Introduction to Computer Vision

By Ali Zaidi

References

- Professor Guang-Zhong Yang's Lectures
- Computer Vision @ Wikipedia
- The Computer Vision Homepage
- Machine Vision (McGraw-Hill Series in Computer Science) by Ramesh Jain, Rangachar Kasturi, Brian G. Schunck
- Image Processing, Analysis, and Machine Vision by Milan Sonka, Vaclav Hlavac, Roger Boyle

What will you learn?

- What is Computer Vision? What are Vision Systems?
- Low level Computer Vision: Preprocessing and How to 'prepare' an image for more intelligent tasks?
- Detecting objects - Part 1: Edge based segmentation
- Detecting objects - Part 2: Region based segmentation (Blob detection)
- Detecting objects - Part 3: Detecting known shapes

Batteries Included !!!

What you will NOT learn

- Fourier based techniques
- Motion and Optical flow based techniques
- 3D reconstruction and its derivatives (Computational Stereo, Photometric Stereo)

Lots of other stuff !!!

What will you learn?

- What is Computer Vision? What are Vision Systems?
- Low level Computer Vision: Preprocessing and How to 'prepare' an image for more intelligent tasks?
- Detecting objects - Part 1: Edge based segmentation
- Detecting objects - Part 2: Region based segmentation (Blob detection)
- Detecting objects - Part 3: Detecting known shapes

Batteries Included !!!

Graphics, Image Processing and Computer Vision

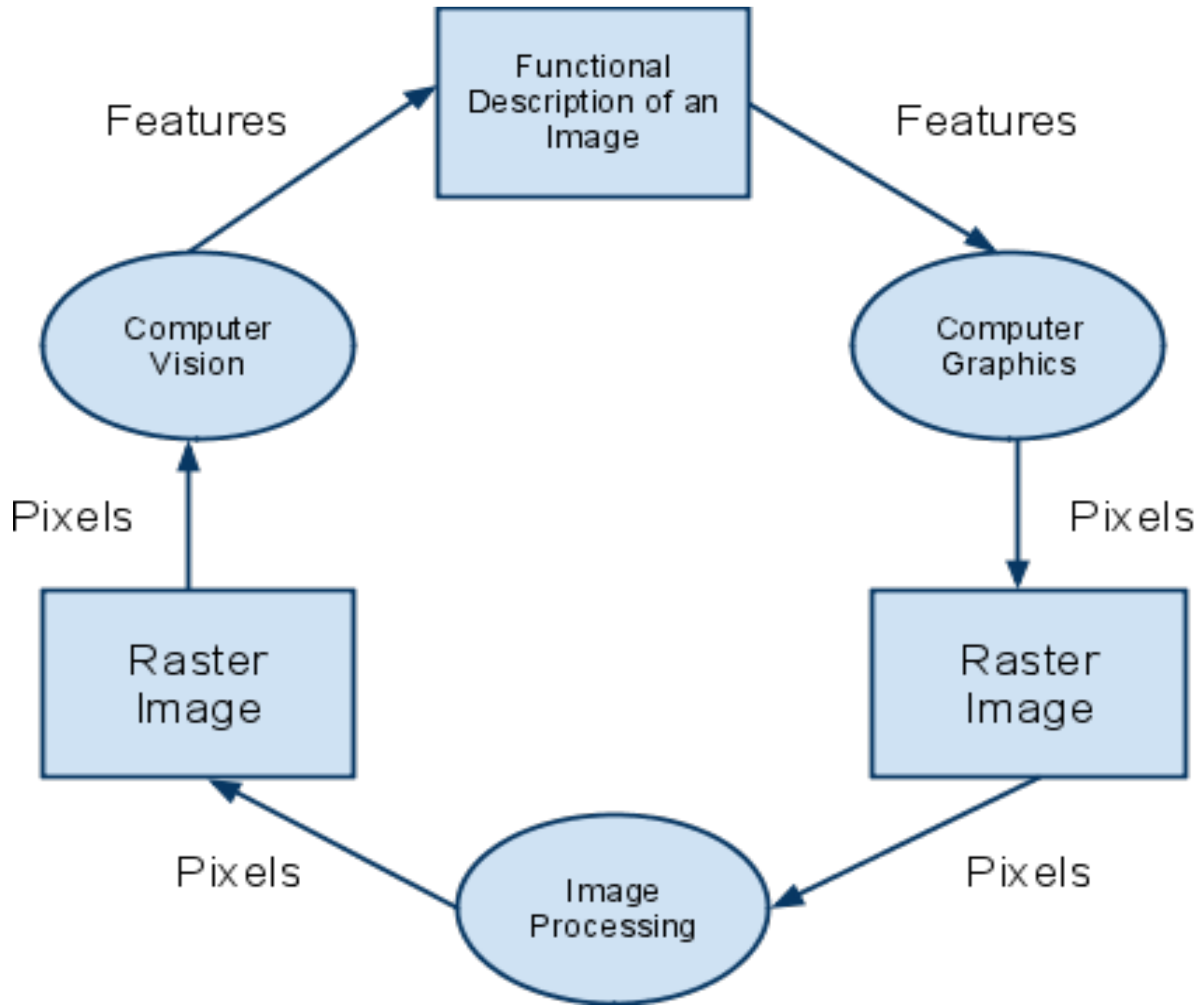


Image Processing

- Removal of defects such as scratches or other noise
- Improvement of contrast
- Removal of camera blur
- Removal of motion blur
- Enhancement of structure
- Enhancement of colour

Computer Vision

- Determining the type of an object in the picture
- Assessing an object for quality
- Breaking a picture into different parts
- Constructing a 3D representation of an object
- Extracting a line representation of an object
- Reasoning about a scene to deduce hidde properties

Vision Systems

Application Oriented Vision Systems

- Vision based robots
- Quality assessment systems

General Purpose Vision Systems

- Model the micro structure of the brain
- High level behavioural models

Questions about Vision Systems

- What information is sought from the image
- How is it manifest in the image
- What a-priori knowledge is required to recover the information
- What is the nature of the computational process
- How should the required information be represented

Intrinsic characteristics

Problem: Locate a house among trees
Straight lines are an intrinsic characteristic of houses, not trees

Problem: Separate land from sea in an aerial image
Uniform reflectivity is an intrinsic characteristic of sea, not land

(Intrinsic characteristics must be matched to known techniques)

Prior Knowledge

Scene Model:
  Object Features, Smoothness, Convexity

Illumination Model:
  Position of Light Source, Surface Reflectance

Sensor Model:
  Camera Position, Optical Performance, Digitisation Error

Other Factors

Resource Limitations (time to process the image)

- Real time
- Interactive
- Batch

Knowledge Representation

- Encoding the prior knowledge
- Interfacing with the user

Levels of Vision

Low Level Vision
  Gradient (edge, corner), depth, optical flow

Intermediate Level Vision
  Contours, regions

High Level Vision
  Objects, and hidden information

Principal of Least Commitment

- From low level to high level of processing, we remove information
- For computational efficiency, we wish to remove as much information as possible
- For effective vision, we need to retain as much information as possible

What will you learn?

- What is Computer Vision? What are Vision Systems?
- Low level Computer Vision: Preprocessing and How to 'prepare' an image for more intelligent tasks?
- Detecting objects - Part 1: Edge based segmentation
- Detecting objects - Part 2: Region based segmentation (Blob detection)
- Detecting objects - Part 3: Detecting known shapes

Batteries Included !!!

Image Processing

Pre-processing is an operation with images at the lowest level of abstraction -both input and output are intensity images

Image pre-processing methods generally involve the use of the considerable redundancy in images

Use of priori information

No knowledge about the nature of the degradation is used, only very general properties of the degradation are assumed.

Knowledge about the properties of the image acquisition device and the conditions under which the image was obtained are employed, the nature of noise is sometimes known.

Knowledge about the objects that are searched for in the image is used to guide and simplify the pre-processing process.

Elementary Noise Reduction

Elementary Noise Reduction

Local Average:

$$g(i,j) = \frac{1}{W} \sum_{(m,\ n) \in \Omega_{ij}} f(m,\ n)$$

Elementary Noise Reduction

Median Filtering

$$g(i,j) = med$$

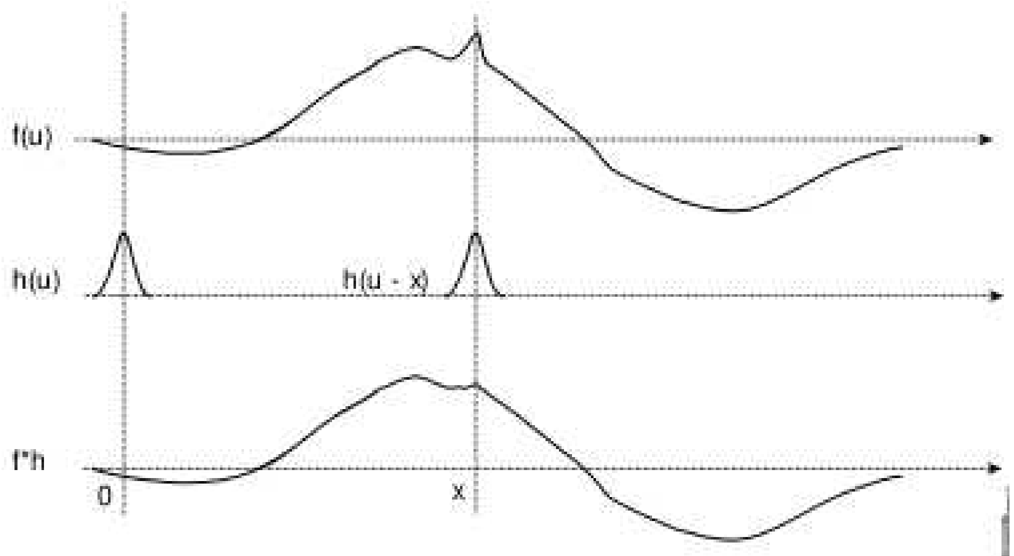
 

Convolution

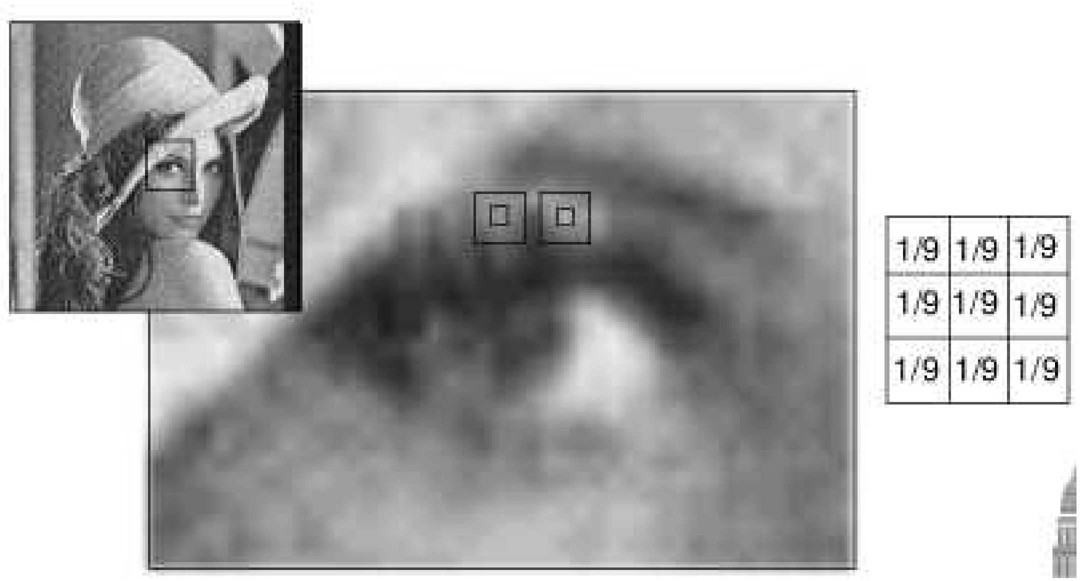
$$g(x)=\int_{-\infty}^{\infty} f(u)h(u-x)du$$

$$g(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(u,v)h(u-x,v-y)\ dudv$$

Convolution



Convolution

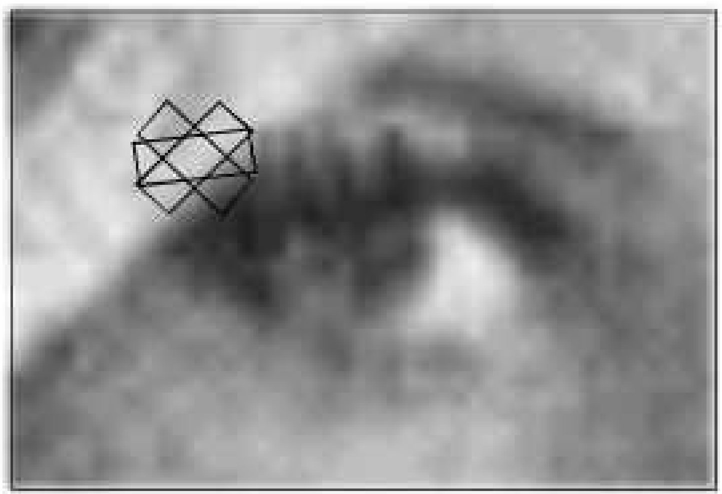


Adaptive Filtering

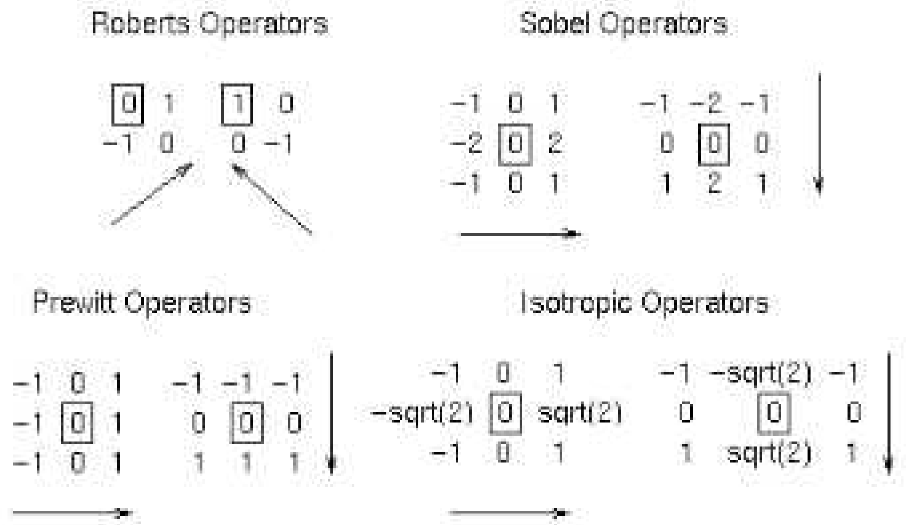
$$\delta(i,j,m,n) = \begin{cases} \dfrac{1}{|f(m,n) - f(i,j)|}, & \text{if } f(m,n) \neq f(i,j) \\ 2, & \text{otherwise.} \end{cases}$$

$$h(i,j,m,n) = 0.5 \frac{\delta(i,j,m,n)}{\sum_{(m,n) \in \Omega} \delta(i,j,m,n)}$$

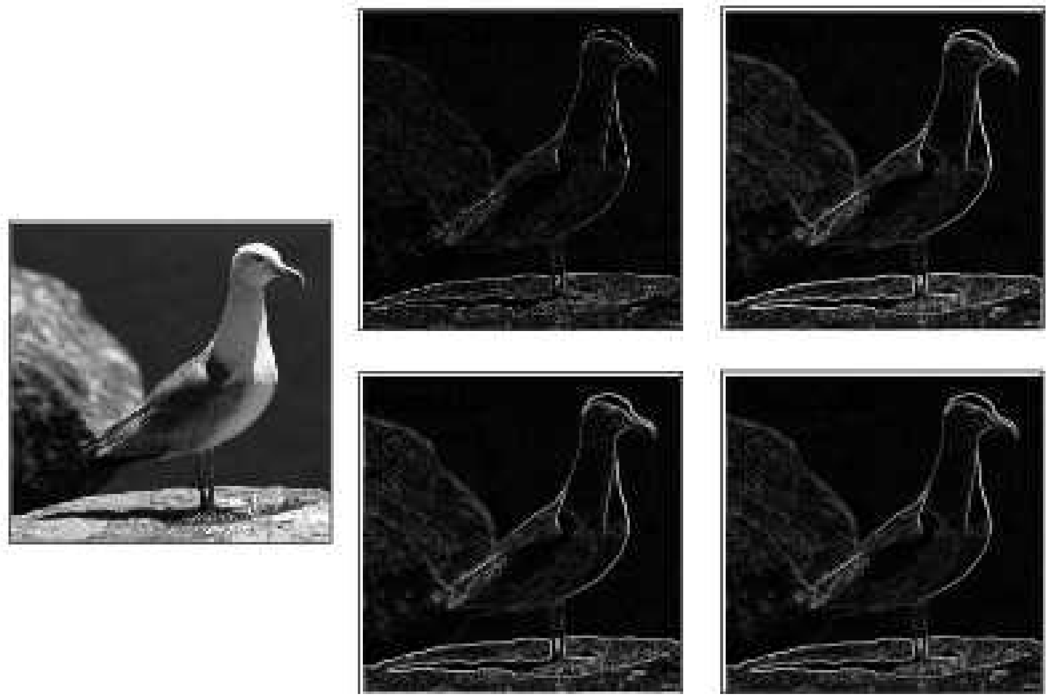
Averaging with roatating masks



Edge Detection



Edge Detection



Second Order Operator

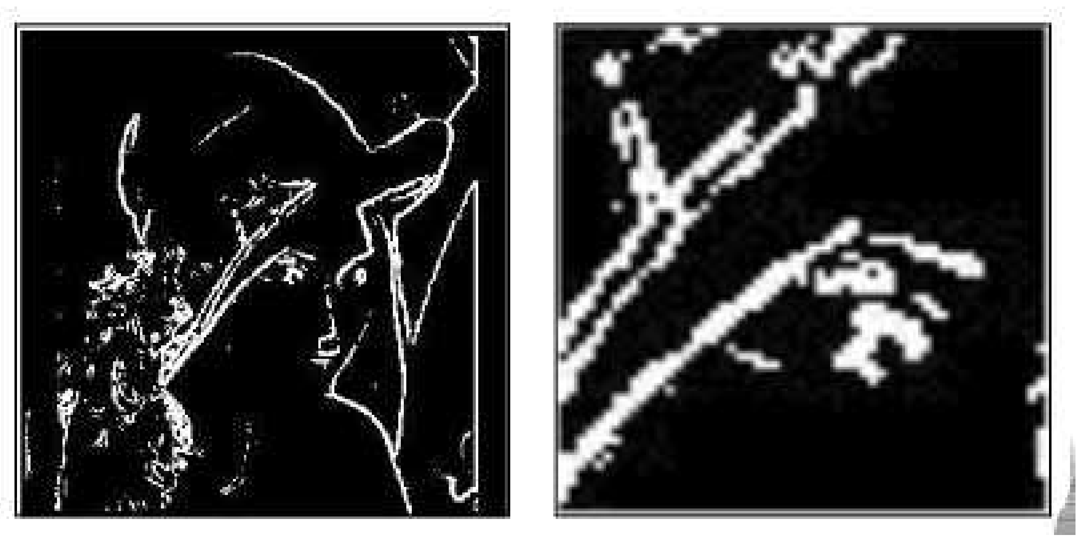
$$h = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \quad h = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -8 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

What will you learn?

- What is Computer Vision? What are Vision Systems?
- Low level Computer Vision: Preprocessing and How to 'prepare' an image for more intelligent tasks?
- Detecting objects - Part 1: Edge based segmentation
- Detecting objects - Part 2: Region based segmentation (Blob detection)
- Detecting objects - Part 3: Detecting known shapes

Batteries Included !!!

Edge based segmentation

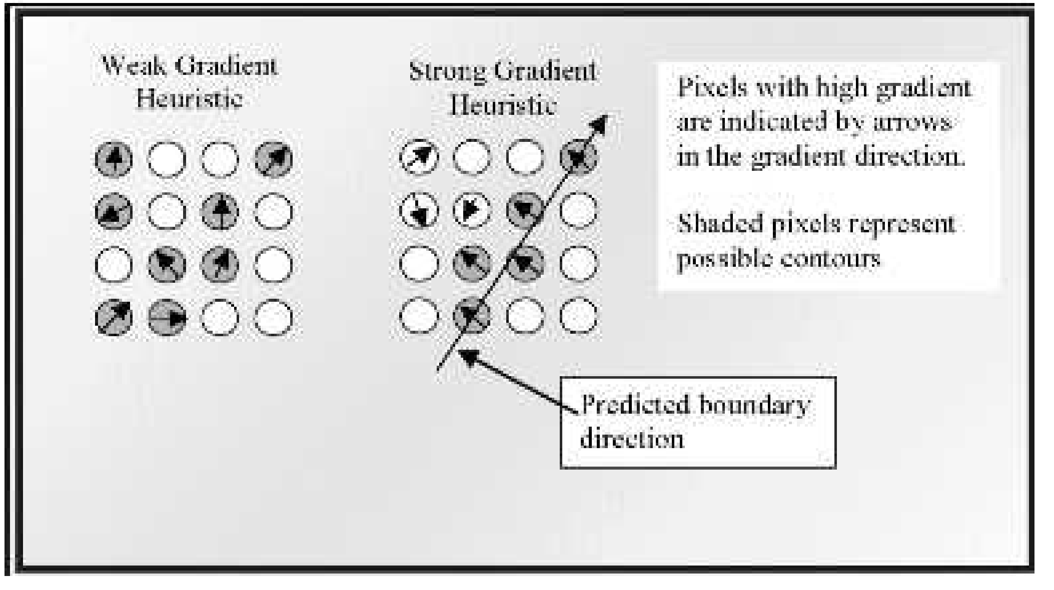


Edge based segmentation

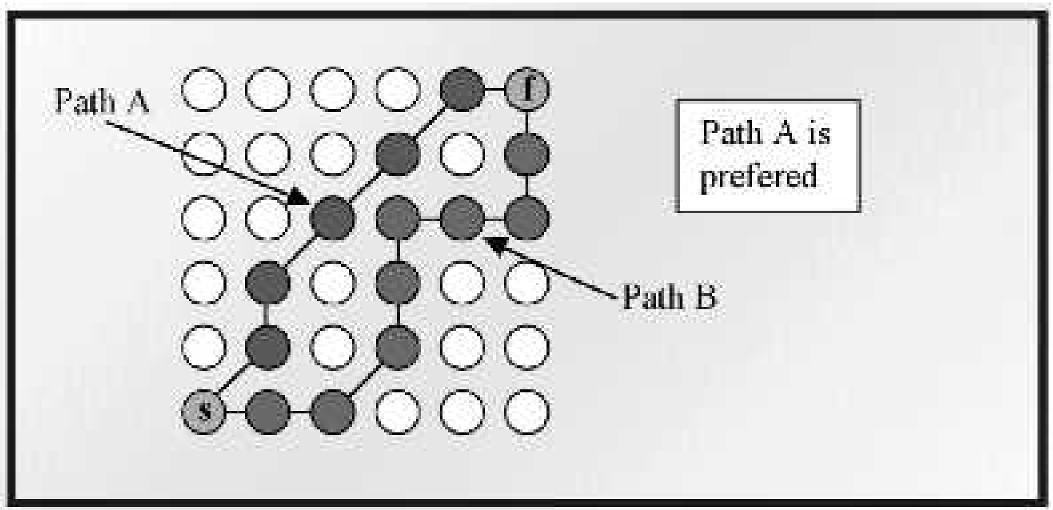
$$\left[\varphi(x_i, y_i) - \varphi(x_i, y_j)\right] \bmod 2\pi < \theta$$

Edge direction and gradient

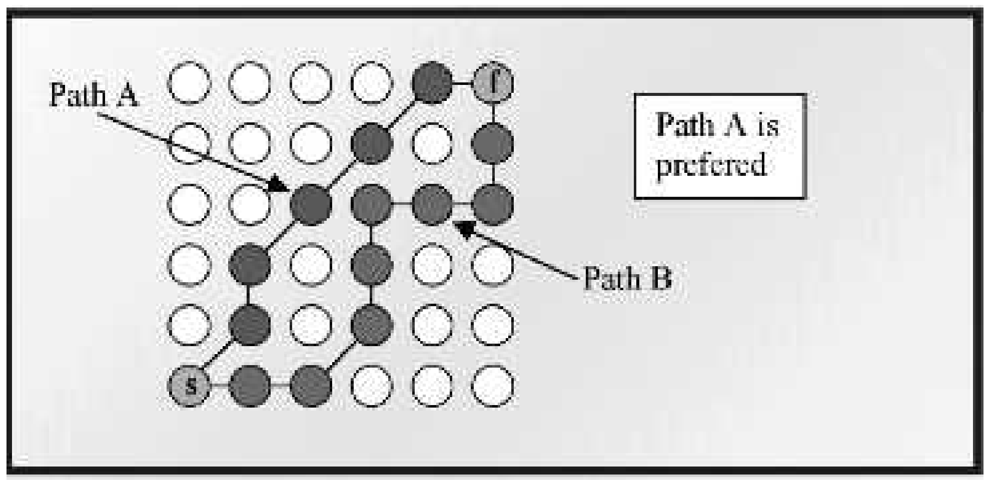
Graph searching



Shortest path heuristics



Gradient Weighting



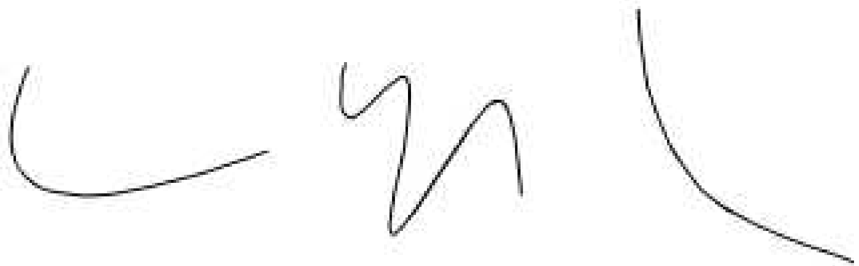
$$w(x, y) = g(x, y) / g_{\max}$$

Active Contours

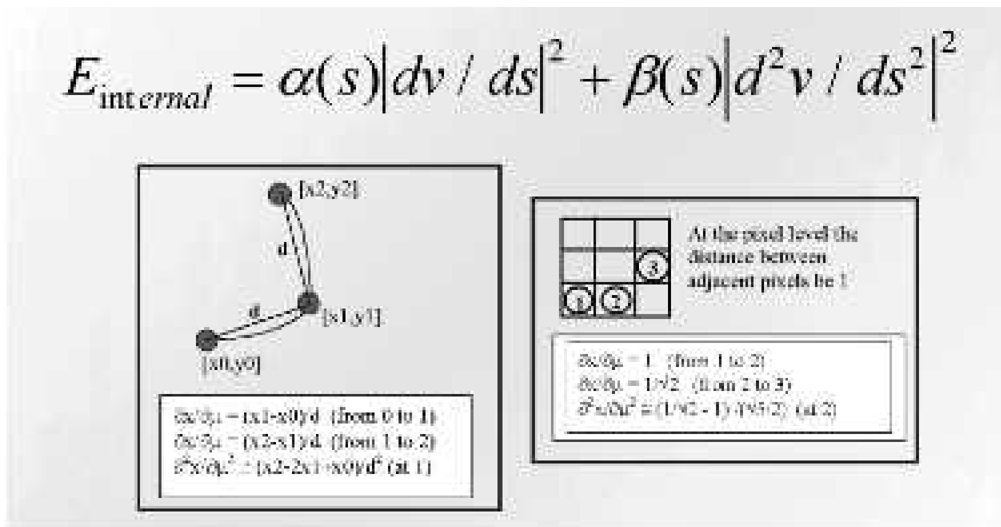
Energy minimising spline

$$E_{snake} = \int_0^1 (E_{internal} v(s) + E_{image} v(s) + E_{constraints} v(s)) ds$$

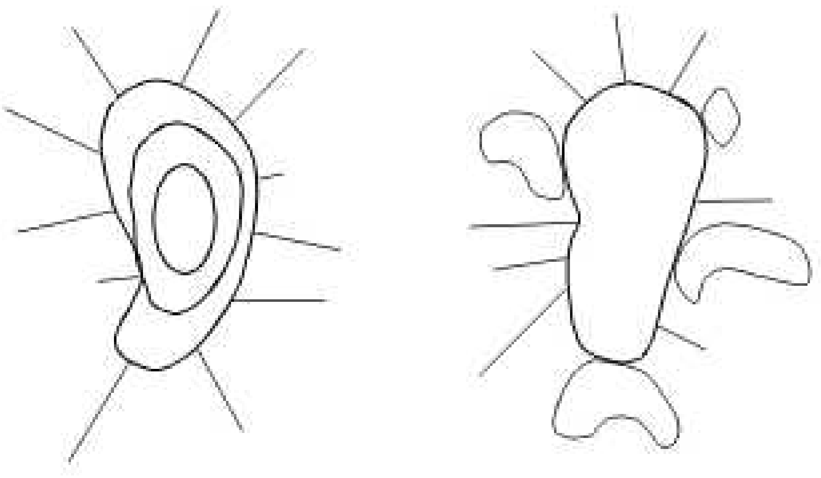
Active Contours



Active Contours

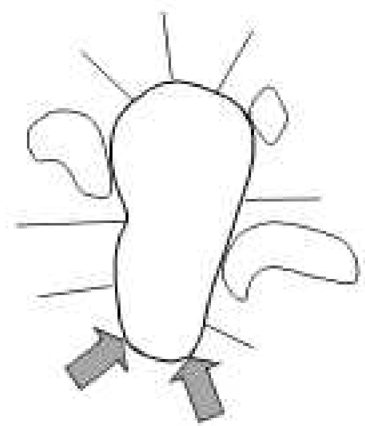


Active Contours

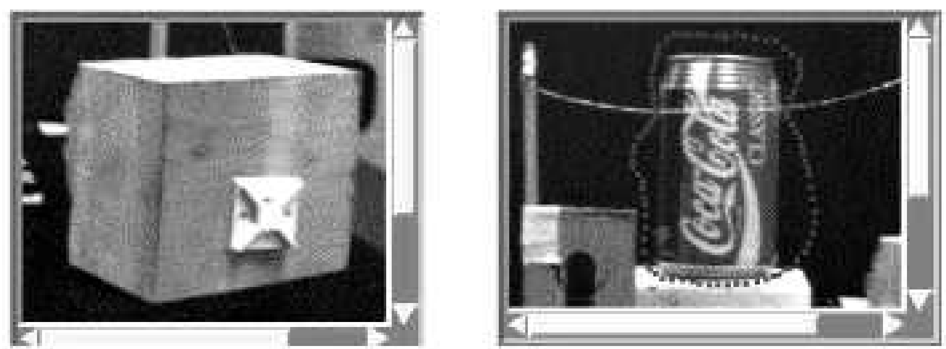


Active Contours

- High level understanding of the general shape
- User-applied energy

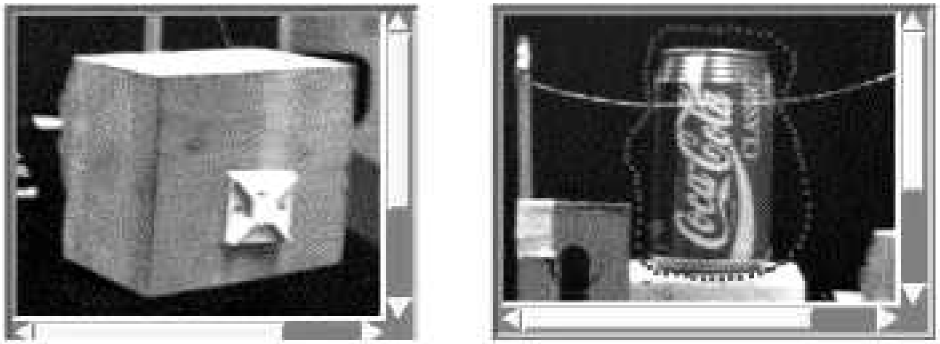


Active Contours



Stereo matching, facial recognition, medical imaging,
object tracking, shape recognition

Active Contours



Stereo matching, facial recognition, medical imaging, object tracking, shape recognition

What will you learn?

- What is Computer Vision? What are Vision Systems?
- Low level Computer Vision: Preprocessing and How to 'prepare' an image for more intelligent tasks?
- Detecting objects - Part 1: Edge based segmentation
- Detecting objects - Part 2: Region based segmentation (Blob detection)
- Detecting objects - Part 3: Detecting known shapes

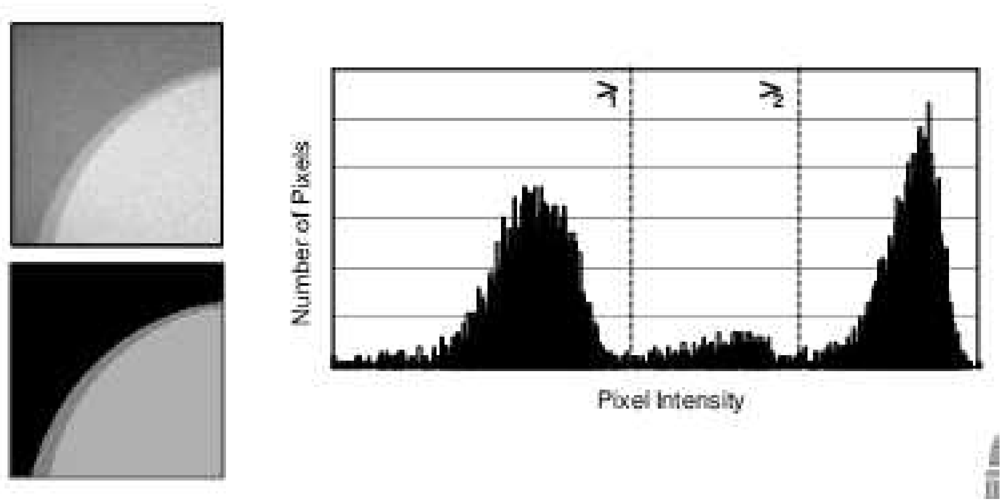
Batteries Included !!!

Region basd segmentation

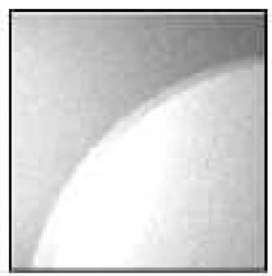
Merging Algorithms: in which neighboring regions are compared and merged if they are close enough in some property.

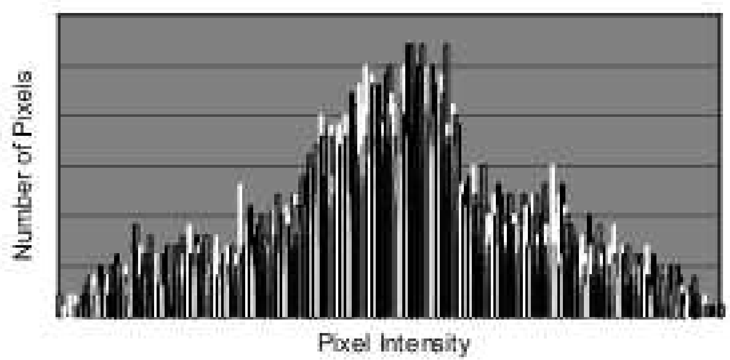
Splitting Algorithms: in which large non-uniform regions are broken up into smaller areas which may be uniform.

Histogram



Histogram





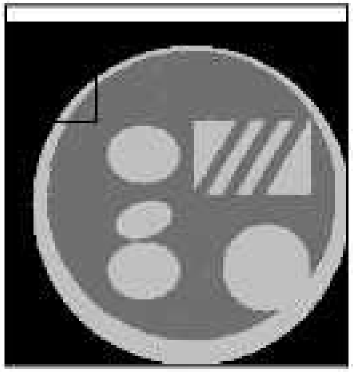
Region Merging

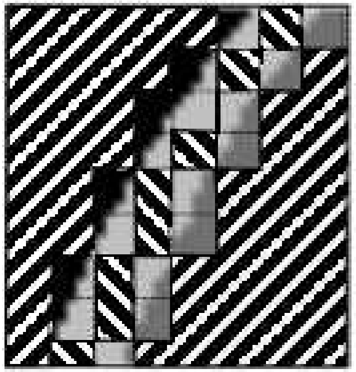
Find some uniform seed region, e.g., divide the image into 2x2 or 4x4 blocks of pixels, check each block for uniformity

Join on any neighboring regions (or pixels) until no further merging can take place

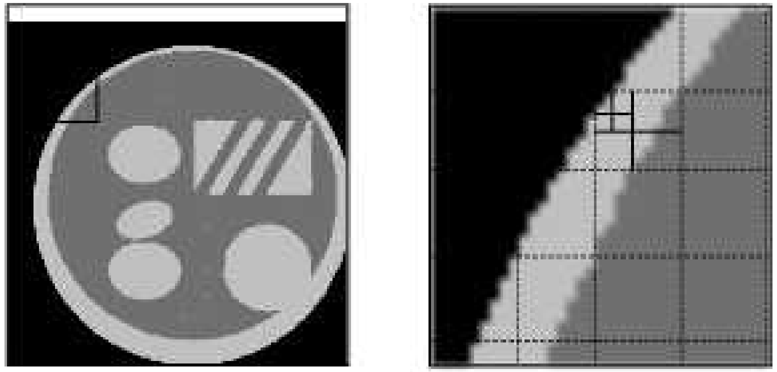
Mark the extracted region and search for others

Region Merging





Region Splitting



Split and merge

Various algorithms combine both processes
either refine a boundary by both techniques. Increasing detail as the boundary becomes more accurate or
split image to find large seeds, then merge (usually through regular decomposition)

Quad-trees for region extraction

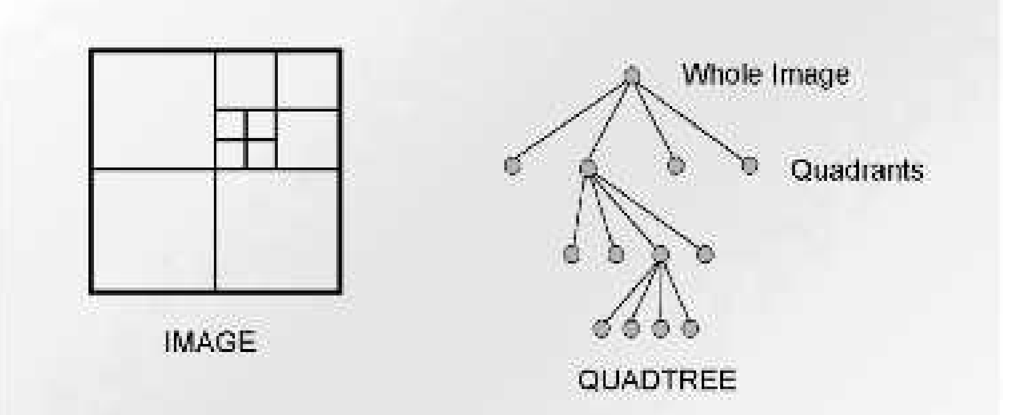


Diagram 4.1: Quadtree Decomposition

Quad-trees for region extraction

For an image with resolution 2mx2m (integer m),

Quardtree has m+1 levels

level 0 - each node corresponds to a pixel

level m - the root (whole image)

level n - each node corresponds to a square region of side 2n pixels

Quad-trees for region extraction

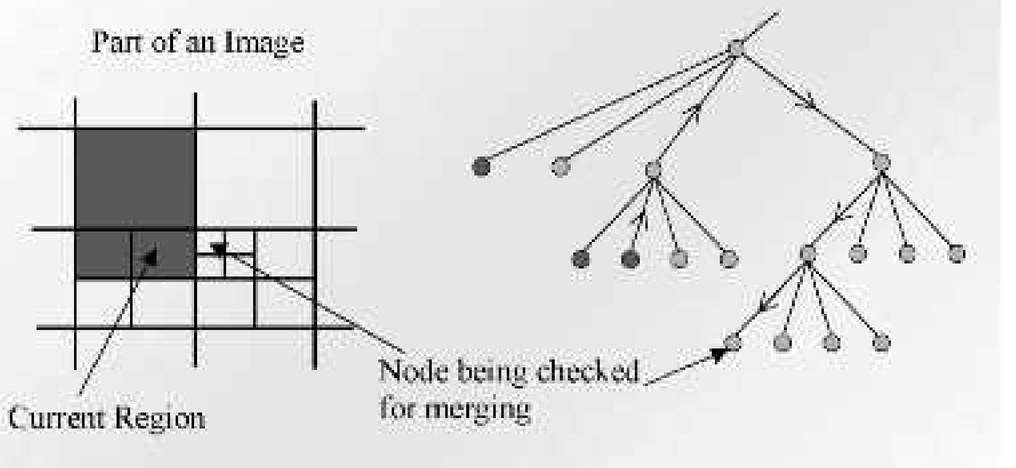


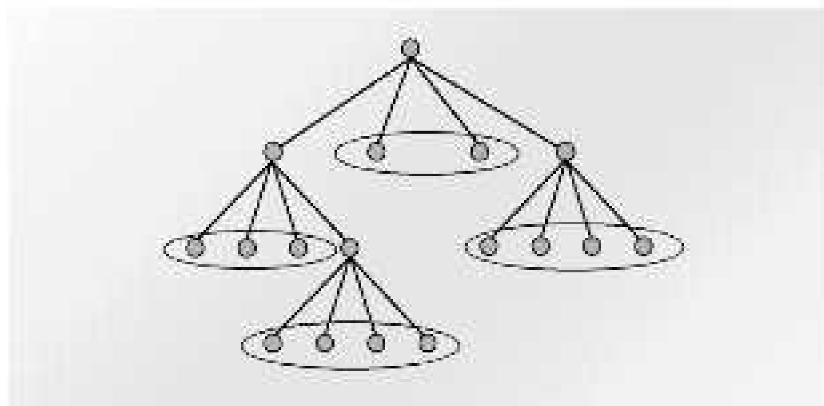
Diagram 4.2: Splitting and Merging with Quadtrees

Quad-trees for region extraction

Image size nxn, $n=2^m$
quardtree nodes = $(2^{2(m+1)} - 1)/3$
1, 5, 21, 85, ...

For a 512x512 image, there are 349,525 nodes, each non-terminal node has 4 pointers, for efficient computatio, each node should also have a pointer its parent. Minimum requirement is about 1.5 Mb.

Quad-trees for region extraction



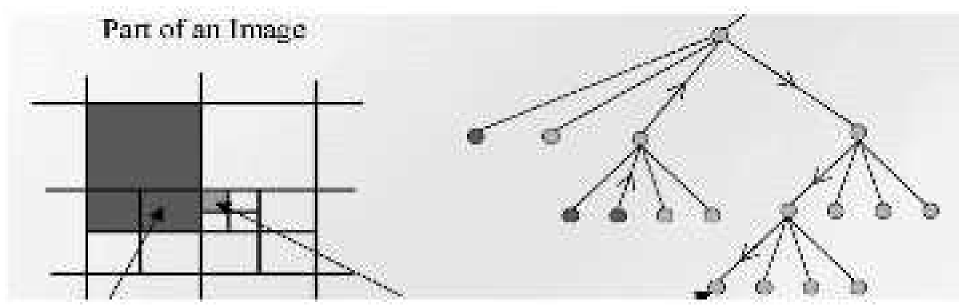
Not the whole tree but cutset is stored. Each cutset node contains { xk, yk, size, Mk, mk}

Position — Level in tree — brightness max, min

Quad-trees for region extraction

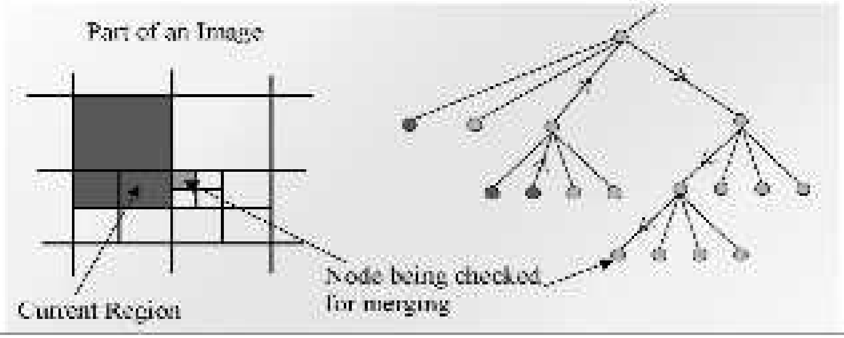
ÿ Define an initial segmentation into regions, a homogeneity criterion, and a pyramid data structure.

Ÿ If an region R in the pyramid data structure is not homogeneous, split it into four child-regions; if any four regions with the same parent can be merged into a single homogeneous region, merge them. If no region can be split or merged, go to step (3).



Quad-trees for region extraction

- If there are any two adjacent regions Ri, Rj (even if they are in different pyramid levels or do not have the same parent) that can be merged into a homogeneous region, merge them.
- Merge small regions with the most similar adjacent region if it is necessary to remove small-size regions.



Quad-tree segmentation

Construct a quadtree storing at each node the mean of its four children.

Search from the root of the tree to find the node highest in the tree whose mean is in the required range.

Use neighbour finding techniques to merge adjacent nodes whose mean is within the desired range.

Uniformity Criterion

Mean

Gives no uniformity information

Min/Max

Works in good images, however random noise can lead to wrong results

Variance

Is a statistical measure of how close to the mean a set of data is

Variance/average pyramid

Mean $\mu = \frac{1}{n}\sum f(x,y)$

Variance $v = \frac{1}{n}\sum \left(f(x,y) - \mu \right)^2 = \frac{1}{n}\sum f(x,y)^2 - \mu^2$

Fischer's Criterion

$$\frac{|\mu_1 - \mu_2|}{\sqrt{v_1^2 + v_2^2}} > \lambda$$

Uniformity Criterion

$$uniformity = 1 - \nu/\mu^2$$

Uniformity 0[0,1] if all pixels have positive values, why ?

$$\frac{|\mu_1 - \mu_2|}{\sqrt{\nu_1^2 - \nu_2^2}} > \lambda_0(1 - \nu/\mu^2)$$

Color space

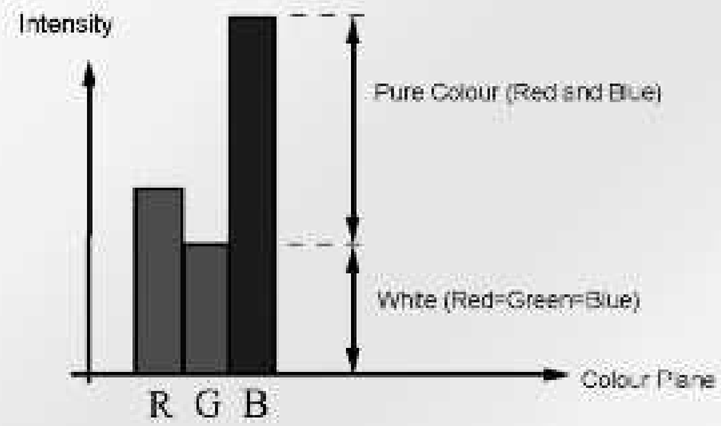
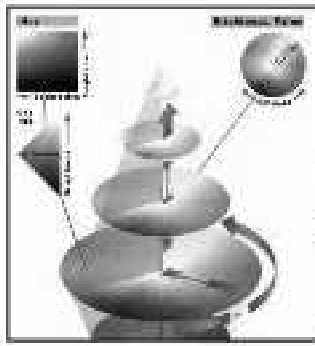


Diagram 5.1 A tri-stimulus colour represented as a pure colour and white

Color space



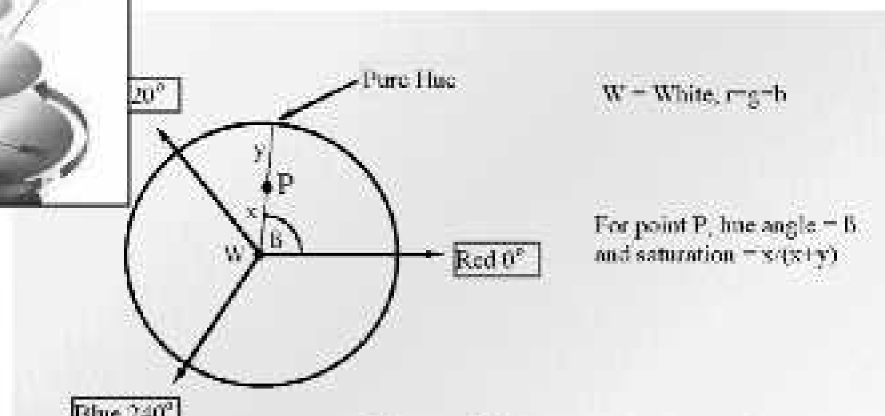


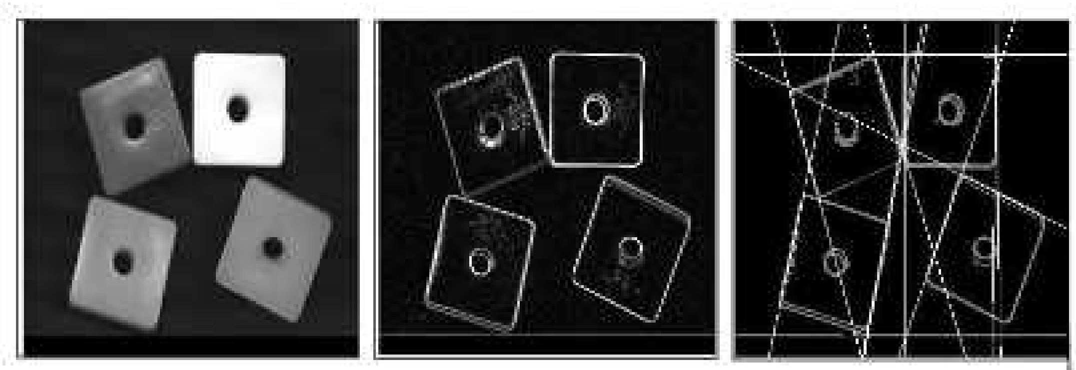
Diagram 5.2: Conceptual colour circle

What will you learn?

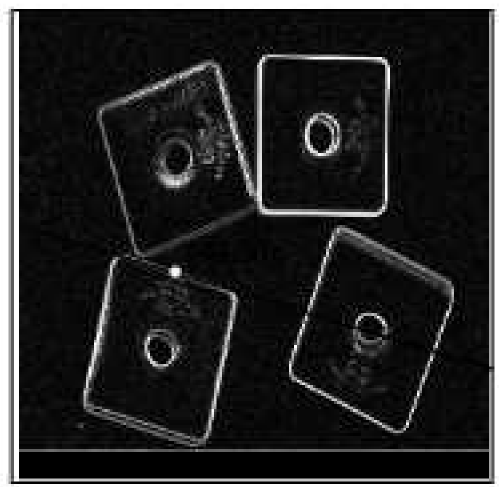
- What is Computer Vision? What are Vision Systems?
- Low level Computer Vision: Preprocessing and How to 'prepare' an image for more intelligent tasks?
- Detecting objects - Part 1: Edge based segmentation
- Detecting objects - Part 2: Region based segmentation (Blob detection)
- Detecting objects - Part 3: Detecting known shapes

Batteries Included !!!

Hough Transform

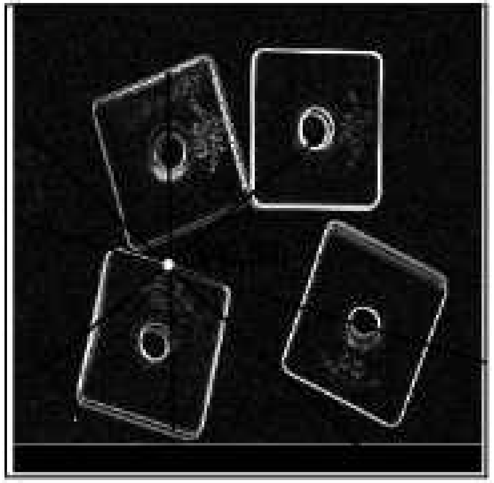


Hough Transform



$$y = mx + c$$

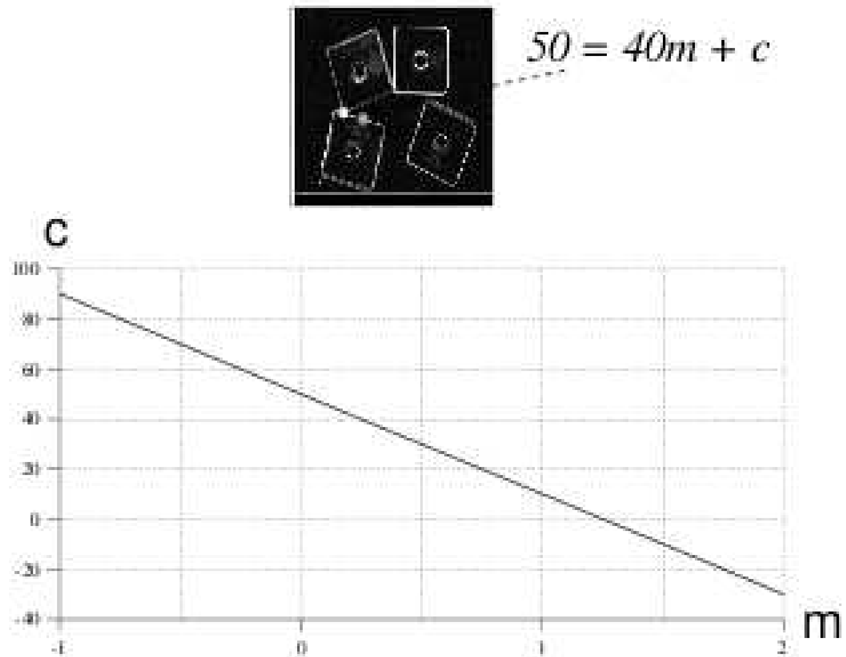
Hough Transform



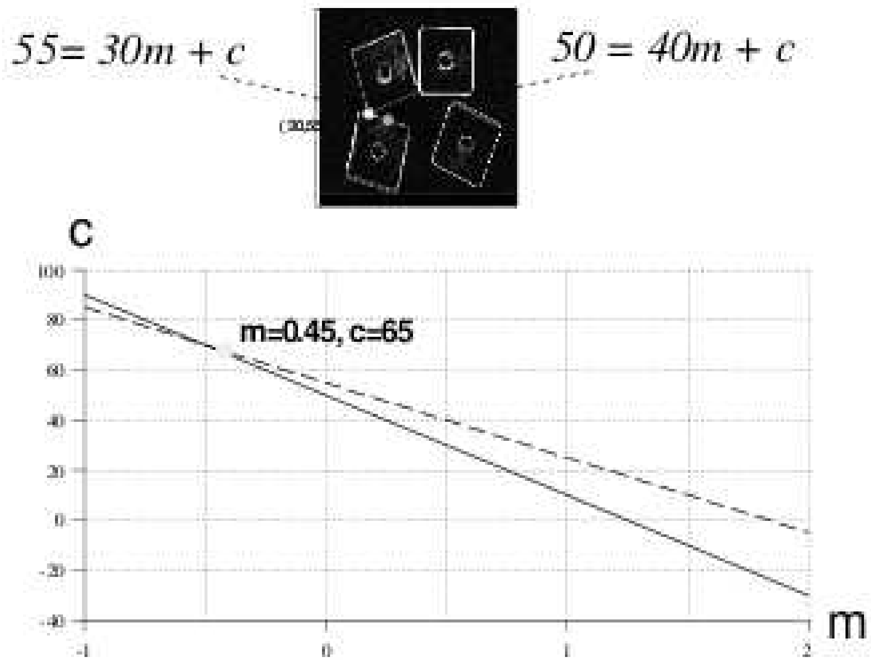
$$y = mx + c$$

$$50 = 40m + c$$

Hough Transform



Hough Transform



Hough Transform

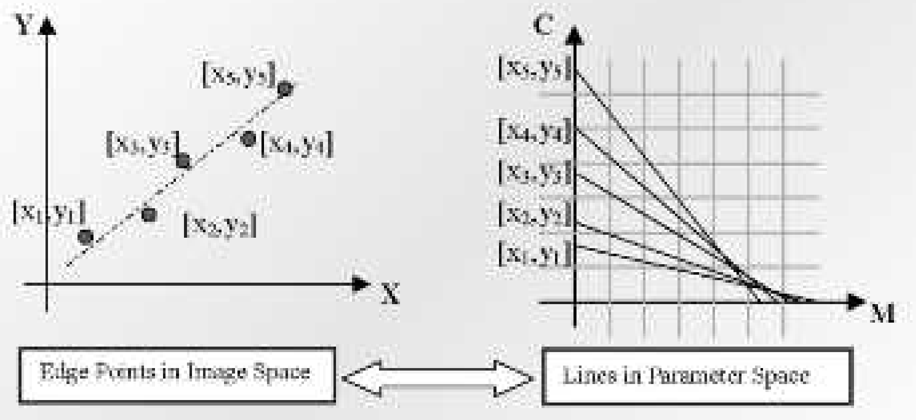
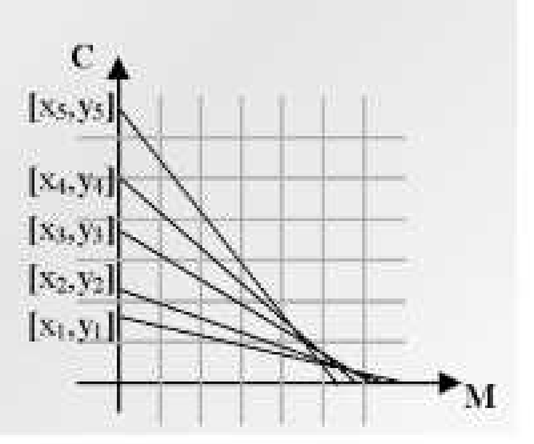


Diagram 7.1: The Hough Transform for Straight Lines

$$y - mx - c = 0$$

Hough Transform

$$y - mx - c = 0$$
$$4 > m > -4$$



Hough Transform

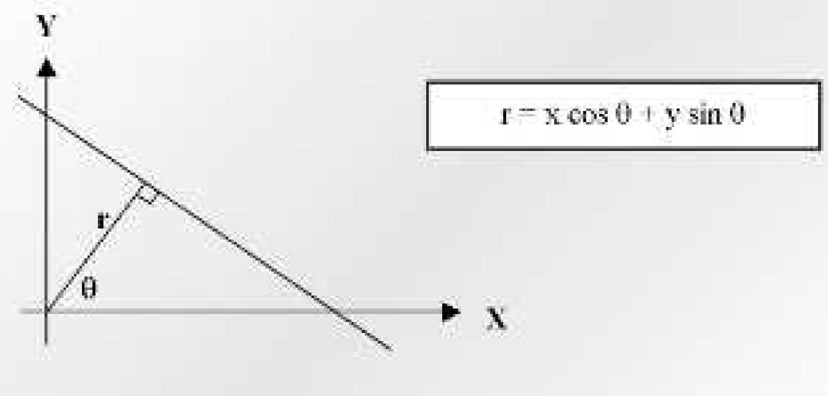


Diagram 7.2: r-θ parameterisation of a line

Hough Transform

$$r = M\cos(\varphi)\cos(\theta) + M\sin(\varphi)\sin(\theta)$$
$$= M\cos(\varphi - \theta)$$

where $M = \sqrt{(x_i^2 + y_i^2)}$,

$\cos(\varphi) = \frac{x_i}{M}$, $\sin(\varphi) = \frac{y_i}{M}$

$$r = x_i \cos\theta + y_i \sin\theta$$

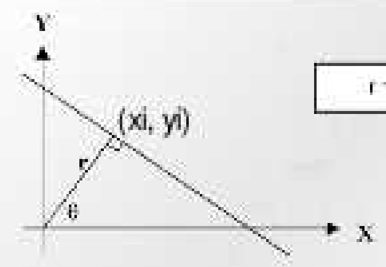


Diagram 7.2 r-θ parameterisation of a line

Hough Transform

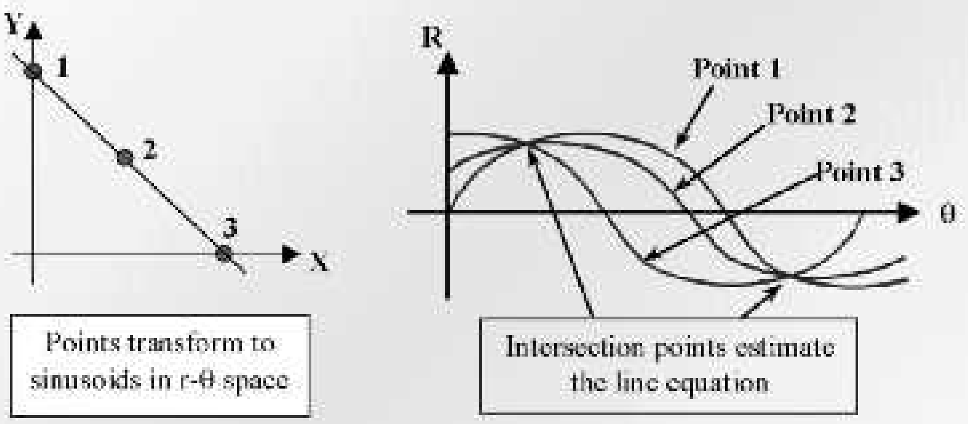
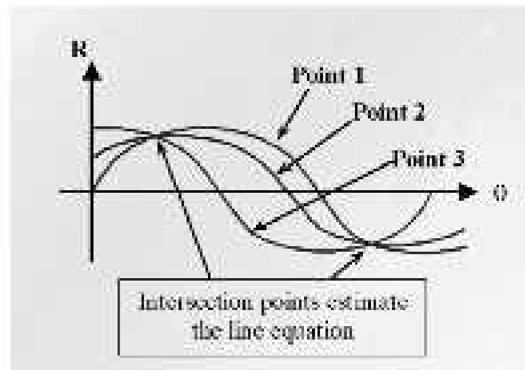


Diagram 7.3: The Hough Transform in r-θ space

Hough Transform



$$r = x\,cos2 + y\,sin2$$

$$-r = xcos(B+2) + ysin(B+2)$$

$$-r = -xcos2 - ysin2$$

Shape Discriminants

- area, perimeter, elongatedness
- moments
- direction
- energy, entropy,
- chord distribution

Moments

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^{p} y^{q} f(x, y)\, dx\, dy$$

$$m_{pq} = \sum_{i=1}^{x_{res}} \sum_{j=1}^{y_{res}} i^{p} j^{q} f(i, j)$$

Moments

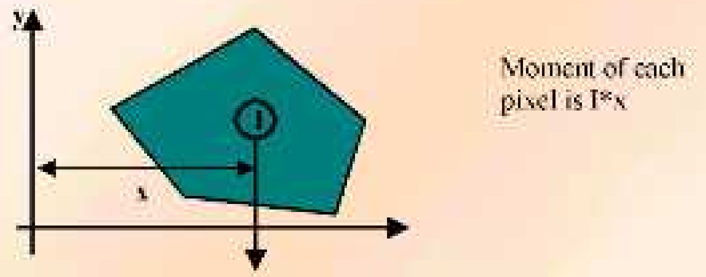
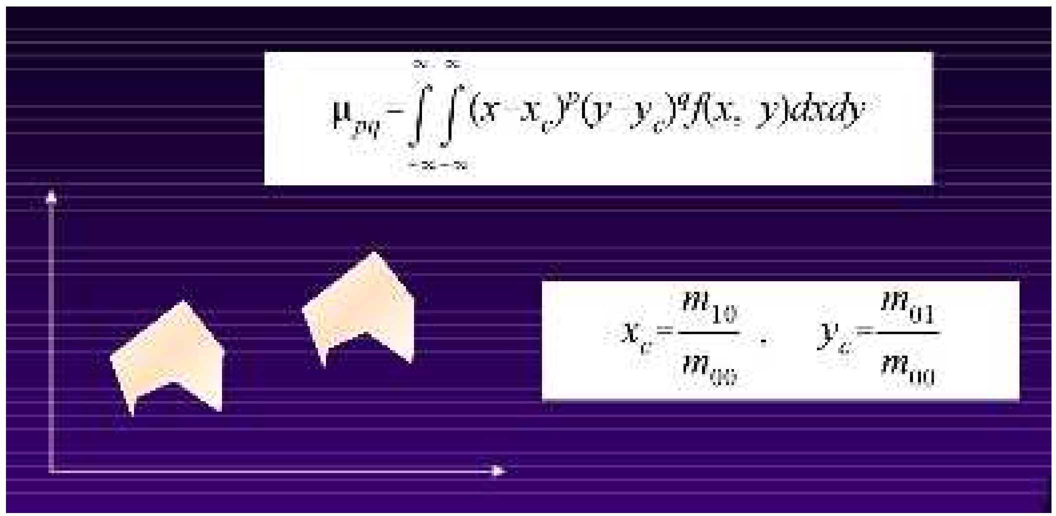


Diagram 10.2. First moments of area about y

Moments


$\mu_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x - x_c)^p (y - y_c)^q f(x, y)\,dx\,dy$
$x_c = \frac{m_{10}}{m_{00}}, \quad y_c = \frac{m_{01}}{m_{00}}$

Second Moments

Second Moments

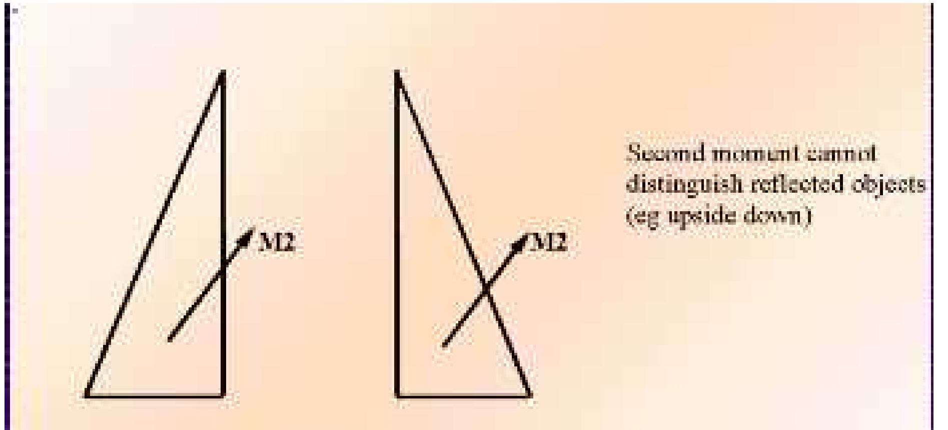
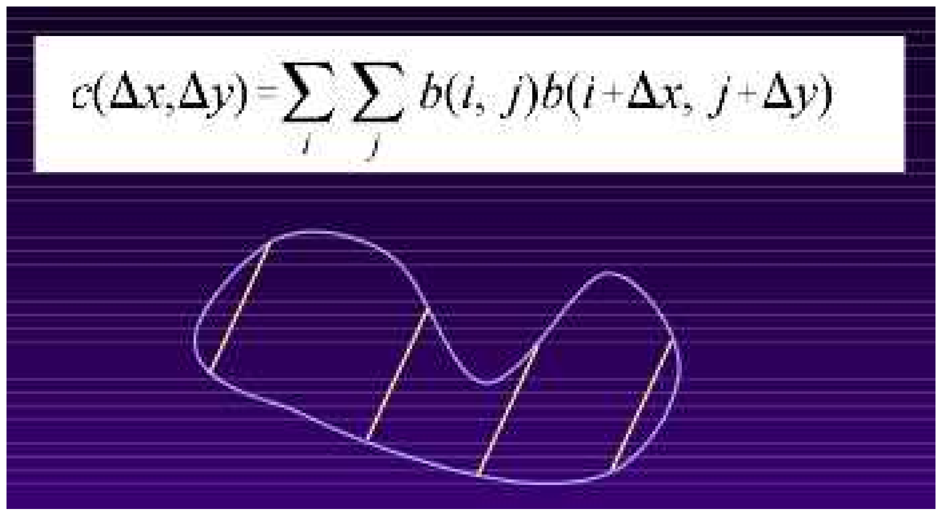
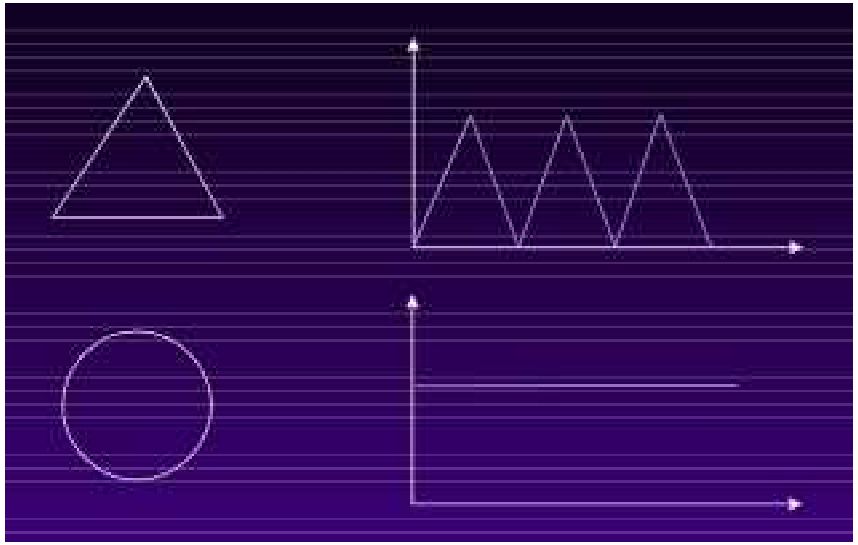


Diagram 10.4: Second Moment Directions

Chord Distribution



Chord Distribution



What will you learn?

- What is Computer Vision? What are Vision Systems?
- Low level Computer Vision: Preprocessing and How to 'prepare' an image for more intelligent tasks?
- Detecting objects - Part 1: Edge based segmentation
- Detecting objects - Part 2: Region based segmentation (Blob detection)
- Detecting objects - Part 3: Detecting known shapes

Batteries Included !!!

Whats in the box?

Python: A programmable language designed for ease of use (you can always use C)

OpenCV: >500 algorithms, DoD Challenge need I say more?

OpenCV python bindings (OpenCV is a C library so you need this for calling it in Python)

Python Standard library ans OpenCV library reference

A few tutorials and lecture slides

Dive In !!!