6. EDGE BASED SEGMENTATION

What is segmentation?

Various edge and line operators produce primitive edge elements. Human vision tends to organise the observed scene into *meaningful units* as a significant step towards image understanding. Further processing is necessary to *group* edge elements or groups of pixels into structures suited to *interpretation*.

One definition of segmentation (Pavlidis) states:

> "The goal of segmentation is to partition an image into disjoined regions which correspond to objects or their parts."

The outcomes of segmentation can be very different depending on definition of "objects". It is clear that some *knowledge* has to be incorporated into a process. One way to incorporate knowledge is to use it as a set of implicit or explicit constraints on the likelihood of a given grouping of image elements. Such constraints can be:

- domain independent - general physical arguments
- based on psychology of human perception
- domain dependent

There are two main approaches to segmentation:

- through extracting boundaries of regions - based on discontinuities (e.g. in grey levels or statistical properties)
- through extracting regions - based on similarities

The two approaches are equivalent - one representation can be converted into the other.

The segmented image is on higher level of abstraction than an intrinsic image - it contains the beginnings of *domain-dependent* interpretation; e.g. choice of kind of uniformity influences the grouping of intrinsic image structures into units meaningful in a given domain.

Representation for segmented image data

Input to a segmentation process is an image: either the original grey level image or an intrinsic image (e.g. edge gradient magnitude and gradient direction).

Output of the segmentation process can have several forms:

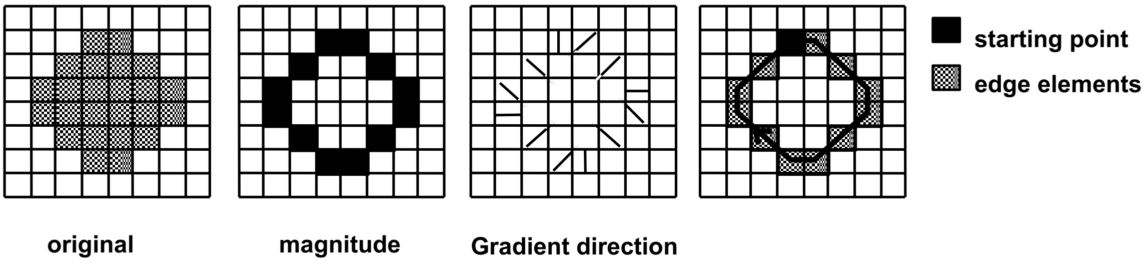
- an image where a pixel value indicates whether the pixel belongs to edge/region or to the background
- an image where a pixel value is a region *label* (e.g. if there are five different regions in the image, there will be five different labels; all pixels belonging to the same region will have the same label)
- a data structure which describes the results of segmentation, for example a linked list of coordinates of the outline of a region.

Segmentation via boundary detection and edge linking

Additional processing is required to group the detected edge elements into structures suited to interpretation. The goal is to make a *coherent* one dimensional edge feature from many individual local edge elements. For example, all edge elements surrounding an object should "belong" together.

Contour following in grey level images

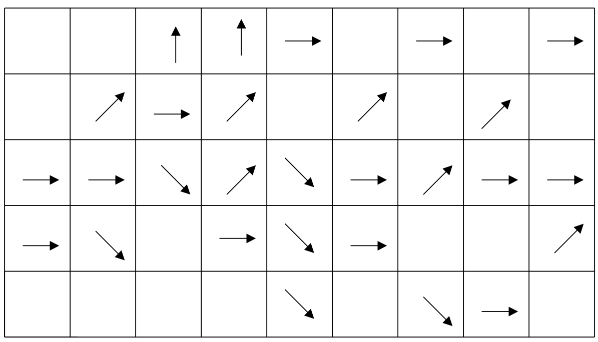
Starting point are intrinsic images showing edge magnitude and gradient.
If a pixel is on a boundary of an object, consequent boundary points should be searched in the direction perpendicular to a local gradient direction (i.e. in the likely direction of the edge contour). All the pixels belonging to the boundary should be either marked as edge elements in the segmented image or their coordinates stored.



Graph searching

Starting point are intrinsic images showing edge magnitude and gradient. Each pixel is considered to be a node in a graph.

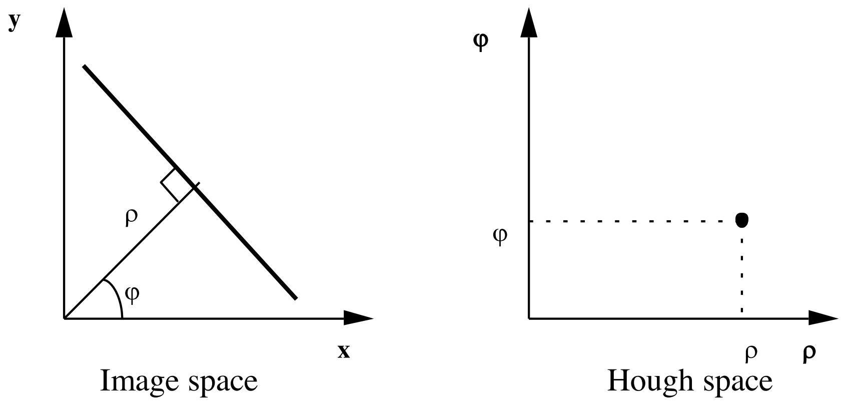
- two endpoints of a curve have to be known (one for closed curves)
- search for a "lowest cost" path between two nodes of a graph
  - magnitude and direction images created by one of the operators
  - direction image interpreted as nodes of the graph
  - magnitude image interpreted as weights associated with nodes
- link two nodes with an arc if
  - two nodes are aligned in the same sense as contour direction
  - second node is "in front" of the first one
  - difference in directions < $\pi/2$
  - weights of both nodes > a threshold
- to generate a path between two nodes
  - heuristic search
  - minimum cost search
- data structures - linked trees



Hough transform

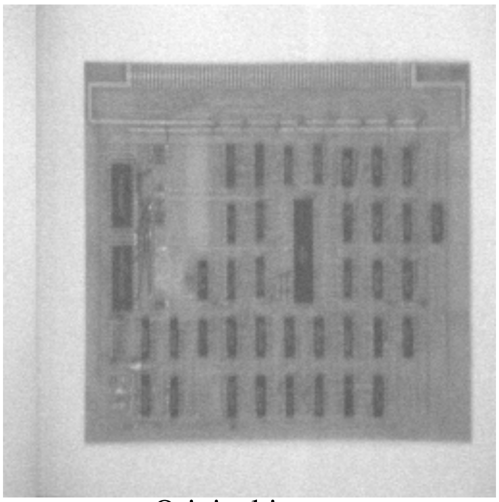
Hough transform is an example of a model-based technique which detects salient features of a particular type by combining local evidence on a global scale. It is used when shape of the boundary is known to have a parametric description.

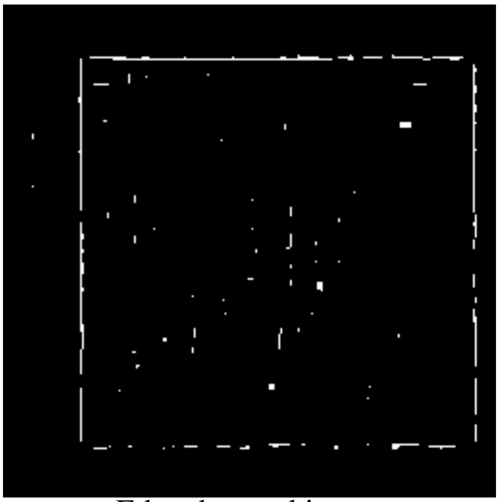
Hough transform for the detection of straight lines

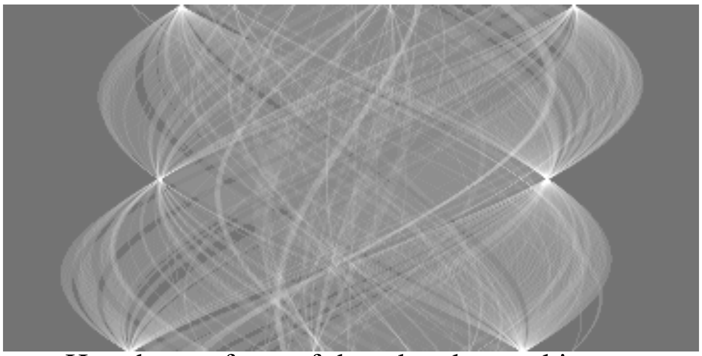


Line equation in polar coordinate system: $\rho = x \cos\varphi + y \sin \varphi$

Notes:


Original image


Edge detected image


Hough transform of the edge detected image

Notes:

Algorithm:

- quantify parameter space between approximate values of ρ and φ (they are bound)
- set elements of A(ρ, φ) to 0
- threshold the edge detected image
- for each point (x,y) in the edge image which lies on a line with the equation
  $\rho = x \cos \varphi + y \sin \varphi$
  perform:
  $A(\rho, \varphi) = A(\rho, \varphi)+1$
- Local maxima in A correspond to collinear points in image array
- Values at maxima are a measure of the line length

Parametrisation of the Hough transform for other curves

- circles - 3 parameters (xc, yc, r)
- ellipses - 5 parameters (xc, yc, a, b, φ)

Other computational approaches to grouping

All the publications listed below are available in the School Library, in the "Book of Readings for Image Understanding and Computer Vision".

Expert system (AI)

Levine, M.D. Nazif, A.M. (1984) Low level image segmentation: and expert system. IEEE Trans PAMI-6, 555-577.

Levine, M.D., Nazif, A.M. (1984) An optimal set of image segmentation rules. Pattern Recognition Letters 2, 243-248.

Hough transforms (mathematics)

Duda, R.O. and Hart, P.E. (1972) Use of the Hough transformation to detect lines and curves in pictures. Communications of the ACM (Graphics and Image Processing) 15, 11-15.

Wahl, F.M. and Biland, H.P. (1986) Decomposition of polyhedral scenes in Hough space. Proceedings of 8th International Conference on Pattern Recognition: IEEE, 78-84.

Ballard, D.H. (1981) Generalising the Hough transform to detect arbitrary shapes. Pattern Recognition 13, 111-122.

Active contours (mechanics)

Kass, M., Witkin, A. and Terzopoulos, D. (1987) Snakes: active contour models. International Journal of Computer Vision, 321-331.

Perception inspired (computational)

Canny, J. (1986) A computational approach to edge detection. IEEE Trans PAMI-8, 679-698. (See also a poplog teach file: vision7)

Sha'ashua, A., Ullman, S. (1988) Structural saliency: the detection of globally salient structures using a locally connected network. Proc IEEE, 321-327.

Neurophysiology and psychophysics based computational modelling

Neumann, H. and Stiehl, H.S. (1989) A competitive/cooperative (artificial neural) network approach to the extraction of n-th order junctions. Proceedings of the 11th DAGM Symposium (Burkhardt, K-H, Neumann, H.B. eds): Springer-Verlag.

Neumann, H. and Stiehl, H.S. (1992) Emergent segmentation of monocular visual invariants for space perception. Proceedings of IJCNN'92 vol III, 266-271: IEEE.

Heitger, F., Rosethaler, L., von der Heydt, R., Peterhans E. (1992) Simulation of neural contour mechanisms: from simple to end-stopped cells. Vision Research 32(5), 963-981.

Exploiting environmental constraints

Winston, P.H. (1992) Artificial Intelligence, Third Edition. Chapter 12: Symbolic constraints and propagation. Addison-Wesley.

Further reading and exploration

Sonka et al Section 5.2. and 8.2
Gonzalez & Woods, Section 7.2.

HIPR

Image Transforms: Hough Transform

CVIP

Analysis -> Edge/Line detection
Hough
Edge Link

Things to think about - discovering "rules" of perceptual grouping

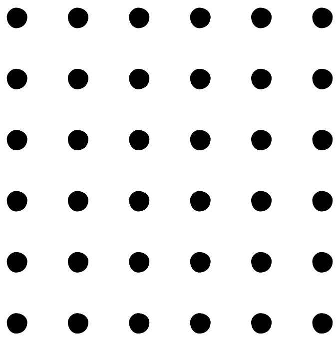
Early in 20th century the "Gestalt" school of psychology proposed a number of "rules" of perceptual organisation, trying to describe principles according to which collections of simple visual primitives are perceived as "belonging together".

In the figures below such organisations are likely to occur. Study the figures and select the interpretation that appears to be more "apparent", more "intuitive", more "immediate" ...

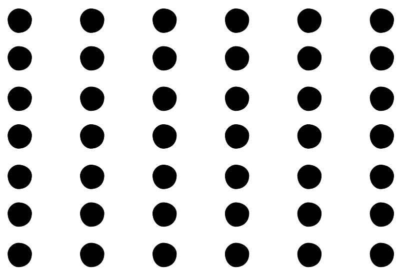
Justify your choice - WHY is one interpretation more apparent than the other?

Try to formulate the "rules" which you would use to instruct a computer vision system so that it groups the parts of the image in the same way as you have done.

1

Figure	Interpretation
	(a) 36 dots (b) 6 rows (c) 6 columns (d) a square

2

Figure	Interpretation
	(a) 42 dots (b) 7 rows (c) 6 columns (d) a rectangle

3

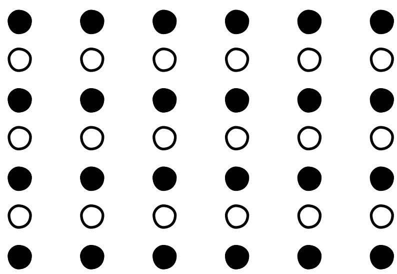
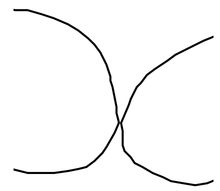
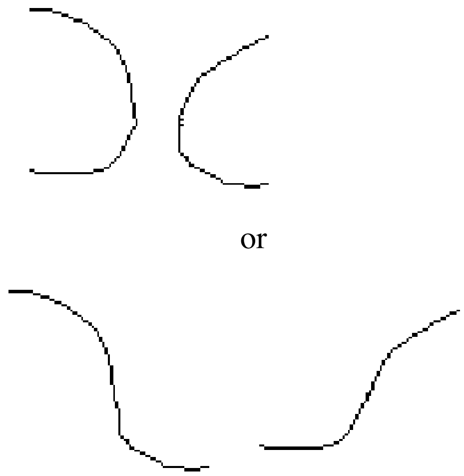
Figure	Interpretation
	(a) 42 dots (b) 7 rows (c) 6 columns (d) a rectangle

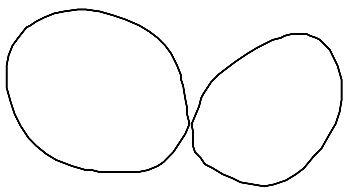
Figure **Interpretation**

4

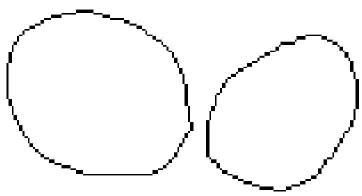




5



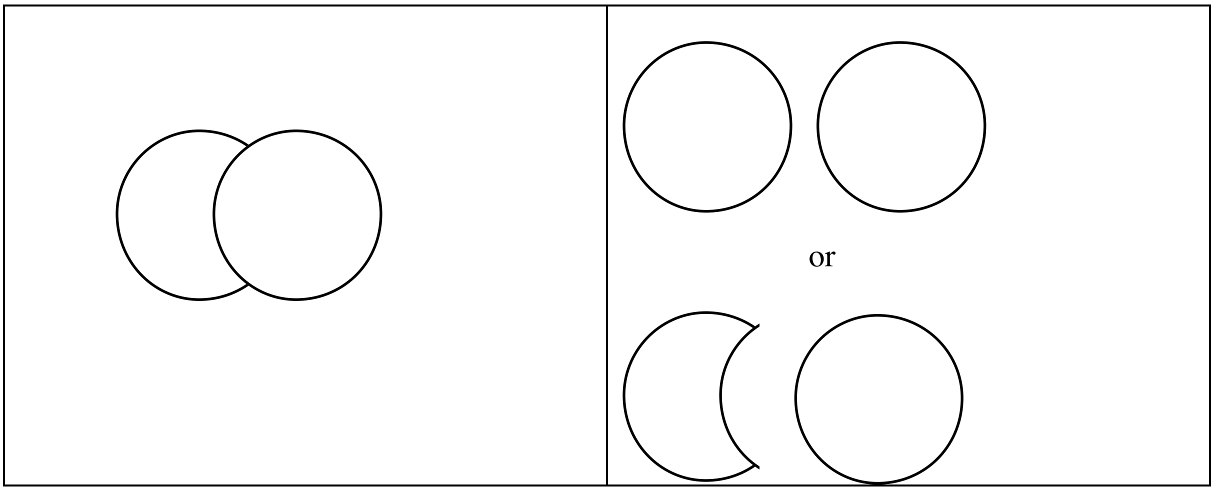
continuous curve



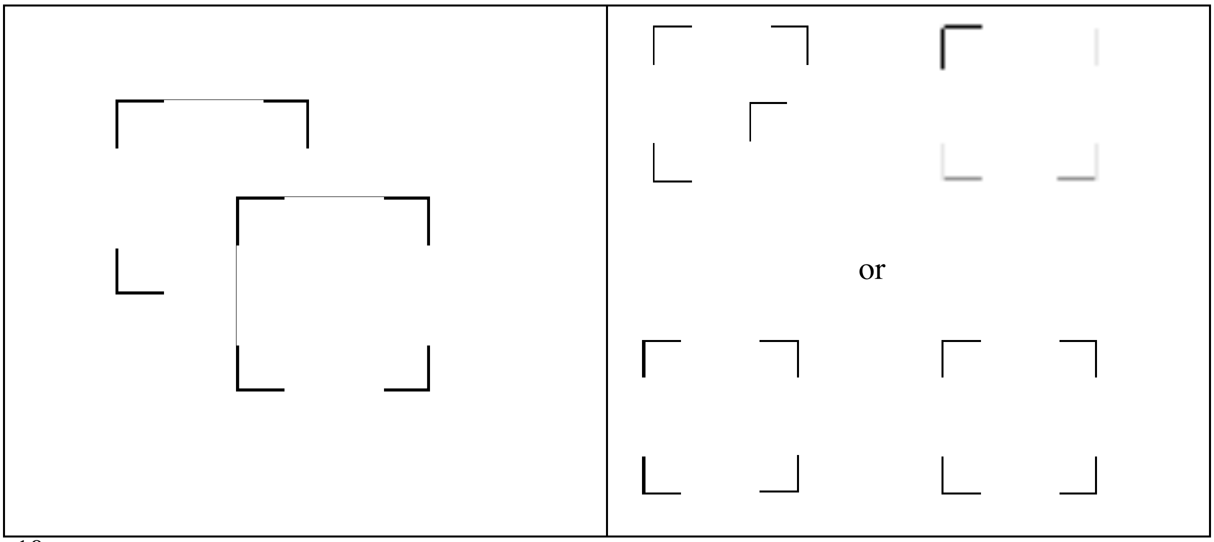
two ovals

Figure **Interpretation**

8



9



10