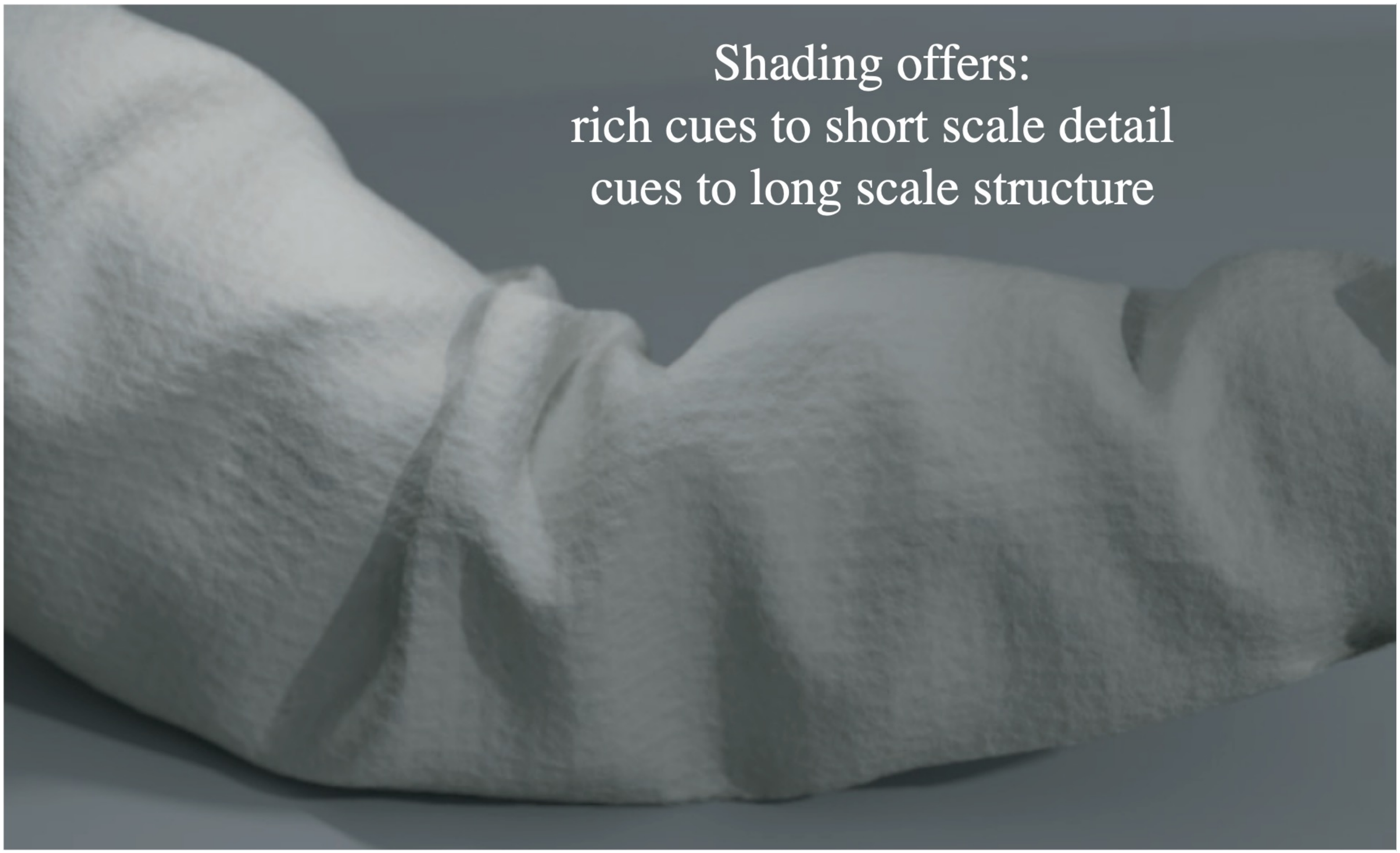
Illumination: Inference

D.A. Forsyth,

University of Illinois at Urbana Champaign

Shading is an amazingly detailed cue



Inferences

- Lightness
  - whether surface is dark or light, independent of shading
- Shape
  - photometric stereo
- Camera response function
  - from images

Lightness

Early algorithms for estimating lightness are mostly variants of an idea referred to as *Retinex*. They remain useful and quite competitive. These algorithms assume that the scene is flat and frontal; that surfaces are diffuse, or that specularities have been removed; and that the camera responds linearly. In this case, the camera response C at a point $\mathbf{x}$ is the product of an illumination term, an albedo term, and a constant that comes from the camera gain:

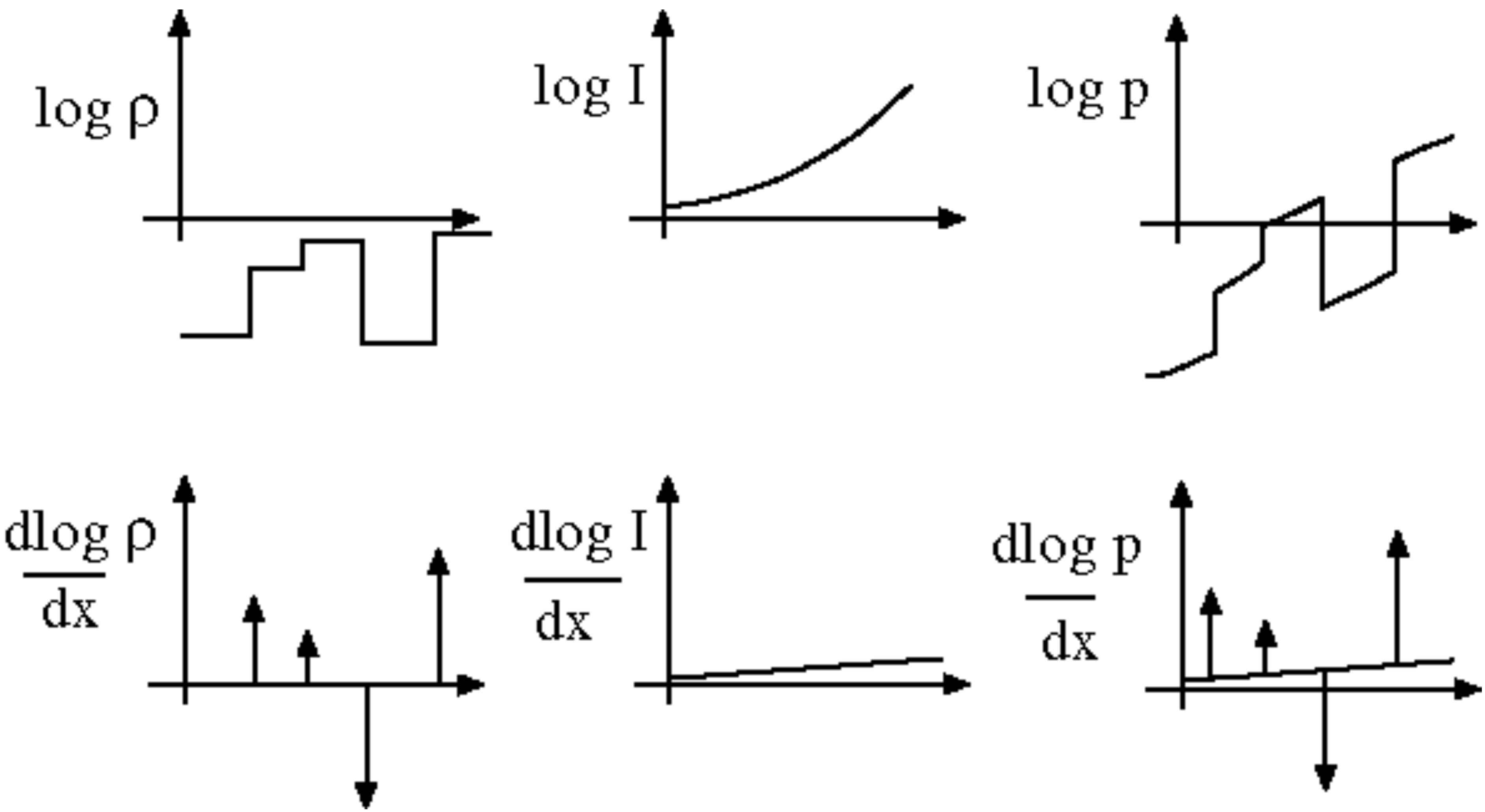
$$C(\mathbf{x}) = k_c I(\mathbf{x}) \rho(\mathbf{x}).$$

If we take logarithms, we get

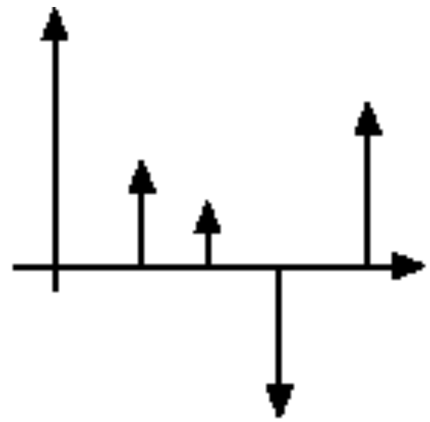
$$\log C(\mathbf{x}) = \log k_c + \log I(\mathbf{x}) + \log \rho(\mathbf{x}).$$

We now assume that:

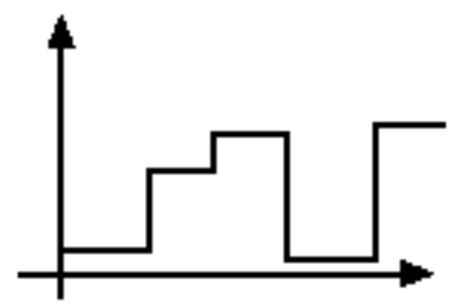
- albedoes are piecewise constant over space;
- and shading changes only slowly over space.


log ρ
log I
log p
dlog ρ / dx
dlog I / dx
dlog p / dx

Thresholded $\frac{d\log p}{dx}$



Integrate
This to get



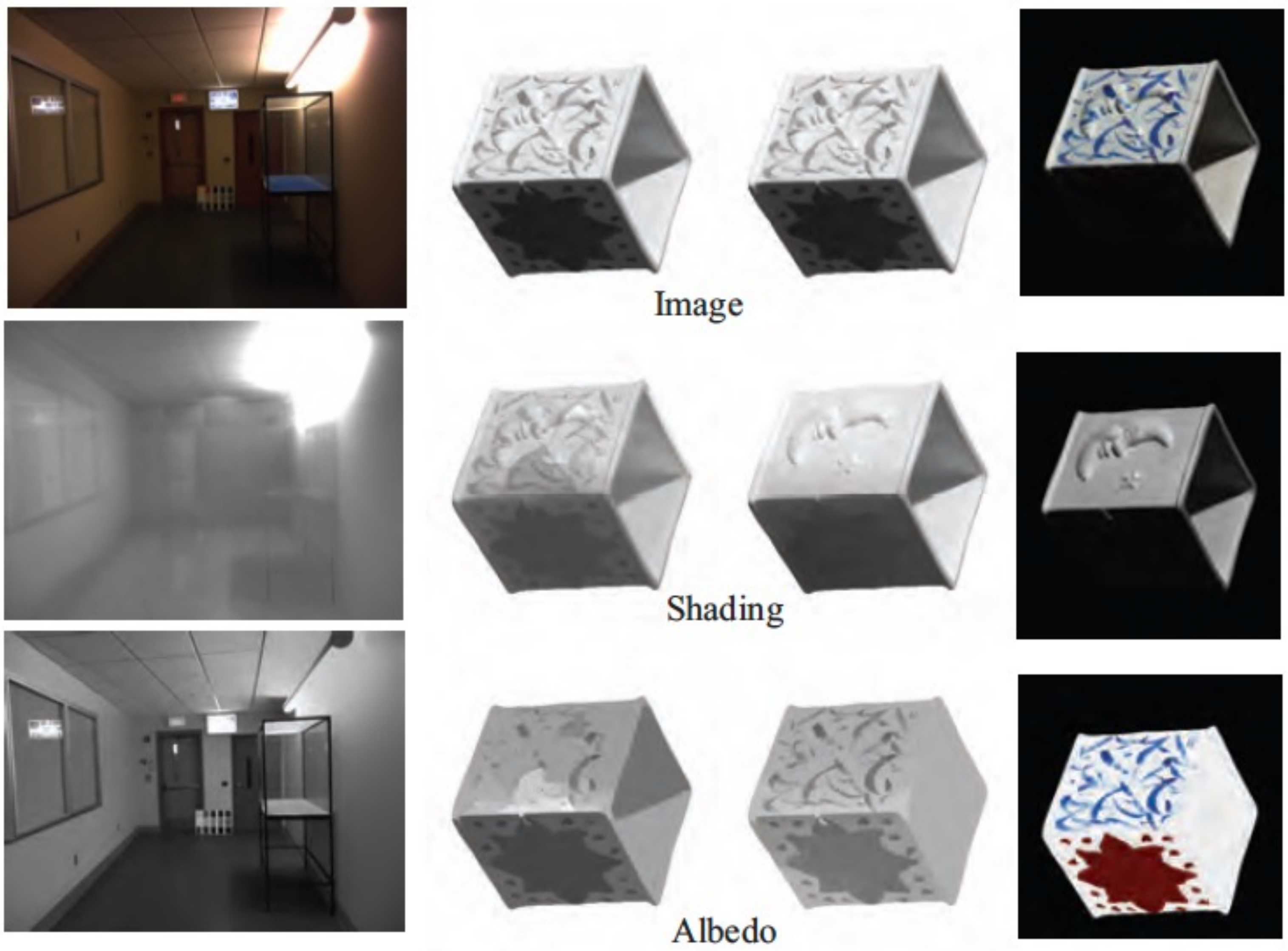


Image
Shading
Albedo

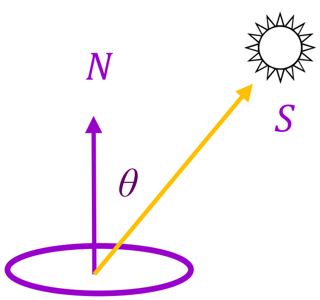
Photometric stereo, or shape from shading

- Can we reconstruct the shape of an object based on shading cues?

Luca della Robbia,
Cantoria, 1438

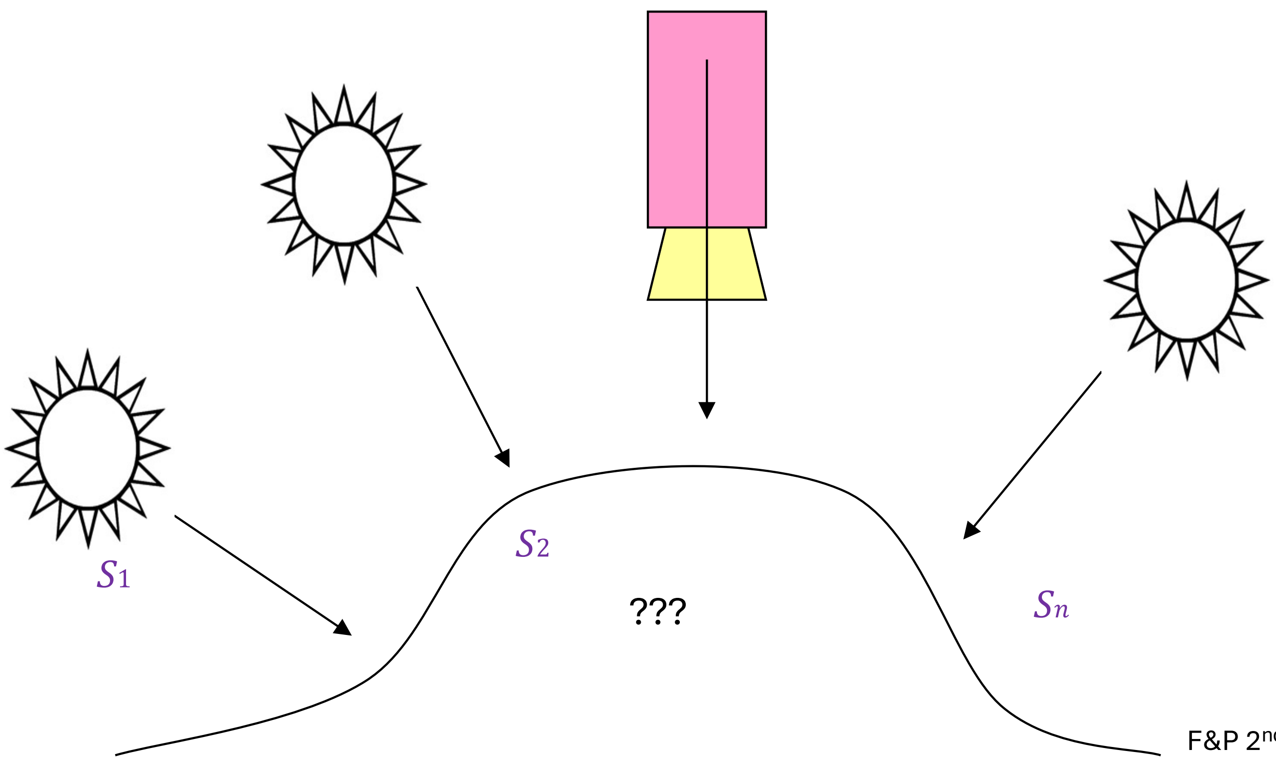
Photometric stereo

- Reconstruct object shape based on shading cues
- Assuming a Lambertian object
  - given the image intensity (I) and light source direction (S)
  - recover the surface normal (N) from a single image



$$I = \rho\,(S \cdot N)$$
$$= \rho \|S\| \cos\theta$$

Photometric stereo

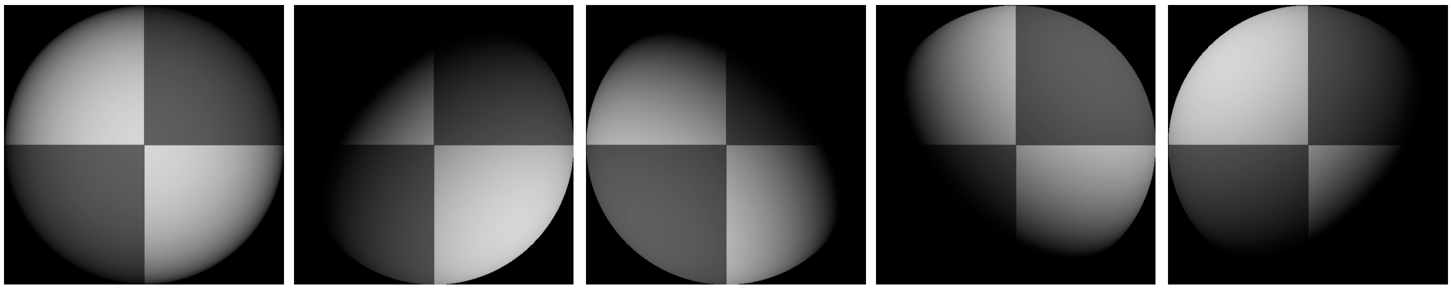


F&P 2nd ed., sec. 2.2.4

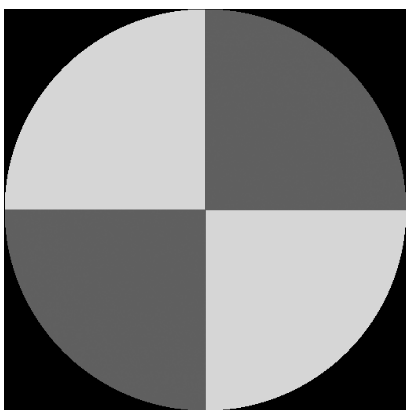
Photometric stereo

- Assume:
  - A Lambertian object
  - A *local shading model* (each point on a surface receives light only from sources visible at that point)
  - A set of *known* light source directions
  - A set of pictures of an object, obtained in exactly the same camera/object configuration but using different sources
  - Orthographic projection
- Goal: reconstruct object shape and albedo

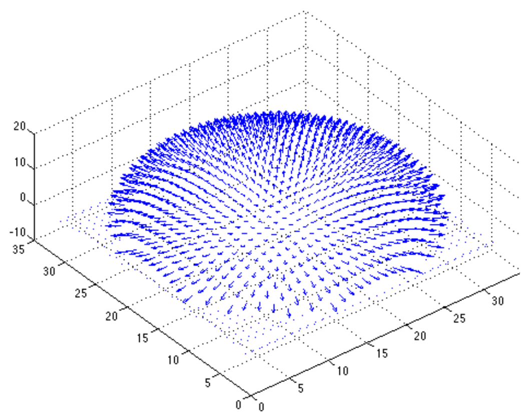
Example 1



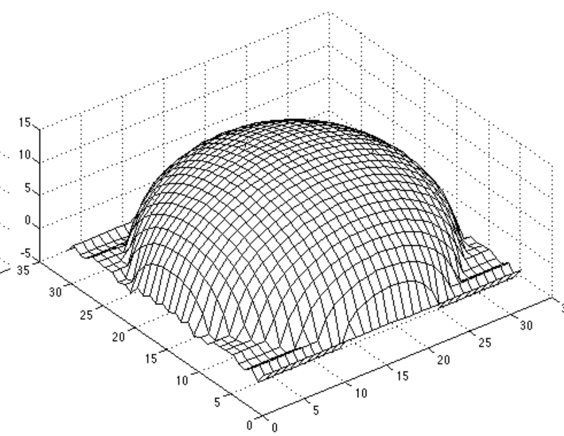
Recovered albedo



Recovered normal field



Recovered surface model



F&P 2nd ed., sec. 2.2.4

Example 2

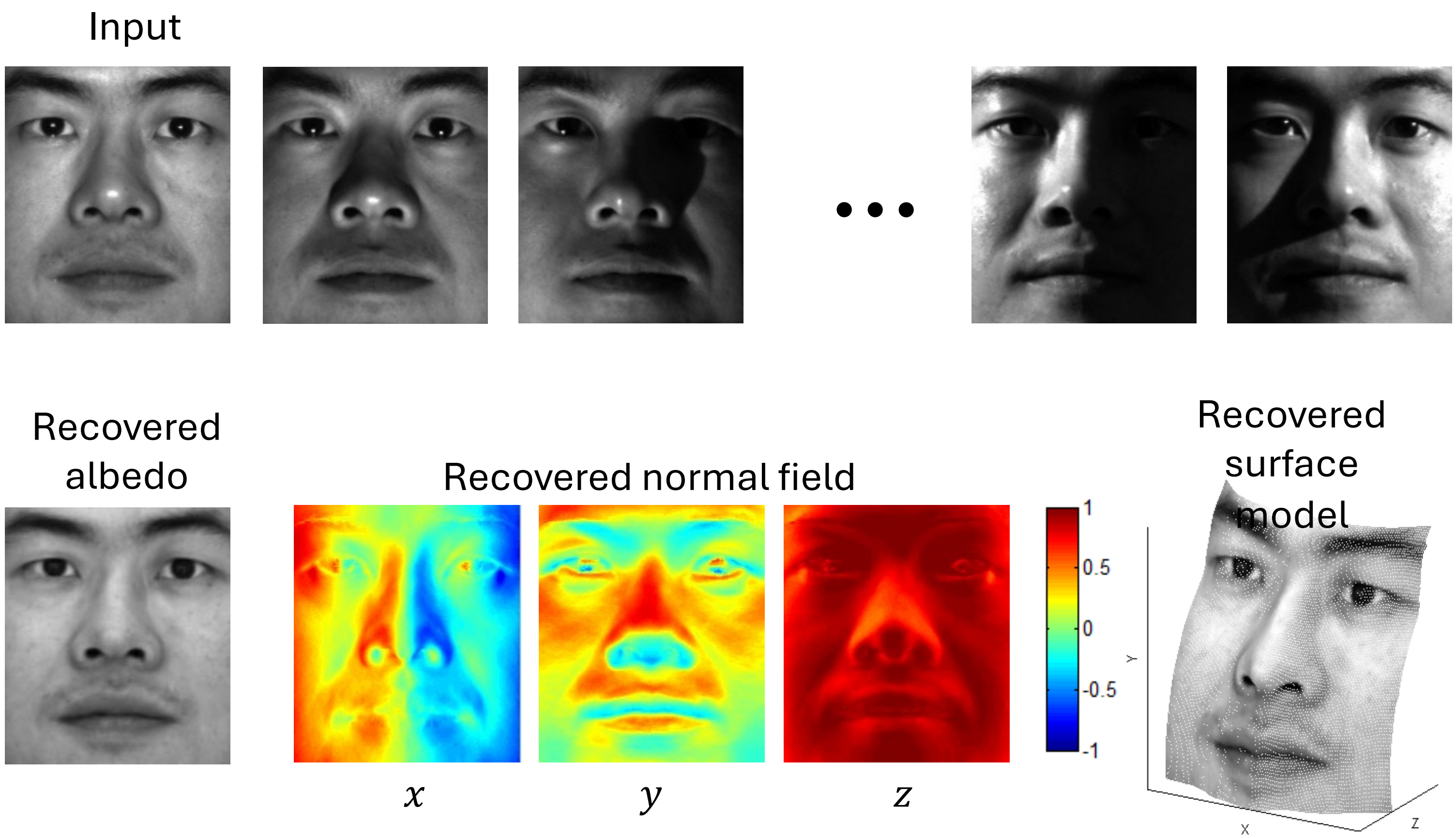
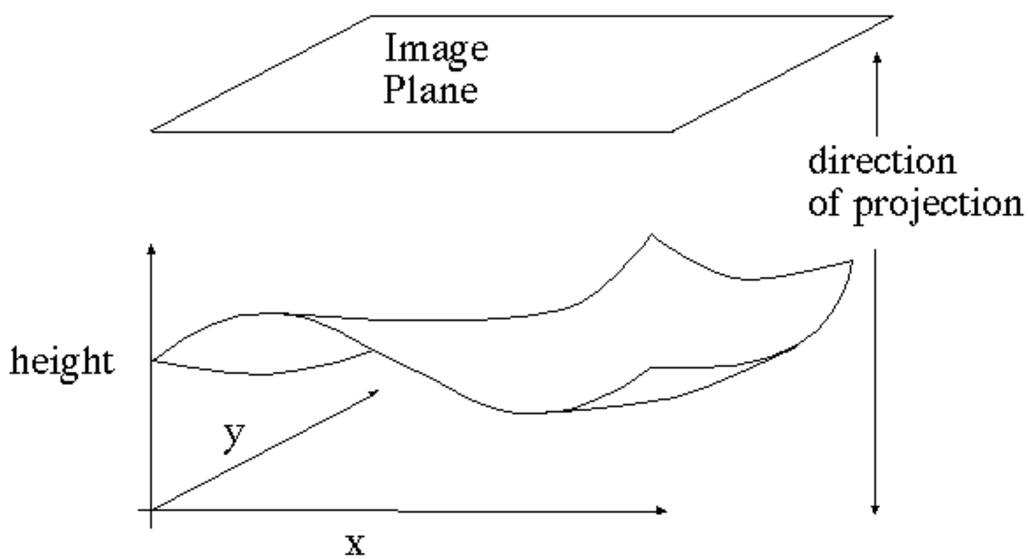

Input
Recovered albedo
Recovered normal field
x
y
z
1
0.5
0
-0.5
-1
Recovered surface model
X
Y
Z

Image model

- **Known:** source vectors S_j and pixel values $I_j(x, y)$
- **Unknown:** surface normal $N(x, y)$ and albedo $\rho(x, y)$



F&P 2nd ed., sec. 2.2.4

Image model

- **Known:** source vectors S_j and pixel values $I_j(x, y)$
- **Unknown:** surface normal $N(x, y)$ and albedo $\rho(x, y)$
- Assume that the response function of the camera is a linear scaling by a factor of k
- Lambert's law:
- $I_j(x, y) = \; k\,\rho(x, y)\big(N(x, y) \cdot S_j\big)$

$$= (\rho(x, y) N(x, y)) \cdot (k\, S_j)$$

$$= g(x, y) \cdot V_j$$

F&P 2nd ed., sec. 2.2.4

Least squares problem

- For each pixel, set up a linear system:

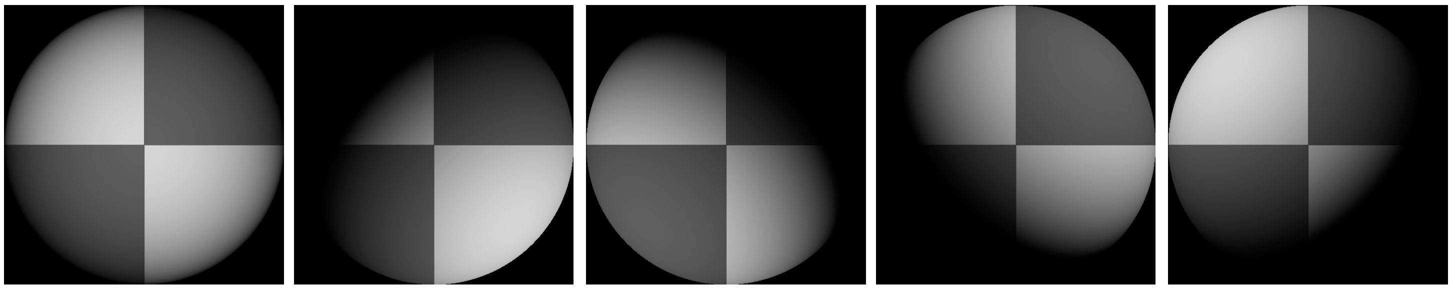
$$\begin{bmatrix} V_1^T \\ V_2^T \\ \vdots \\ V_n^T \end{bmatrix} g(x,y) = \begin{bmatrix} I_1(x,y) \\ I_2(x,y) \\ \vdots \\ I_n(x,y) \end{bmatrix}$$

$n \times 3$ known | 3×1 unknown | $n \times 1$ known

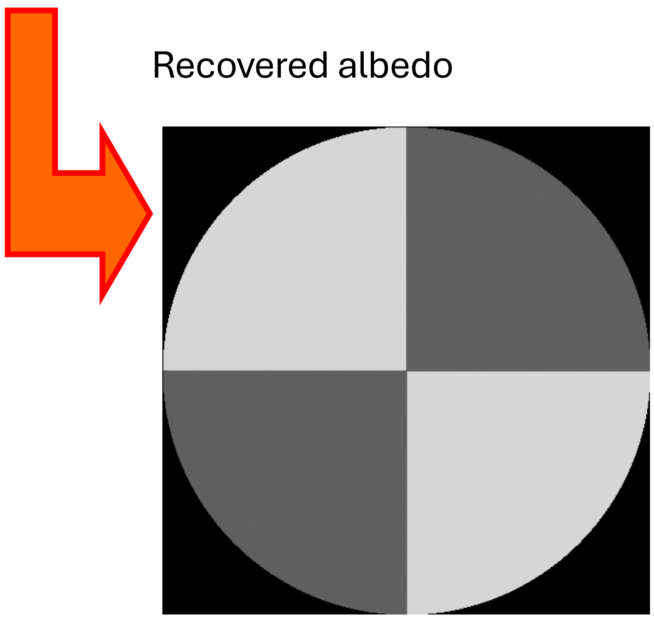
- Obtain least-squares solution for $g(x,y)$, which we defined as $\rho(x,y)N(x,y)$
- Since $N(x,y)$ is the *unit* normal, $\rho(x,y)$ is given by the magnitude of $g(x,y)$
- Finally, $N(x,y) = \frac{1}{\rho(x,y)} g(x,y)$

F&P 2nd ed., sec. 2.2.4

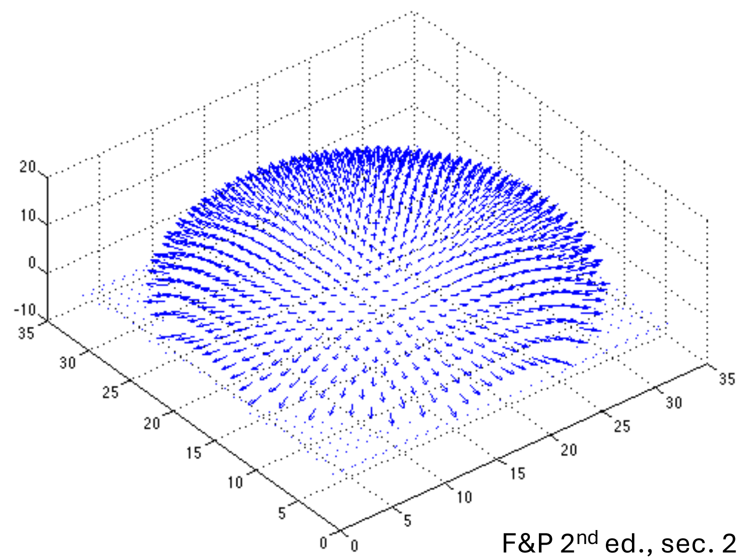
Synthetic example



Recovered albedo



Recovered normal field



F&P 2nd ed., sec. 2.2.4

Recovering a surface from normals

- Recall: the surface is written as

- $(x, y, f(x, y))$

- This means the unit normal has the following form:

- $N(x, y) = \frac{1}{\sqrt{f_x^2 + f_y^2 + 1}} \begin{bmatrix} f_x \\ f_y \\ 1 \end{bmatrix}$

- Write the estimated vector g as

- $g(x, y) = \begin{bmatrix} g_1(x, y) \\ g_2(x, y) \\ g_3(x, y) \end{bmatrix}$

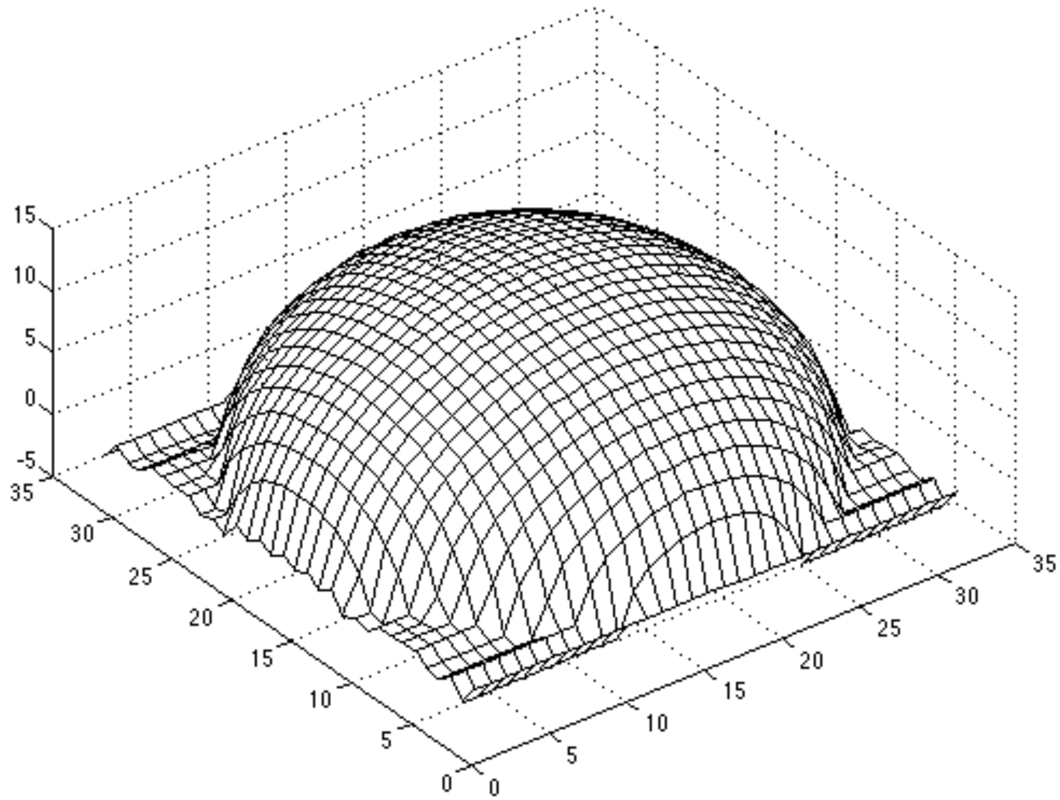
- Then we obtain values for the partial derivatives of the surface:

- $f_x(x, y) = \frac{g_1(x,y)}{g_3(x,y)}$

- $f_y(x, y) = \frac{g_2(x,y)}{g_3(x,y)}$

F&P 2nd ed., sec. 2.2.4

Surface recovered by integration

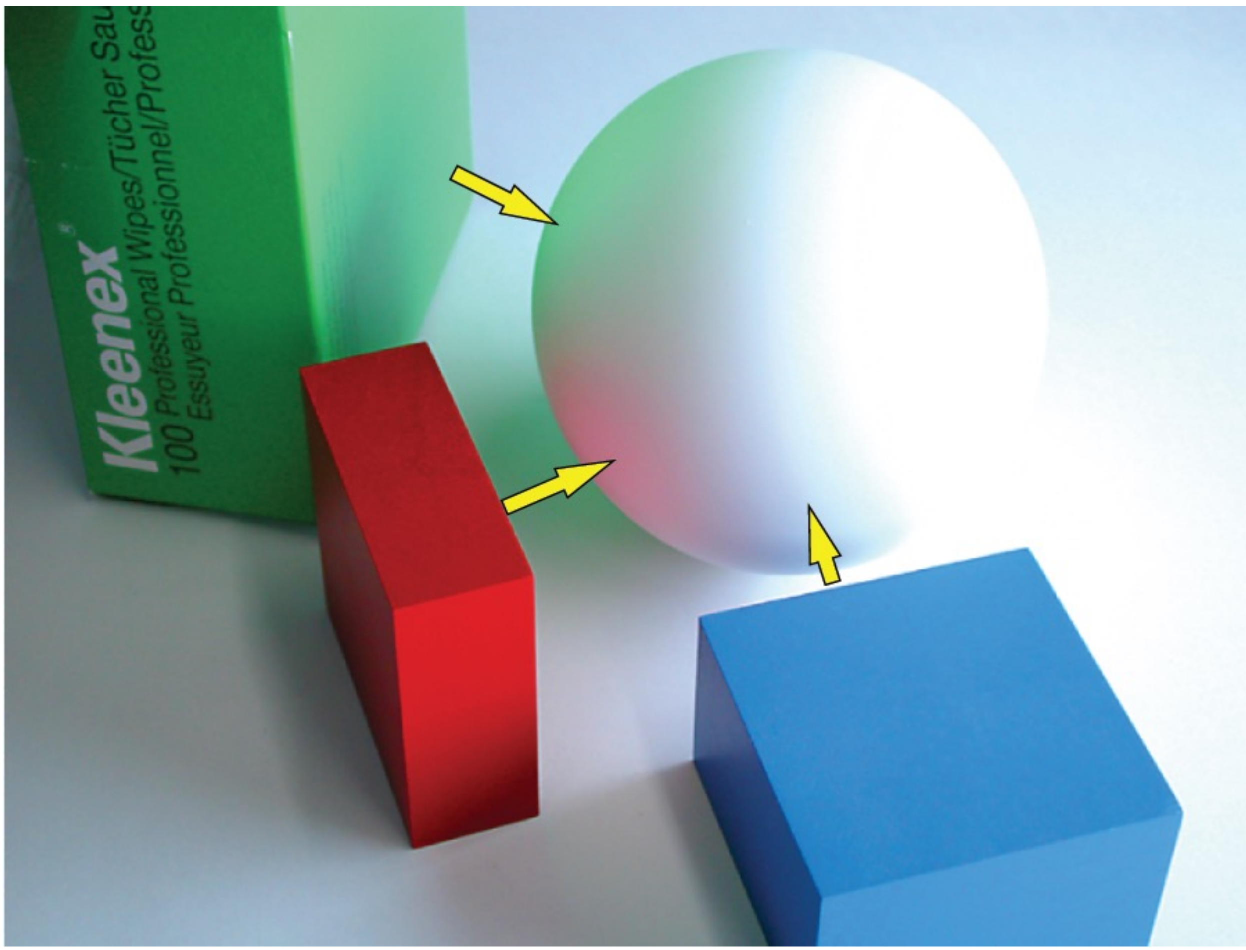


F&P 2nd ed., sec. 2.2.4

Variants

- With enough images, etc.
  - you can do this if the illuminants aren’t known
- You can do this with one image:
  - if there are three illuminants, each of a different color
- With enough images, etc.
  - interreflections are not a major problem

Interreflections are a problem



From Koenderink slides on image texture and the flow of light

- but not a major one for photometric stereo – why?

CRF Estimation

Assume you have N images of exactly the same scene under exactly the same lighting, but obtained with different exposure times. Turn each image into a vector, so there are fewer indices to bother with. Write p_i for the power arriving at the i'th location. This is the same for each image, but the energy is different because the exposure time is different. The energy arriving at the i'th location for the k'th exposure is

$$E_{ik} = p_i \left[\Delta t\right]_k .$$

If you write $\mathbf{p}$ for a vector of powers, $\mathbf{t}$ for a vector of exposure times, and arrange these unknown energies into a matrix $\mathcal{E}$, it will have the form

$$\mathcal{E} = \mathbf{p}\mathbf{t}^T$$

and so rank one (**exercises**).

CRF Estimation

Write $F(\cdot;\theta)$ for the CRF (so $F(\mathcal{M};\theta)$ is the matrix of values obtained by applying the CRF to the elements of $\mathcal{M}$, and θ are the parameters of the CRF). Choose some loss to compare pixel predictions from the model with observations, and write $\mathcal{L}(\cdot)$ for that loss. Choose a smoothness penalty for the CRF **exercises** , and write $\mathcal{L}_s(\theta)$ for that penalty. You want to find θ, $\mathbf{e}$ and $\mathbf{t}$ to minimize

$$\mathcal{L}(\mathcal{I} - F(\mathbf{e}\mathbf{t}^T;\theta)) + \mathcal{L}_s(\theta).$$

CRF estimation