CSE585/EE555 Lecture 15

1. Basic image modeling
2. The Marr paradigm

III. IMAGE MODELING, ANALYSIS, AND SEGMENTATION

A. Basic image modeling
$\longrightarrow$ Human Visual System (Marr)

B. *Texture: more complex than gray scale only

a) Preattentive (human) perception (Julesz)

b) *HVS and the filter-bank approach (Malik & Perona, Bovik, Randen)

c) *Simple features and texture classification (Siew)

d) Gauss–Markov random fields

C. *Fractals: self–similar structures (Pentland, Barnsley)

a) Synthetic (deterministic)

b) Natural scenes (random)

*Emphasized topics

Image Modeling important for intelligent choice
of methods for:

a) image enhancement

b) image segmentation

c) image interpretation

d) image compression

References (*required readings):

1. D. Marr, *Vision*, pp. 41–87, Freeman, San Francisco, CA, 1982. Also, MIT Press, Cambridge, MA, 2010.
2. B. Julesz and J. R. Bergen, "Textons, the fundamental elements in preattentive vision and perception of textures," *Bell Syst. Tech. Journal*, vol. 62, no. 6, 1983.
3. S. Zhu, C. Guo, Y. Wang, and Z. Xu, "What are textons?" *Int. J. Computer Vision,* vol. 62, no. 1-2, pp. 121-143, 2005.
4. W. K. Pratt, *Digital Image Processing*, 3rd. Edition, ch. 16, "Image Feature Extraction," 2001.
5. *A. C. Bovik, M. Clark, W. S. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Trans. PAMI*, vol. 12, no. 1, pp. 55–73, Jan. 1990.
6. T. Randen and J. H. Husoy, "Filtering for texture classification: a comparative study," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 4, pp. 291-310, April 1999.
7. *L. H. Siew, R. M. Hodgson, E. J. Wood, "Texture measures for carpet wear assessment," *IEEE Trans. Pattern Analysis Machine Intell.*, vol. 10, no. 1, pp. 92–105, Jan. 1988.
8. *A. Pentland, "Fractal–based description of natural scenes," *IEEE Trans. Pattern Analysis Machine Intell.*, vol. PAMI-6, no. 6, pp. 661–674, Nov. 1984.
9. M. F. Barnsley, *Fractals Everywhere,* 2nd. Edition, Academic Press, London, UK 1993.
10. A. Bovik, Ed.,*Handbook of Image and Video Processing,* 2nd. Edition, Chapters 4.1, 4.2, 4.9, Academic Press, 2005.

III.A. BASIC IMAGE MODELING

Two fundamental problems:

a) Image Segmentation

Process of partitioning an image into "meaningful" units

(i.e., "uniform" regions)

"meaningful" units all share certain common *properties*

b) Image Classification

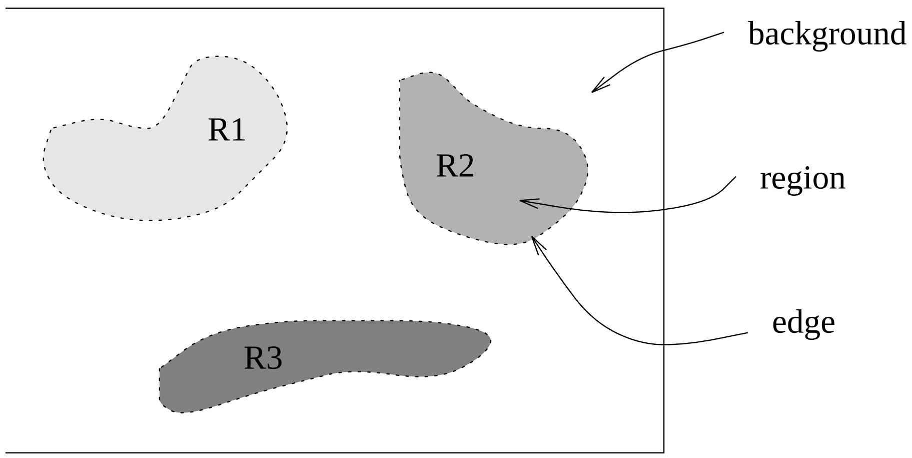
Characterize "patches" of an image,

based on intrinsic features.

To tackle the above problems, *preprocessing* needed (coming up: L16).

Image Modeling

(1) General situation:



1. regions differ from each other in terms of smaller “patches”
2. regions not necessarily uniform in gray scale

(2) Similar to (1) but 3-D cues in 2-D image:

1. perspective distortion
2. lighting cues (depth)
3. occlusion (objects blocking each other)

< May violate general idea that regions have uniform structure >

Influences on Image Segmentation:

1) Image–formation process

    lighting, viewpoint

2) Object properties

    a) reflectance, occlusion (viewpoint)

    b) texture, edges

    c) size, geometry

"Segment" types:

1) *edges* : (borders, boundaries)

    "discontinuities" between different regions.

2) *regions* :

    collection of pixels homogeneous

    in some property.

Simple Image-Segmentation Paradigm (CMPEN/EE455):

1. Filter image
   (a) Reduce noise, artifacts
   (b) Sharpen (bring out) important features
2. Define regions
   (a) Threshold out "bright" regions
   (b) Detect region edges

But what if regions are <u>not</u> "uniform" (homogeneous)?

Consider historical context.

Gestalt psychologists (1920's) — study human perception

People organize/group shapes in images (in their vision) based on certain *features* :

1. proximity
2. similarity
3. continuity

David Marr — Primal sketch: based on human visual system (HVS)

D. Marr, *Vision*, pp. 41–87, Freeman, San Francisco, CA, 1982; also, MIT Press, Cambridge, MA, 2010.

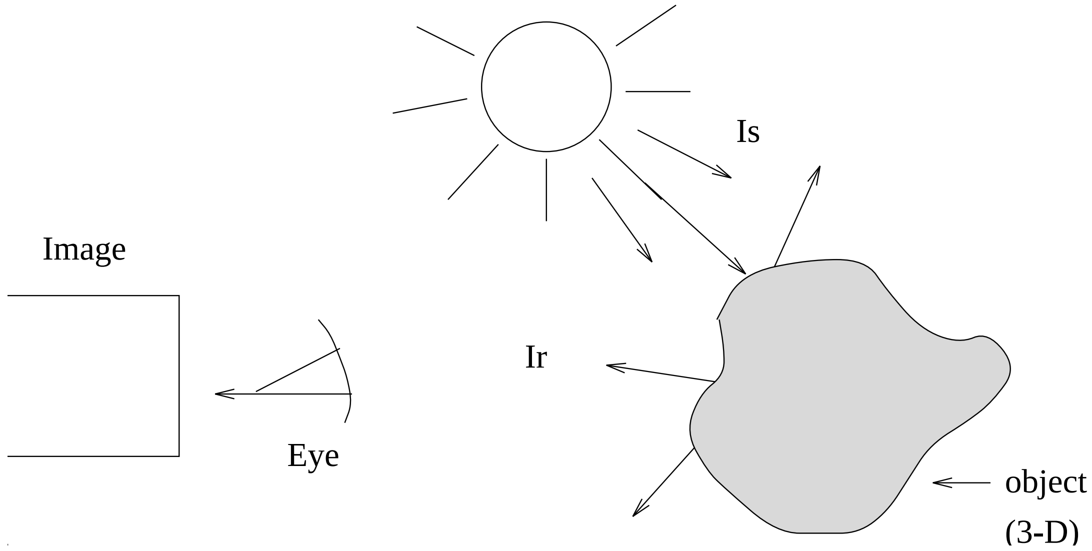
$\Longrightarrow$ paradigm for *human* visual (image) perception

$\Longrightarrow$ aid understanding of image segmentation, classification

More recent modification/extension:

J. Aloimonos, D. Shulman, *Integration of Visual Modules: An extension of the Marr Paradigm*, Academic Press, 1989.

$\longrightarrow$ Development below from Marr.



$\approx$ *eye/camera model* : 3–D world $\longrightarrow$ 2–D image formation

I_s illumination

I_r reflected light

Marr : Four factors influencing image appearance:

1. scene geometry

2. reflectance of visible surfaces

3. scene illumination

4. viewpoint (of observer) $\longleftarrow$ "eye"

Goal : (Vision problem)

Construct a meaningful description ("understanding") of image's content.

Method:

1. Determine object–to–image characteristics
   ("underlying physical assumptions")

2. Devise a scene *representation* mechanism $\Longrightarrow$ *Primal Sketch*

3. Extract primitives necessary to construct representation.

4. Intelligently group primitives to form larger–scale tokens.

5. Make inferences on scene content.

   Note : Huge jump from 4. to 5.!

$\longrightarrow$ Figures on pages 15-11 — 15-18 from Marr.

1. *Six Underlying Physical Assumptions* :

(I.) *Existence of Surfaces*

having reflectance functions

(II.) *Spatial organization* (of reflectance)

generated by processes of multiple scales

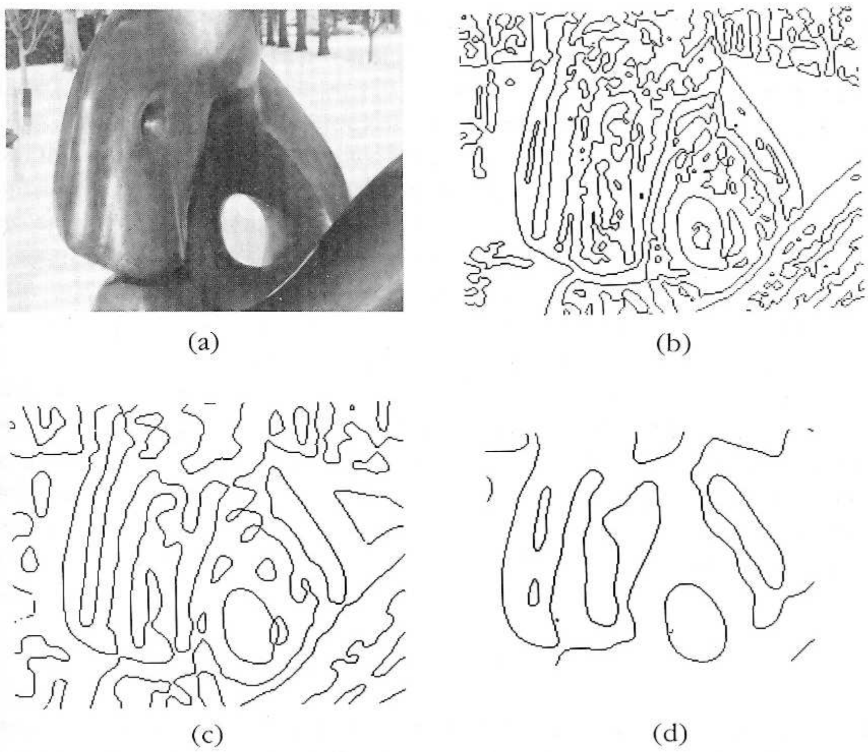


Figure 2–20. The image (a) has been convolved with $\nabla^2 G$ having $w_{2-D} = 2\sqrt{2}\,\sigma = 6$, 12, and 24 pixels. These filters span approximately the range of filters that operate in the human fovea. (b), (c), and (d) show the zero-crossings thus obtained. Notice the fine detail picked up by the smallest. This set of figures neatly poses the next problem—How do we combine all this information into a single description? (Reprinted by permission from D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond. B 204,* pp. 301–328.)

Notes:

(III.) Items (from a given surface) are more likely to appear *similar* (size, local contrast, etc.) to items of the *same scale* than to items at other scales.

(IV.) *Spatial Continuity*

Items often appear spatially organized to create lines, curves, ...

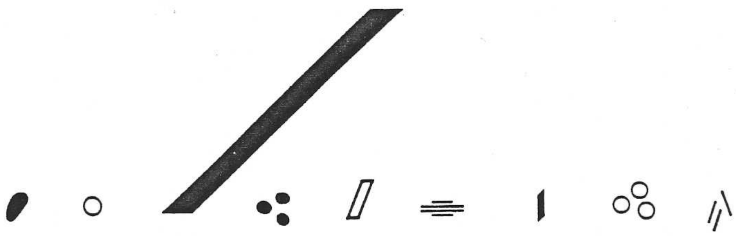


Figure 2–4. More evidence for place tokens. In this diagram every subgroup is defined differently, yet the collinearity of all of them is immediately apparent. This suggests that each group causes a place token to be created, whose collinearity is detected almost independently of the way the token is defined, provided that the tokens represent sufficiently similar items (compare Fig. 2–3d). (Reprinted by permission from D. Marr "Early processing of visual information," *Phil. Trans. R. Soc. Lond. B 275* 1976, fig. 10.)

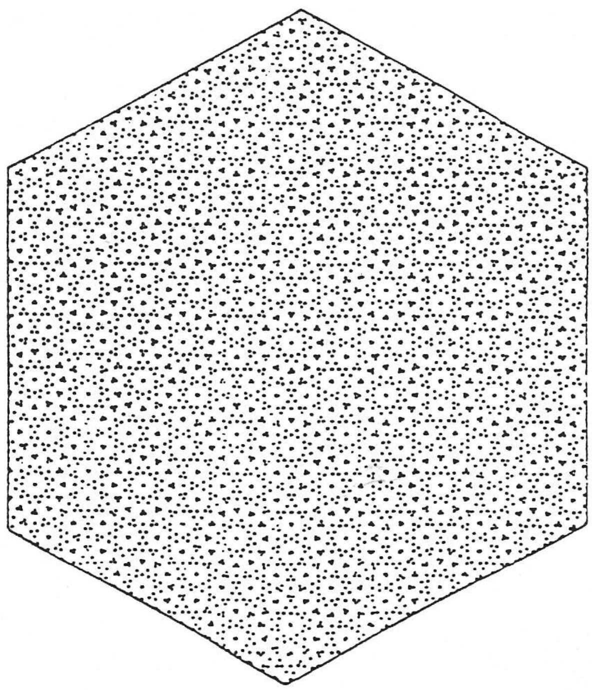


Figure 2–5. Evidence for the existence of active grouping processes. This pattern

(V.) *Continuity of Discontinuities*

Discontinuities (edges) appear smooth (continuous).

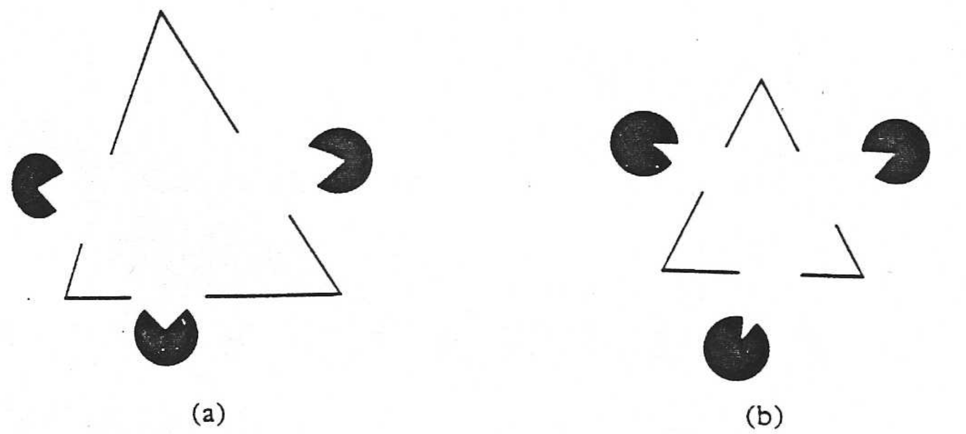


Figure 2–6. Subjective contours. The visual system apparently regards changes in depth as so important that they must be made explicit everywhere, including places where there is no direct visual evidence for them.

(VI.) *Continuity of Flow*

If direction of motion ever discontinuous at more than one point (along a line), then an object boundary is present.

2. Representational Scheme : Raw *Primal Sketch*

Capture structure of surfaces in an image using simple *tokens* :

a) edge, boundary segments.

b) bars (parallel edge pairs)

c) terminations (of edges, bars)

d) blobs (doubly terminated bars) with $l \leq 3w$

($w \longrightarrow$ width of main excitatory lobe of LoG)

Associated with each token are *attributes*:

a) position

b) orientation

c) contrast

d) length

e) width

End theory (steps 1-2 above). Now, make representation (steps 3-5).

3. To construct tokens,

detect edges ("zero crossings") at varying image scales.

$\downarrow$

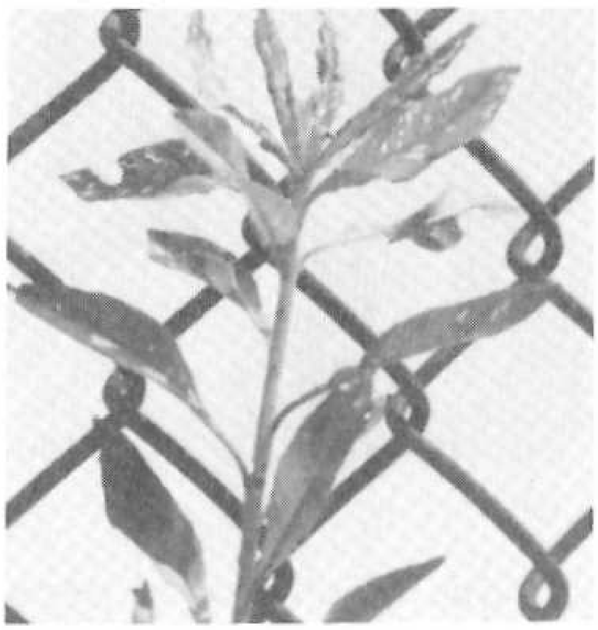
a) Apply Marr (& Hildreth) method :

$$I(x,y) \quad ** \quad \underbrace{\nabla^2 G(x,y)}_{\text{Laplacian of a Gaussian (LoG)}} \quad \longrightarrow \quad \hat{I}(x,y)$$

where $G(x,y) =$ Gaussian; $\nabla^2 G(x,y)$ depends on σ (scale).

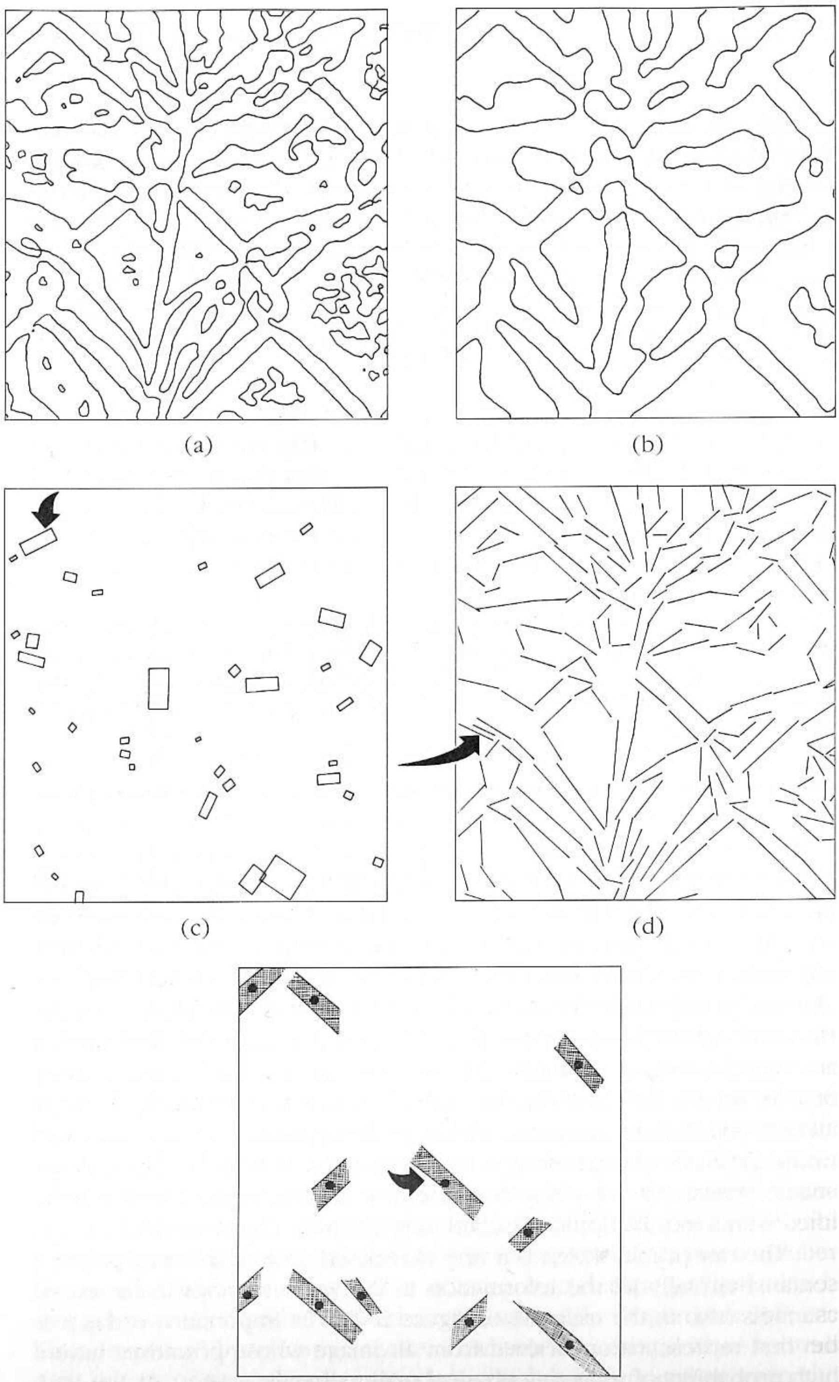
b) Detect zero crossings in $\hat{I}(x,y) \implies$ edges

Example — Get raw primal sketch of image below:



(flower in front of a fence)

Raw primal sketch:



(a) (b) (c) (d) (e)

Caption associated with previous page:

Figure 2–21. *(opposite)* The raw primal sketch as computed from two channels. (a), (b) The zero-crossings obtained from the image of Figure 2–12 by using masks with w_{2-D} = 9 and 18 pixels. Because there are no zero-crossings in the larger channel that do not correspond to zero-crossings in the smaller channel, the locations of the edges in the combined description also correspond to (a). (c), (d), and (e) Symbolic representations of the descriptors attached to the zero-crossing locations shown in (a). (c) Blobs. (d) Local orientations assigned to the edge segments. (e) Bars. The diagrams show only the spatial information contained in the descriptors. Typical examples of the full descriptors are:

BLOB	EDGE	BAR
(POSITION 146 21)	(POSITION 184 23)	(POSITION 118 134)
(ORIENTATION 105)	(ORIENTATION 128)	(ORIENTATION 128)
(CONTRAST 76)	(CONTRAST −25)	(CONTRAST −25)
(LENGTH 16)	(LENGTH 25)	(LENGTH 25)
(WIDTH 6)	(WIDTH 4)	(WIDTH 4)

The descriptors to which these correspond are marked with arrows. The resolution of this analysis of the image of Figure 2–12 roughly corresponds to what a human would see when viewing it from a distance of about 6 ft. (Reprinted, by permission, from D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond. B 204,* pp. 301–328.)

4. Image Properties to Detect

$\Longrightarrow$ aid in decoding surface geometry

$*$ Similarity
$\star$ Grouping
(spatial arrangement)

- $*$ 1. local average intensity
- $*$ 2. average size of items
- $\star$ 3. local intensity of items in 2.
- $* \star$ 4. local orientation of items in 2.
- $\star$ 5. Distances between items
- $\star$ 6. Orientation of groupings of similar items

$\downarrow$
VIRTUAL LINE

Use similarity and grouping concepts to ascertain properties above

$\Longrightarrow$ gives low–level spatial organization of information
in (raw) primal sketch

Between steps 4. and 5.

make $2\frac{1}{2}$ –D *sketch* (Marr, pg. 37, pg. 42)

$\uparrow$

Information on surfaces

(more explicit than raw primal sketch)

Then, make 3–D model representation

$\longrightarrow$ describe shapes , scene organization

5. Then, do recognition, inferencing!

* Marr gave a general theoretical paradigm that has

influenced <u>many</u> since 1982 !