lecture 15

- probability review

- maximum likelihood method

- cue combinations (sketch only - see Exercises)

Thursday Nov. 5, 2015
Scene \xrightarrow{formation} \text{image} I(x, y) \xrightarrow{vision} \text{perceived scene}

- contrast $\frac{\Delta I}{I_0}$
- depth
- disparity
- $(V_x, V_y)$
- surface orientation $(\text{slant \& tilt})$

- contrast $\frac{\Delta I}{I_0}$
- depth
- disparity
- $(V_x, V_y)$
- surface orientation $(\text{slant \& tilt})$
What are the sources of uncertainty in a perceived scene?

- images may contain noise
- images may contain insufficient information (many possible scenes map to one image)
- neural computations are noisy

Goal of today's lecture and next one:

To sketch out a probabilistic model of uncertainty in perception of scene parameters, such as parameters listed on previous slide.

I am assuming here that you are familiar with the basic definitions of probability. I will give a brief review of the definitions, for our particular context. But this is not meant to replace the offline review that you should do if you are rusty.
Review of Probability (Sketch)

\[ p(I) \] \{ marginal \}

\[ p(S) \] \{ probabilities \}

\[ p(I, S) \] - joint probability

\[ p(I | S) = \frac{p(I, S)}{p(S)} \]

\[ p(S | I) = \frac{p(I, S)}{p(I)} \]  \{ Conditional probabilities \}
scene variables (not directly measured)

\[ S = s \]

\[ p(S = s) \]

\[ I = \hat{i} \]

image intensities or Gabor outputs (directly measured)
\[ p(I, S) \]

\[ p(I=i, S=s) \]

\[ p(I, S) \] is a “joint” probability function.
\[ p(S = s_j) = \sum_{i \in I} p(I = i, S = s_j) \]
$S = s_j$

$I = i$

$p(I, S) = p(I = i, S = s_j)$

$p(I = i) = \sum_{s_j \in S} p(I = i, S = s_j)$

"$p(I)$ is a marginal probability function"
\[ P(I = i \mid S = s_j) = \frac{P(I = i, S = s_j)}{\sum_{i \in I} P(I = i, S = s_j)} \]

\[ p(I \mid S) = \frac{p(I, S)}{p(S)} \]

is a conditional probability function.
\[ p(S = s_j | I = i) = \frac{\sum_{\substack{s_j' \in S}} p(I = i, S = s_j')}{p(I = i)} \]

\[ p(S | I) = \frac{p(S, I)}{p(I)} \]

is a conditional probability function.
Independent Random Variables

We say two random variables $X_1$ and $X_2$ are independent if, for all $i, j$,

$$p(X_1 = i, X_2 = j) = p(X_1 = i) \cdot p(X_2 = j)$$

For example, the noise at two pixels is generally assumed to be independent.
Example 1: image contrast

For simplicity, assume there is no noise in surround

Task: Given $I(x,y)$ and $I_0$, estimate $\Delta I$.

Pixel noise $n(x,y)$ is independent

---

Image not random

Scene
Assume noise $n(x,y)$ is Gaussian.

$$P(n(x,y) = n_i) = \frac{1}{\sqrt{2\pi} \sigma_n} e^{-\frac{n_i^2}{2\sigma_n^2}}$$

$$= e^{-\frac{(I(x,y) - I_0 - \Delta I)^2}{2\sigma_n^2}}$$
Assume noise at different pixels is independent.

\[ p(n) = \prod_{(x, y) \in \text{center}} p(n(x, y)) \]

\[ p(n) = \prod_{\text{numpixels in center}} \left( \frac{1}{\sqrt{2\pi} \sigma_n} \right) e^{-\frac{(I(x, y) - I_0 - \Delta I)^2}{2 \sigma_n^2}} \]

\[ p(n) = \prod_{(x, y)} e^{-\sum_{(x, y)} \frac{(I(x, y) - I_0 - \Delta I)^2}{2 \sigma_n^2}} \]
We want to infer $\Delta I$, given $I(x,y), I_0$. $\Delta I$ is random, in the sense that we are uncertain what it is. It also happens to be fixed i.e. same value for each $(x,y)$.

\[
p(n) \quad \text{not random}
\]

\[
= \quad p(I \mid \Delta I, I_0)
\]

\[
= \quad c \cdot e^{-\frac{1}{2\sigma^2} \sum_{x,y} (I(x,y) - I_0 - \Delta I)^2}
\]
The condition probability function \( p(I|S) \) is known as the likelihood function of I, given S.

One often writes \( p(I|S) = \mathcal{L}(S|I) \).

CAREFUL WITH NOTATION:

This likelihood function \( \mathcal{L}(S|I) \) is not the conditional probability of S, given I.

I will avoid this \( \mathcal{L}(S|I) \) notation.
Maximum Likelihood Method:

Given an image instance $I = i$, choose scene $S = s$ that maximizes $p(I = i | S = s)$.
Example

Choose \( \Delta I \) that maximizes \( p(I(x,y) \mid \Delta I) \)

\[ p(I \mid \Delta I) \]

\( \Delta I \)

choose this one
To find $\Delta I$ that maximizes

$$p(I | \Delta I) = e^{-\frac{1}{2\sigma^2} \sum_{x,y} (I(x,y) - I_0 - \Delta I)^2}$$

we find $\Delta I$ that minimizes

$$\sum_{x,y} (I(x,y) - I_0 - \Delta I)^2.$$

**Solution:** Exercise

ML estimate of $\Delta I = \text{mean}(I(x,y)) - I_0$
N = 400, Δ I = 2, sigNoise= 3 mean is 2.1969

CAREFUL
Not the correct estimate.
Example 2: motion estimation

$I(x,y,t)$

$\begin{pmatrix} v_x & v_y \end{pmatrix}$

unique solution

add noise

NOT unique solution
motion constraint equation

$$\frac{\partial\epsilon}{\partial x} v_x + \frac{\partial\epsilon}{\partial y} v_y + \frac{\partial\epsilon}{\partial t} = 0$$

motion constraint equation with "noise"

(assume independent at different (x,y,t))

$$\frac{\partial\epsilon}{\partial x} v_x + \frac{\partial\epsilon}{\partial y} v_y + \frac{\partial\epsilon}{\partial t} = n(x,y,t)$$

(assume \(G(0, \sigma_n^2)\))
Maximum likelihood estimate of \((v_x, v_y)\)?

\[
p(I \mid v_x, v_y, \sigma^2) = \prod_{x,y,t} \frac{1}{\sqrt{2\pi} \sigma_n} e^{-\frac{1}{2} \sum \left( \frac{\partial I}{\partial x} v_x + \frac{\partial I}{\partial y} v_y + \frac{\partial I}{\partial t} \right)^2 / 2 \sigma_n^2}
\]

Maximizing the conditional probability is achieved by choosing \((v_x, v_y)\) that minimizes the sum of squares. (Further details omitted.)
Maximum likelihood methods are commonly used! (and Bayesian methods too -- next lecture)

- "Ideal Observers" (Computer Vision)
  e.g. what I just did: set up and solve a minimization problem
  (image contrast, image velocity)

- Population Coding (Computational Neuroscience)

  Given responses I(x,y) * f(x,y) of model neurons e.g. V1 cells
  estimate attributes of image I(x,y) such as orientation, disparity,
  depth, image velocity, .....

- Psychophysics (Behavioral Psychology)

  Use psychometric functions and thresholds to model how human observers
  are using available information to perform task.
Example 1: Population Coding (sketch only)

Assignment 2 Question 1: Given responses of a population of orientation tuned cells, estimate "most likely" orientation at each pixel.

e.g. If we interpret orientation responses at each (x,y) as a likelihood function, then the dominant response would give the maximum likelihood orientation.

recall color coding
Assignment 2 Question 2: Given responses of a population of disparity tuned cells to a pair of images, estimate "most likely" disparity at each \((x,y)\).

e.g. If we interpret disparity responses at each \((x,y)\) as a likelihood function, then the dominant response would give the maximum likelihood disparity.
Psychophysics (Behavioral Psychology)

Task: Judge if the test is closer or further than the reference.

percent respond "test is further"

100
50
0

reference distance

test distance
Fit a curve to the data to get a psychometric function. We can either force the 50% point to be at the reference distance in the fit, or we can allow it to be a parameter and fit it also. The slope of the psychometric function around 50% performance is an indication of how good one is at the task: a large slope means one is good at it, and a small slope means one is poor at it. We take the change in test distance from 50% to 75% performance as a measure of how good one is at the task. This is sometimes called the "just noticeable difference" or JND.
One commonly fits a cumulative Gaussian function to the data. The underlying Gaussian is then *assumed* to be a likelihood function for the perceived depth of the test, when the test depth = reference depth. i.e. The maximum likelihood is the test value at the fitted *50% point*, and the standard deviation of the Gaussian is the fitted *JND*. (Note that this interpretation/assumption is just a convenient model. It does not follow directly from mathematical reasoning.)
How is the above used?

Consider the task of judging surface slant.

You can use various sources of information to do this:
- Texture (size, density, foreshortening)
- Stereo disparity
- ...
Consider two "conditionally independent" cues

\[ p(I_1 I_2 | s) = p(I_1 | s) p(I_2 | s) \]

**Example:** texture \( I_1 \) and stereo disparity \( I_2 \)

Each cue gives (independent) information about slant.
Two key questions:

- Given \( p(I_1 | s) \) and \( p(I_2 | s) \), what is the maximum likelihood estimate of \( s \) and what is the uncertainty when both stereo and texture are present?

  (Exercises)

- Does the model fit human percepts?

  (Oral Presentations)
Announcements

- exam 2
  I plan to grade them this weekend, and email you to let you know which questions you should do again.

- surveys
  - pace generally good
  - typeset lecture notes would be appreciated (I will try)

- oral presentation topics
  - only one week left