CS 696i, Spring 2005, part III

Computer Vision

(Making Machines See)

and other miscellaneous stuff

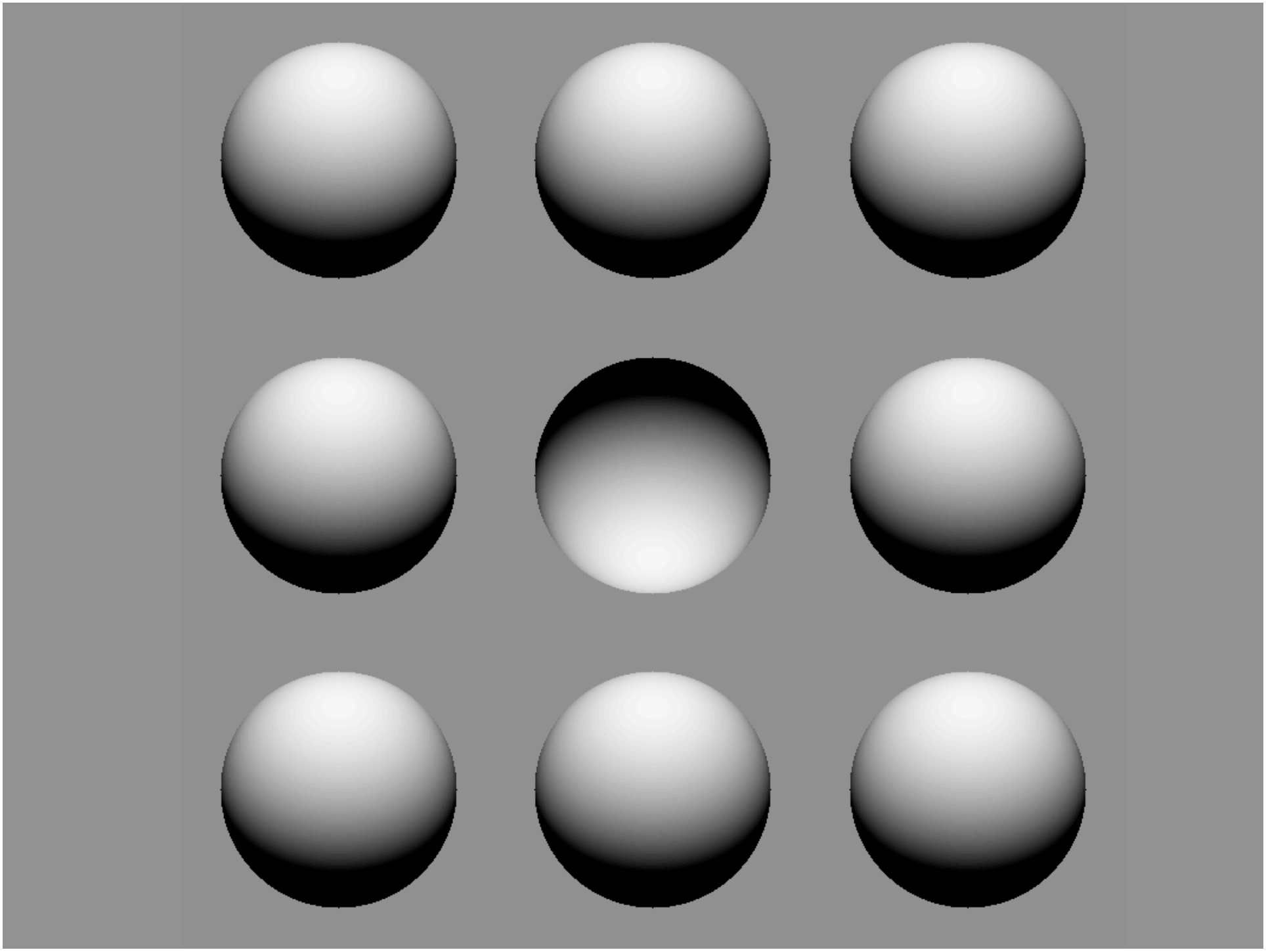
Kobus Barnard

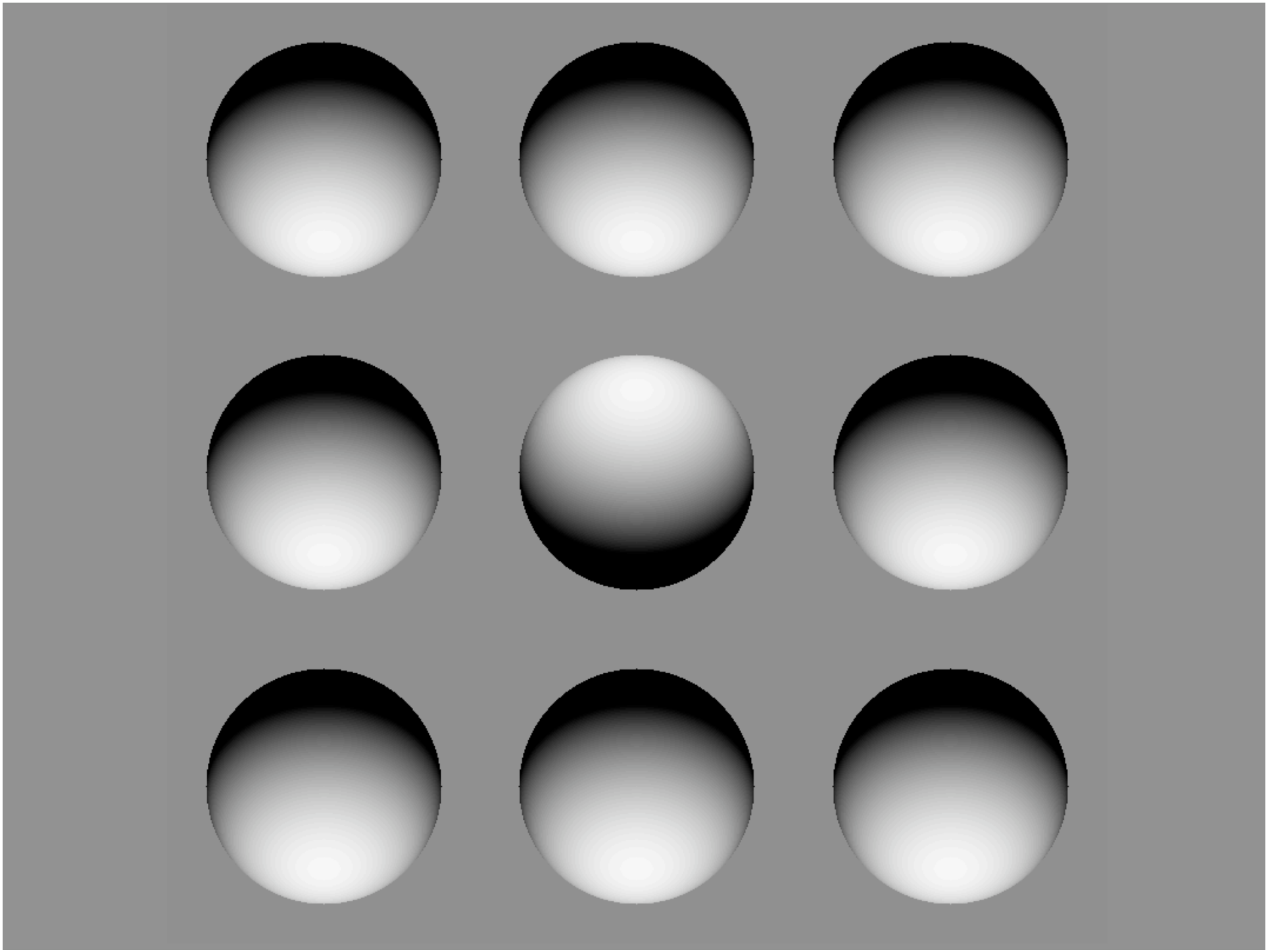
Computer Science, University of Arizona

Upside down / inverted recognition

(Fun upside down images excluded from web version)

Priors in vision?





Ambiguity

Vision is inherently ambiguous

The preceeding images had multiple possible interpretations

To resolve the ambiguity we choose an explanation consistent with world knowledge---provides a good guess.

Ambiguity

Discussion point:

An artificial system need not commit to a solution until one is required. However, we always "see" something if we look.

> Is it a computational advantage to provide an answer that may be wrong instead of probalities over a reasonable set of answers?
>
> Does this mean probablistic inference is not a good model of the brain?

Ambiguity

In the preceeding case the world knowlegde seems explanable by a “prior” on illumination direction

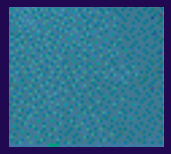
Alternatively, belief about the world belief is part of the model

(The distinction between model choice and prior is often blurred, but usually clear in context).

More examples of ambiguity

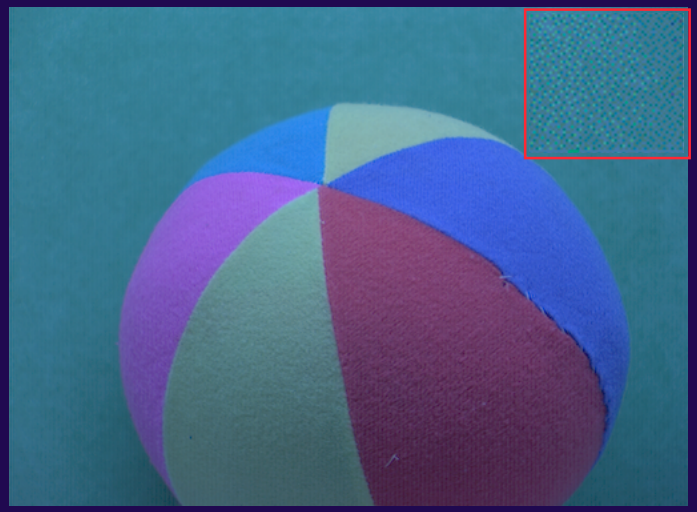
Top part is further away, assuming that in the world, the crests are roughly equidistant (not the case in the image!).

More examples of ambiguity



Surface color as a function of observed color is ambiguous

More examples of ambiguity



(Same scene, but different illuminant)

Computer vision in 100 easy minutes

Be on the lookout for ambiguityand dealing with it.

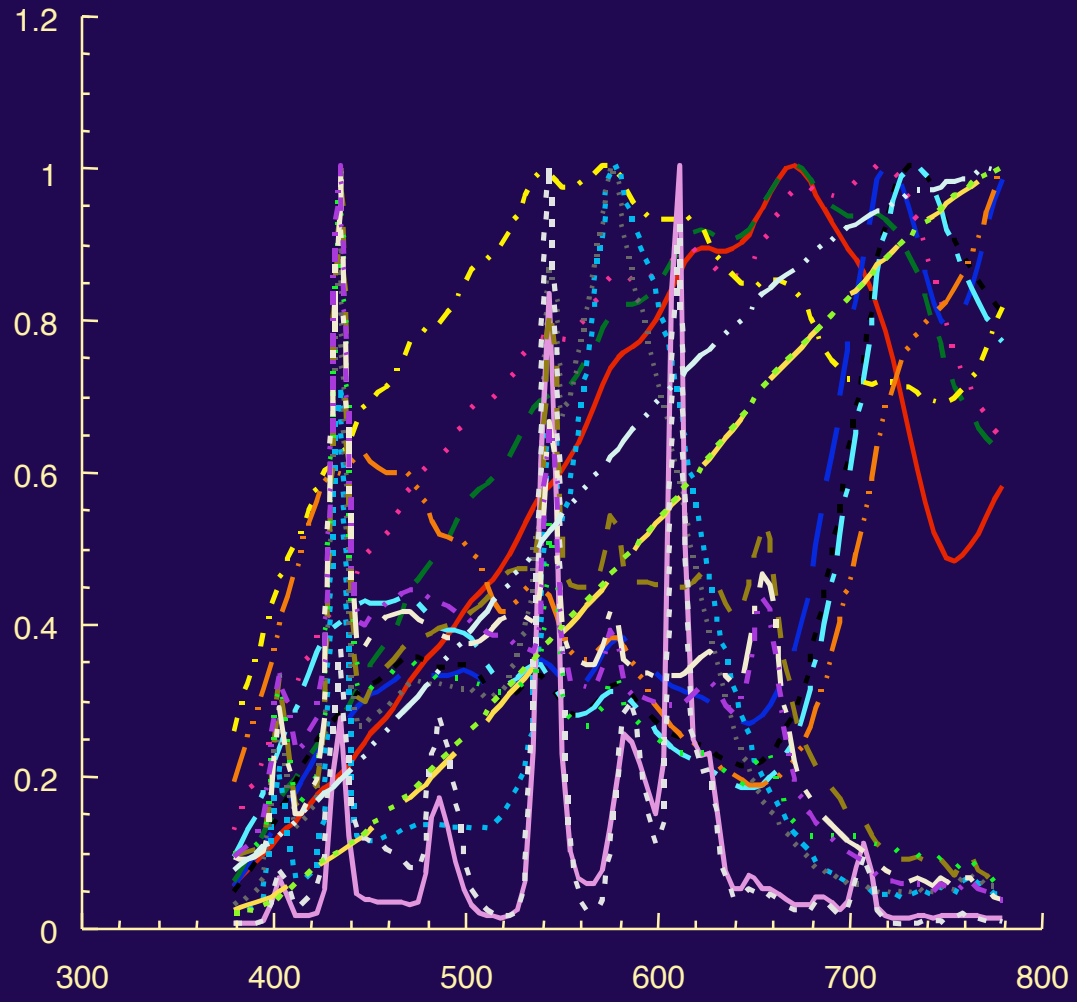
Bottom up processing---assemble an understanding of images from the pixels upwards: Reason about the pixels in terms of image formation, find edges, regions, estimate distance, group things based on low level inference.

Top down---you have a model of what you are looking for. Can constrain the bottom up activities.

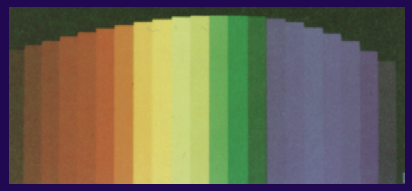
First there was light

Light energy reaching a camera sensor has a distribution over wavelength, □.
(*Recall from physics that wavelength is inversely related to photon energy)

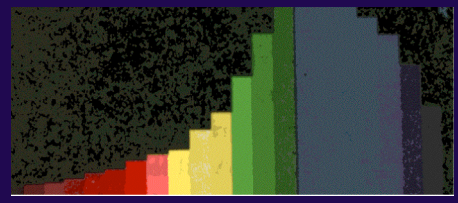
Light takes many forms

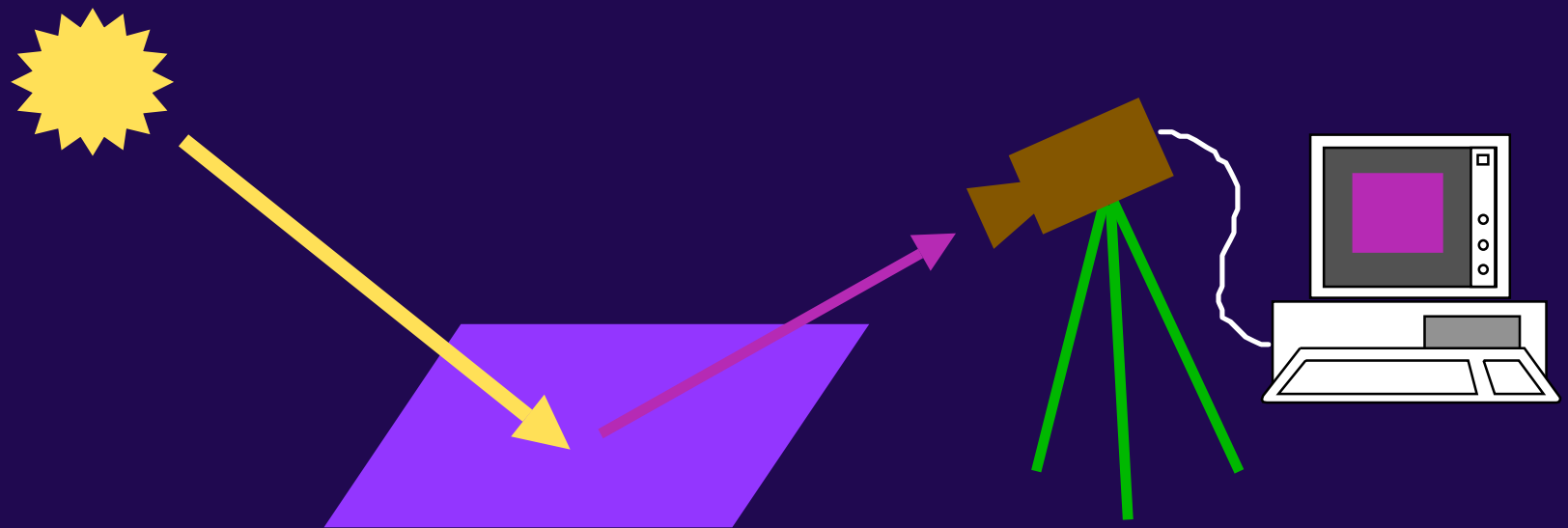


Then there was substance, and it modified the light



x



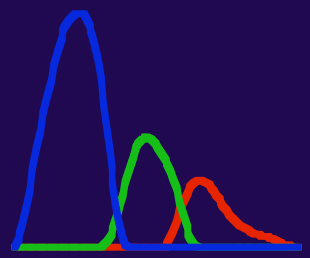


Light, camera, ...

Image Formation (Spectral)

$$(R,G,B) = \int_{380}^{780} \; * \; d\lambda$$

Light reaching camera



Camera sensitivities

Image Formation (Geometric)

Abstract camera model--box with a small hole in it (pinhole camera)

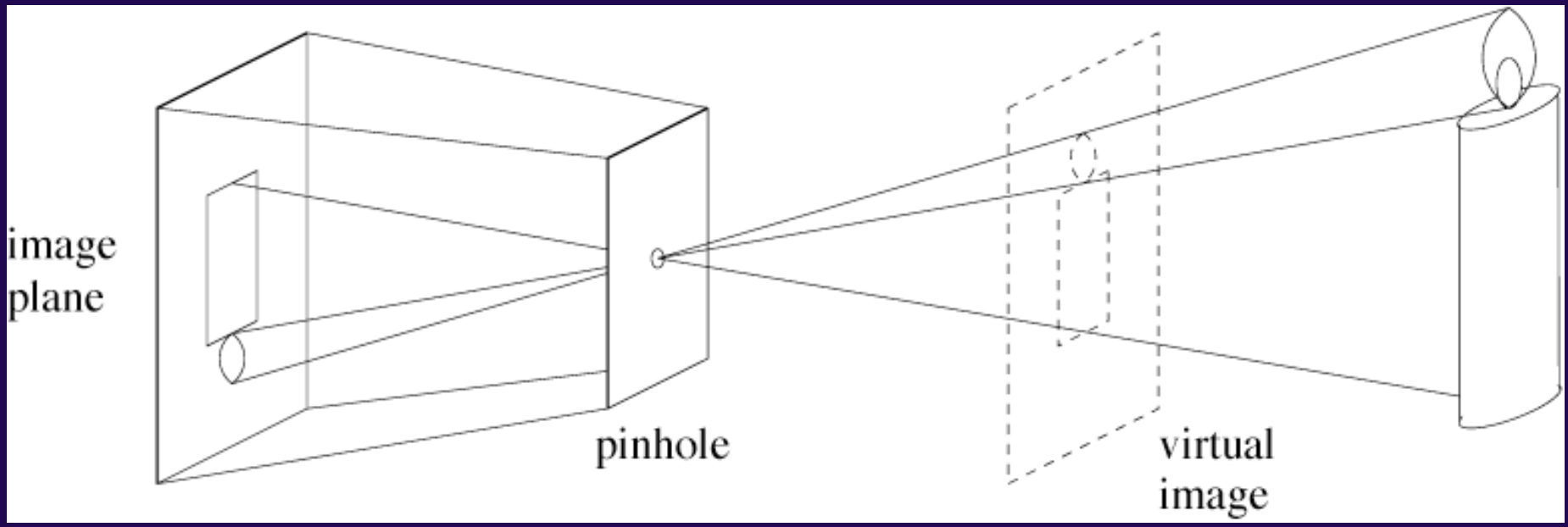
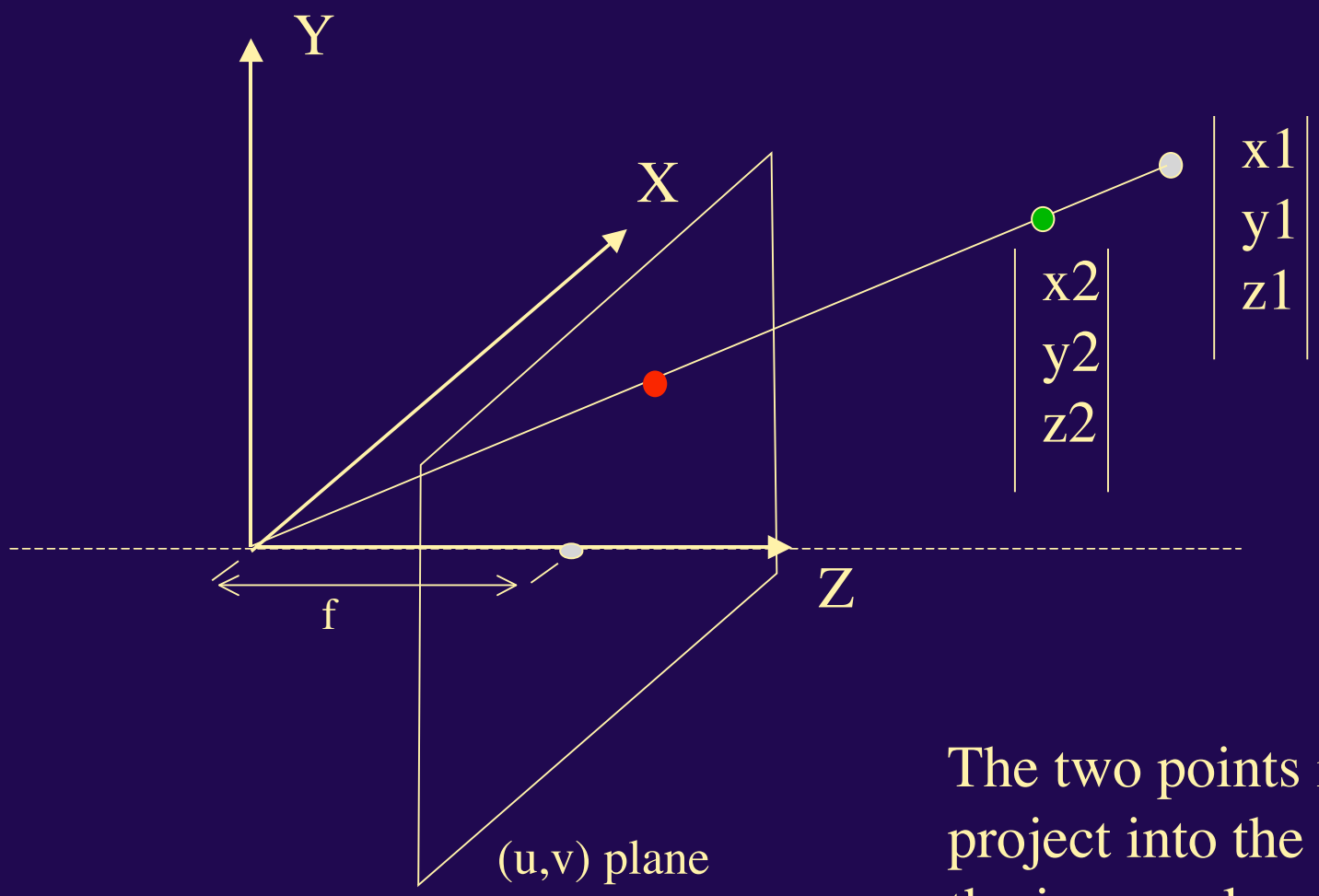


Image Formation (Geometric)



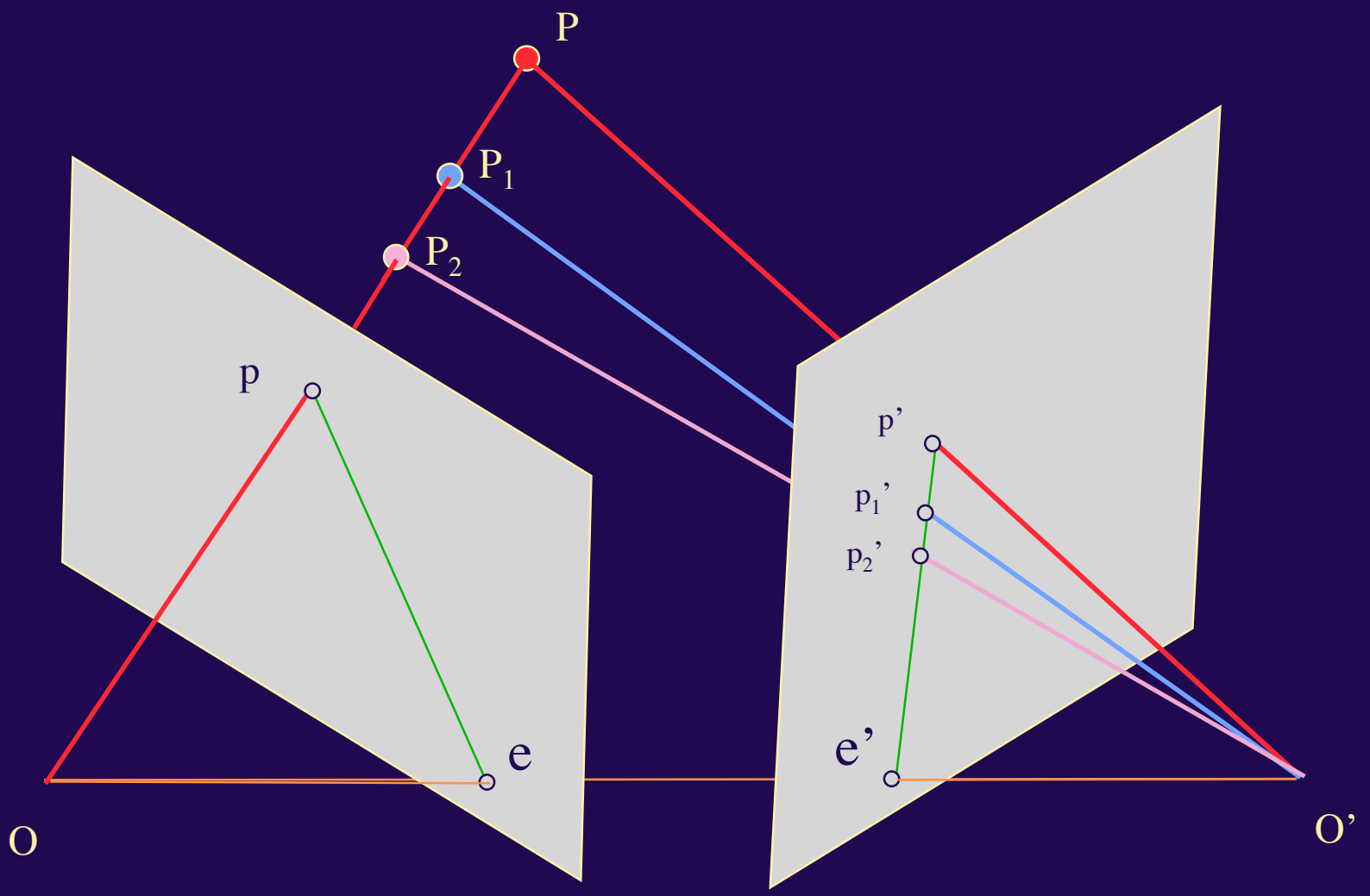
The two points in 3D space project into the same point on the image place.

Multiple View Geometry

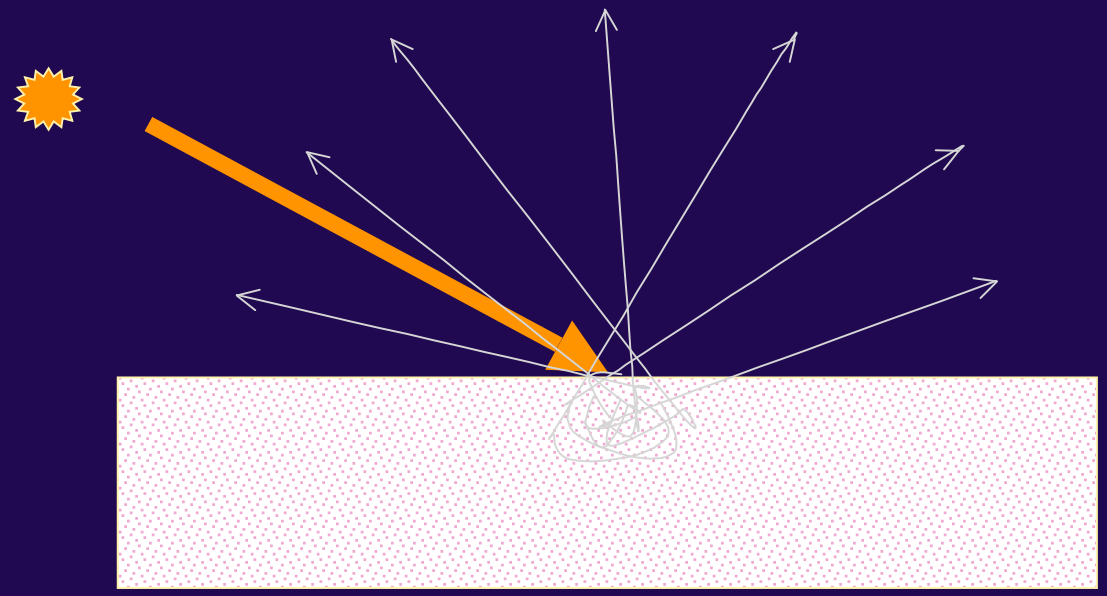
Includes “stereo”, “structure from motion”

If we have more than one camera, then we can deal with the depth ambiguity.

Multiple View Geometry



Shading (physics)



Idealized, perfect diffuse reflector (Lambertian surface)

Shading (physics)

In the Lambertian model, the surface brightness varies as the cosine of the angle.

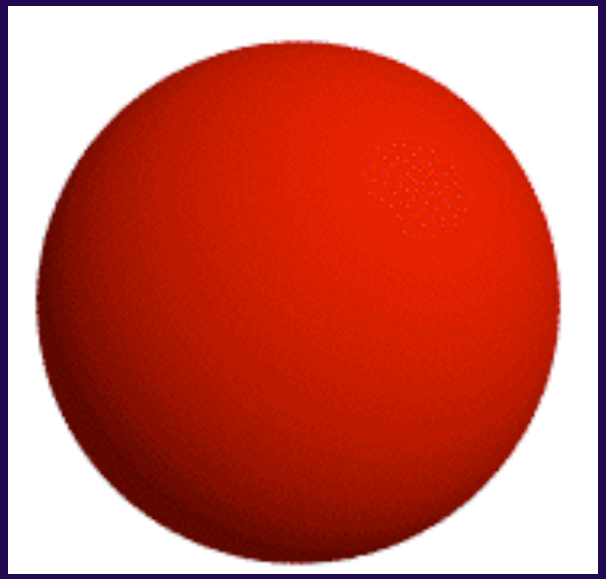
Thus we just take the dot product of the surface normal with the light direction.

IE, attenuate brightness by

$$\mathbf{n} \cdot \mathbf{s}$$

Surface normal

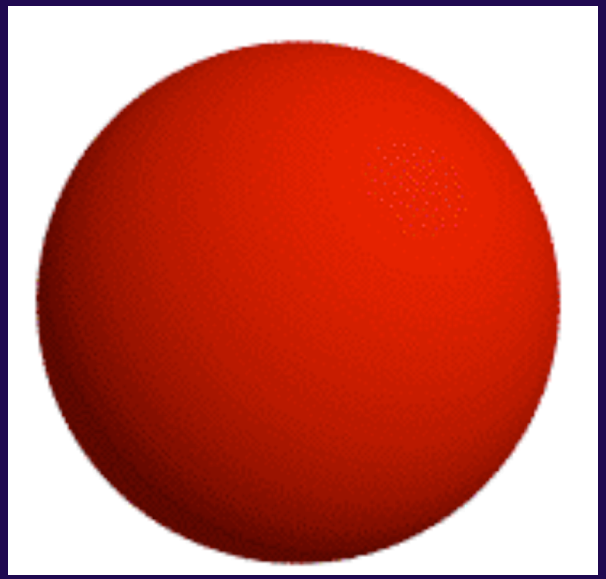
Light source direction



Shape from shading

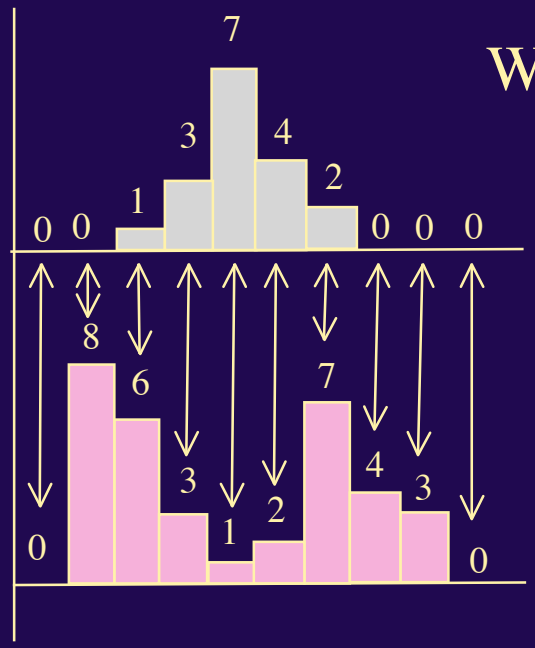
If the surface normal affects brightness in the image, then we might be able to use the brightness to say something about the normals.

Normals with smoothing and integratability constraints can give the surface under good conditions.



Features

Typically detect features such as edges by “linear filtering”,

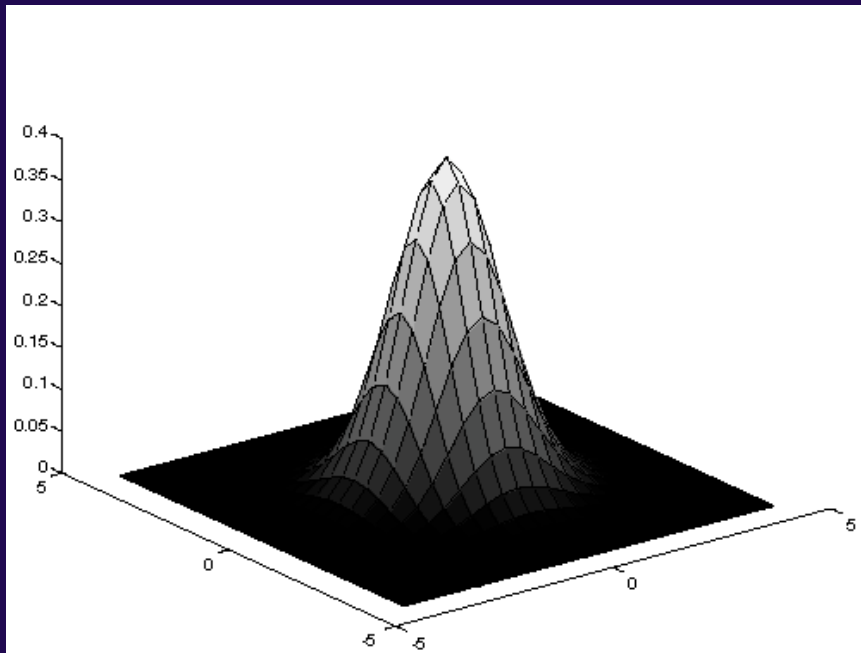


Smoothing Filters

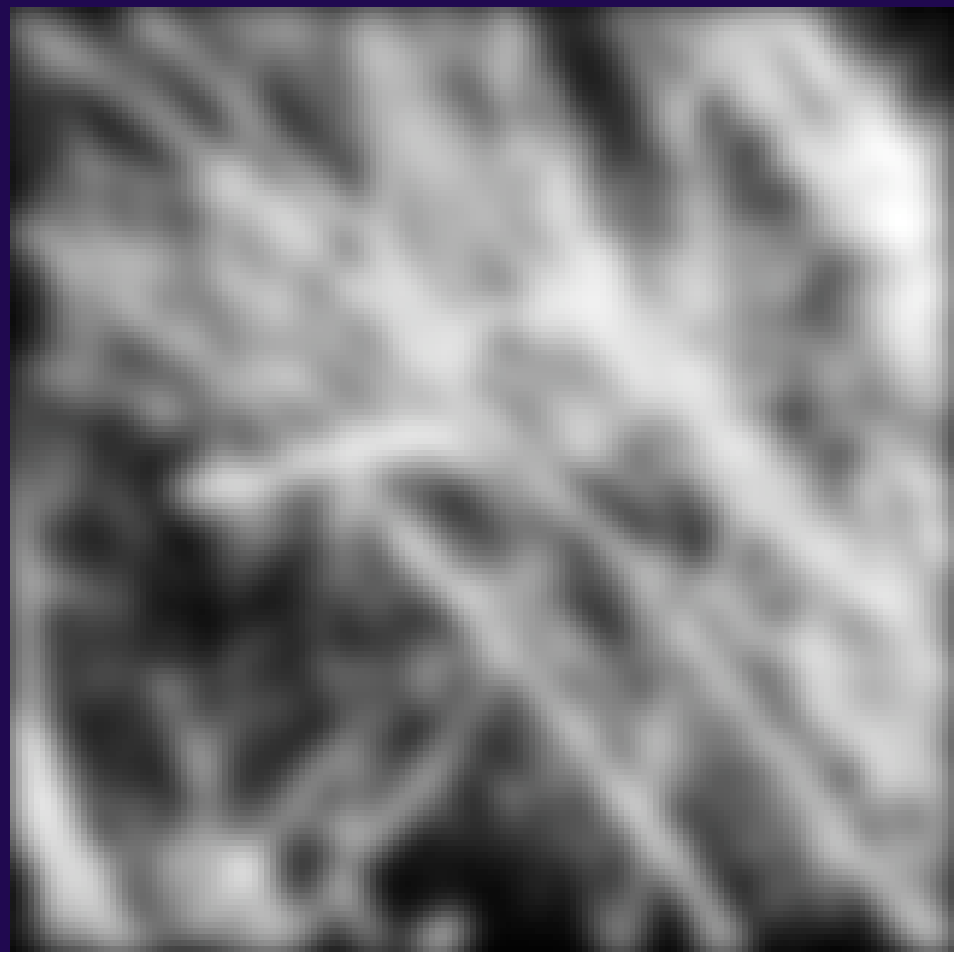
Reduce noise

Analyze structure at different scales

A Gaussian gives a good model of a fuzzy blob

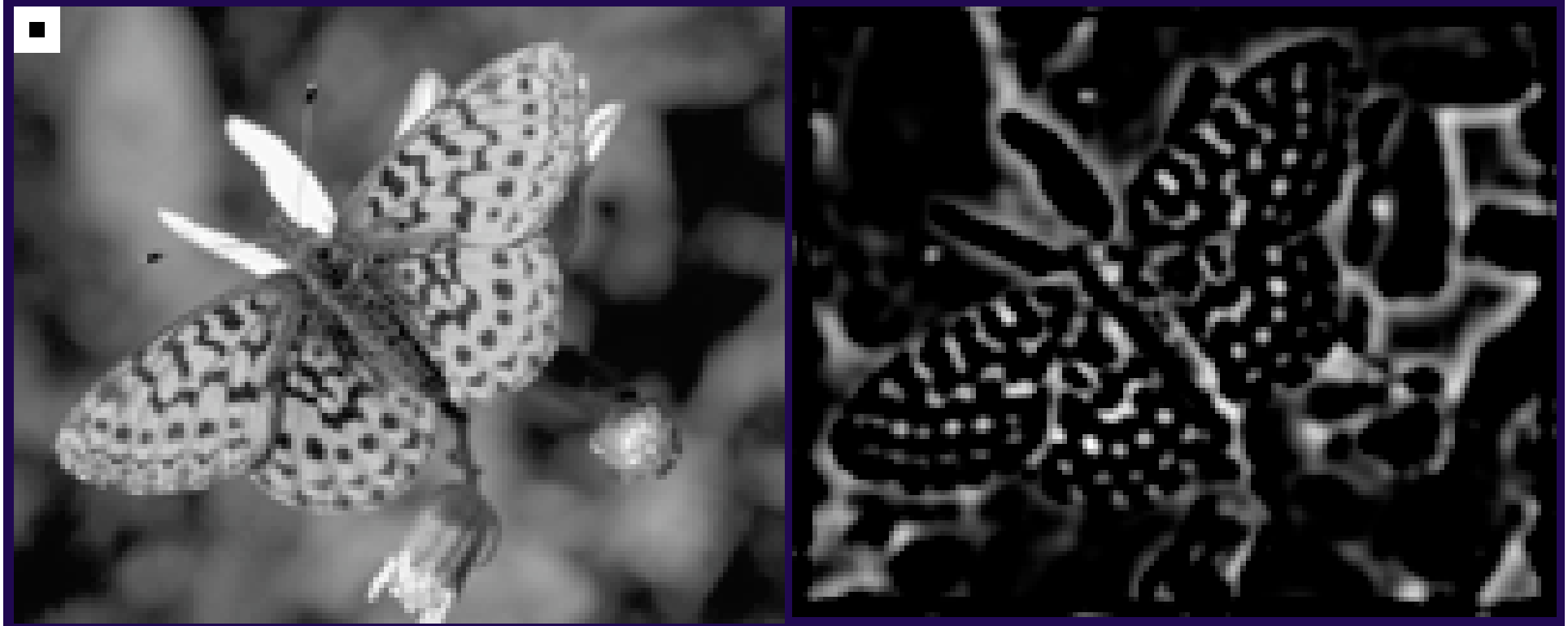


Smoothing with a Gaussian



Filters as templates

Filters respond to structures that “look” like the filter



Filters as templates

Filters respond to structures that “look” like the filter