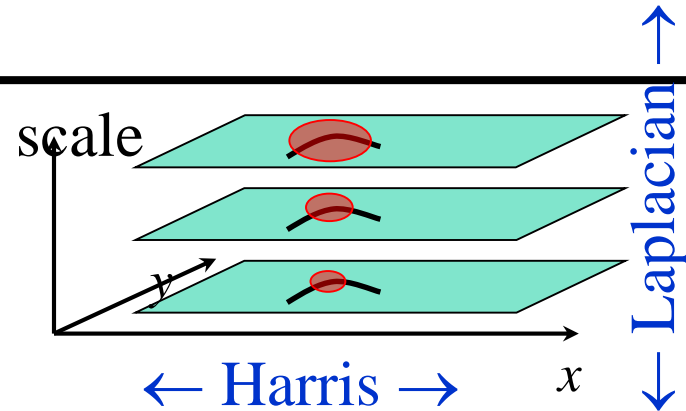
$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad \begin{aligned} \text{Tr}(\mathbf{H}) &= D_{xx} + D_{yy} = \alpha + \beta, \\ \text{Det}(\mathbf{H}) &= D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta. \end{aligned} \quad \frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} < \frac{(r+1)^2}{r}$$

# Scale Invariant Detectors

## Harris-Laplacian<sup>1</sup>

*Find local maximum of:*

- Harris corner detector in space (image coordinates)

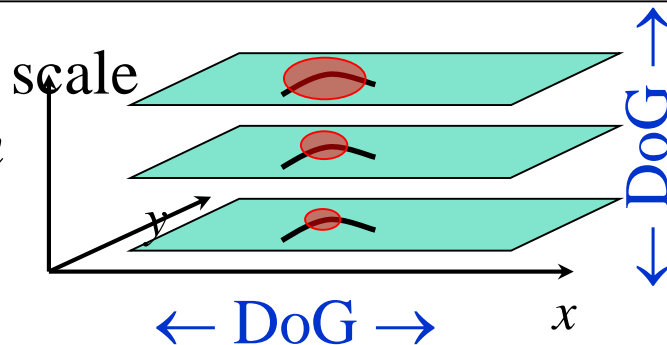


- Laplacian in scale

## • SIFT (Lowe)<sup>2</sup>

*Find local maximum of:*

- Difference of Gaussians in space and scale



<sup>1</sup> K.Mikolajczyk, C.Schmid. “Indexing Based on Scale Invariant Interest Points”. ICCV 2001

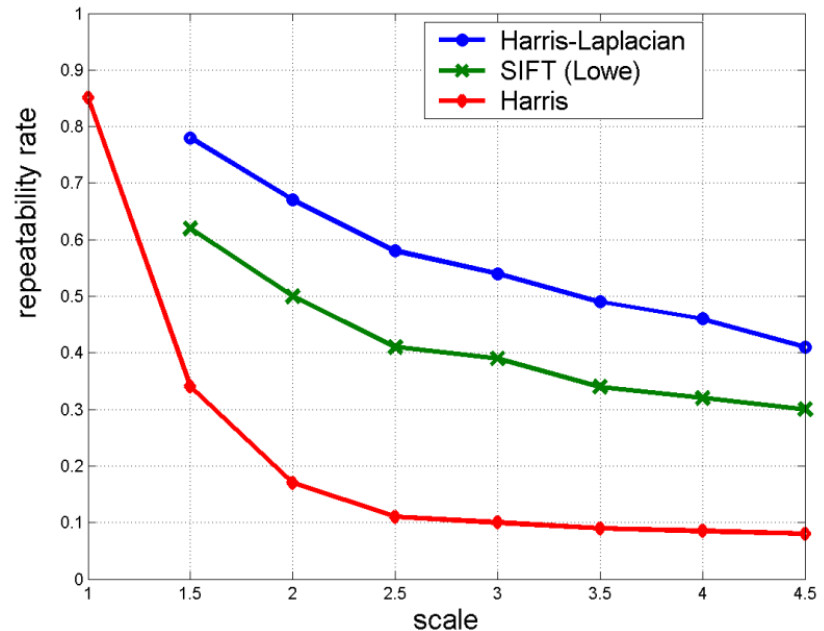
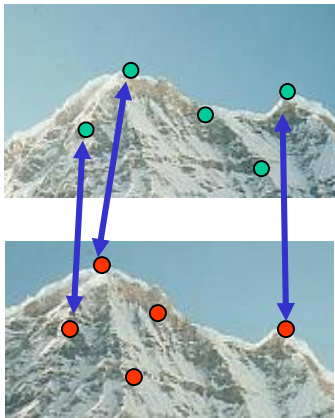
<sup>2</sup> D.Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. Accepted to IJCV 2000

# Scale Invariant Detectors

## Experimental evaluation of detectors w.r.t. scale change

Repeatability rate:

$$\frac{\# \text{ correspondences}}{\# \text{ possible correspondences}}$$



Affine Covariant Features - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Search Favorites Media Print AutoFill Options mikolajczyk

Address <http://www.robots.ox.ac.uk/~vgg/research/affine/> Go Links

Google mikolajczyk Search Web

## Affine Covariant Regions



Image 1 Image 2

## Publications

### Region detectors

- *Harris-Affine & Hessian Affine*: [K. Mikolajczyk](#) and [C. Schmid](#), Scale and Affine invariant interest point detectors. In IJCV 1(60):63-86, 2004. [PDF](#)
- *MSER*: [J. Matas](#), [O. Chum](#), [M. Urban](#), and [T. Pajdla](#), Robust wide baseline stereo from maximally stable extremal regions. In BMVC p. 384-393, 2002. [PDF](#)
- *IBR & EBR*: [T. Tuytelaars](#) and [L. Van Gool](#), Matching widely separated views based on affine invariant regions. In IJCV 1(59):61-85, 2004. [PDF](#)
- *Salient regions*: [T. Kadir](#), [A. Zisserman](#), and [M. Brady](#), An affine invariant salient region detector. In ECCV p. 404-416, 2004. [PDF](#)

### Region descriptors

- *SIFT*: [D. Lowe](#), Distinctive image features from scale invariant keypoints. In IJCV 2(60):91-110, 2004. [PDF](#)

### Performance evaluation

- [K. Mikolajczyk](#), [T. Tuytelaars](#), [C. Schmid](#), [A. Zisserman](#), [J. Matas](#), [F. Schaffalitzky](#), [T. Kadir](#) and [L. Van Gool](#), A comparison of affine region detectors. Technical Report, accepted to IJCV. [PDF](#)
- [K. Mikolajczyk](#), [C. Schmid](#), A performance evaluation of local descriptors. Technical Report, accepted to PAMI. [PDF](#)

## Software

- [Region detectors](#) - Linux binaries for detecting affine covariant regions.
- [Region descriptors](#) - Linux binaries for computing region descriptors.
- [Detectors evaluation](#) - Matlab files to compute the repeatability.
- [Descriptors evaluation](#) - Matlab files to compute the matching score.

## Test Data

The packages contain images in PPM format and homographies between image pairs. [Data description](#).

Done Internet



Affine Covariant Features - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://www.robots.ox.ac.uk/~vgg/research/affine/>

Google mikolajczyk Search Web AutoFill Options mikolajczyk

### Region descriptors

- [SIFT](#): [D. Lowe](#), Distinctive image features from scale invariant keypoints. In IJCV 2(60):91-110, 2004. [PDF](#)

### Performance evaluation









- [K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir and L. Van Gool](#), A comparison of affine region detectors. Technical Report, accepted to IJCV. [PDF](#)
- [K. Mikolajczyk, C. Schmid](#), A performance evaluation of local descriptors. Technical Report, accepted to PAMI. [PDF](#)

### Software

- [Region detectors](#) - Linux binaries for detecting affine covariant regions.
- [Region descriptors](#) - Linux binaries for computing region descriptors.
- [Detectors evaluation](#) - Matlab files to compute the repeatability.
- [Descriptors evaluation](#) - Matlab files to compute the matching score.

### Test Data

The packages contain images in PPM format and homographies between image pairs. [Data description](#).

<p><b>Blur</b></p>  <p>1000x700 6 images</p>	<p><b>Blur</b></p>  <p>1000x700 6 images</p>	<p><b>Viewpoint</b></p>  <p>800x640 6 images</p>	<p><b>Viewpoint</b></p>  <p>1000x700 6 images</p>
<p><b>Zoom+rotation</b></p>  <p>765x512 6 images</p>	<p><b>Zoom+rotation</b></p>  <p>800x640 6 images</p>	<p><b>Light</b></p>  <p>921x614 6 images</p>	<p><b>JPEG compression</b></p>  <p>800x640 6 images</p>

Page maintained by [km@robots.ox.ac.uk](mailto:km@robots.ox.ac.uk)  
Last updated 26th November 2004

Internet

# Scale Invariant Detection: Summary

---

**Given:** two images of the same scene with a large *scale difference* between them

**Goal:** find *the same* interest points *independently* in each image

**Solution:** search for *maxima* of suitable functions in *scale* and in *space* (over the image)

## Methods:

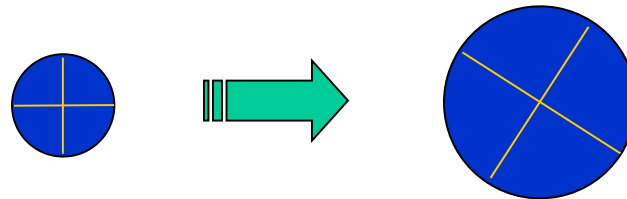
1. **Harris-Laplacian** [Mikolajczyk, Schmid]: maximize Laplacian over scale, Harris' measure of corner response over the image
2. **SIFT** [Lowe]: maximize Difference of Gaussians over scale and space

# Affine Invariant Detection

---

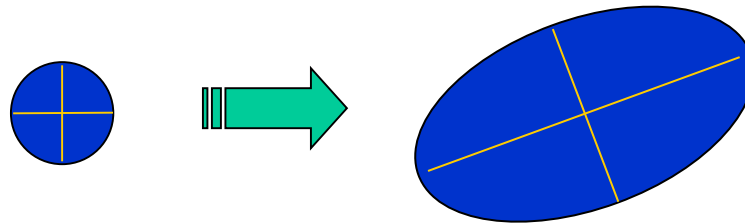
Above we considered:

Similarity transform (rotation + uniform scale)



- Now we go on to:

Affine transform (rotation + non-uniform scale)



# Affine invariant descriptors

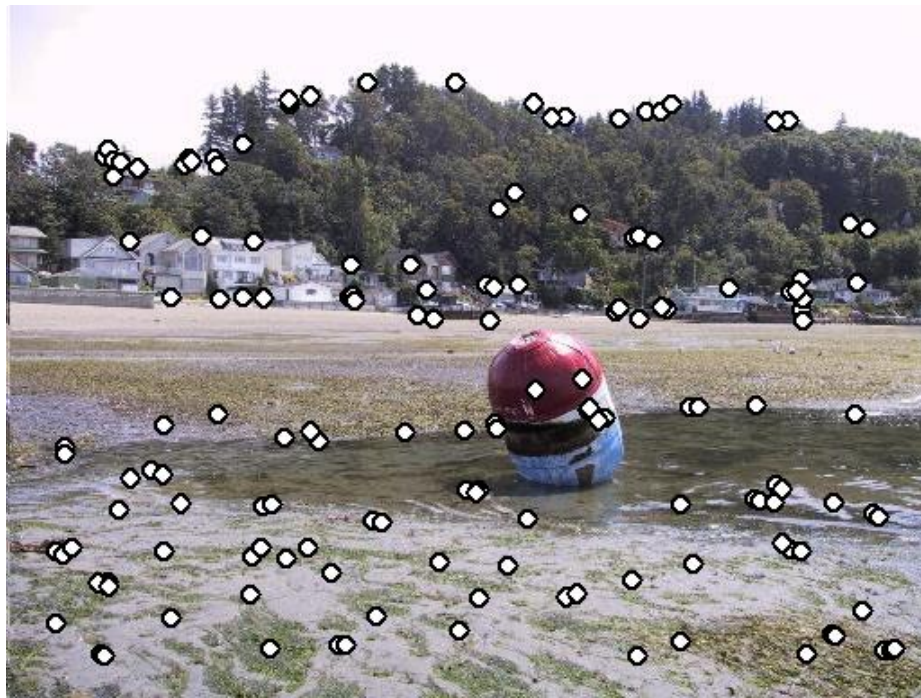
---

... skip these slides (see handouts and on-line class notes) ...

# Feature selection

---

Distribute points evenly over the image



# Adaptive Non-maximal Suppression

---

Desired: Fixed # of features per image

- Want evenly distributed spatially...
- Search over non-maximal suppression radius  
[Brown, Szeliski, Winder, CVPR'05]



$r = 8, n = 1388$



$r = 20, n = 283$



# Today's lecture

---

- What is computer vision?
  - Scale-space and pyramids
  - What are good features?
  - Feature detection
  - **Feature descriptors**
  - (Next lecture: feature matching)
- 
- Project 1 description and demo [Ian Simon]

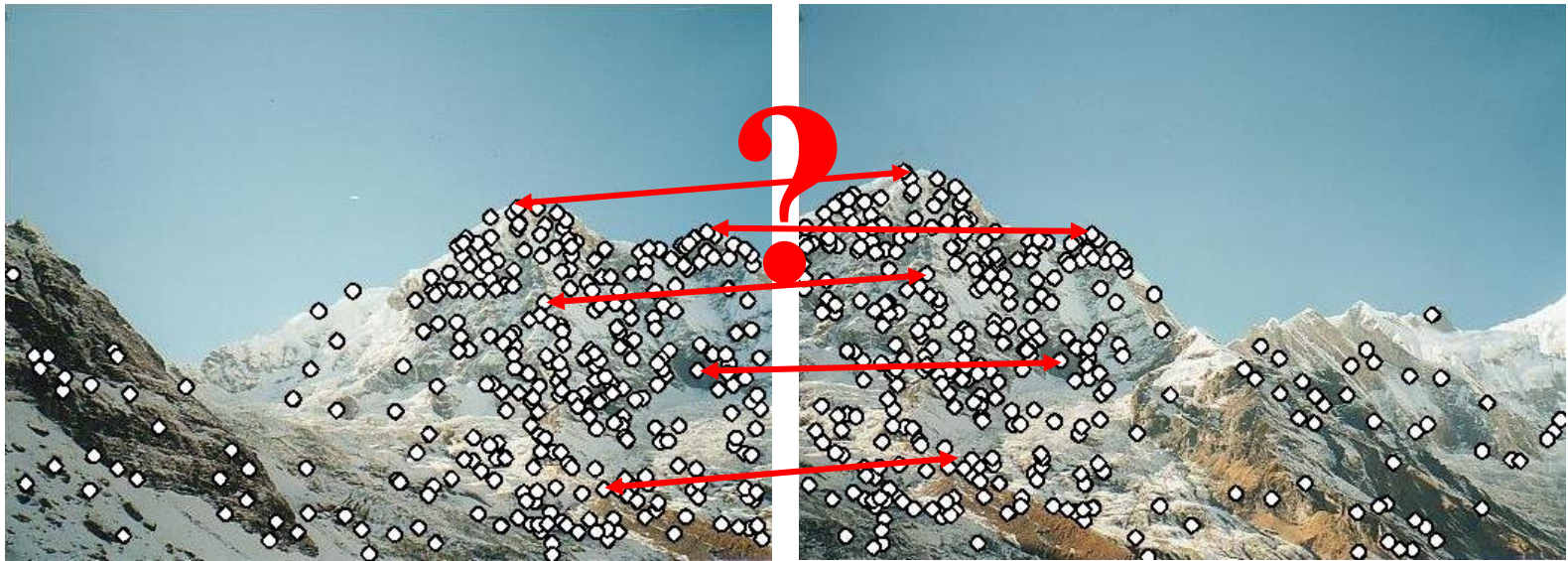
# Point Descriptors

We know how to detect points

---

Next question:

**How to match them?**



Point descriptor should be:

1. Invariant
2. Distinctive

CSE 551 Computer Vision

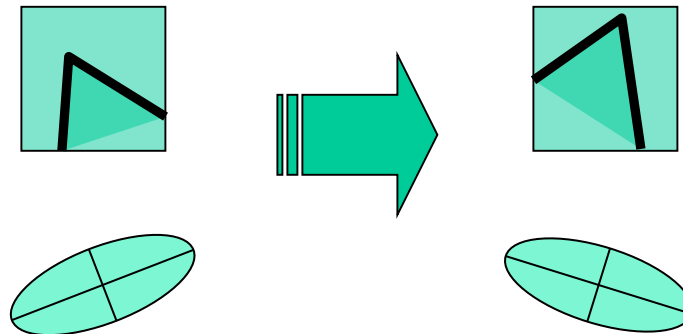


# Descriptors Invariant to Rotation

Harris corner response measure:

depends only on the eigenvalues of the matrix  $M$

$$M = \sum_{x,y} w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$



# Multi-Scale Oriented Patches

---

## Interest points

- Multi-scale Harris corners
- Orientation from blurred gradient
- Geometrically invariant to similarity transforms

## Descriptor vector

- Bias/gain normalized sampling of local patch (8x8)
- Photometrically invariant to affine changes in intensity

# Descriptor Vector

---

Orientation = blurred gradient

Similarity Invariant Frame

- Scale-space position  $(x, y, s)$  + orientation  $(\theta)$



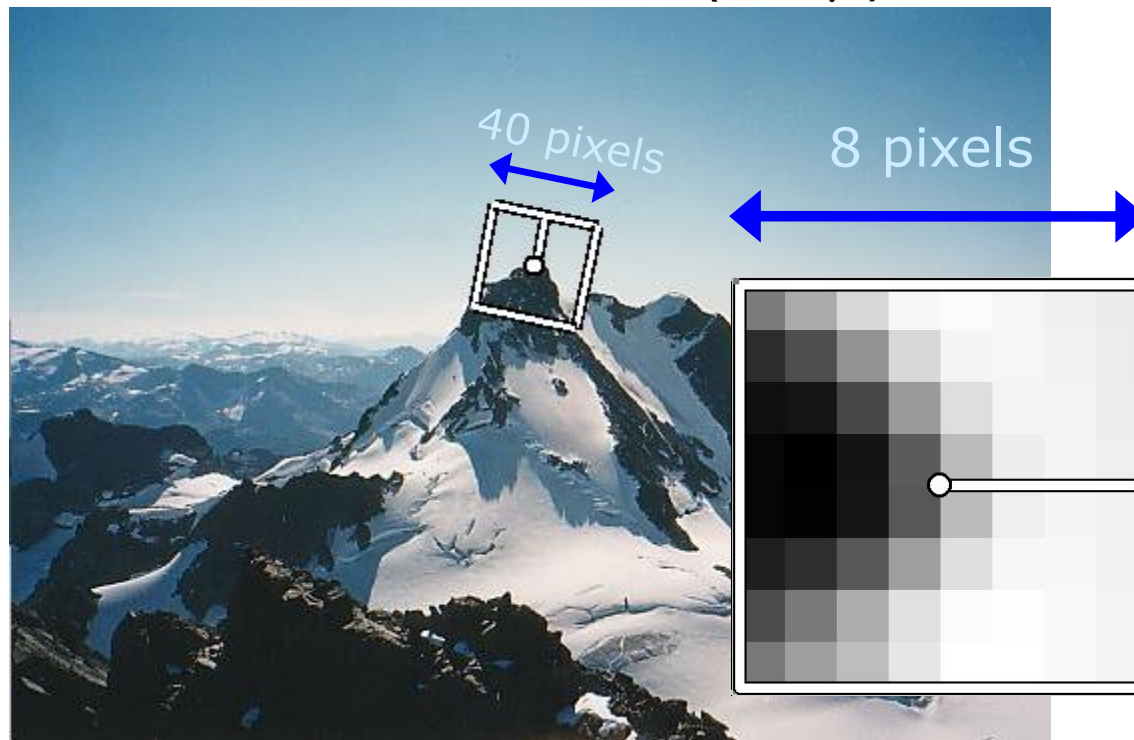
# MOPS descriptor vector

---

8x8 oriented patch

- Sampled at 5 x scale

Bias/gain normalisation:  $I' = (I - \mu)/\sigma$



# Descriptors Invariant to Rotation

## Image moments in polar coordinates

---

$$m_{kl} = \iint r^k e^{-i\theta l} I(r, \theta) dr d\theta$$

Rotation in polar coordinates is translation of the angle:

$$\theta \rightarrow \theta + \theta_0$$

This transformation changes only the phase of the moments, but not its magnitude

Rotation invariant descriptor consists of magnitudes of moments:

$$|m_{kl}|$$

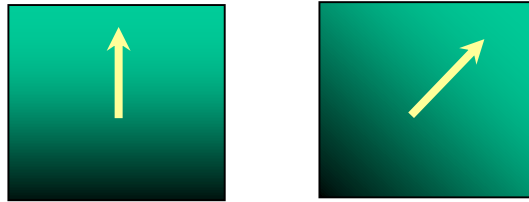
Matching is done by comparing vectors  $[|m_{kl}|]_{k,l}$

# Descriptors Invariant to Rotation

Find local orientation

---

Dominant direction of gradient



- Compute image derivatives relative to this orientation

<sup>1</sup> K.Mikolajczyk, C.Schmid. “Indexing Based on Scale-Invariant Interest Points”. ICCV 2001

<sup>2</sup> D.Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. Accepted to IJCV 2004

# Descriptors Invariant to Scale

Use the scale determined by detector to compute descriptor in a normalized frame

For example:

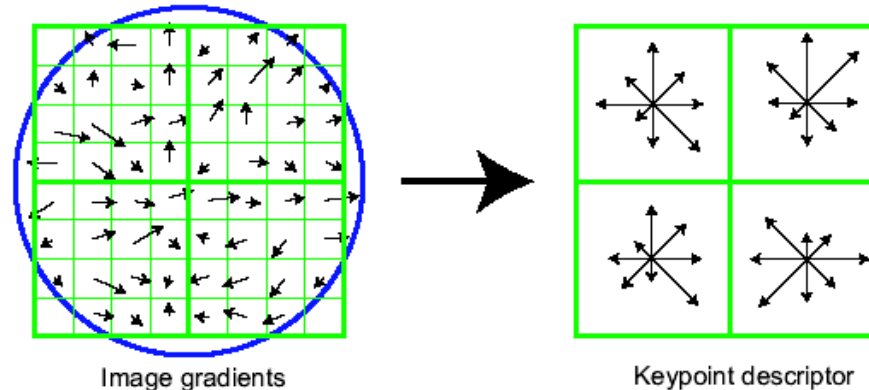
- moments integrated over an adapted window
- derivatives adapted to scale:  $sI_x$

# SIFT – Scale Invariant Feature Transform

---

## Descriptor overview:

- Determine **scale** (by maximizing DoG in scale and in space), **local orientation** as the dominant gradient direction. Use this scale and orientation to make all further computations invariant to scale and rotation.
- Compute **gradient orientation histograms** of several small windows (128 values for each point)
- Normalize the descriptor to make it invariant to intensity change





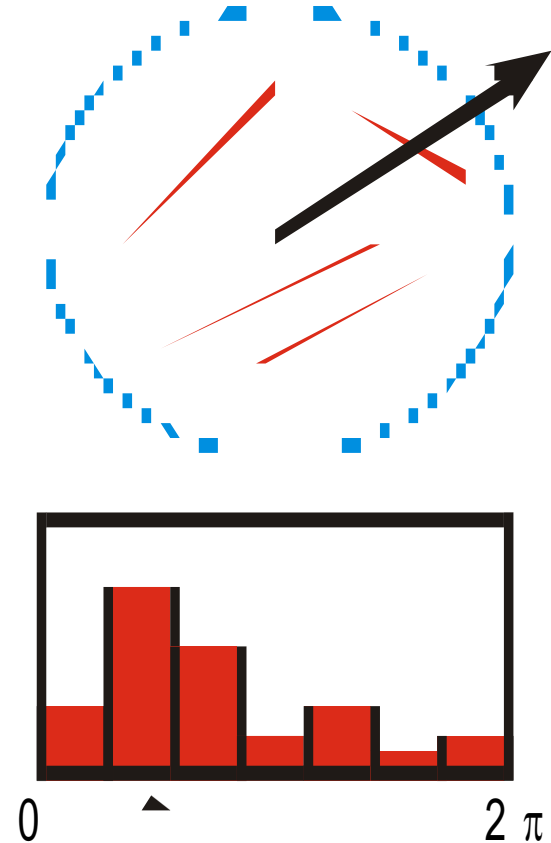
# Select canonical orientation

---

Create histogram of local  
gradient directions  
computed at selected scale

Assign canonical orientation  
at peak of smoothed  
histogram

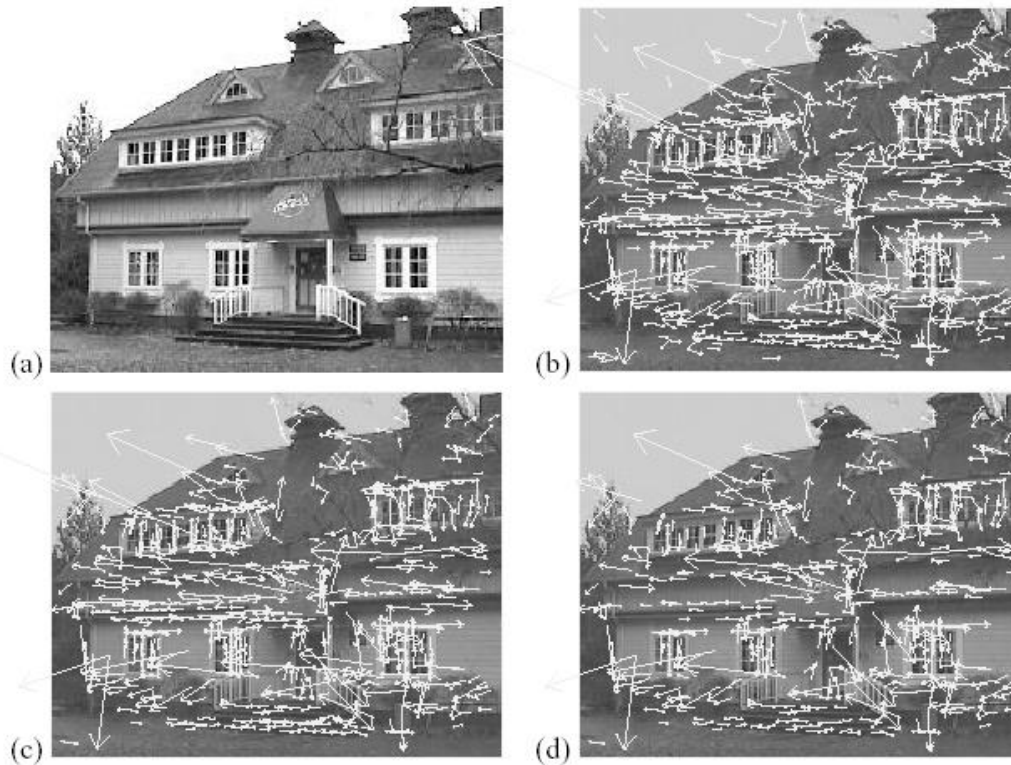
Each key specifies stable 2D  
coordinates (x, y, scale,  
orientation)



# Example of keypoint detection

---

Threshold on value at DOG peak and on ratio of principle curvatures (Harris approach)



- (a)** 233x189 image
- (b)** 832 DOG extrema
- (c)** 729 left after peak value threshold
- (d)** 536 left after testing ratio of principle curvatures

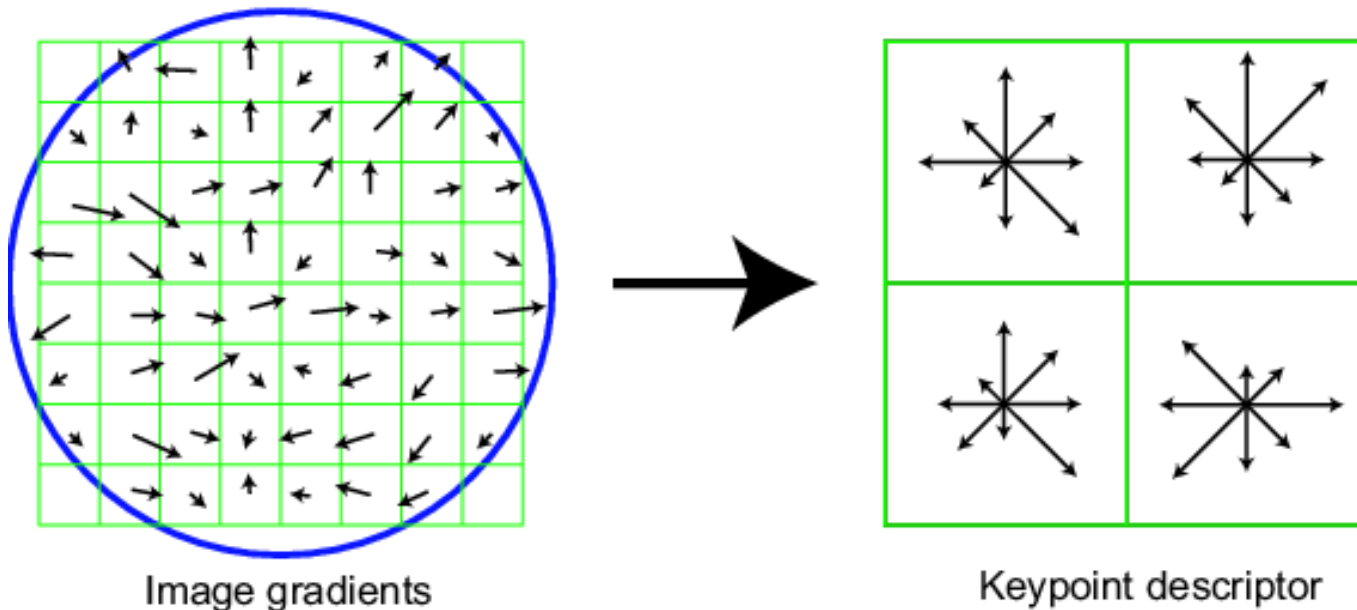
# SIFT vector formation

---

Thresholded image gradients are sampled over 16x16 array of locations in scale space

Create array of orientation histograms

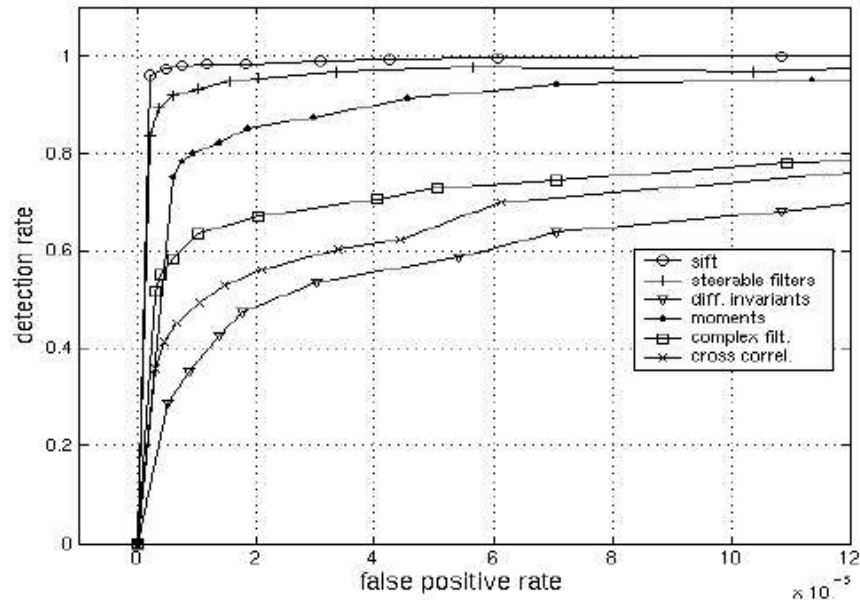
8 orientations x 4x4 histogram array = 128 dimensions



# SIFT – Scale Invariant Feature Transform<sup>1</sup>

Empirically found<sup>2</sup> to show very good performance, invariant to *image rotation, scale, intensity change*, and to moderate *affine* transformations

Scale = 2.5  
Rotation = 45<sup>0</sup>



<sup>1</sup> D.Lowe. “Distinctive Image Features from Scale-Invariant Keypoints”. Accepted to IJCV 2004

<sup>2</sup> K.Mikolajczyk, C.Schmid. “A Performance Evaluation of Local Descriptors”. CVPR 2003

# Invariance to Intensity Change

---

## Detectors

- mostly invariant to affine (linear) change in image intensity, because we are searching for *maxima*

## Descriptors

- Some are based on derivatives => invariant to intensity shift
- Some are normalized to tolerate intensity scale
- Generic method: pre-normalize intensity of a region (eliminate shift and scale)

# Today's lecture

---

- What is computer vision?
  - Scale-space and pyramids
  - What are good features?
  - Feature detection
  - Feature descriptors
  - (Next lecture: feature matching)
- 
- Project 1 description and demo [Ian Simon]