Lecture 40 of 42

A Brief Survey of Computer Vision and Robotics

William H. Hsu
Department of Computing and Information Sciences, KSU

KSOL course page: http://snipurl.com/v9v3
Course web site: http://www.kddresearch.org/Courses/CIS730
Instructor home page: http://www.cis.ksu.edu/~bhsu

Reading for Next Class:

Chapter 24, p. 863 – 894, Russell and Norvig 2e
Chapter 25, p. 901 – 938, Russell and Norvig 2e

Lecture Outline

- This Week: Chapter 26, Russell and Norvig 2e
- Today: Chapter 23, R&N 2e
- Wednesday (Last Lecture!): Chapter 24, R&N 2e
- References
  - *Robot Vision*, B. K. P. Horn
  - Courses: http://www.palantir.swarthmore.edu/~maxwell/visionCourses.htm
  - UCB CS 280: http://www.cs.berkeley.edu/~efros/cs280/
- The Vision Problem
  - Early vs. late vision
  - Marr's 2 ½ - D sketch
  - Waltz diagrams
- Shape from Shading
  - Ikeuchi-Horn method
  - Subproblems: edge detection, segmentation
- Optical Flow

Line Drawing Interpretation

2.1 Definitions

Contour generator: is a 3D space curve.
Contour: is its projection on an image.
Limbs: these occur when a surface curves smoothly and hides itself.
Edges: are when there are discontinuities in surface normals.

2.2 Symbols

+ convex
- concave
> sharp changes in visibility (i.e. occluding edges)–these can only arise from convex edges
>> occluding contours (smooth surface that occludes itself)

The convention with the arrows (>,>>) is that if you are walking like an ant on an edge, following the direction of the arrow, then the belonging surface is on your right.

2.3 Discontinuities

z - changes in depth arise from occluding limbs/edges.
p(rho) - variations in reflectances from surface markings etc.
L(Illumination) - shadows.

Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

Line Labeling [1]: Solid Polyhedra and Other Shapes

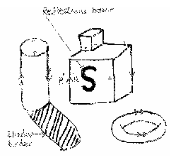


Figure 1: Some sample diagrams

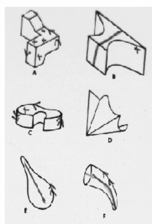


Figure 2: More sample diagrams

Line labelling: Classifying the image curves as depth or orientation discontinuities (and then further sub-classifying them).

Waltz, others

Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

Line Labeling [2]: Junctions

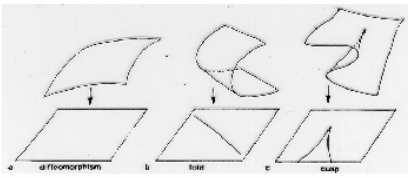


Figure 3: Junctions

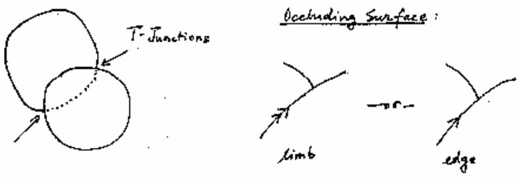
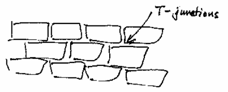


Figure 4: T-Junctions



False
T-junctions

Junctions occur at tangent discontinuities

Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

Orientation and Texture Discrimination (Textons) [1]

Textons and Julesz' Theory
Julesz' theory for finding boundaries in texture:
Julesz proposed that texture is composed of elementary units called textons. Differences in these textons help us in finding boundaries in texture.

Textons are classified in terms of:
- length, width, & orientation (for elongated blobs)
- # of crossings
- # of terminators

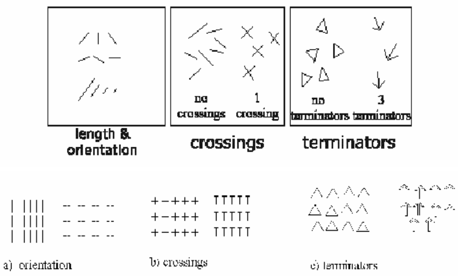


Figure 1: Textons

Adapted from slides © 1999 T. Leung, UC Berkeley (CS 280 Computer Vision)

Orientation and Texture Discrimination (Textons) [2]

Using textons, we can distinguish between different textures:
Examples:
1) Triangles versus arrow (see image#3): distinguishable through terminators (triangle has 0 terminators, whereas arrows have three terminators)
2) T's versus slanted T's (see image#5): distinguishable through orientation
3) R's and mirror R's (see image#1): indistinguishable because textons are the same

Julesz said that we can distinguish texture if there is a density difference of texton distribution in different areas of an image.

Segmentation (Grouping) [1]: Definition

Boundaries of image regions defined by a number of attributes.

- Brightness and color
- Texture
- Motion
- Stereoscopic depth
- Familiar configuration

Segmentation (Grouping) [2]: Physical Factors

Figure 7: Example picture – bedroom

Factors that lead to grouping:

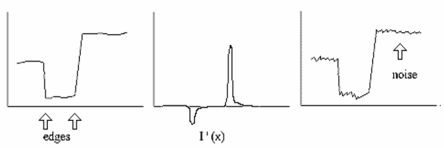
- Similarity of brightness, color, texture, disparity, motion, ...
- Proximity
- Good continuation of boundary contour
- Closure
- Symmetry and parallelism
- Familiar configuration

Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

Edge Detection [1]: Convolutional Filters and Gaussian Smoothing

Smoothing and Edge Detection:

Consider an intensity function I(x) shown in the graph below. The values of x for which || I'(x) || is large corresponds to edges.



I(x) [leftmost], I'(x) [center], and noisy I(x) [rightmost]

PROBLEM: When we introduce noise to I(x), we can produce undesirable edges. Consider: $I(x) + \alpha \sin(\omega x)$, where I(x) is our original intensity function and "$\alpha \sin(\omega x)$" is the noise. After differentiation of this function, we get $I'(x) + \alpha\omega \cos(\omega x)$. When ω is large (i.e. high frequency noise), we can get a lot of undesirable edges.

IDEA: Do smoothing before differentiation. This can be done by convolving I(x) with a Gaussian $G_\sigma(x)$ before differentiation:

$$I(x) * G_\sigma(x) \text{ for smoothing}$$

Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

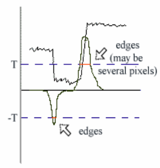
EDGE DETECTION [2]: DIFFERENCE OF GAUSSIAN

Thresholding and Non-maximum suppression:

Consider edge detection by convolution with the derivative of a Gaussian:

$(I(x) * G_\sigma(x))' = I(x) * G_\sigma'(x)$

PROBLEM: Just using ($I(x) * G_\sigma'(x)$) > 0 for edge detection does not work well because the edges can spread over several neighboring pixels. This means that too many pixels will be classified as edges.



IDEA#1: Pick a threshold T and mask edges where $\| I * G_\sigma' \| > T$
IDEA#2: (Non-maximum suppression) Take only the pixels with local maximum values in $\| I * G_\sigma' \|$; i.e., mark only those pixels where $\| I * G_\sigma' \|$ is greater than its neighbors.

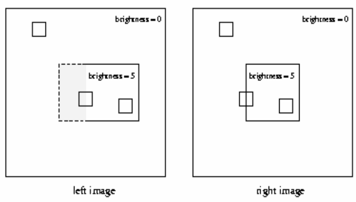
Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

BINOCULAR STEREO [1]: STEREO CORRESPONDENCE – PROPERTIES

1 Stereo Correspondence

Corresponding points have

- Similar values of brightness: result in high cross correlation score.
- Piecewise smoothness of disparity field.
- Uniqueness constraint: less than 1 corresponding point.
- Ordering constraint.
- Epipolar constraint: corresponding point lies on the epipolar lines.



Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

Binocular Stereo [2]: Stereo Correspondence – Open Problems

1.3 Stereo correspondence – still open

There are still some unsolved problems.

- regions without texture. Certain regions like a white wall make it impossible to find out the disparity.
- depth discontinuities, that is, regions of half-occlusion. We met this kind of problem when choosing the size of matching window. See Figure 2, 3.

Also, current continuity constraint is only one-dimensional case. Ohta and Kanade extend dynamic programming and use inter-scanline constraint. Figure 6 shows the idea.

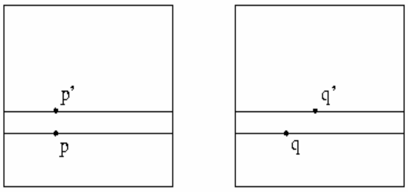


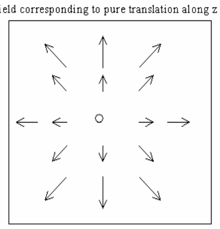
Figure 6: Extended DP by Ohta and Kanade, using the information that if p and p' are close, q and q' should also be close.

Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

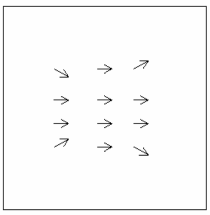
Optical Flow

you can create optic flow diagrams by plotting the (u,v) vector
at the appropriate (x,y) locations

This is the optic flow field corresponding to pure translation along z



Rotation about y – axis



Adapted from slides © 1999 J. Malik, UC Berkeley (CS 280 Computer Vision)

Terminology

- Vision Problem
  - Early vs. late vision
  - Marr's 2 ½ - D sketch
  - Waltz diagrams
- Shape from Shading
  - Ikeuchi-Horn method
  - Subproblems: edge detection, segmentation
- Optical Flow

Summary Points

- References
  - *Robot Vision*, B. K. P. Horn
  - http://www.palantir.swarthmore.edu/~maxwell/visionCourses.htm
- The Vision Problem
  - Early vs. late vision
  - Marr's 2 ½ - D sketch
  - Waltz diagrams
- Shape from Shading
  - Ikeuchi-Horn method
  - Subproblems: edge detection, segmentation
- Optical Flow
- Next Week
  - Natural Language Processing (NLP) survey
  - Final review