Computer Vision

Computer Science Tripos Part II

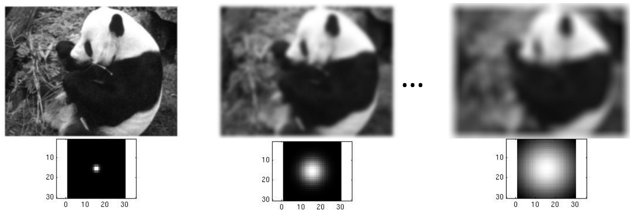
Dr Christopher Town

5. Multi-scale feature detection and matching.

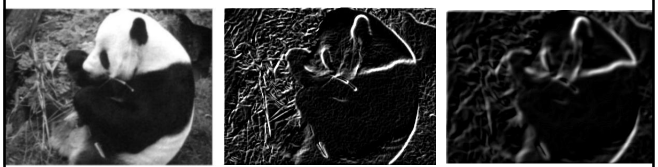
UNIVERSITY OF CAMBRIDGE

Recap: Smoothing with a Gaussian

Recall: parameter σ is the “scale” / “width” / “spread” of the Gaussian kernel, and controls the amount of smoothing.



Recap: Effect of σ on derivatives



σ = 1 pixel σ = 3 pixels

The apparent structures differ depending on Gaussian’s scale parameter.

Larger values: larger scale edges detected
Smaller values: finer features detected

Multi-scale feature detection and matching

- An interesting property of edges as defined by the zero-crossings of multi-scale operators whose scale is determined by convolution with a Gaussian, is that as the Gaussian is made coarser (larger), new edges (new zero-crossings) can never appear. They can only merge and thus become fewer in number. This property is called causality. It is also sometimes called ‘monotonicity,’ or ‘the evolution property,’ or ‘nice scaling behaviour.’

- One reason why causality is important is that it ensures that features detected at a coarse scale of analysis were not spuriously created by the blurring process (convolution with a low-pass filter, which is the normal way to create a multi-scale image pyramid using a hierarchy of increasing kernel sizes). One would like to know that image features detected at a certain scale are “grounded” in image detail at the finest resolution.

Multi-scale feature detection and matching

- For purposes of edge detection at multiple scales, a plot showing the evolution of zero-crossings in the image after convolution with a linear operator, as a function of the scale of the operator which sets the scale (i.e. the width of the Gaussian), is called scale-space.

- Scale-space has a dimensionality that is one greater than the dimensionality of the signal. Thus a 1D waveform projects into a 2D scale-space. An image projects into a 3D scale space, with its zero-crossings (edges) forming surfaces that evolve as the scale of the Gaussian changes. The scale of the Gaussian, usually denoted by σ, creates the added dimension.

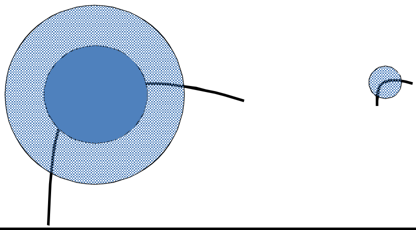
Multi-scale feature detection and matching

- A mapping of the edges in an image (its zero-crossings after such filtering operations, evolving with operator scale) is called a scale-space fingerprint. Several theorems exist called “fingerprint theorems” showing that the Gaussian blurring operator uniquely possesses the property of causality. In this respect, it is a preferred edge detector when combined with a bandpass or differentiating kernel such as the Laplacian.

- However, other non-linear operators have advantageous properties, such as reduced noise-sensitivity and greater applicability for extracting features that are more complicated (and more useful) than mere edges.

Scale Invariant Detection

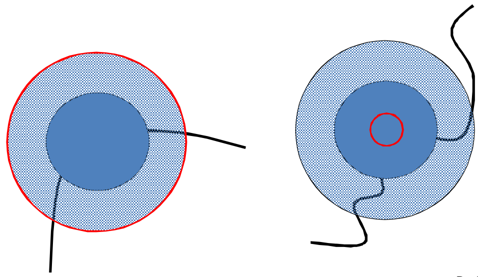
- Consider regions (e.g. circles) of different sizes around a point
- Regions of corresponding sizes will look the same in both images



Dr Chris Town

Scale Invariant Detection

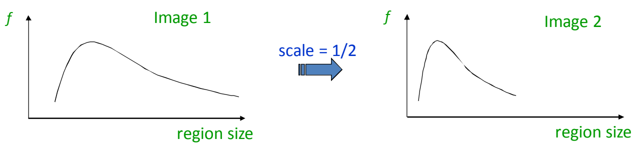
- The problem: how do we choose corresponding circles ***independently*** in each image?



Dr Chris Town

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)



Dr Chris Town

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 *extrema* of suitable functions in *scale* and in *space* (over the image)

Methods:

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

Dr Chris Town

Image Matching

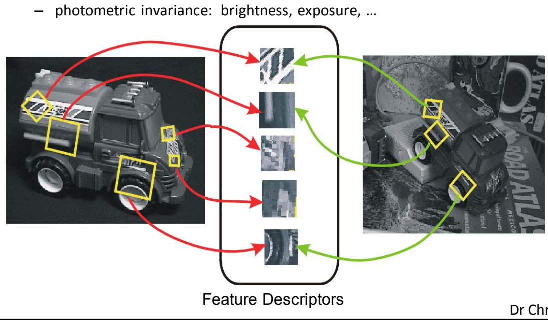
Dr Chris Town

Invariant local features

-Algorithm for finding points and representing their patches should produce similar results even when conditions vary

-Buzzword is "invariance"

- geometric invariance: translation, rotation, scale
- photometric invariance: brightness, exposure, …

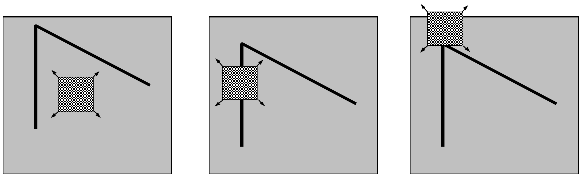


Dr Chris Town

Feature detection

Local measure of feature uniqueness

– How does the window change when you shift it?



"flat" region: no change in all directions

"edge": no change along the edge direction

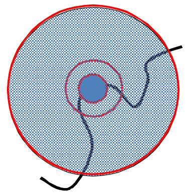
"corner": significant change in all directions

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

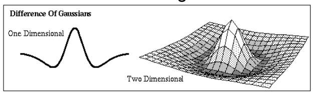
Dr Chris Town

Scale invariant detection

Suppose you're looking for corners

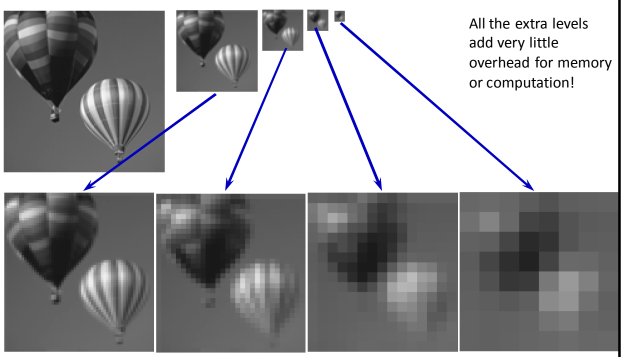


Key idea: find scale that gives local maximum response in both position and scale: use a Laplacian approximated by difference between two Gaussian filtered images with different sigmas)



Dr Chris Town

Gaussian Pyramid

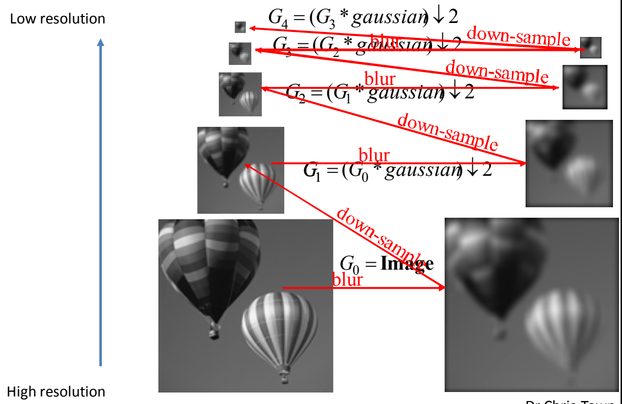


All the extra levels add very little overhead for memory or computation!

Source: Irani & Basri

Dr Chris Town

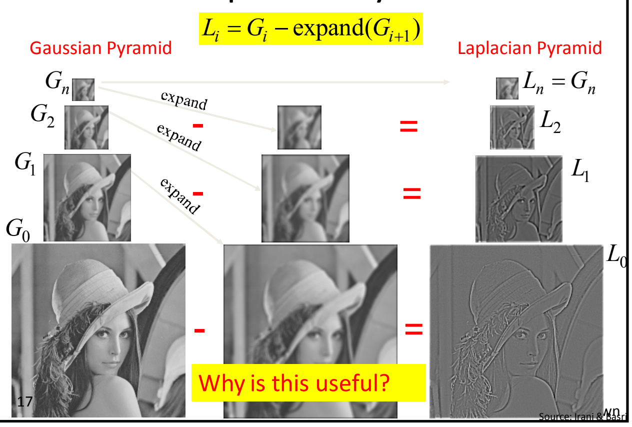
The Gaussian Pyramid



Source: Irani & Basri

Dr Chris Town

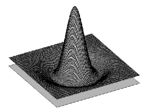
The Laplacian Pyramid



Laplacian ~ Difference of Gaussian

DoG = Difference of Gaussians

Cheap approximation – no derivatives needed.

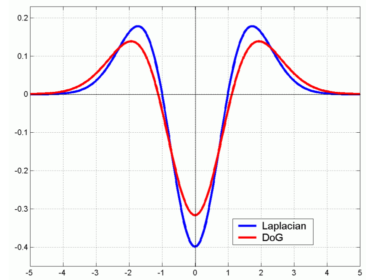


B. Leibe

Dr Chris Town

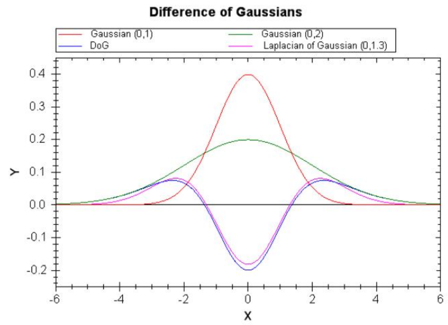
DoG approximation to LoG

- We can efficiently approximate the (scale-normalised) Laplacian of a Gaussian with a difference of Gaussians:



B. Leibe

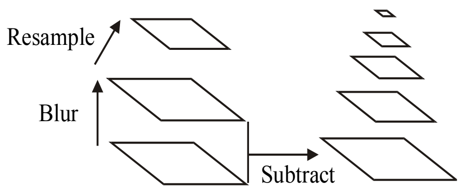
Dr Chris Town



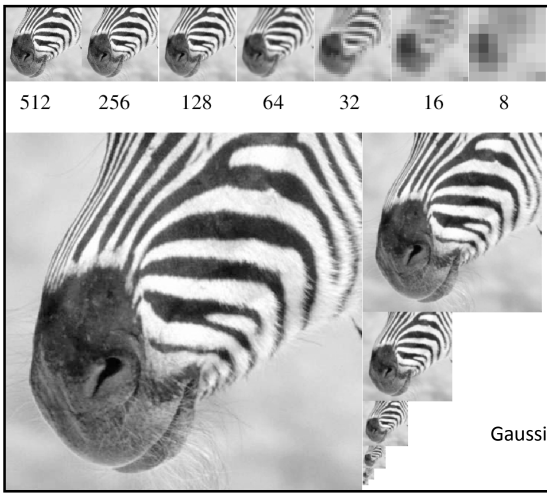
Dr Chris Town

Scale-Space Pyramid

- Multiple scales must be examined to identify scale-invariant features
- An efficient function is to compute the Difference of Gaussian (DOG) pyramid (Burt & Adelson, 1983)

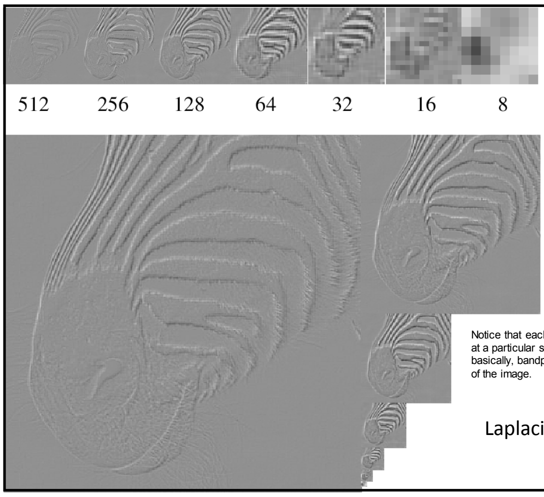


Dr Chris Town



Gaussian pyramid

Dr Chris Town

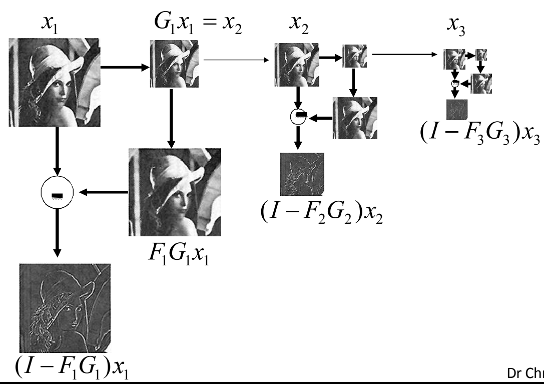


Notice that each layer shows detail at a particular scale --- these are, basically, bandpass filtered versions of the image.

Laplacian pyramid

Dr Chris Town

Laplacian pyramid algorithm



Dr Chris Town

Showing, at full resolution, the information captured at each level of a Gaussian (top) and Laplacian (bottom) pyramid.

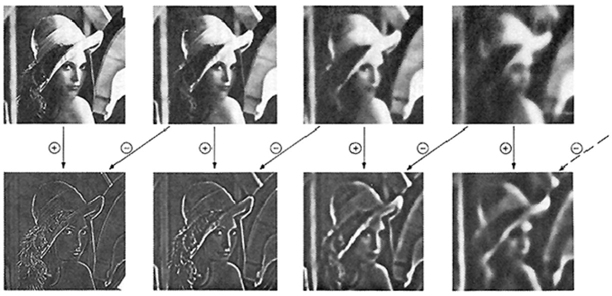
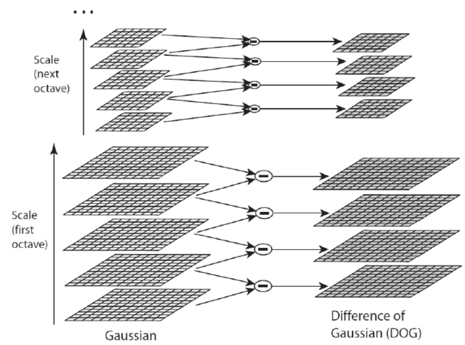


Fig 5. First four levels of the Gaussian and Laplacian pyramid. Gaussian images, upper row, were obtained by expanding pyramid arrays (Fig. 4) through Gaussian interpolation. Each level of the Laplacian pyramid is the difference between the corresponding and next higher levels of the Gaussian pyramid.

http://www-bcs.mit.edu/people/adelson/pub_pdfs/pyramid83.pdf IEEE TRANSACTIONS ON COMMUNICATIONS, VOL. COM-31, NO. 4, APRIL 1983

SIFT – Scale Invariant Feature Transform



From: David Lowe (2004)

Dr Chris Town

DoG approximates scale-normalised Laplacian of a Gaussian

$$\sigma^2 \nabla^2 G$$

$$DoG(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

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

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G \qquad \text{(heat diffusion equation)}$$

If we consider the finite difference approximation to $\frac{\partial G}{\partial \sigma}$ at neighbouring scales $k\sigma$ and σ

$$\frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}$$

then by multiplying by $k\sigma - \sigma = (k-1)\sigma$ we get

$$G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^2 \nabla^2 G \tag{15}$$

Dr Chris Town

- The σ of the Gaussian filters smoothes the image by blurring it, which helps to eliminate noise but also eliminates detail (low-pass filter in the Fourier domain). Convolution with a Gaussian followed by re-sampling is the standard technique for downsampling images, for reasons discussed at the start of this section.
- The constant k is a multiplicative factor between neighbouring Gaussian-blurred images whose difference we wish to compute to extract stable features. SIFT does this by comparing each pixel in the DoG images to its eight neighbours at the same scale and nine corresponding neighbouring pixels in each of the adjacent scales (pyramid levels).

Dr Chris Town

Octave increment in scale of the Gaussian Pyramid

$$\sigma_{i+1} = 2\sigma_i$$

followed by factor-of-two downsampling (for efficiency). To achieve better performance, each octave **i** is further divided into **s** intervals.

Remember that we defined neighbouring scales as

$$DoG(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

So starting with some σ_0, the next scale parameter will be $k\sigma_0$, followed by $kk\sigma_0$ etc., so that after **s** sub-levels of the pyramid we have a complete octave with

$$\overset{s}{k}\sigma_0 = 2\sigma_0$$

Therefore $\quad k = 2^{1/s}$

Dr Chris Town

$$k = 2^{1/s}$$

and the value of σ at octave i and interval n of the pyramid is given by

$$\sigma(i, n) = \sigma_0 \, 2^{i+n/s}; \; n \in [0, s-1]$$

A value of $s = 3$ was found by Lowe to provide a good accuracy vs efficiency trade-off. The number of octaves depends on original image resolution.

Dr Chris Town

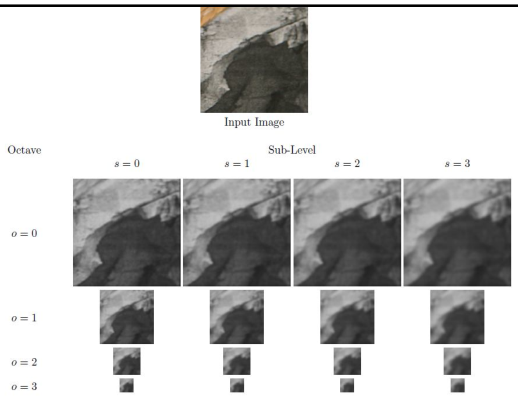


Figure 2: An example of a Gaussian scale-space generated from the input image. *The smoothing is increased from left to right by increasing the variance of the Gaussian kernel which is convolved with the grayscale image. Each level, from top to bottom, is produced by downsampling by two from the previous level before further smoothing.*

Neill Campbell

Dr Chris Town

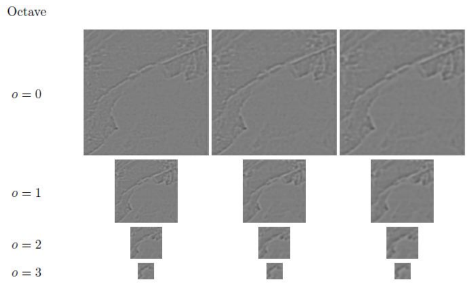


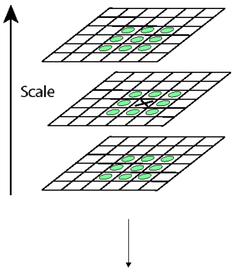
Figure 3: The 'Difference of Gaussians' (DoG) scale-space corresponding to the Gaussian scale-space of Figure 2. *At each level in the Gaussian scale-space, the difference between successive images provides the DoG scale-space.*

Neill Campbell

Dr Chris Town

Key point localization with DoG

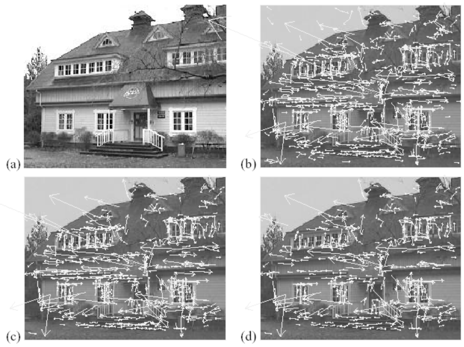
- Detect extrema of difference-of-Gaussian (DoG) in scale space
- Then reject points with low contrast (threshold)
- Eliminate edge responses



Slide credit: David Lowe

Dr Chris Town

Example of Keypoint Detection



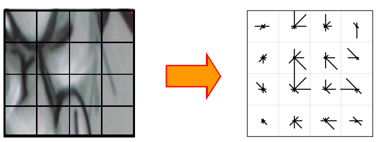
(a) 233x189 image
(b) 832 DoG extrema
(c) 729 left after peak value threshold
(d) 536 left after testing ratio of principle curvatures (removing edge responses)

Slide credit: David Lowe

Dr Chris Town

Feature Descriptors: SIFT

- Scale Invariant Feature Transform
- Descriptor computation:
  - Divide patch into 4x4 sub-patches: 16 cells
  - Compute histogram of gradient orientations (8 reference angles) for all pixels inside each sub-patch
  - Resulting descriptor: 4x4x8 = 128 dimensions



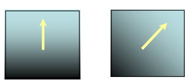
David G. Lowe. "Distinctive image features from scale-invariant keypoints." *IJCV* 60 (2), pp. 91-110, 2004.

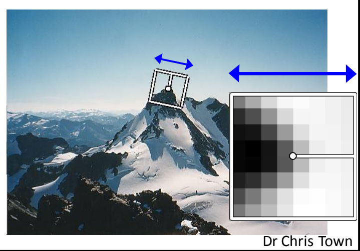
Slide credit: Svetlana Lazebnik

Dr Chris Town

Rotation Invariant Descriptors

- Find local orientation
  - Dominant direction of gradient for the image patch
- Rotate patch according to this angle
  - This puts the patches into a canonical orientation.





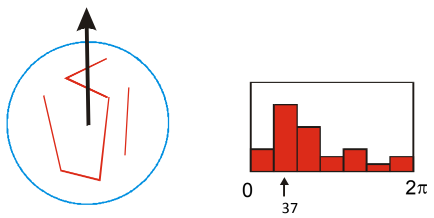
Slide credit: Svetlana Lazebnik, Matthew Brown

Dr Chris Town

Orientation Normalisation: Computation

[Lowe, SIFT, 1999]

- Compute orientation histogram
- Select dominant orientation
- Normalise: rotate to fixed orientation

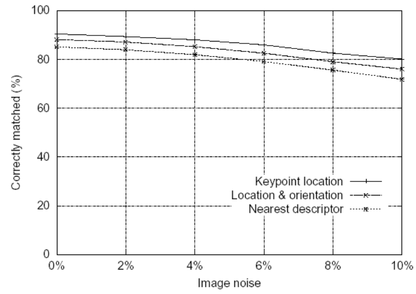


Slide adapted from David Lowe

Dr Chris Town

Feature stability to noise

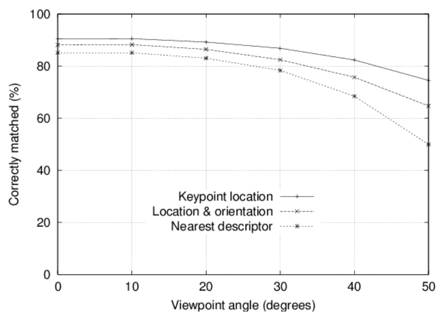
- Match features after random change in image scale & orientation, with differing levels of image noise
- Find nearest neighbor in database of 30,000 features



Dr Chris Town

Feature stability to affine change

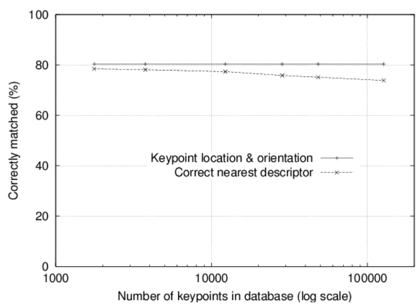
- Match features after random change in image scale & orientation, with 2% image noise, and affine distortion
- Find nearest neighbor in database of 30,000 features



Dr Chris Town

Distinctiveness of features

- Vary size of database of features, with 30 degree affine change, 2% image noise
- Measure % correct for single nearest neighbor match



Dr Chris Town

Working with SIFT Descriptors

- One image yields:
  - *n* 128-dimensional descriptors: each one is a histogram of the gradient orientations within a patch
    - [n x 128 matrix]
  - *n* scale parameters specifying the size of each patch
    - [n x 1 vector]
  - *n* orientation parameters specifying the angle of the patch
    - [n x 1 vector]
  - *n* 2D points giving positions of the patches
    - [n x 2 matrix]

Slide credit: Steve Seitz

Dr Chris Town

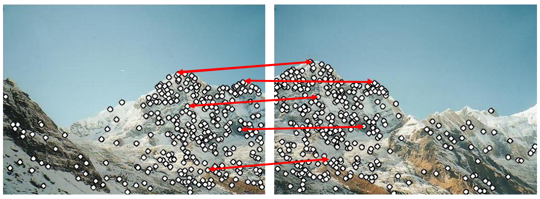
SIFT

Figure 12: The training images for two objects are shown on the left. These can be recognized in a cluttered image with extensive occlusion, shown in the middle. The results of recognition are shown on the right. A parallelogram is drawn around each recognized object showing the boundaries of the original training image under the affine transformation solved for during recognition. Smaller squares indicate the keypoints that were used for recognition.

D. Lowe, 2004

Dr Chris Town

Feature matching



Slides from Steve Seitz and Rick Szeliski

Dr Chris Town

Image stitching

Brown, Lowe, 2007

Dr Chris Town

Nearest-neighbor matching

- Solve following problem for all feature vectors, **x**:

$$\forall j \; NN(j) = \arg\min_i ||\mathbf{x}_i - \mathbf{x}_j||, \; i \neq j$$

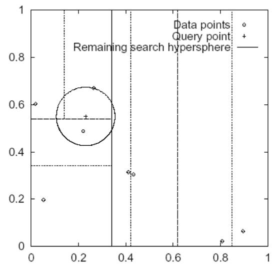
- Nearest-neighbour matching is the major computational bottleneck
  - Linear search performs dn^2 operations for n features and d dimensions
  - No exact methods are faster than linear search for d>10
  - Approximate methods can be much faster, but at the cost of missing some correct matches. Failure rate gets worse for large datasets.

Dr Chris Town

Approximate k-d tree matching

Key idea:

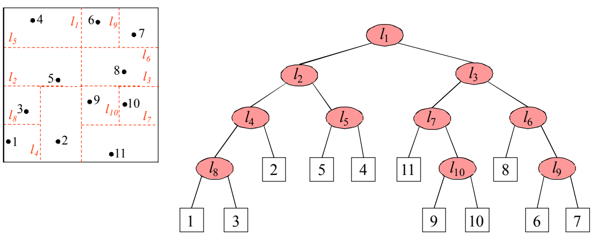
- Search k-d tree bins in order of distance from query
- Requires use of a priority queue



Dr Chris Town

K-d tree construction

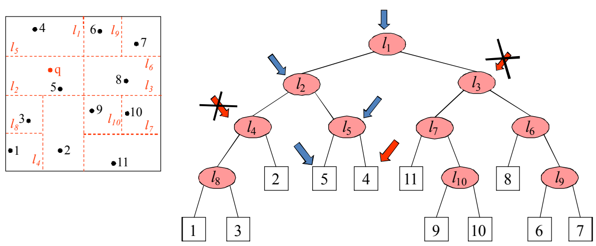
Simple 2D example



Slide credit: Anna Atramentov

Dr Chris Town

K-d tree query

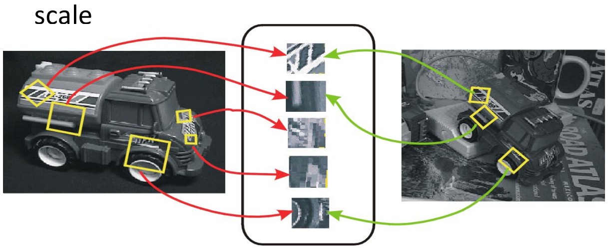


Slide credit: Anna Atramentov

Dr Chris Town

Recognition with Local Features

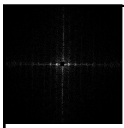
- Image content is transformed into local features that are invariant to translation, rotation, and scale

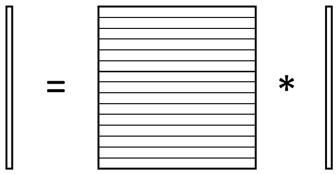


Local Features, e.g. SIFT

Slide credit: David Lowe

Fourier transform



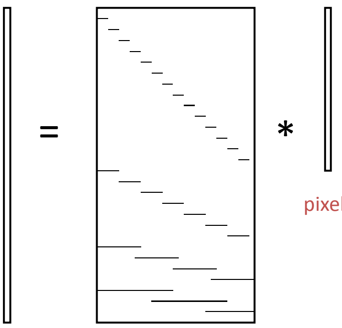


Fourier transform

Fourier bases are global: each transform coefficient depends on all pixel locations.

pixel domain image

Gaussian pyramid

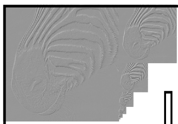


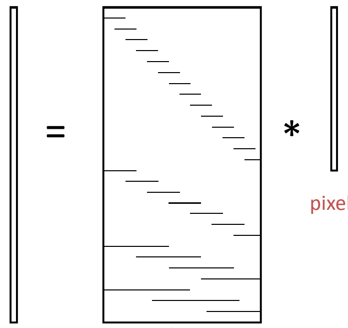
Gaussian pyramid

pixel image

Overcomplete representation.
Low-pass filters, sampled appropriately for their blur.

Laplacian pyramid





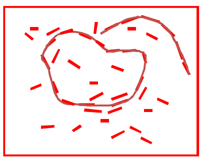
Laplacian pyramid

pixel image

Overcomplete representation.
Transformed pixels represent bandpassed image information.

Edge Fitting

- Edge Detection:
  - The process of labeling the locations in the image where the gray level's "rate of change" is high.
    - OUTPUT: "edgels" locations, direction, strength



- Edge Integration, Contour fitting:
  - The process of combining "local" and perhaps sparse and non-contiguous "edgel"-data into meaningful, long edge curves (or closed contours) for segmentation
    - OUTPUT: edges/curves consistent with the local data

Active Contours ("snakes")

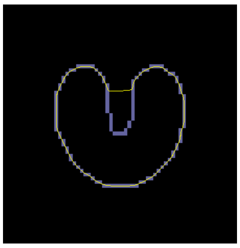
$$\arg\min \int \left((M - I)^2 + \lambda (M_{xx})^2\right) dx$$

where M is the shape model, and I is the image data

Thus we have the combination of two factors: a *data term* and a *cost term* (the latter sometimes also called a *smoothness term* or an *energy term*), which are in contention, in the following sense: we could fit the available edge data with arbitrarily high precision, if we used a model with enough complexity; but simpler models are generally more useful and credible than overly complex models (which "over-fit" the data).

Framework for snakes

- A higher level process or a user initialises any curve *close to the object boundary.*
- The snake then starts *deforming* and moving towards the desired object boundary.
- In the end it completely "shrink-wraps" around the object.



(Diagram courtesy "Snakes, shapes, gradient vector flow", Xu, Prince)

Dr Chris Town

Modeling

- The contour is defined in the (x, y) plane of an image as a parametric curve

  $\mathbf{v}(s)=(x(s),\ y(s))$

- Contour is said to **possess an energy** ($\boldsymbol{E_{snake}}$) which is defined as the sum of the three energy terms.

$$E_{snake} = E_{internal} + E_{external} + E_{constraint}$$

- The energy terms are defined in a way such that the final position of the contour will have minimum energy (E_{min})
- Therefore our problem of detecting objects reduces to an energy minimisation problem.

A. Poonawala

Dr Chris Town

Internal Energy (E_{int})

- Depends on the intrinsic properties of the curve.
- Sum of **elastic energy** and **bending energy**.

Elastic Energy ($\boldsymbol{E_{elastic}}$):

- The curve is treated as an elastic rubber band possessing elastic potential energy.
- It discourages stretching by introducing tension.

$$E_{elastic} = \frac{1}{2}\int_s \alpha(s)\,|v_s|^2\,ds \qquad v_s = \frac{dv(s)}{ds}$$

- Weight α(s) allows us to control elastic energy along different parts of the contour. Considered to be constant α for many applications.
- Responsible for shrinking of the contour.

A. Poonawala

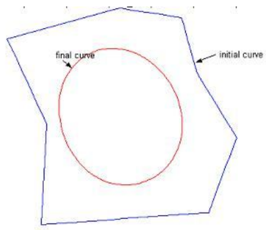
Dr Chris Town

Elastic force

- Generated by elastic potential energy of the curve.

$$F_{elastic} = \alpha v_{ss}$$

- Characteristics (refer diagram)



A. Poonawala

Dr Chris Town

Bending Energy ($E_{bending}$):

- The snake is also considered to behave like a thin metal strip giving rise to bending energy.
- It is defined as sum of squared curvature of the contour.

$$E_{bending} = \frac{1}{2}\int_s \beta(s)\,|v_{ss}|^2\,ds$$

- β**(s)** plays a similar role to α**(s).**
- Bending energy is minimum for a circle.
- Total internal energy of the snake can be defined as

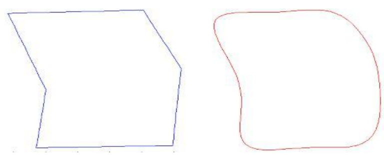
$$E_{int} = E_{elastic} + E_{bending} = \int_s \frac{1}{2}\left(\alpha|v_s|^2 + \beta|v_{ss}|^2\right)ds$$

A. Poonawala

Dr Chris Town

Bending force

- Generated by the bending energy of the contour.
- Characteristics (refer diagram):



Initial curve
(High bending energy)

Final curve deformed by bending force. (low bending energy)

- Thus the bending energy tries to smooth out the curve.

A. Poonawala

Dr Chris Town

External energy of the contour (E_{ext})

- Image fitting

$$E_{ext} = \int_s E_{image}(v(s))ds$$

For example

- $E_{edge} = -|\nabla I(x,y)|^2$
- $E_{edge} = -\left|G_\sigma * \nabla^2 I\right|^2$

A. Poonawala

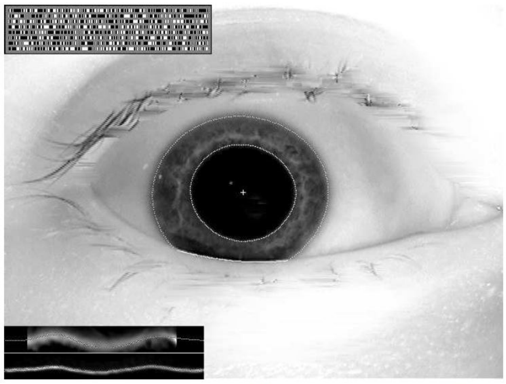


Figure 5 *Active contours are deformable yet constrained shape models. The "snakes" in the box show radial edge gradients at the iris boundaries, and active contour approximations (dotted curves).*

leafmv.mpg

dancemv.mpg

http://www.robots.ox.ac.uk/~ab/dynamics.html

2D Gabor "Logons;" Quadrature pair wavelets

$$f(x) = \exp(-i\mu_0(x - x_0))\exp(-(x - x_0)^2/\alpha^2)$$

$$F(\mu) = \exp(-ix_0(\mu - \mu_0))\exp(-(\mu - \mu_0)^2\alpha^2)$$

Note that for the case of a wavelet $f(x)$ centred on the origin ($x_0 = 0$), its Fourier Transform $F(\mu)$ is simply a Gaussian centred on the modulation frequency $\mu = \mu_0$, and whose width is $1/\alpha$, the reciprocal of the wavelet's space constant. This shows that it acts as a bandpass filter, passing only those frequencies that are within about $\pm\frac{1}{\alpha}$ of the wavelet's modulation frequency μ_0.

Generalisation of wavelet Logons to 2D for image analysis

Two-dimensional Gabor wavelets have the functional form:

$$f(x,y) = e^{-\left[(x-x_0)^2/\alpha^2 + (y-y_0)^2/\beta^2\right]} e^{-i[u_0(x-x_0)+v_0(y-y_0)]}$$

where (x_0, y_0) specify position in the image, (α, β) specify effective width and length, and (u_0, v_0) specify modulation, which has spatial frequency $\omega_0 = \sqrt{u_0^2 + v_0^2}$ and direction $\theta_0 = \arctan(v_0/u_0)$. (A further degree-of-freedom not included above is the relative orientation of the elliptic Gaussian envelope, which creates cross-terms in xy.)

2D Fourier transform $F(u,v)$

$$F(u,v) = e^{-\left[(u-u_0)^2\alpha^2 + (v-v_0)^2\beta^2\right]} e^{-i[x_0(u-u_0)+y_0(v-v_0)]}$$

The real part of one member of the 2D Gabor filter family, centred at the origin $(x_0, y_0) = (0,0)$ and with unity aspect ratio $\beta/\alpha = 1$

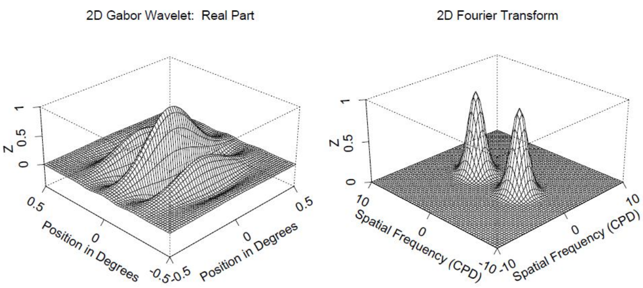


Figure 6 *The real part of a 2D Gabor wavelet, and its 2D Fourier transform.*

Generating Functions

By appropriately parameterising them for dilation, rotation, and translation, 2D Gabor wavelets can form a complete self-similar (but non-orthogonal) expansion basis for images.

$$\Psi_{mpq\theta}(x, y) = 2^{-2m}\Psi(x', y')$$

where the substituted variables (x', y') incorporate dilations in size by 2^{-m}, translations in position (p, q), and rotations through orientation θ:

$$x' = 2^{-m}[x\cos(\theta) + y\sin(\theta)] - p$$
$$y' = 2^{-m}[-x\sin(\theta) + y\cos(\theta)] - q$$

Since the wavelets are dilates, translates, and rotates of each other, such a transform seeks to extract image structure in a way that may be invariant to dilation, translation, and rotation of the original image or pattern.

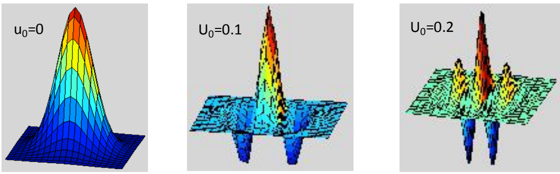
Dr Chris Town

Reconstruction of Lena: 25, 100, 500, and 10,000 Two-Dimensional Gabor Wavelets

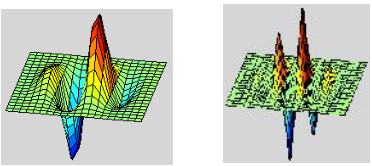
Dr Chris Town

Gabor wavelets

$$\psi_c(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \cos(2\pi u_0 x)$$

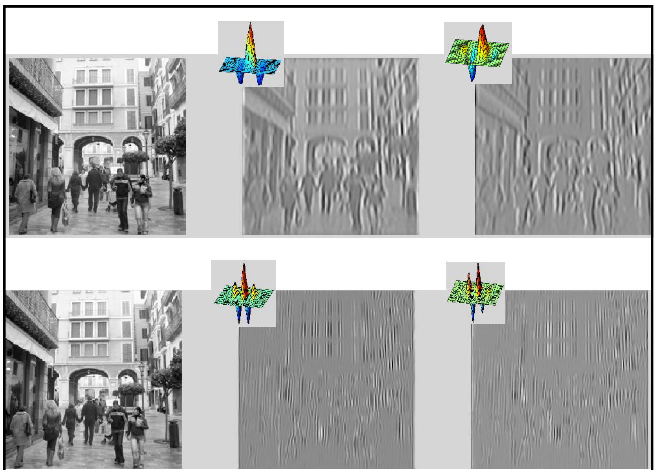


$$\psi_s(x,y) = e^{-\frac{x^2+y^2}{2\sigma^2}} \sin(2\pi u_0 x)$$



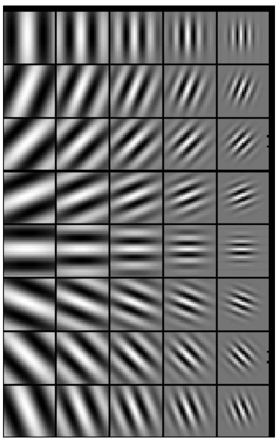
A. Torralba

Dr Chris Town



A. Torralba

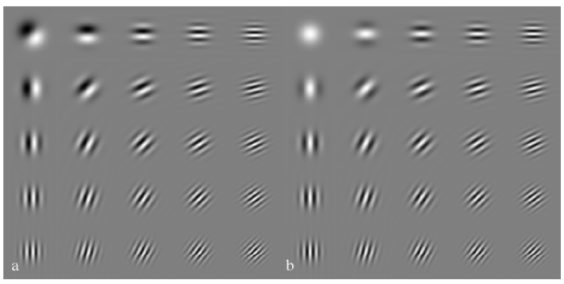
Dr Chris Town



Dilation and rotation

Dr Chris Town

Frequency, orientation and symmetry (phase)



Dr Chris Town

Gabor-Heisenberg-Weyl Uncertainty Principle

If we define the "effective support" of a function $f(x)$ by its normalized variance, or the normalized second-moment

$$(\Delta x)^2 = \frac{\int_{-\infty}^{+\infty} f(x)f^*(x)(x-x_0)^2 dx}{\int_{-\infty}^{+\infty} f(x)f^*(x)dx}$$

where x_0 is the mean value, or first-moment, of the function

$$x_0 = \int_{-\infty}^{+\infty} x f(x)f^*(x)dx$$

and if we similarly define the effective support of the Fourier Transform $F(\mu)$ of the function by its normalized variance in the Fourier domain

$$(\Delta\mu)^2 = \frac{\int_{-\infty}^{+\infty} F(\mu)F^*(\mu)(\mu-\mu_0)^2 d\mu}{\int_{-\infty}^{+\infty} F(\mu)F^*(\mu)d\mu}$$

where μ_0 is the mean value, or first-moment, of the Fourier transform $F(\mu)$

$$\mu_0 = \int_{-\infty}^{+\infty} \mu F(\mu)F^*(\mu)d\mu$$

Dr Chris Town

Gabor-Heisenberg-Weyl Uncertainty Principle

then it can be proven (by Schwartz Inequality arguments) that there exists a fundamental lower bound on the product of these two "spreads," *regardless* of the function $f(x)$!

$$\boxed{(\Delta x)(\Delta\mu) \geq \frac{1}{4\pi}}$$

The unique family of signals that actually achieve the lower bound in the Gabor-Heisenberg-Weyl Uncertainty Relation are the complex exponentials multiplied by Gaussians. These are sometimes referred to as "Gabor wavelets:"

$$f(x) = e^{-i\mu_0 x}e^{-(x-x_0)^2/a^2}$$

Dr Chris Town

Unification of Domains

$$f(x) = e^{-i\mu_0 x}e^{-(x-x_0)^2/a^2}$$

The single parameter a (the space-constant in the Gaussian term) actually builds a continuous bridge between the two domains: if the parameter a is made very large, then the second exponential above approaches 1.0, and so in the limit our expansion basis becomes

$$\lim_{a\to\infty} f(x) = e^{-i\mu_0 x}$$

the ordinary Fourier basis! If the parameter a is instead made very small, the Gaussian term becomes the approximation to a delta function at location x_o, and so our expansion basis implements pure space-domain sampling:

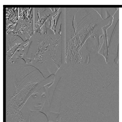
$$\lim_{\mu_0, a\to 0} f(x) = \delta(x-x_0)$$

Hence the Gabor expansion basis "contains" both domains at once. It allows us to make a continuous deformation that selects a representation lying anywhere on a one-parameter continuum between two domains that were hitherto distinct

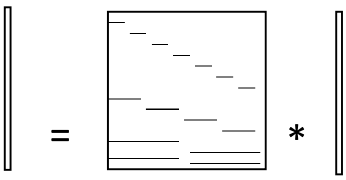
Dr Chris Town

Illustration of Facial Feature Detection by Quadrature Filter Energy. Left panel: original image. Right panel (clockwise from top left): the real part after 2D Gabor wavelet convolution; the imaginary part; the modulus energy; and this energy superimposed on the original (faint) image, illustrating successful feature localization.

Dr Chris Town



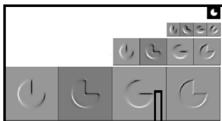
Wavelet (QMF) transform



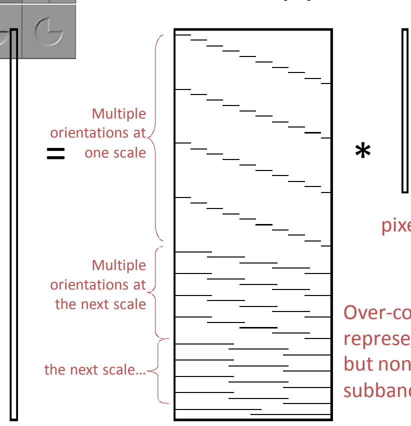
Wavelet pyramid = * pixel image

Ortho-normal transform (like Fourier transform), but with localized basis functions.

Dr Chris Town



Steerable pyramid



Steerable pyramid = * pixel image

Multiple orientations at one scale

Multiple orientations at the next scale

the next scale…

Over-complete representation, but non-aliased subbands.

Dr Chris Town