Computer Vision (I)

Course Number: 526 U1090
Credits: 3
Time: Tuesday 6, 7, 8 (2:10PM-5:00PM)
Classroom: New CSIE Classroom 309
Classification: Elective for junior, senior, and graduate students
Prerequisite: None
Instructor: Chiou-Shann Fuh
Office: New Computer Science and Information Engineering 327
Phone: 23625336 ext. 327, 23630231 ext. 3232 ext. 327
Office Hours: Tuesday 9AM-11AM
Objective: To learn computer and robot vision through extensive course projects.

Textbook: R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision,* Vol. I, Addison Wesley, Reading, MA, 1992.
Reference: L. G. Shapiro and G. C. Stockman, *Computer Vision,* Prentice-Hall, Upper Saddle River, NJ, 2001.
Reference: R. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision,* McGraw-Hill, New York, 1995.

Reference: R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* Addison Wesley, Reading, MA, 1992.
Reference: R. C. Gonzalez and R. E. Woods, *Digital Image Processing,* 2nd Ed., Prentice-Hall, Upper Saddle River, NJ, 2002.
Projects: will be assigned every week or every other week (30%)
Examinations: one midterm (30%) and one final (40%)

Content:
This is the first semester of a fast pace course which covers robot and computer vision. This semester covers low-level vision and mostly no reference to the third dimension:

1. Computer Vision: Overview
2. Binary Machine Vision: Thresholding and Segmentation
3. Binary Machine Vision: Region Analysis
4. Statistical Pattern Recognition
5. Mathematical Morphology
6. Neighborhood Operators
7. Conditioning and Labeling
8. The Facet Model
9. Texture
10. Image Segmentation
11. Arc Extraction and Segmentation

Next semester covers higher-level techniques.

Chapter 1 Computer Vision: Overview

1.1 Introduction

computer vision is the science that develops the theoretical and algorithmic basis by which useful information about the world can be automatically extracted and analyzed from an observed image, image set, or image sequence from computations made by special-purpose or general-purpose computers.
computer vision: to emulate human vision with computers
computer vision: dual process of computer graphics: 2D $\rightarrow$ 3D

Information:

- recognition of a generic object
  =====Nalwa, *A Guided Tour of Computer Vision,* Cover=====
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 2.4=====
- three-dimensional description of an unknown object
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 6.9=====
- position and orientation of the observed object
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 6.1=====
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 6.6=====

- measurement of any spatial property of an object
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 5.1=====

=====Oldie 34:10=====

Applications:

- vision-guided robot assembly
- inspection tasks: mensuration, verification that all parts are present, determination that surfaces have no defects
- pattern recognition: geographic information system
  =====C. K. S. International Airport=====
- alignment: printed circuit board drilling
- measurement: length, area
- stereo: 3-D reconstruction
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 7.9=====
- motion and surface structure recovery
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 8.1=====
- interpretation
  =====Nalwa, *A Guided Tour of Computer Vision,* Fig. 10.1=====

image: spatial representation of object, 2D or 3D scene, or another image
intensity image: optic or photographic sensors, radiant energy
range image: line-of-sight distance
image $I(r, c)$: intensity value at row r and column c of the matrix
pixel: picture element: has properties of position and value
gray levels: pixel values of intensity images,
0 (black) – 255 (white) for 8-bit integers

factors determining the difficulty of computer vision problem:

- kind of object
- way objects are lit
- background
- kind of imaging sensor
- viewpoint of the sensor
=====Fig. 1.2=====

(a) white (gray tone 255) square on black (gray tone 0) background
(b) corners from corner-feature extractor, N: noncorner, C: corner
=====Fig. 1.1=====

atomic image features:

- edge
- corner
- hole
- topographic labelings of the gray tone intensity surface e.g. peaks, pits, ridges, valleys

Composite features: atomic features merged

- arcs: edge or ridge pixels linked together
- regions: connected sets of pixels with similar properties

1.2 Recognition Methodology

Recognition methodology must pay attention to:

1. image formation e.g. perspective or orthographic projection
2. conditioning
3. labeling

4. grouping
5. extracting
6. matching

1.2.1 Conditioning

Conditioning is based on a model that suggests that the observed image is composed of an informative pattern modified by uninteresting variations that typically add to or multiply the informative pattern.
e.g. noise suppression, background normalization

1.2.2 Labeling

Labeling is based on a model that suggests that the informative pattern has structure as a spatial arrangement of events, each spatial event being a set of connected pixels.
e.g. thresholding, edge detection, corner finding

1.2.3 Grouping

The grouping operation identifies the events by collecting together or identifying maximal connected sets of pixels participating in the same kind of event.
before grouping: pixels, after grouping: sets of pixels
e.g. segmentation, edge linking

1.2.4 Extracting

The extracting operation computes for each group of pixels a list of its properties.
example properties: centroid, area, orientation, spatial moments
e.g. region holes, arc curvature

1.2.5 Matching

Matching operation determines the interpretation of some related set of image events, associating these events with some given three-dimensional object or two-dimensional shape.
e.g. template matching
=====Garfield 17:82=====

1.3 Outline of Book

This text describes those aspects of computer vision that are needed in robotics and other real-world applications such as industrial-part inspection, medical diagnosis, aerial-image interpretation, and space station maintenance.

Journals

- *IEEE Transactions on Pattern Analysis and Machine Intelligence*
- *IEEE Transactions on Image Processing*
- *IEEE Transactions on Robotics and Automation*
- *IEEE Transactions on Systems, Man, and Cybernetics*
- *Computer Vision and Image Understanding*
  formerly *CVGIP (Computer Vision, Graphics, and Image Processing): Image Understanding*
- *CVGIP: Graphical Models and Image Processing*
- *International Journal of Computer Vision*
- *Pattern Recognition*
- *Pattern Recognition Letters*

- *Image and Vision Computing*
- *Machine Vision and Applications*

Conferences

- Asian Conference on Computer Vision
- IEEE Conference on Computer Vision and Pattern Recognition
- Image Understanding Workshop
- International Conference on Computer Vision
- International Conference on Image Processing
- International Conference on Pattern Recognition
- Scandinavian Conference on Image Analysis
- SPIE (The International Society for Optical Engineering) Conferences

Bibliography

- D. H. Ballard and C. M. Brown, *Computer Vision,* Prentice-Hall, Englewood Cliffs, NJ, 1982.

- G. A. Baxes, *Digital Image Processing,* Wiley, New York, 1994.
- K. Castleman, *Digital Image Processing,* Prentice-Hall, Englewood Cliffs, NJ, 1996.
- E. R. Davies, *Machine Vision: Theory, Algorithms, Practicalities,* 2nd Ed., Academic Press, San Diego, CA, 1997.
- R. O. Duda, P. E. Hart and D. G. Stork, *Pattern Classification,* 2nd Ed., Wiley, New York, 2001.
- E. Gose, R. Johnsonbaugh, and S. Jost, *Pattern Recognition and Image Analysis,* Prentice-Hall, Englewood Cliffs, NJ, 1996.
- R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision,* Vol. II, Addison Wesley, Reading, MA, 1993.
- B. K. P. Horn, *Robot Vision,* MIT Press, Cambridge, MA, 1986.
- A. K. Jain, *Fundamentals of Digital Image Processing,* Prentice-Hall, Englewood Cliffs, NJ, 1989.
- J. S. Lim, *Two-Dimensional Signal and Image Processing,* Prentice-Hall, Englewood Cliffs, NJ, 1990.
- D. Marr, *Vision,* W. H. Freeman, San Francisco, 1982.

- V. S. Nalwa, *A Guided Tour of Computer Vision,* Addison Wesley, Reading, MA, 1993.
- M. Petrou and P. Bosdogianni, *Image Processing: The Fundamentals,* Wiley, New York, 1999.
- W. K. Pratt, *Digital Image Processing,* 2nd ed., Wiley-Interscience, New York, 1991.
- R. J. Schalkoff, *Digital Image Processing and Computer Vision: An Introduction to Theory and Implementations,* Wiley, New York, 1989.
- R. J. Schalkoff, *Pattern Recognition: Statistical, Structural, and Neural Approaches,* Wiley, New York, 1992.
- M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis, and Machine Vision,* 2nd Ed., Brooks/Cole, Pacific Grove, CA, 1999.
- E. Trucco and A. Verri, *Introductory Techniques for 3-D Computer Vision,* Prentice-Hall, Upper Saddle River, NJ, 1998.
- S. E. Umbaugh, *Computer Vision and Image Processing: A Practical Approach Using CVIPtools,* Prentice-Hall, Englewood Cliffs, NJ, 1998.
- B. A. Wandell, *Foundations of Vision,* Sinauer, Sunderland, MA, 1995.

=====this file: ∼fuh/vcourse/haralick/chapter1.tex=====
man latex
latex chapter1
dvips -f chapter1 >! t.ps
ghostview t.ps
=====∼fuh/vcourse/haralick/programs/style/lena.im=====
=====pseudo.c=====
=====lena.b.im=====
Add /usr/local/vision/man to $MANPATH
Add /usr/local/vision/solaris2 to $PATH
Add /usr/local/SUNWspro/bin to $PATH
Add /usr/local/vision/solaris2 to $LD_LIBRARY_PATH
man hvision: image processing functions.
cc -o pseudo pseudo.c -lhvision -lm
pseudo lena.im lena.r.im lena.g.im lena.b.im
hview lena.im
laserhs lena.im
(to print: change textwidth ⇒ 7.2in, textheight ⇒ 9.75in,
Huge ⇒ normalsize, LARGE ⇒ normalsize)
dvips -t landscape -f chapter1 >! t.ps
lpr t.ps

copy ∼fuh/.tkinit to your home directory before invoking tk
tk
$>$ list command
$>$ read image : lena.im lena.im
$>$ affine transform : lena.im lena.aff 0 2 2 0 0 0 0
$>$ list image : ??
$>$ list image
$>$ view image : lena.aff
$>$ write image : lena.aff lena.aff
$>$ quit
=====lena.aff=====
=====joke=====

Project due Oct. 4:

1. Use B_PIX to write a program to generate
   (a) upside-down lena.im
   (b) right-side-left lena.im
   (c) diagonally mirrored lena.im
2. Use tk to

(a) rotate lena.im 45 degrees clockwise

(b) shrink lena.im in half

(c) binarize lena.im at 128 to get a binary image (hint: binarize)