Prior probabilities and Bayes
- line drawings and 3D regularities
- motion prior
- surface orientation priors
- depth reversal ambiguities

For a given image \( I(x,y) \), there are many scenes that could have produced this image. Last lecture we looked at one way to choose the "best" scene parameter(s) for a given image, namely choose the parameter(s) that maximize the likelihood.

Example: interpreting a line drawing
Q: Are there possible 3D interpretations here other than a box?
A: Yes, a flat drawing. What else?

We can capture the preference for a cube interpretation by having a non-uniform marginal probability \( p(S) \) on the set \( S \) of scenes.

The marginal \( p(S) \) is often called a "prior", since it doesn't depend on the image instance.

We have a strong preference for the regular 3D cube over any of the wireframe objects with 'random' depth.

- We see a lot of cubes in the world.
- We like our shapes to be simple
- If we chose one of the random configurations, we would need to explain why \( z \) was random but the \( xy \) positions accidently gave rise to parallel lines in the image.

Maximum likelihood scheme \( p(I(x,y) | S) \) would not favour any one of these scenes over others.

Maximum likelihood method: given an image \( I = i \), find the scene parameters \( S = s_j \) that maximize \( p(I = i | S = s_j) \).

But... what we really want is to find the scene parameters that maximize \( p(S = s_j | I = i) \).

How are \( p(I = i | S = s_j) \) and \( p(S = s_j | I = i) \) related?

Bayes Rule

The "image prior" and "scene prior" are marginal distributions.
Maximum Posterior (MAP)
Given an image $I=i$, find the scene $S=s_j$ that maximizes $p(S = s_j | I = i)$.

If the prior $p(S)$ is uniform then the maximum likelihood gives the same solution as MAP.

Note that we don't care about $p(I = i)$ since it is fixed.

But often the prior is non-uniform.

Example 1

Related Example: Ames Room
Check out this video! [http://www.youtube.com/watch?v=Ttd0YjXF0no](http://www.youtube.com/watch?v=Ttd0YjXF0no)

Example 2: motion estimation

Recall likelihood depends on image noise (for fixed image contrast).

Motion is fast during saccades, but brain does not process this motion.

Example 3: surface orientation from texture

What is the likelihood function $p(I(x,y) | \theta, \tau)$?
What is the prior $p(\theta, \tau)$?
The prior is independent of whether cue is texture or stereo.

What about the prior on $(v_x, v_y)$?

(Retinal) image motion is typically slow.
Motion is fast during saccades, but brain does not process this motion.
What is the likelihood function $p( l(x,y) \mid \theta, \tau )$?

- $p(\text{foreshortening cue} \mid \theta, \tau )$
- $p(\text{density cue} \mid \theta, \tau )$
- $p(\text{size cue} \mid \theta, \tau )$

Models of these likelihood functions were developed in the 1990's and early 2000's. These papers can be technically challenging. Maybe we will see some of these in Oral Presentations.

Two-fold ambiguities: how to resolve them?

Sometimes the likelihood function $p( l(x,y) \mid S )$ has more than one local maximum.

What is the prior $p(\theta, \tau)$ on surface orientation?

For any slant, $p(\text{left wall}) = p(\text{right wall})$

$p(\text{frontoparallel})$ has a local maxima.

The foreshortening cue likelihood function has two local maxima.

We tend to perceive a rightward slant because other cues affect the likelihood function too:
- the visible sides of the coins
- the (slight) size gradient

Here we have a strong perception of upward slant. Why?

There are only weak texture cues e.g. size.

The prior for floor orientation may play a strong role here.

Compare the shapes. Roughly the same, yes?

Once you interpret the shapes as table tops, you apply your prior for seeing the shape in a "floor" rather than "ceiling" orientation. Then, you see the 3D shapes quite differently.

A valley illuminated from the right produces the same shading as a hill illuminated from the left.
Do you perceive the center point as a hill or a valley? (You should be able to flip it and see either.)

Note that when you see it as a hill, the overall slant is left. But when you see it as a valley, the overall slant is right.

For the center, we tend to see a hill in the left image and a valley in the right image.

Why? The images are the same (except for a flip), so surely the likelihood functions of the perceived shape are the same too.

The only difference is in the prior.

The visual system uses three priors to resolve the depth reversal ambiguity:
- surface orientation: \( p(\text{floor}) > p(\text{ceiling}) \)
- surface curvature: \( p(\text{convex}) > p(\text{concave}) \)
- light source direction: \( p(\text{from above}) > p(\text{from below}) \)

Convex shape, illuminated from above the line of sight

Concave shape, illuminated from below the line of sight

Lighting from below looks weird, because it competes with other priors (convexity) and other sources of information (shadows, ...).

In [Langer and Buelthoff, 2001], I showed how people combined the three different "priors":

Percent correct in judging points to be on local "hill" or "valley":

\[
\begin{align*}
\text{percent correct} &= \\
&= 50 + 10 \times \left\{ 1, -1 \right\}^{\text{floor} / \text{ceiling}} \\
&+ 10 \times \left\{ 1, -1 \right\}^{\text{light from above/below}} \\
&+ 10 \times \left\{ 1, -1 \right\}^{\text{globally convex/concave}}
\end{align*}
\]

Announcements
- A2 grades should be posted later today
- midterm solutions tomorrow
- A3 by end of week (hopefully)
- choose your Oral Presentation paper by end of week

http://www.cim.mcgill.ca/~langer/546/CourseOutline.html