Grouping features
So far in the course

- Images formed by cameras
  - Image intensity related to scene brightness
  - 2D images related to 3D world by projection

- Image understanding/Computer vision is Inference

- Inference uses various cues
  - Geometry
  - Patterns in the distribution of intensities “Features”

- Extraction of “low-level” features
  - Edges, Corners, Texture

- Linear Filters
  - Fourier Analysis, Convolution
  - Noise smoothing
Effect of image analysis

- Image analysis produces a “feature image”
  - Outputs per pixel, or per region
- Next step
  - Use these for performing inference
- Higher level questions
  - Which object(s)?
  - What is the foreground?
  - What is in the background?
- Lower level analysis provides inputs
Collect features together to form objects

Figure 14-1

As these images suggest, an important component of vision involves organizing image information into meaningful assemblies. The human vision system seems to do so surprisingly well. In each of these three images, blobs are organized together to form textured surfaces that appear to bulge out of the page (you may feel that they are hemispheres). The blobs appear to be assembled “because they form surfaces,” hardly a satisfactory explanation and one that begs difficult computational questions. Notice that saying that they are assembled because together they form the same texture also begs questions (How do we know?). In the case of the surface on the left, it might be quite difficult to write programs that can recognize a single coherent texture. This process of organization can be applied to many different kinds of input. Reprinted from “Shape
Mid-Level Vision

• Combine these features in some way to perform inference
• Integrate information over distance.
• Use Gestalt cues
  – Smoothness
  – Closure
• Perceptual grouping is about putting parts together into a whole:
  – Finding regions with a uniform property
  – Linking edges into object boundaries
Surfaces and objects are critical.
Also, simpler “objects” such as lines
• Previous class: saw how it is possible to “group” edge features into lines using the “Hough transform”
• Computational technique
• How do humans do it?
Human perceptual grouping

• This has been significant inspiration to computer vision.

• Why?
  – Perceptual grouping seems to rely partly on the nature of objects in the world.
  – This is hard to quantify, we hypothesize that human vision encodes the necessary knowledge.

• Reading
  – In coming classes, 16, then rest of 14.

• Extra Reading:


  http://psy.ed.asu.edu/~classics/Wertheimer/Forms/forms.htm
Gestalt Principles of Grouping: some history

- Behaviorists were dominant psychological theorists in early 20\textsuperscript{th} century.
  - To make psych scientific, wanted to view it as rules describing relation between stimulus and response, described as atomic elements.
  - Not role for “mind”.
  - Influential early behaviorist was Pavlov
• Gestalt movement claimed atomic stimulus and response don’t exist.

  - The mind perceives world as objects, as wholes, not as atomic primitives.

  - Can’t understand psych without understanding how we perceive the world.

• ge·stalt or Ge·stalt  (g -shtält , -shtôlt , -stält , -stôlt )
  n. pl.
  A physical, biological, psychological, or symbolic configuration or pattern of elements so unified as a whole that its properties cannot be derived from a simple summation of its parts.
I stand at the window and see a house, trees, sky.

Theoretically I might say there were 327 brightnesses and nuances of colour. Do I have "327"? No. I have sky, house, and trees. It is impossible to achieve "327 " as such. And yet even though such droll calculation were possible and implied, say, for the house 120, the trees 90, the sky 117 -- I should at least have this arrangement and division of the total, and not, say, 127 and 100 and 100; or 150 and 177.

Max Wertheimer, 1923
I. A row of dots is presented upon a homogeneous ground. The alternate intervals are 3 mm. and 12 mm.

. . . . . . . . . . (i)

Normally this row will be seen as $ab/cd$, not as $a/bc/de$. As a matter of fact it is for most people impossible to see the whole series simultaneously in the latter grouping.

Max Wertheimer
Gestalt Movement

- Perceptual organization was a big issue.
  - How we perceive the world in terms of things/objects, not pixels.
- This was part of broader attack on behaviorism.
  - Gestalt viewed mind as constructing representations of the world, no learning/behavior could be understood without understanding this.
Issues in Perceptual Organization

• What is the role of an edge in an image? To what object (if any) does it belong?
Figure 5.5.16 Four kinds of edges. This scene contains four different kinds of luminance edges: orientation edges (O) due to abrupt changes in surface orientation, depth edges (D) due to gaps between surfaces at different distances, reflectance edges (R) due to different surface pigments or materials, and illumination edges (I) due to shadows.
(Bregman)
The famous Muller–Lyer illusion; the horizontal lines are in fact the same length, although that belonging to the upper figure looks longer. Clearly, this effect arises from some property of the relationships that form the whole (the gestaltqualität), rather than from properties of each separate segment.
Gestalt Principles

Lists of cues are “grouped” according to various “Gestalt principles”
Gestalt Principles

Parallelism

Symmetry

Continuity

Closure
Issues in Perceptual Organization

• What factors determine which parts of an image are combined in the same object?
Good Continuation
Good Continuation
Common Form: (includes color and texture)
Connectivity
Symmetry
Symmetry

Figure 7.25
Symmetry and figure ground. Look to the left and to the right, and observe which colors become figure and which become ground. (Adapted from Hochberg, 1971.)
Convexity (stronger than symmetry?)
Good continuation also stronger than symmetry?

Figure 5.8  Not two symmetrical figures juxtaposed, but two overlapping less regular figures.

Figure 5.9  Continuity of direction prevails over maximal symmetry.
Closure
Closure
• Some more reading (for fun)
• V.S. Ramachandran “The science of art”
• http://www.bbc.co.uk/radio4/reith2003/
• See especially
  – Professor Ramachandran's suggested 10 universal laws of art:
    – Peak shift
    – Grouping
    – Contrast
    – Isolation
    – Perception problem solving
    – Symmetry
    – Abhorrence of coincidence/generic viewpoint
    – Repetition, rhythm and orderliness
    – Balance
    – Metaphor
Computer Vision Again

Divide P.O. approaches into two groups.

- **Parametric:** We have a description of what we want, with parameters:
  - Examples: lines, circles, constant intensity, constant intensity + Gaussian noise.

- **Non-parametric:** We have constraints the group should satisfy, or optimality criteria.
  - Example: Find the closed curve that is smoothest and that also best follows strong image gradients.
The General Algorithm

• Define what it means for a group to be good.
  – Usually this involves simplifications
• Search for the best group.
• One version of this algorithm
  – Hough Transform
    • Good group … features lie on straight lines
    • Searched for best group in Hough space
    • Found peaks.
Smoothness

• Discrete Curvature: if you go from $p(j-1)$ to $p(j)$ to $p(j+1)$ how much did direction change?
  – Be careful with discrete distances.

• Change of direction of gradient from $p(j-1)$ to $p(j)$
Finding contours of objects: SNAKES

- We have concentrated on finding boundary points ("edges") or simple curves ("lines")
- How about general regions?
- Silhouettes of humans, animals, organs, tumors ...
Snakes

• variously known as *snakes, dynamic contours, active contours*

• mechanism for bringing degree of prior knowledge to bear on low-level visual problems

• Build cost-functions with prior information and then find configuration that minimizes this cost
  – in particular object boundaries tend to be continuous, smooth
  – tell us information about object shape
Snakes

Consider a contour

A snake is a deformable contour \( \mathbf{r}(s) \) parameterised by arc-length \( s \) (and possibly by time, \( \mathbf{r}(s, t) \)) that minimises an energy equation

\[
\mathcal{E}(\mathbf{r}) = \mathcal{E}_{\text{int}}(\mathbf{r}) + \mathcal{E}_{\text{ext}}(\mathbf{r})
\]

where

- \( \mathcal{E}_{\text{ext}}(\mathbf{r}) \) is some “external potential” (from the image), counterbalanced by
- \( \mathcal{E}_{\text{int}}(\mathbf{r}) \) “internal energy” tending to preserve its smoothness.

The equilibrium equation is:

\[
\left( \frac{\partial (w_1 \mathbf{r})}{\partial s} - \frac{\partial^2 (w_2 \mathbf{r})}{\partial s^2} \right) + \nabla F = 0
\]

where \( w_1(s) \) controls the “tension”, \( w_2 \) controls the “stiffness” of the contour, and \( \nabla F \) is the external potential gradient or force field.
Combine into a cost function

• Path: p(1), p(2), … p(n).

\[ \sum_{j=1}^{n} d(p(j), p(j+1)) \times [g(p(j)) + \lambda f(p(j-1), p(j))] \]

Where

• d(p(j), p(j+1)) is distance between consecutive grid points (i.e., 1 or \( \sqrt{2} \)).

• g(p(j)) measures strength of gradient

• \( \lambda \) is some parameter

• f measures smoothness, curvature.
One Example cost function

$$\sum_{j=1}^{n} d(p(j), p(j+1)) \times [g(p(j)) + \lambda f(p(j-1), p(j))]$$

$$g(p(j)) = \frac{1}{l_j^2 + \rho} \quad l_j = \frac{\|\nabla I_j\|}{\max_j \|\nabla I_j\|} \quad \rho \text{ is some constant}$$

$f$ is the angle between the gradient at $p(j-1)$ and $p(j)$.

Or it could more directly measure curvature of the curve.

Sometimes it’s not enough.