Stat 238:
Treat as an Advanced Course
Grade: Projects / Homework
Vision as Bayesian Inference.

Note: But mainly hand-outs (papers)
      Link from Course to Research

Topics: Images $\rightarrow$ Segment
        Detect objects $\rightarrow$ Recognize
        \quad and other aspects like Motion / Stereo / Depth

Philosophy: $\rightarrow$ Consult to Dr. Zhu.
            (2 covers)

Bayesian $P(I \mid W) \rightarrow$ Generative Model
        $P(W) \rightarrow$ Prior Model

Inference $W^* = \arg \max_{W} P(W \mid I) = \frac{P(I \mid W) P(W)}{P(I)}$

Issues: (i) Representation
        (ii) Learning
        (iii) Inference

This issue will fill out by Dr. Zhu
Strategy → Tech best work on these topics & Current State of the Art (Technique)

List of Topics:

(1) Edge Detection (Statistical)
(2) Marginal Shaf / Other Generalization / Prior
(3) Max-Flow / Min-Cut Algorithm
(4) Reprint Computer / Image Processing
(5) Neighborhood City → Sharon's Queen
(6) Discriminant Random Fields
(7) AdaBoost for Detection
(8) Perona's Anisotropy Model
(9) Essman's Video Cogid
(10) Deconvoluted Object Detection → Cognido
(11) Priit Shlay
(12) Activ Appraiser Models
(13) Discriminative w/ MRF's & Discriminative Kemp
(14) Niranjan Erkopyr
(15) Ultra Otherran
Statistical Region Segmentation

Images: Handlabelled
6 classes: $x \in A$
edge, vegetation, air, road, building, other.

$\text{Images } \rightarrow \text{ set of pixels}$
$I(i,j)$
$\gamma = 1 \text{ to } 255$
$\delta = 1 \text{ to } 255$
$0 \leq I(i,j) \leq 255$
$256 = 2^8$

Colour $\rightarrow$ three images
$R(i,j)$, $B(i,j)$, $G(i,j)$ red, green, blue.

Fitters: $\phi_N(i,j)$
$\phi_N + I(i,j) = \sum_{k,l} \phi_{kl}(i-k,j-l) I(k,l)$
convolution.
For continuous $\phi_d * I(x,y) = \int \phi_d(x-u,y-v) I(u,v) du dv$

Strategy used labelled images to learn
conditional distribution:
$P(\phi_a(i,j) \mid (i,j) \in \text{ Region class } d)$

Problem: how to represent the distributions?
how to generalize to new data?
Cross-validation: learn on part of the dataset, evaluate on the rest.
Statistical Region Segmentation

Cure-of-Dimensionality

\[ \implies \text{not enough data.} \]

Represent probability distribution better:

(a) parametrically - e.g. Gaussian with parameters \( \mu, \sigma \)

(b) non-parametrically - e.g. Histogram

Parametric means choosing a model for the data.

What model? Gaussians are non-robust due to outliers.

Non-parametric makes less model choice, but requires more data.

Typically, parametric has \( O(N^2) \) parameters, where \( N \) is dimension of filter.

Non-parametric has \( O(M^{n^2}) \) parameters.

Require \( \rightarrow \) more training data than this, by factor of 5-10.

We choose non-parametric, but with adaptively chosen bin boundaries.

Add \( c \) up to \( N = 9 \) bin filters.
Statistical Region Segmentation

Filter: \[ |\nabla I| \quad \text{image gradient} \quad \frac{1}{\sqrt{(dI)^2 + (dI)^2}} \]

Laplacian: \[ \nabla^2 I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \]

Gaussian: \[ N = \int G(x) \cdot VI \cdot VI^T \, dx \quad \text{Nitzberg-Harits} \]

Intuition: in flat region of image, \( N \) has 2 zero-eigenvectors and 2 big eigenvalues. At edges, \( N \) has 1 big eigenvalue, 1 small eigenvalue. At corners, \( N \) has 2 big eigenvalues.

In addition, filters can be run at various scales by smoothing the image with a Gaussian:
\[ G(x; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2} \frac{x^2}{\sigma^2}} \]

Larger \( \sigma \) \( \Rightarrow \) larger scale.

Apply filters: learn \( P(\phi_\mu \mid x) \)
Classify \( \hat{a}(x, y) = \text{arg} \max_{a} P(\phi_\mu(x, y) \mid a) \)

Error: \( P(\hat{a} \neq a \mid a) \)

Results: can classify pixels surprisingly well as road, vegetation, sky, from either texture or colour cues.