COMP 546

Lecture 15

Cue combinations, Bayes (priors)

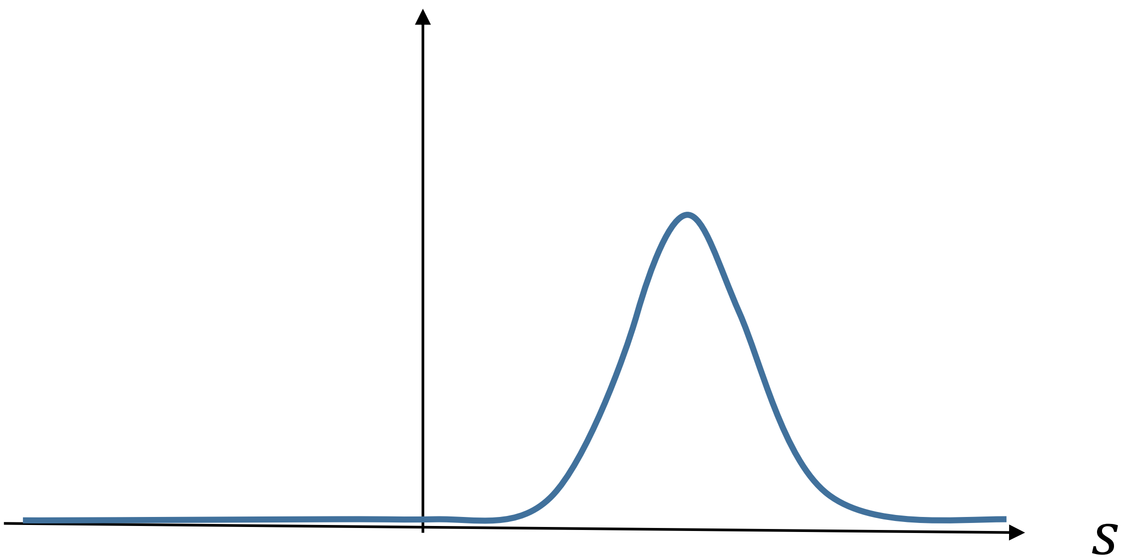
Thurs. Feb. 28, 2019

Recall from last lecture: Likelihood

Let I and S be two random variables, representing some image and scene property, respectively. The conditional probability

$$p(\, I = i \mid S = s \,)$$

is known as the "likelihood" of scene $S = s$, for that image $I = i$.



How to combine image cues ?

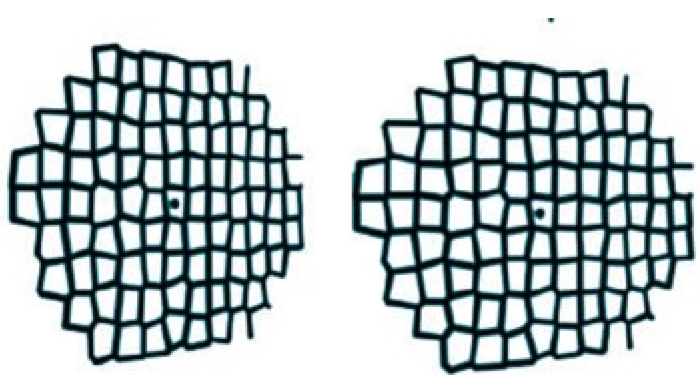
$$p(\, I_1,\ I_2\ ,\ I_3\ , \ldots \mid S\ )\quad = \quad ?$$

- binocular disparity ("stereo")
- image orientation
- 2D motion
- shading
- etc

Example:

texture only
(monocular)

stereo only



texture and stereo

[Hillis 2004]

Assume likelihood function is “conditionally independent”:

$$p(\, I_1, I_2 \mid S) \;=\; p(\, I_1 | \; S)\; p(\, I_2 | \; S)$$

e.g. I_1 is texture.

I_2 is binocular disparity.

Assume likelihood function is "conditionally independent":

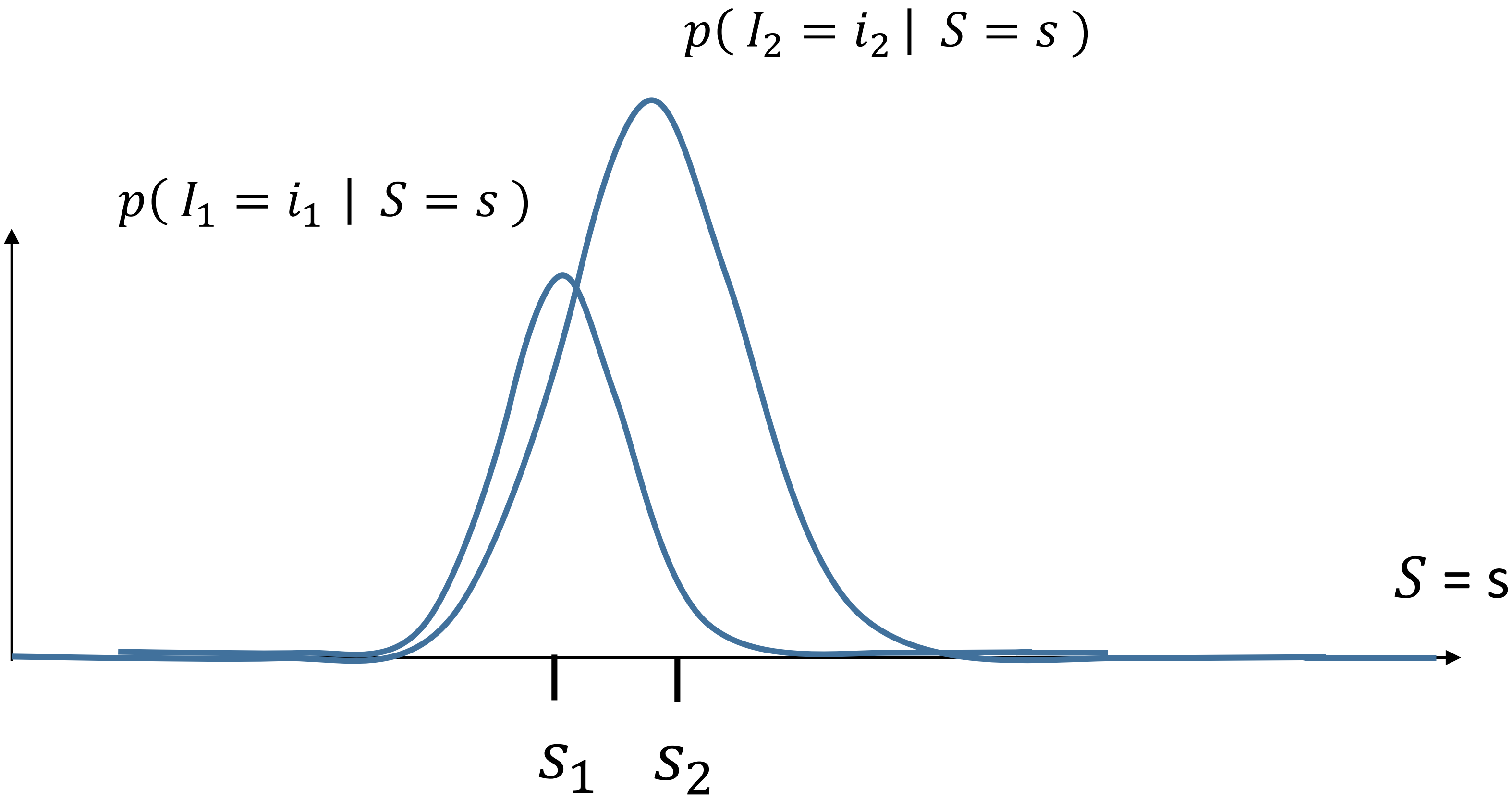
$$p(\, I_1, I_2 \ \mid S) \ = \ p(\, I_1 \mid S) \ p(\, I_2 \mid S)$$

That is, for any $I_1 = i_1, \ \ I_2 = i_2 \,, \ \ S = s$:

$$p(\, I_1 = i_1, I_2 = i_2 \ \mid S = s\,) \ = \ p(\, I_1 = i_1 \mid S = s\,) \ \ p(\, I_2 = i_2 \mid S = s\,)$$

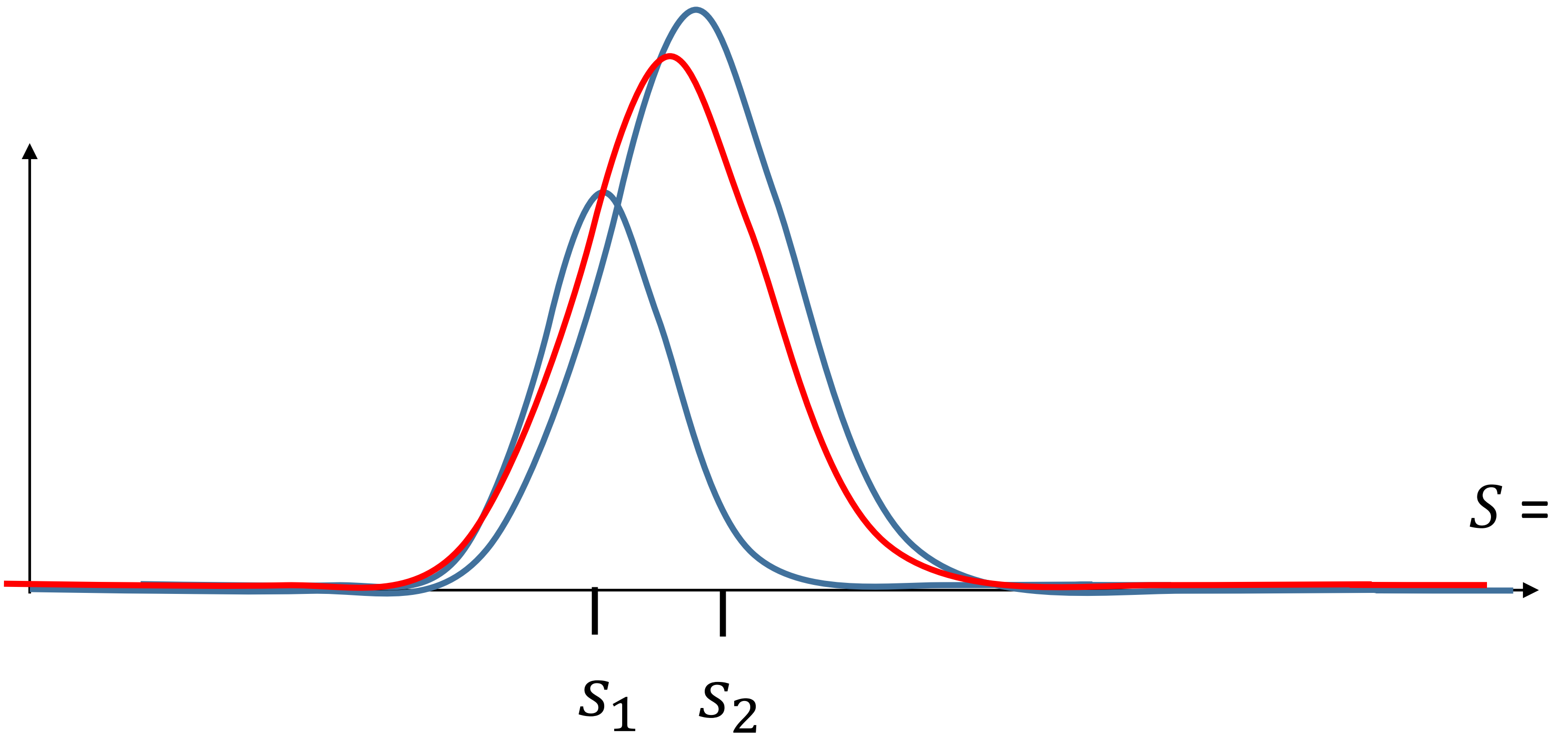
e.g. I_1 is texture.

I_2 is binocular disparity.



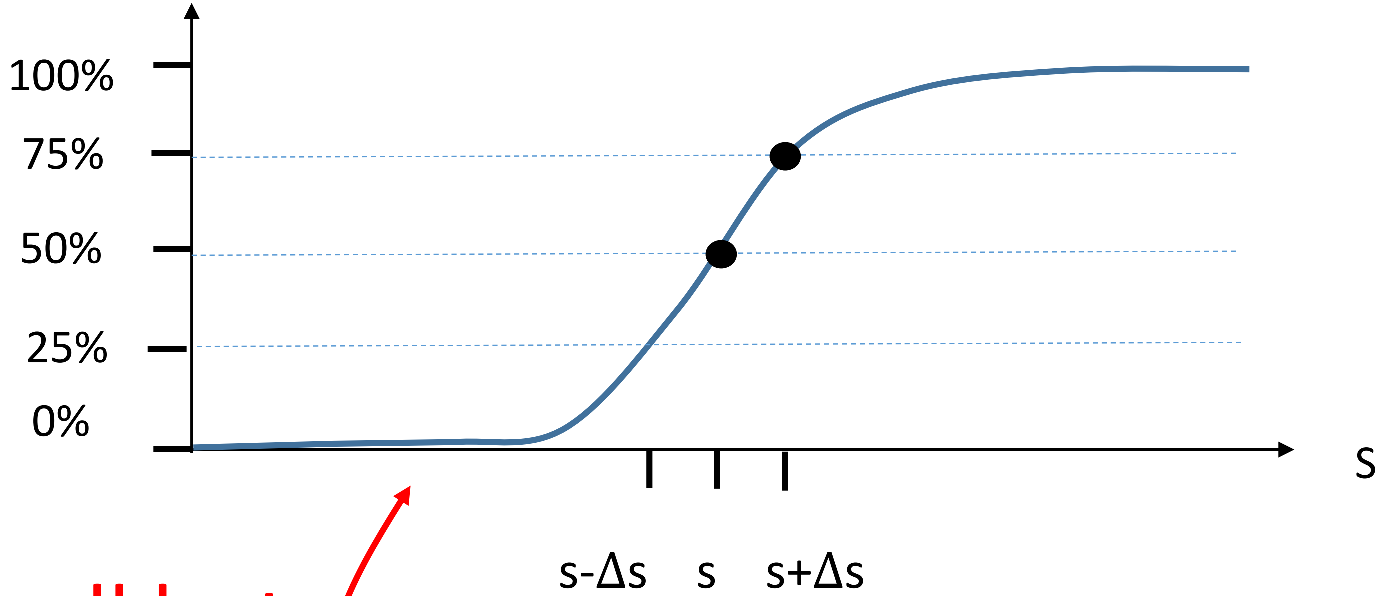
Their likelihood maxima might occur at different values of s and their variance (spread) might be different too.

$$p(\,I_1 = i_1, I_2 = i_2 \mid S = s\,) \;=\; p(\,I_1 = i_1 \mid S = s\,)\; p(\,I_2 = i_2 \mid S = s\,)$$

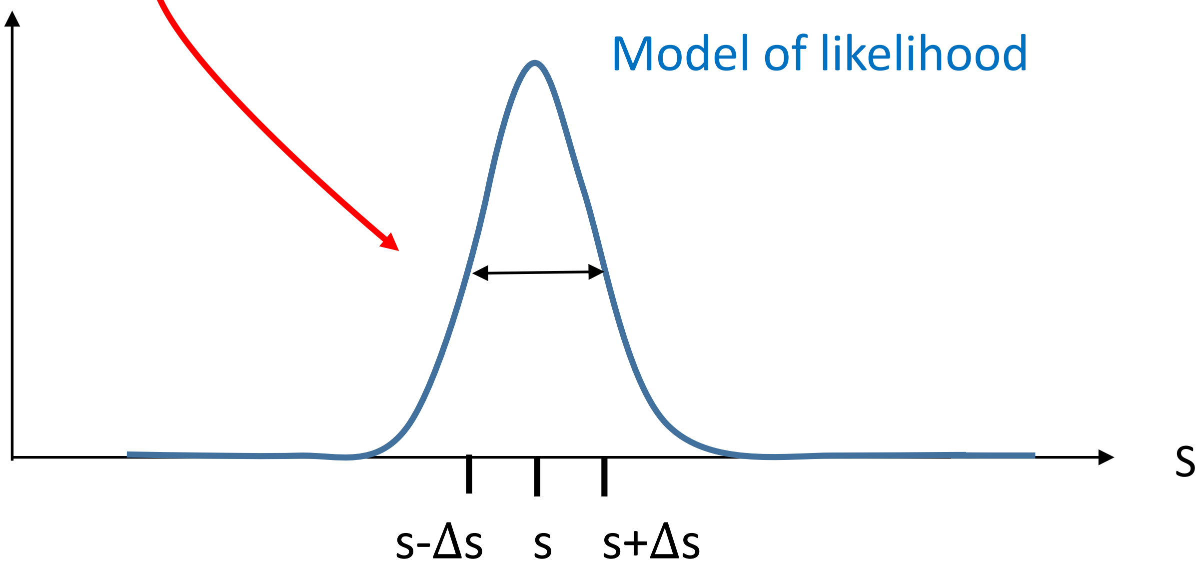


The conditional independence assumption gives us a model of what the likelihood function is when both cues are present.

Psychometric function



Recall last lecture



Q: What value of s maximizes the product of the likelihoods?

$$p(I_1 = i_1 \mid S = s) \;\; p(I_2 = i_2 \mid S = s) \; = \; e^{-\frac{(s - s_1)^2}{2\,\sigma_1{}^2}} \;\; e^{-\frac{(s - s_2)^2}{2\,\sigma_2{}^2}}$$

Q: What value of s maximizes the product of the likelihoods?

$$p(I_1 = i_1 \mid S = s) \;\; p(I_2 = i_2 \mid S = s) \;=\; e^{-\frac{(s - s_1)^2}{2\,\sigma_1{}^2}} \;\; e^{-\frac{(s - s_2)^2}{2\,\sigma_2{}^2}}$$

A: ("Linear Cue Combination") (see lecture notes)

$$s = w_1 s_1 + w_2 s_2$$

where

$$w_1 + w_2 = 1 \qquad\qquad 0 < w_i < 1$$

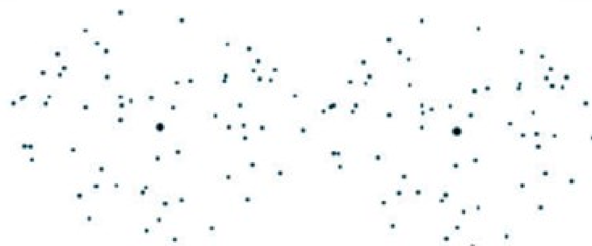
The more reliable cue gets more weight.

Psychophysical Method:

Measure discrimination thresholds (e.g. slant) for cues *in isolation*.

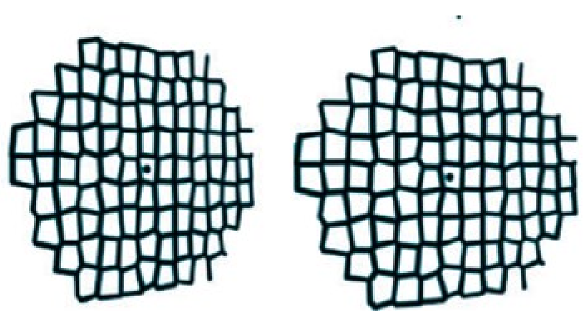
texture only

(monocular)



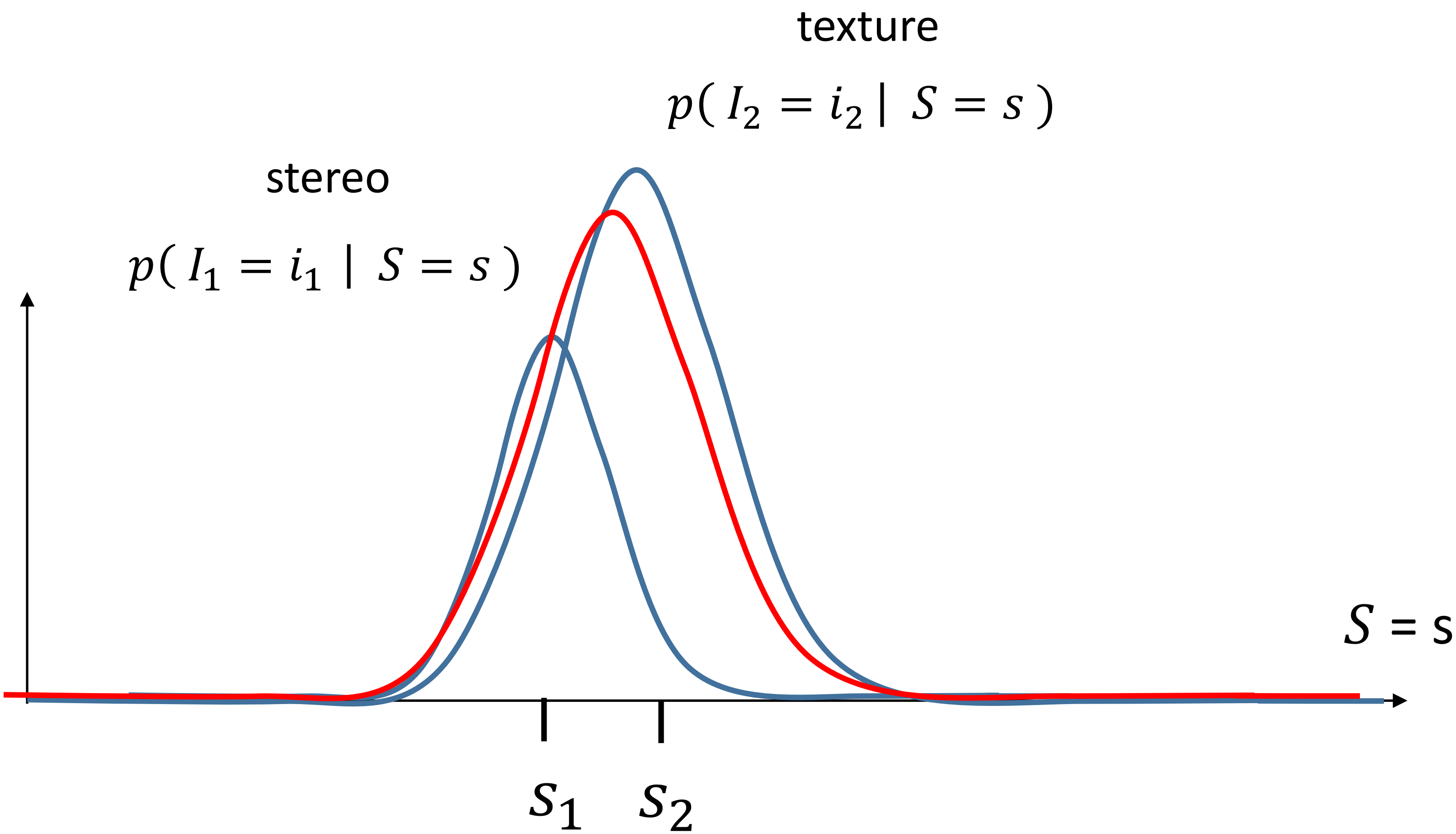
stereo only

Then, present cues together and re-measure the thresholds and check if they are consistent with the linear cue combination model.



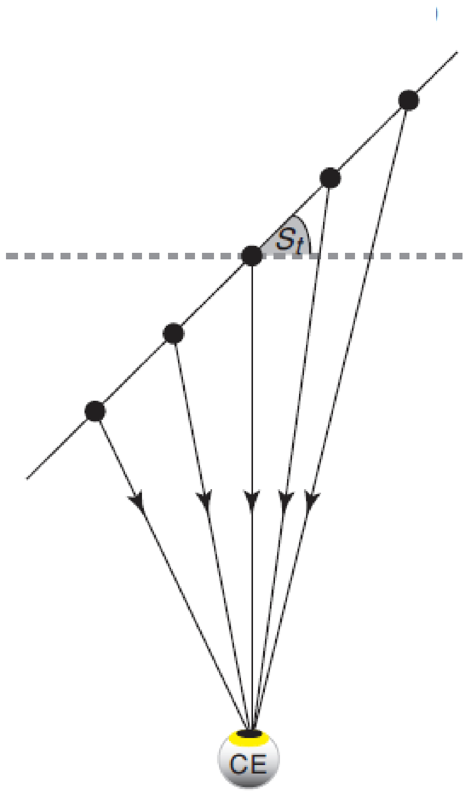
texture and stereo

[Hillis 2004]



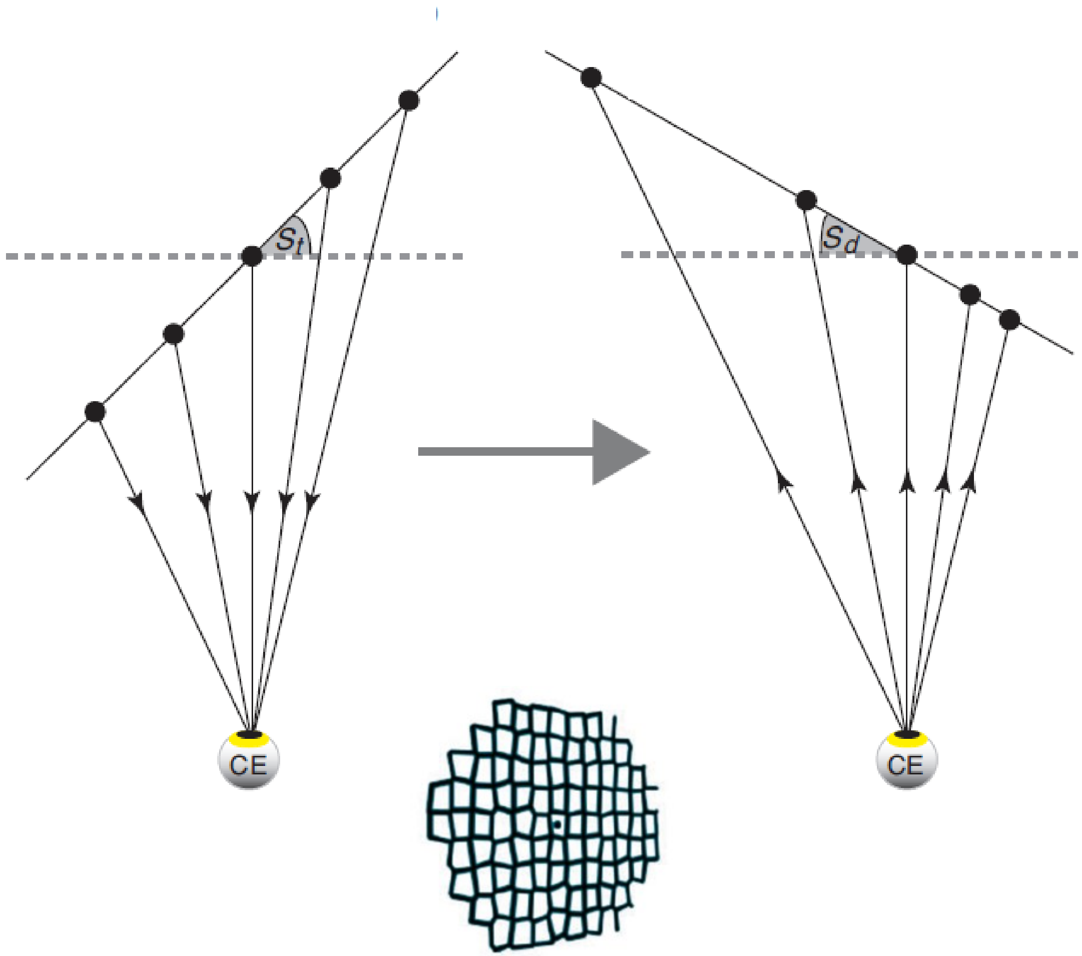
Experimenter can manipulate s_1 , s_2 , σ_1 , σ_2 and predict effect on perception of slant.

e.g. Creation of a "cue conflict" stimulus [Hillis 2004]



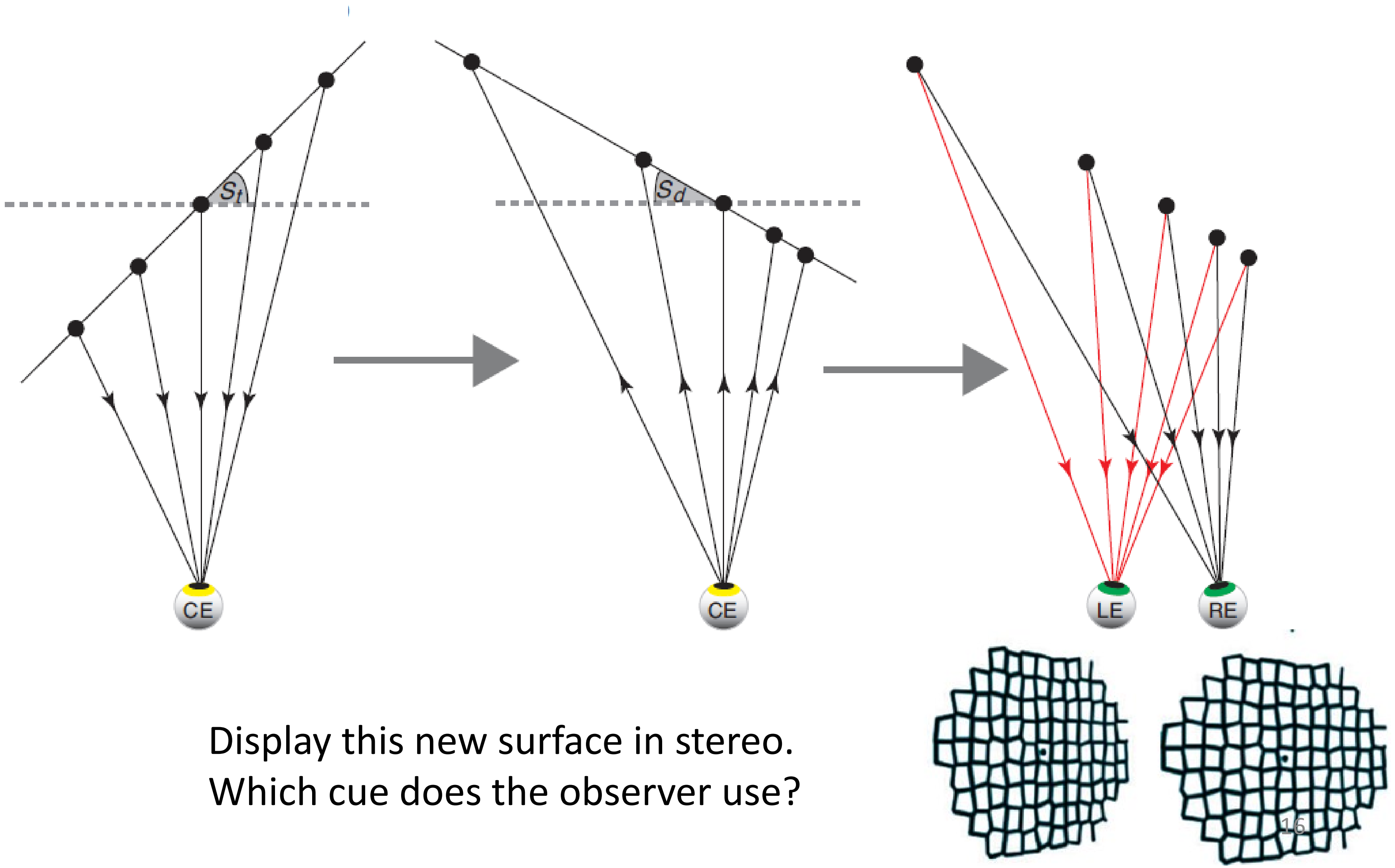
texture only
(monocular)

e.g. Creation of a "cue conflict" stimulus [Hillis 2004]

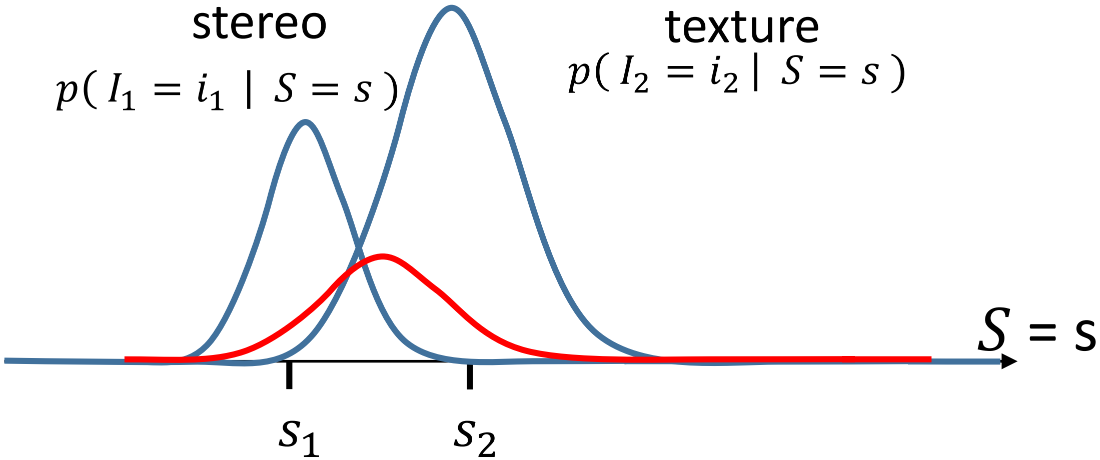


Compute where the surface markings (lines) would appear on a surface that has some different slant, but that produces the same image.

e.g. Creation of a "cue conflict" stimulus [Hillis 2004]

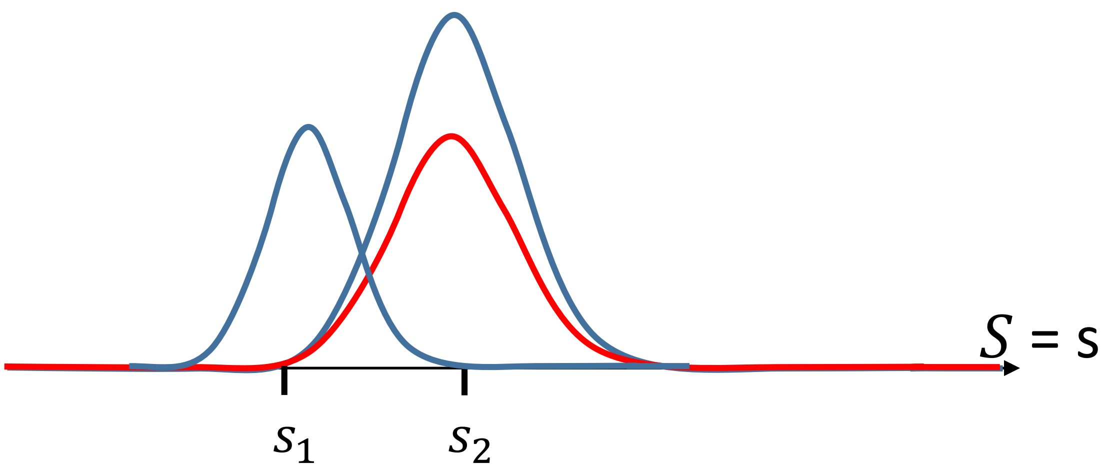


Display this new surface in stereo.
Which cue does the observer use?



Cue conflict situation:

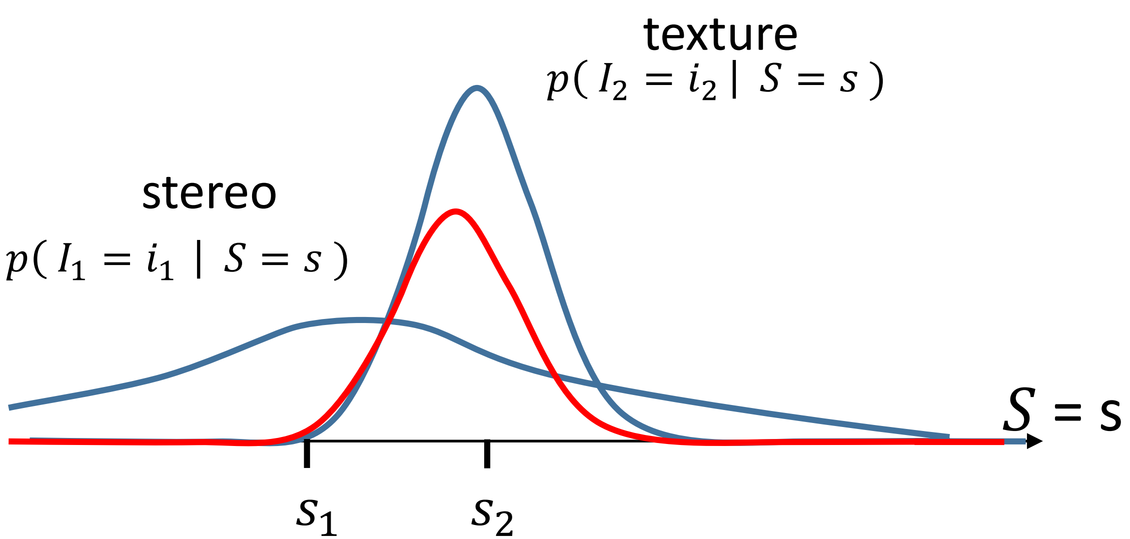
Do human observers linearly combine stereo and texture cue?



Or do human observers just ignore one of the cues? (because the conflict is too great) e.g. Only use the texture cue.

Cues can be made less reliable.

e.g.   Stereo is less reliable for objects that are farther away.



Linear cue combination theory says less reliable cue will have less weight.

This is what is found. [Hillis 2004]

COMP 546

Lecture 15

Cue combinations, Bayesian models (priors)

Thurs. Feb. 28, 2019

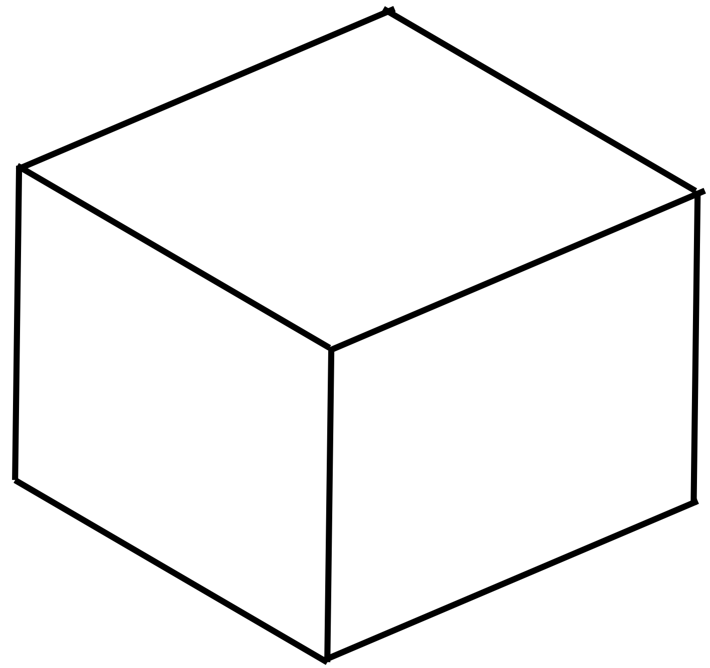
$$p(I = i | S = s) \neq p(S = s | I = i)$$

Likelihood of scene s, for image i.

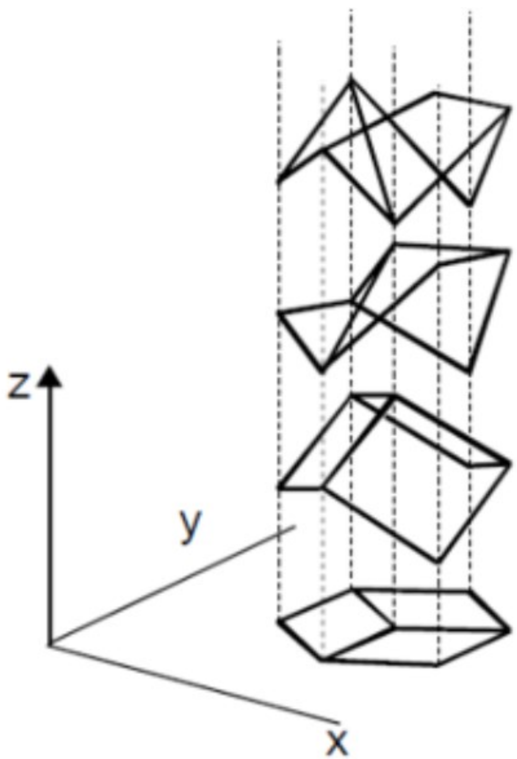
Condition probability of scene s, given image i

What is the crucial difference ?

Example: interpreting a line drawing



There are many scene geometries that can account for this image equally well.   So, why do we prefer the regular solid cube?

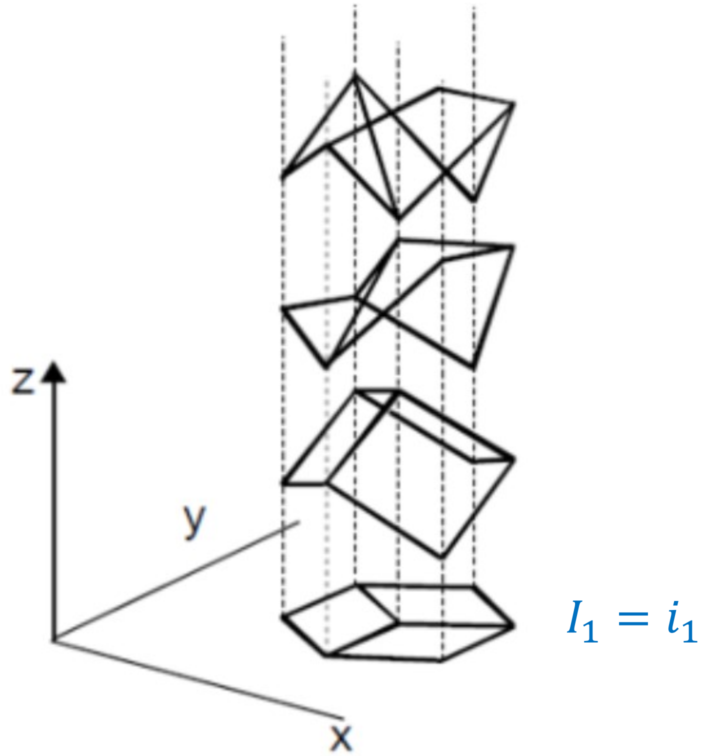


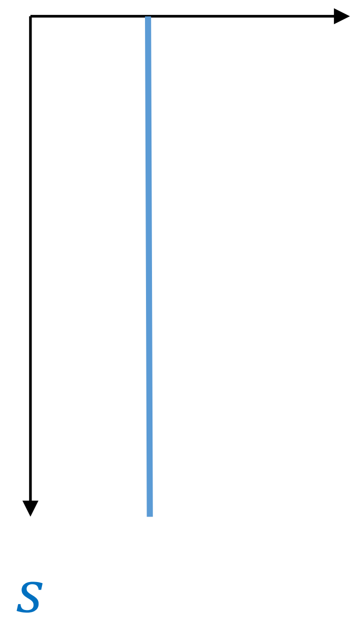
There are many objects that can account for this image equally well.

[Kersten & Yuille 2003]

likelihood

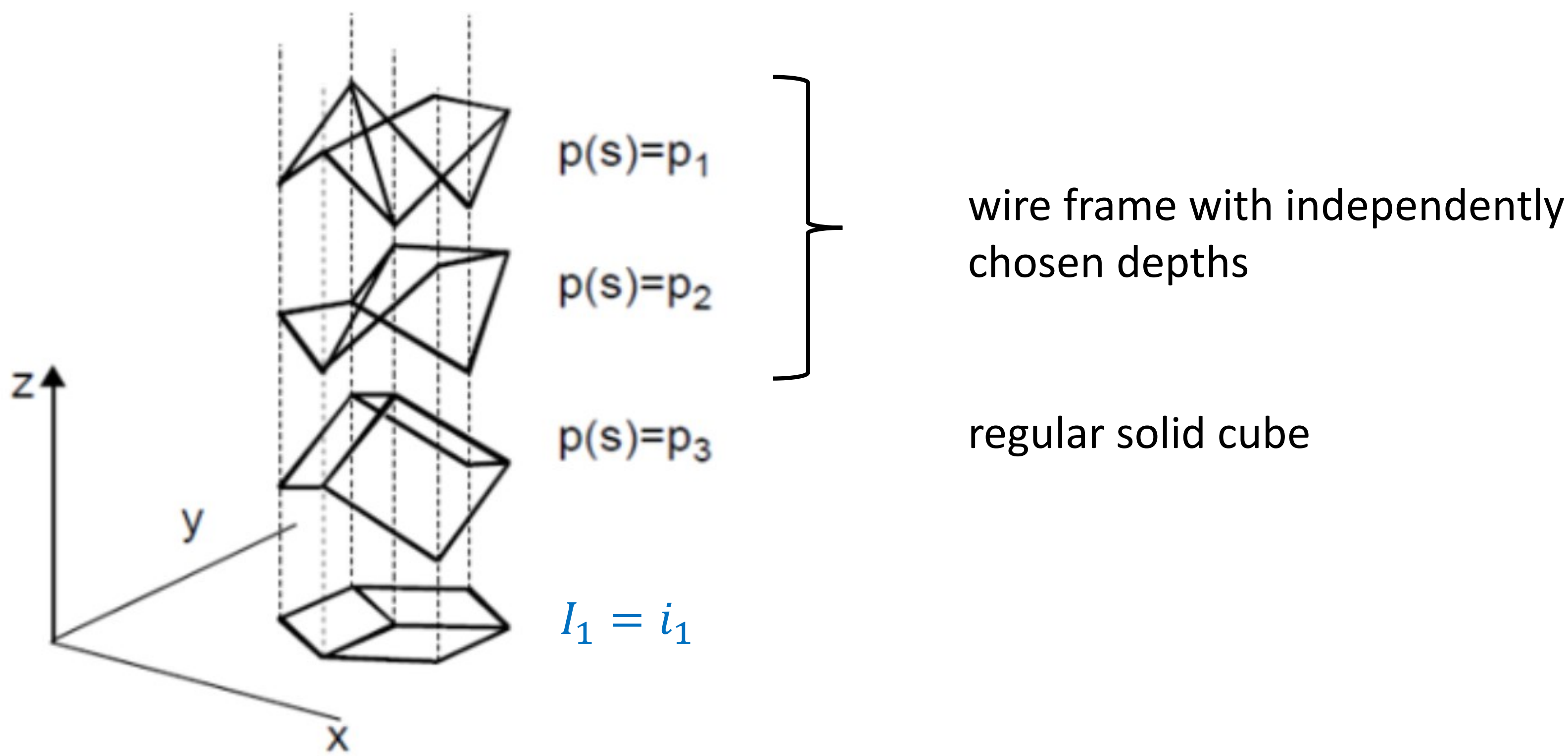
$$p(\, I_1 = i_1 |\; S = s \,)$$





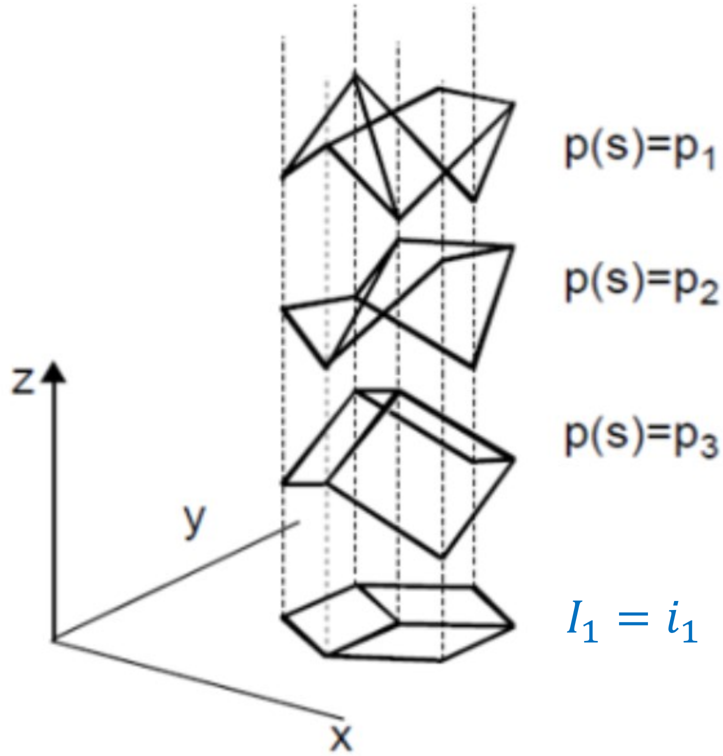
There are many objects that can account for this image equally well.
Thus the likelihood function over these objects is uniform.

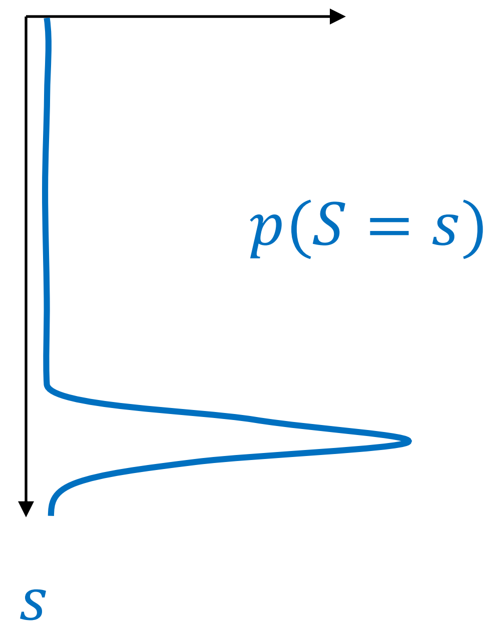
However, these scenes do not occur with equal frequency (probability) in the world. The *marginal* probability $p(S = s)$ is not uniform.



The marginal probably $p(S = s)$ is called the "prior“ probability.

prior





The reason we see the regular solid cube is that (we believe) it has greater probability of occurring than *each of the instances* of non regular geometries.

The more interesting cases arise when we need to consider *both* the likelihoods and priors.

How do we combine them ?

“Bayes Rule”

“likelihood” “scene prior”

$$p(S \mid I) = \frac{p(I \mid S)\ \ p(S)}{p(I\ )}$$

“posterior” “image prior”

MATH 323 Probability (3 credits)

Offered by: Mathematics and Statistics (Faculty of Science)

Overview

Mathematics & Statistics (Sci) : Sample space, events, conditional probability, independence of events, Bayes' Theorem. Basic combinatorial probability, random variables, discrete and continuous univariate and multivariate distributions. Independence of random variables. Inequalities, weak law of large numbers, central limit theorem.

Maximum *'a Posteriori'* (MAP)

"likelihood" "prior"

$$p(\,S = s \mid I = i\,) \;=\; \frac{p(\,I = i \mid S = s\,)\quad p(\,S = s\,)}{p(\,I = i\,)}$$

"posterior"

Given an image, $I = i$, choose the scene $S = s$ that maximizes the posterior $p(\,S = s \mid I = i\,)$.

We don't care about $p(\,I = i\,)$. Why not ?

“likelihood” “prior”

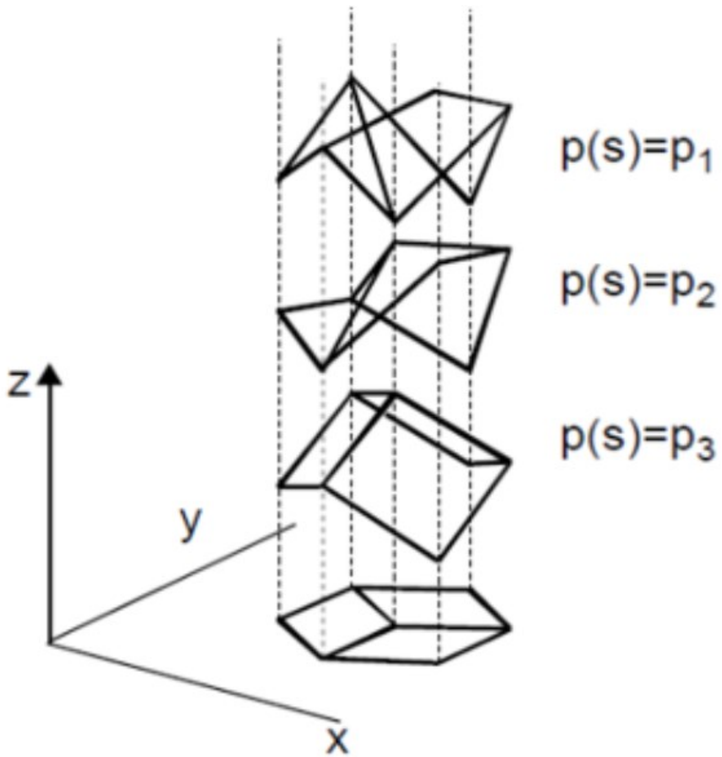
$$p(S = s \mid I = i) = \frac{p(I = i \mid S = s)\ \cancel{p(S = s)}}{p(I = i)}$$

“posterior”

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

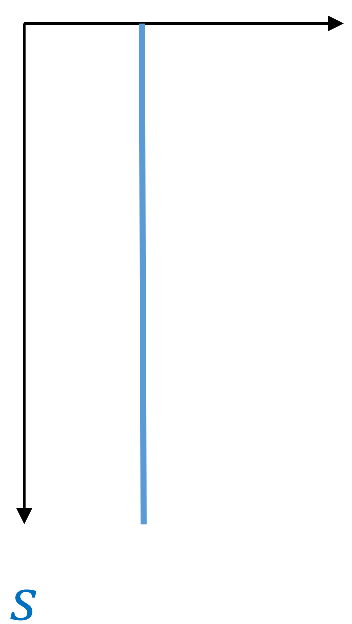
Interesting cases arise when the prior is non-uniform.

$$p(\,S = s \mid I = i\,) \;=\; \frac{p(\,I = i \mid S = s\,)\quad p(\,S = s\,)}{p(\,I = i\,)}$$



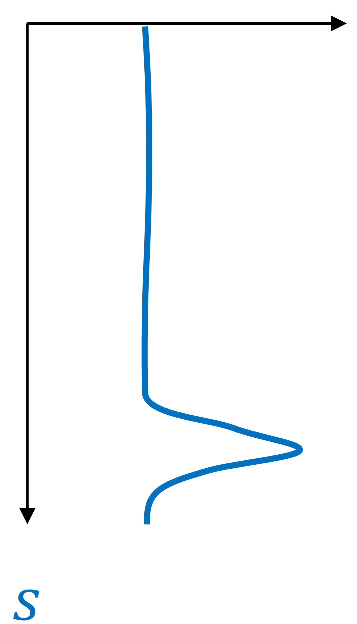
likelihood

$p(\,I_1 = i_1 \mid\; S = s\,)$

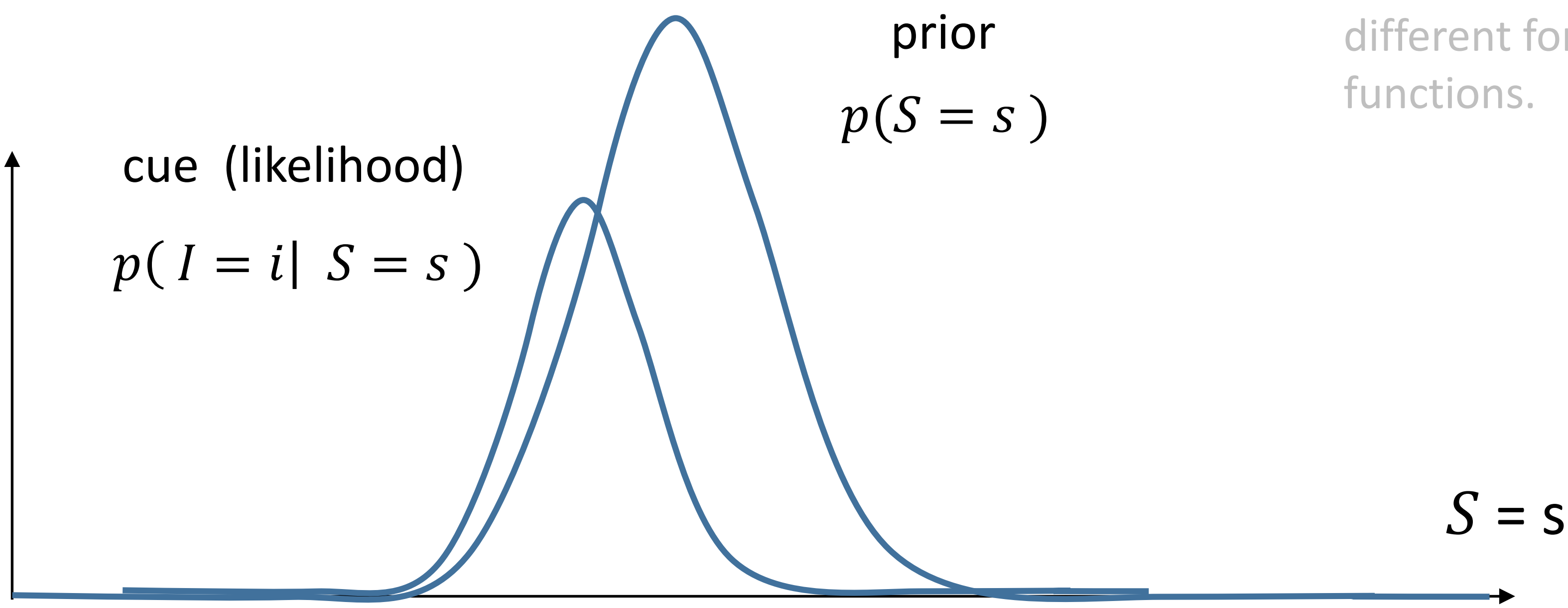


prior

$p(S = s)$

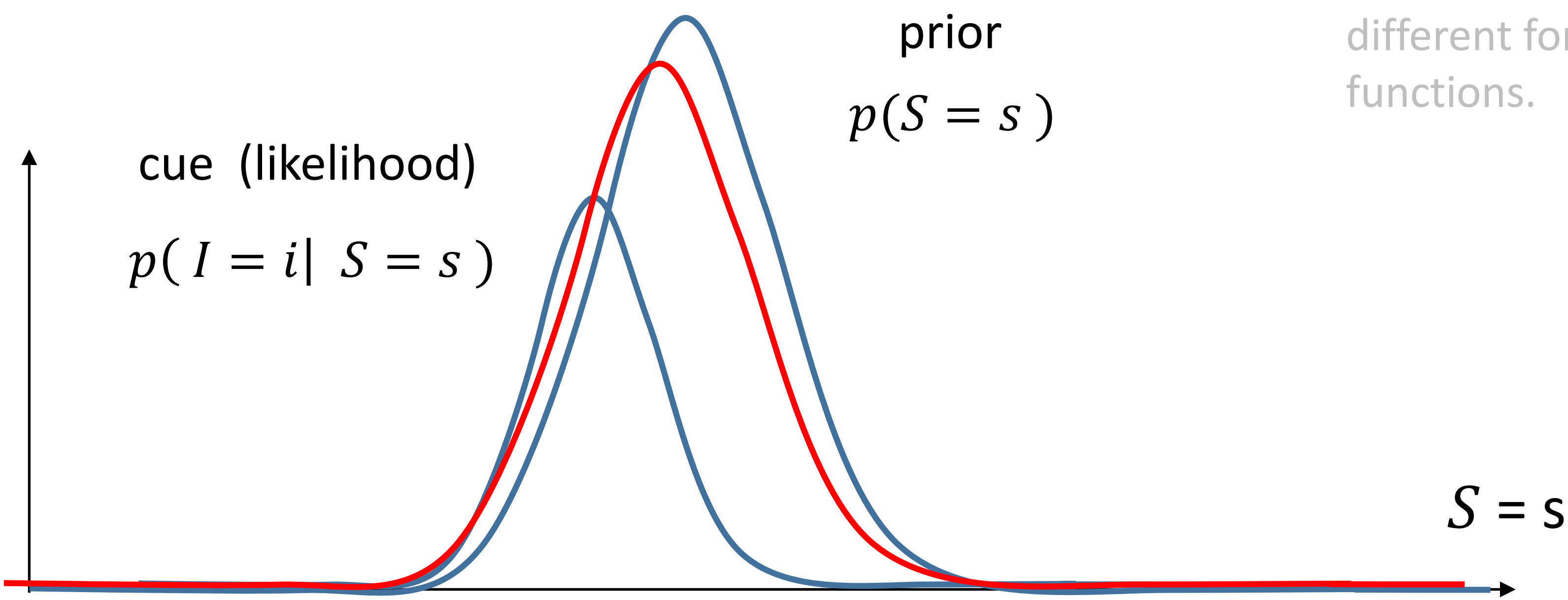


Note the 'y axis' scales are very different for the two functions.



Combining priors and cues using Bayes Rule:

$$p(\ S = s\ |\ I = i\ )\ =\ \frac{p(\ I = i\ |\ S = s\ )\quad p(\ S = s\ )}{p(\ I = i\ )}$$



Note the 'y axis' scales are very different for the two functions.

Combining priors and cues using Bayes Rule:

(The same linear combination theory works here.)

$$p(S = s \mid I = i) = \frac{p(I = i \mid S = s)\ \ p(S = s)}{p(I = i)}$$

In real word situations, our perceptual systems combine many priors and cues:

$$p(S \mid \; I_1, I_2, I_3, \ldots)$$

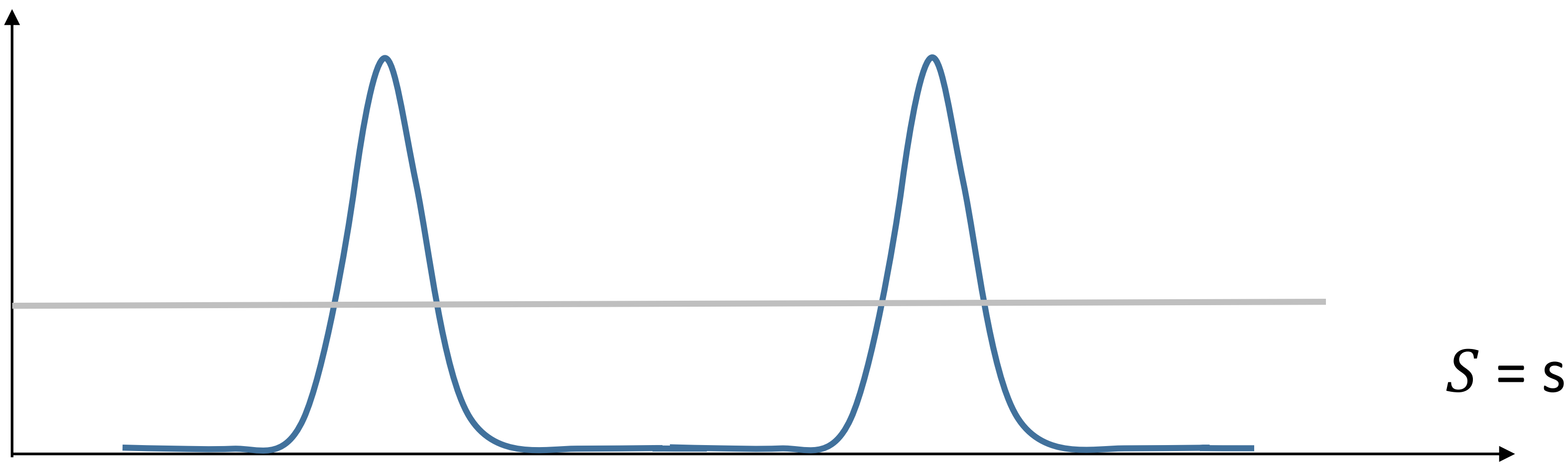
$$= \quad p\,(\,I_1, I_2, I_3, \ldots \mid S) \quad p_a(\,S\,)\;\; p_b(\,S) \ldots$$

“natural scene probabilities”

cues such as
texture, motion,
shading, ….

An interesting case:
a flat likelihood, and a prior with two maxima

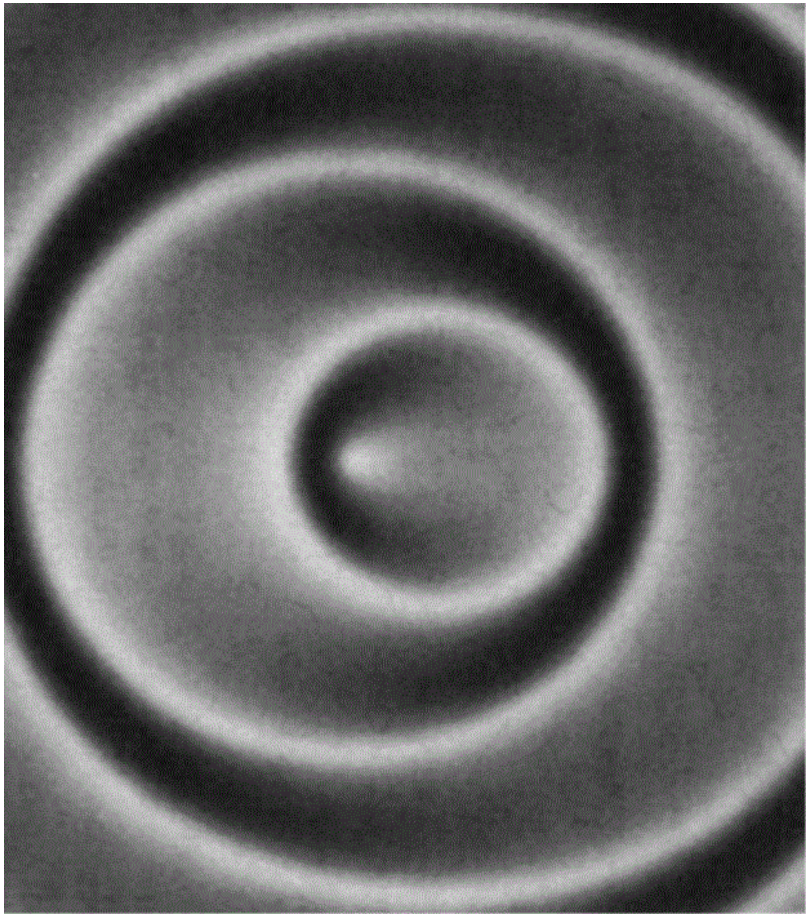
$$p(I = i \mid S = s) \; p(S = s)$$



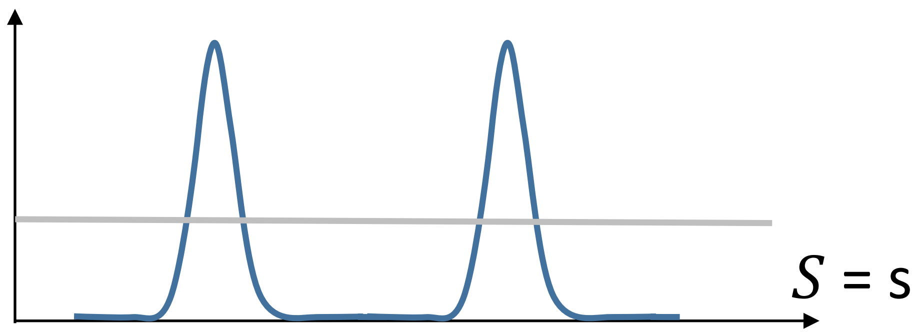
This situation can arise from certain symmetries.

What “priors” does the visual system use to resolve such twofold ambiguities ?

Let’s look at an example.



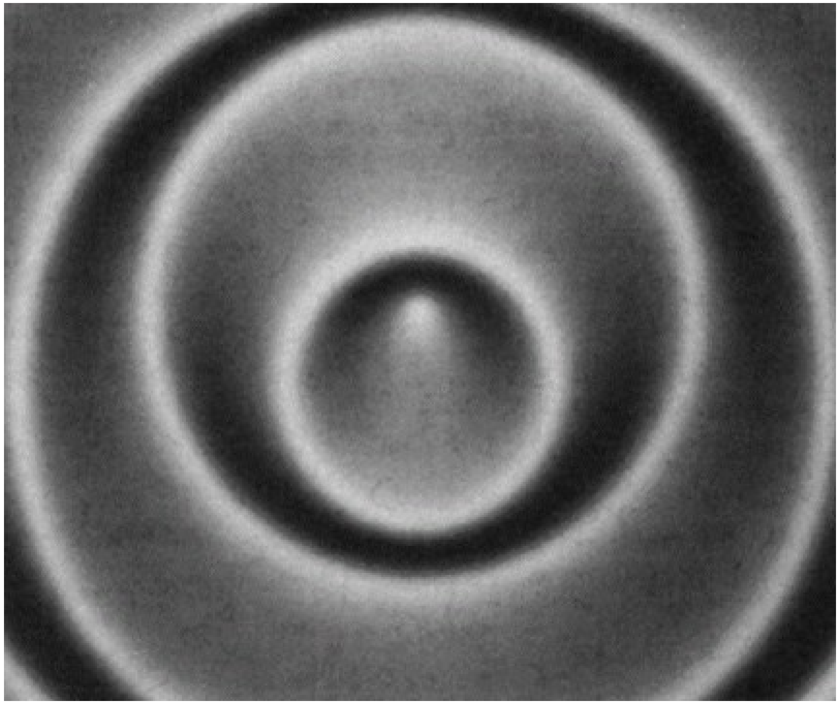
$p(\, I = i \mid S = s \,)\ p(S = s)$



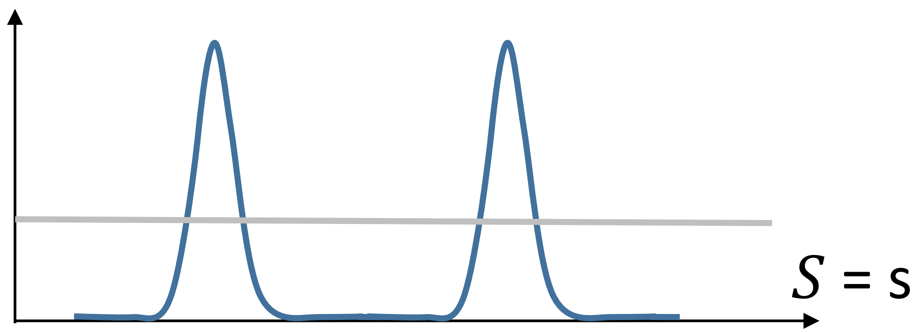
You can perceive the center point as a hill or a valley.

When you see it as a hill, you perceive the overall surface slant to be leftward.   But when you see it as a valley, the slant is rightward.

Rotate the image by 90 degrees.

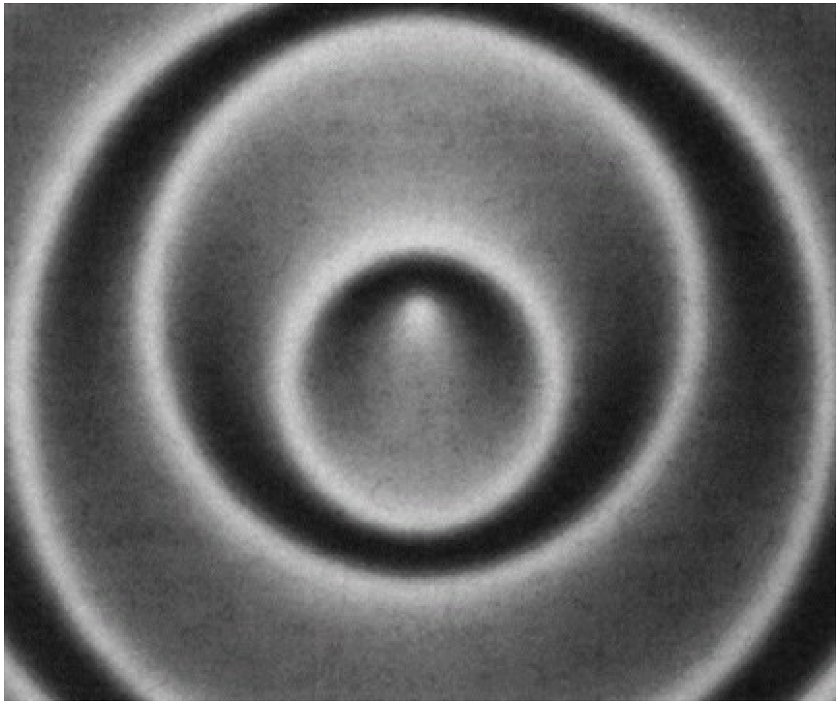


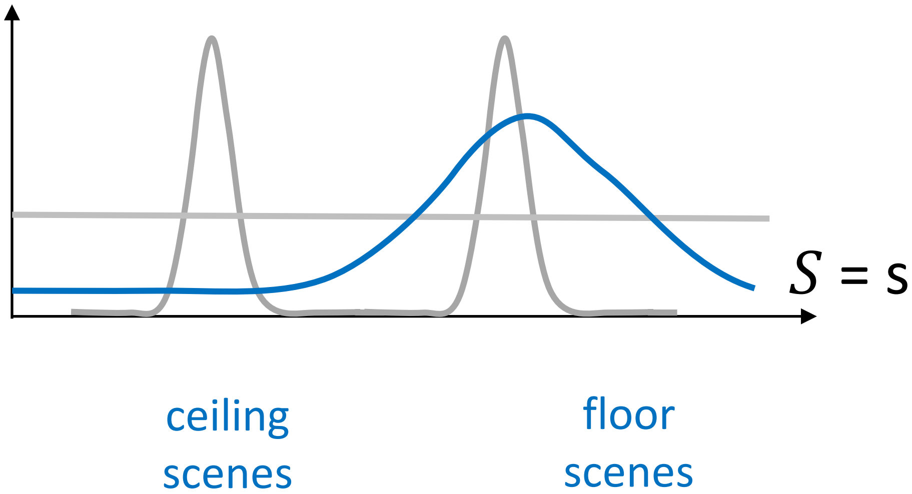
$$p(\, I = i \mid S = s \,) \; p(S = s)$$



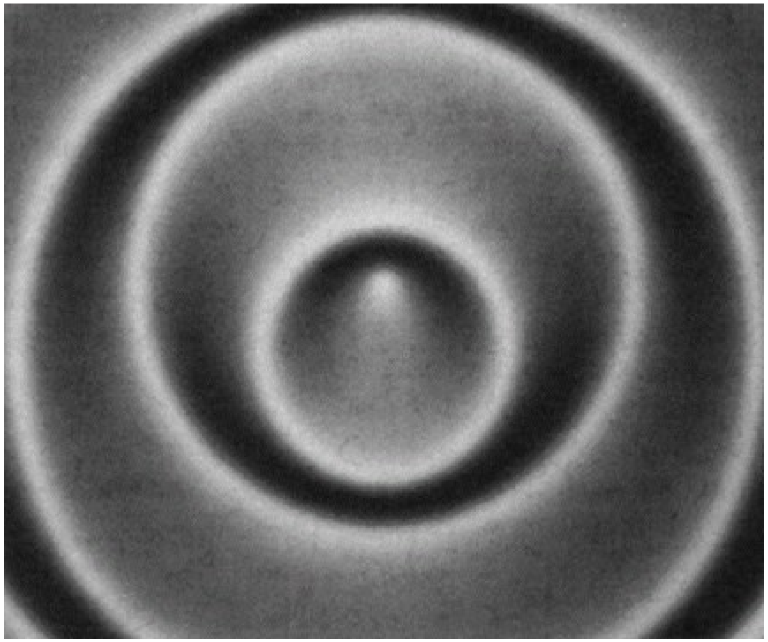
You *can* perceive the center either as a hill or valley.
However, we tend to perceive the center as a hill.
Why ?

$$p(\,I = i \mid S = s\,)\,p(\,S = s\,)\,p_{slant}(\,S = s\,)$$

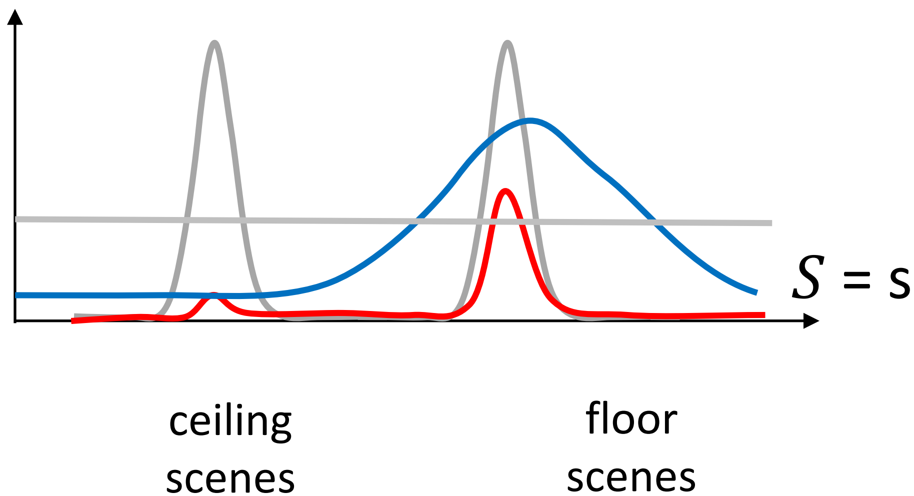




When you see it as a hill, you perceive the overall slant upward (like a floor).
When you see it as a valley, the overall slant is downward (like a ceiling).
We have a prior for floors over ceilings. (Reichel and Todd 1990)



$$p(S = s \mid I = i) = p(I = i \mid S = s)\, p(S = s)\, p_{slope}(S = s)$$



When you see it as a hill, you perceive the overall slant upward (like a floor).
When you see it as a valley, the overall slant is downward (like a ceiling).
We have a prior for floors over ceilings. (Reichel and Todd 1990)
A prior with a single maximum would then produce a higher maximum in the posterior.

Midterm exam is Tuesday after Study Break.

Multiple choice and short answer questions only.

Review the Lecture Notes and Exercises.