CS 294-2 Visual Grouping and Object Recognition

Jitendra Malik, CS Division, Dept. of EECS
malik @ cs.berkeley.edu

Fall 1999
310 Soda Hall, MW 230-4

Course Content

This course examines three problems in vision: grouping, figure-ground and recognition. Grouping is the partitioning of the image or video stream into regions/spatiotemporal volumes corresponding to a single object or part of an object. The figure-ground problem is to label contours as belonging to one or the other of the two regions it borders. Recognition is identifying in images, objects, surfaces or movements as instances of previously seen classes.

We develop a unified framework for these visual capabilities, modeled as a set of probabilistic inference problems. Relevant psychophysical/perceptual phenomena will be reviewed, many of which were first discovered by the Gestaltists, back in the 1920s and 30s. For modeling, we take a twin track approach (1) uncover the ecological optics constraints and statistical regularities that underly any solution of these problems, whether by a biological or computer vision system (2) develop algorithms that solve specific problems based on these constraints.

The course could be of interest to several audiences. For students of biological or computer vision, grouping, figure-ground and recognition are central research problems. For students in artificial intelligence, machine learning, or statisticians interested in image analysis, these problems offer an opportunity to develop their methodology in the context of the grand challenge of vision. In addition, various embedded graph partitioning and matching problems, and the application of MCMC, may be of interest to CS theoreticians.

Lecture Topics

1. The 3 problems: grouping, figure-ground and recognition, and their inter-relationship. Quick survey of the principal factors that lead to grouping, figure-ground, and recognition.

2. Visual field properties: brightness, color, texture, disparity, motion.

3. Segmentation of piecewise-smooth brighness images in 3 different formulations. ML approach using EM with mixture of Gaussians model. Bayesian approach using spatial smoothness prior in MRFs; introduction to MCMC. Spectral graph theory approach using brightness similarity.

4. Multiscale, multi-orientation filtering; orientation energy maxima; textons.

5. Statistics of natural images in wavelet and texton representations.
6. Combined use of brightness, color and texture for grouping.
7. Grouping based on good continuation.
8. Grouping based on common fate. Tracking.
9. Grouping based on familiar configuration/movement. Image templates and point sets. Simultaneous grouping and recognition illustrated on handwritten digits, faces and Johansson's moving light displays.
10. Depth ordering information available at junctions, contour neighborhoods monocularly, binocularly and during motion.
11. Symmetry as a grouping and figure-ground cue. Variants of the medial axis transform.
12. Convexity and closure as grouping and figure-ground cues.
13. Ecological statistics of different grouping and figure-ground cues.
14. Graphical models for probabilistic inference of grouping and figure-ground relationships.
15. Review of key perceptual findings on recognition. Importance of color, texture, shape, material and movement attributes. Role of context. View dependence.
16. Bayesian classification. Nearest neighbor classification.
17. Modeling similarity. Goldmeier's studies. Prototypes. PCA, factor analysis, MDS. Shape spaces.
18. Discriminative approach to classification. Support Vector Machines.
19. Grand Unified Theory?

There is no required text for this course. Selected papers will be distributed in class or made available as a reader.

Some previous exposure to human or machine vision would be very helpful. Any previous class in computer vision (e.g. CS 280) or biological vision would do. If you have zero background in vision, as a minimum read lecture the Perception chapter in Russell and Norvig's book on Artificial Intelligence. It is also advisable to review your undergraduate probability material (e.g. concepts such as random variables, conditional and joint probabilities, Gaussian distributions, Bayes theorem).

We will use a scribe system to make course notes available through the semester. Each lecture, one or two students will take turns taking notes and typing them up. I'll edit and make the notes available on the web.

The grade will be determined by a combination of home assignments, scribe notes, and a final project. The project could be the mathematical/statistical analysis of a visual task or the implementation of some interesting algorithm or some psychophysical experiment.

You'll be encouraged to work in teams for the projects and for the home assignments.

I hope you enjoy the course!