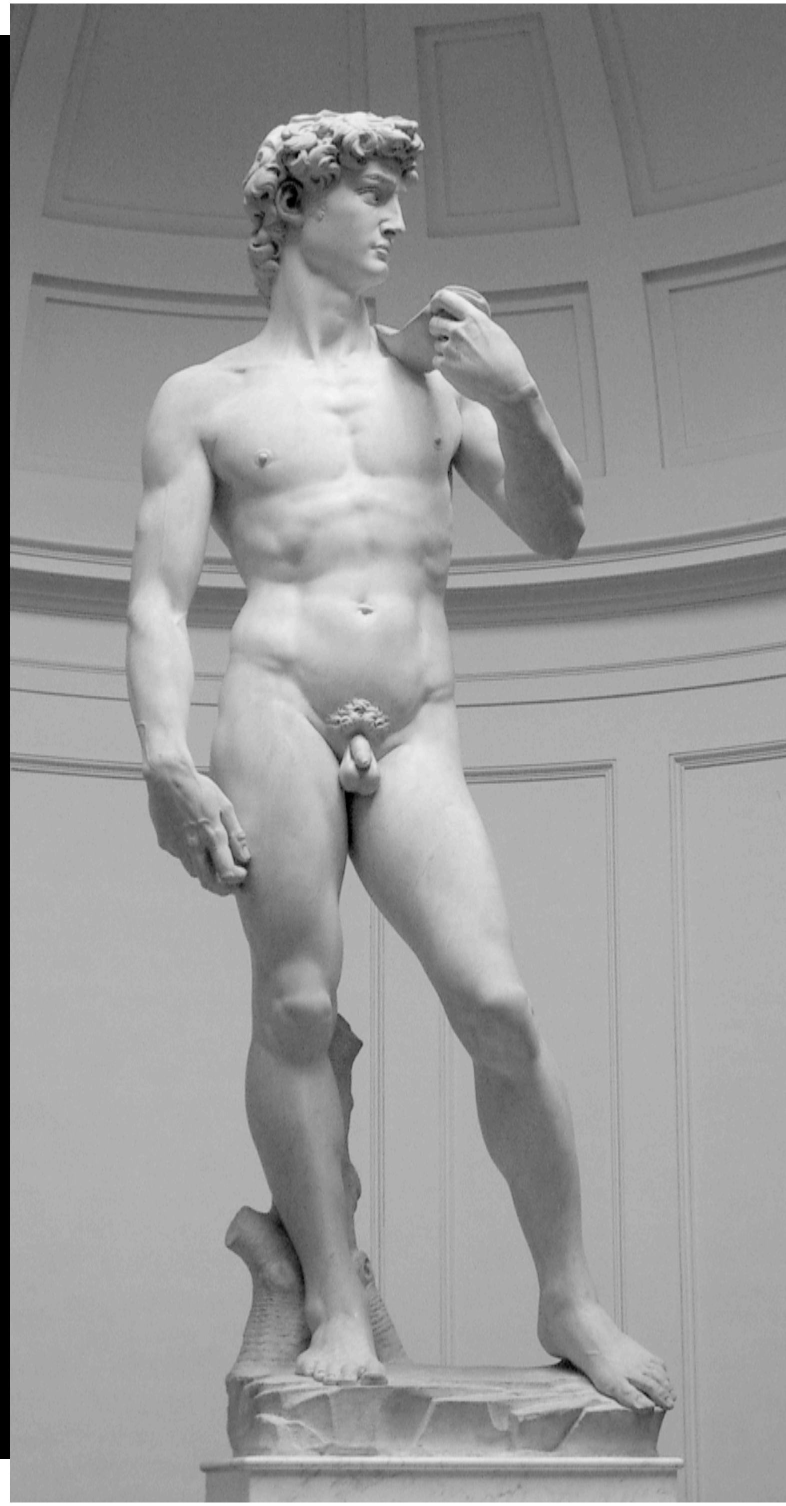
Presentation of “Shape Estimation in Natural Illumination,” by Johnson and Adelson

Daniel J Butler
University of Washington
11/29/11

Question:

Question:

Does Michaelangelo’s David have a six-pack?

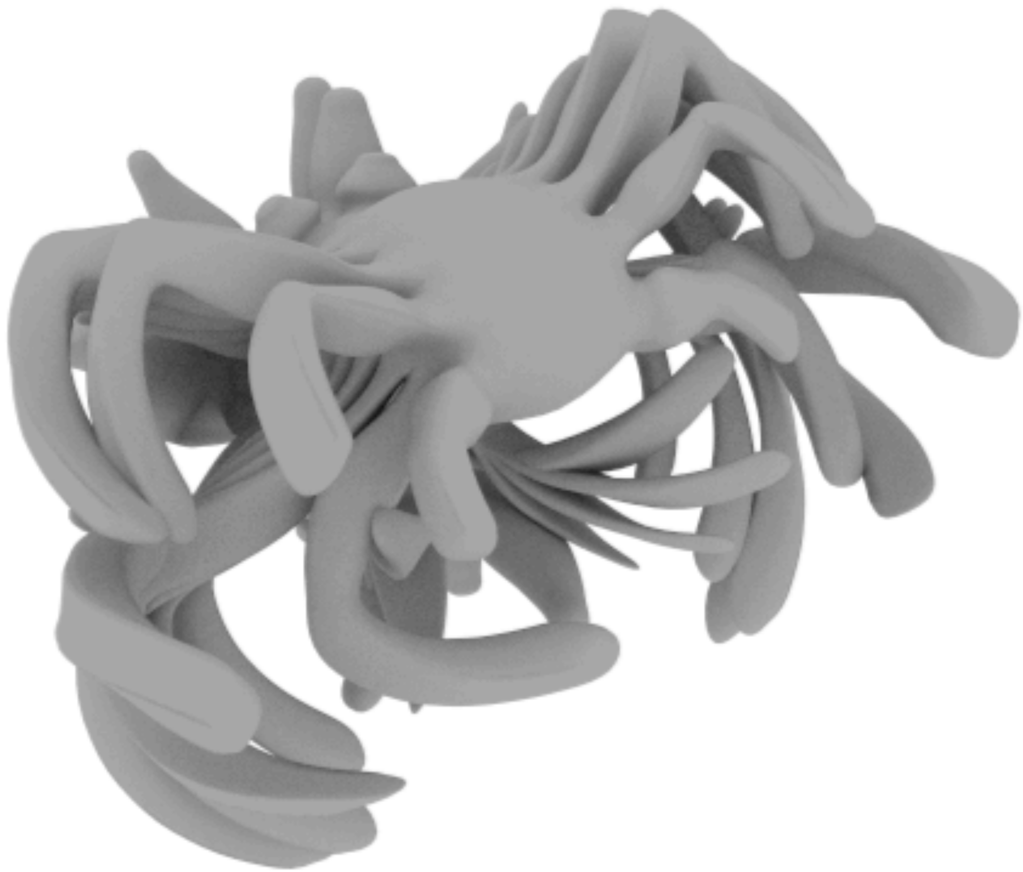


How do you know?

How do you know?

Shape from shading.

You can even do unfamiliar objects!



entoforms, by macuono

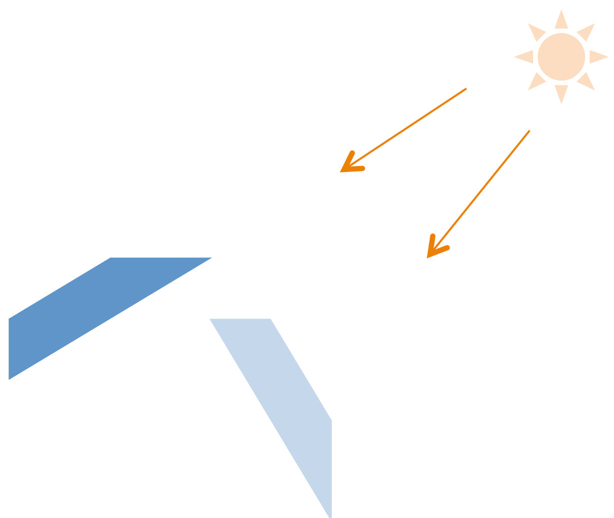
… on other planets!

Martian geology

Each **pixel value**

contains information about

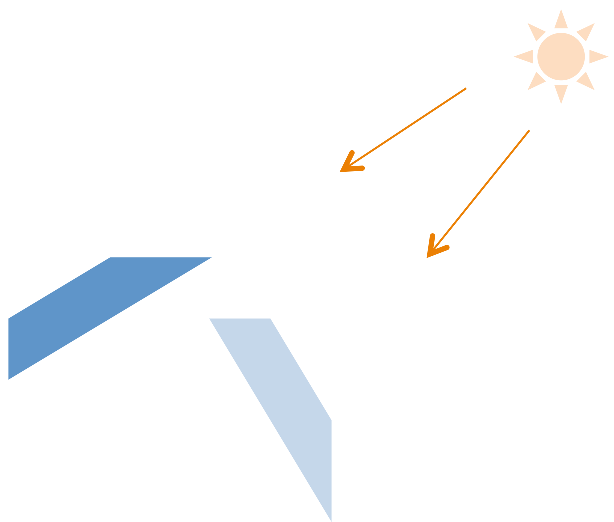
the **surface normal** at that point.



Each **pixel value**

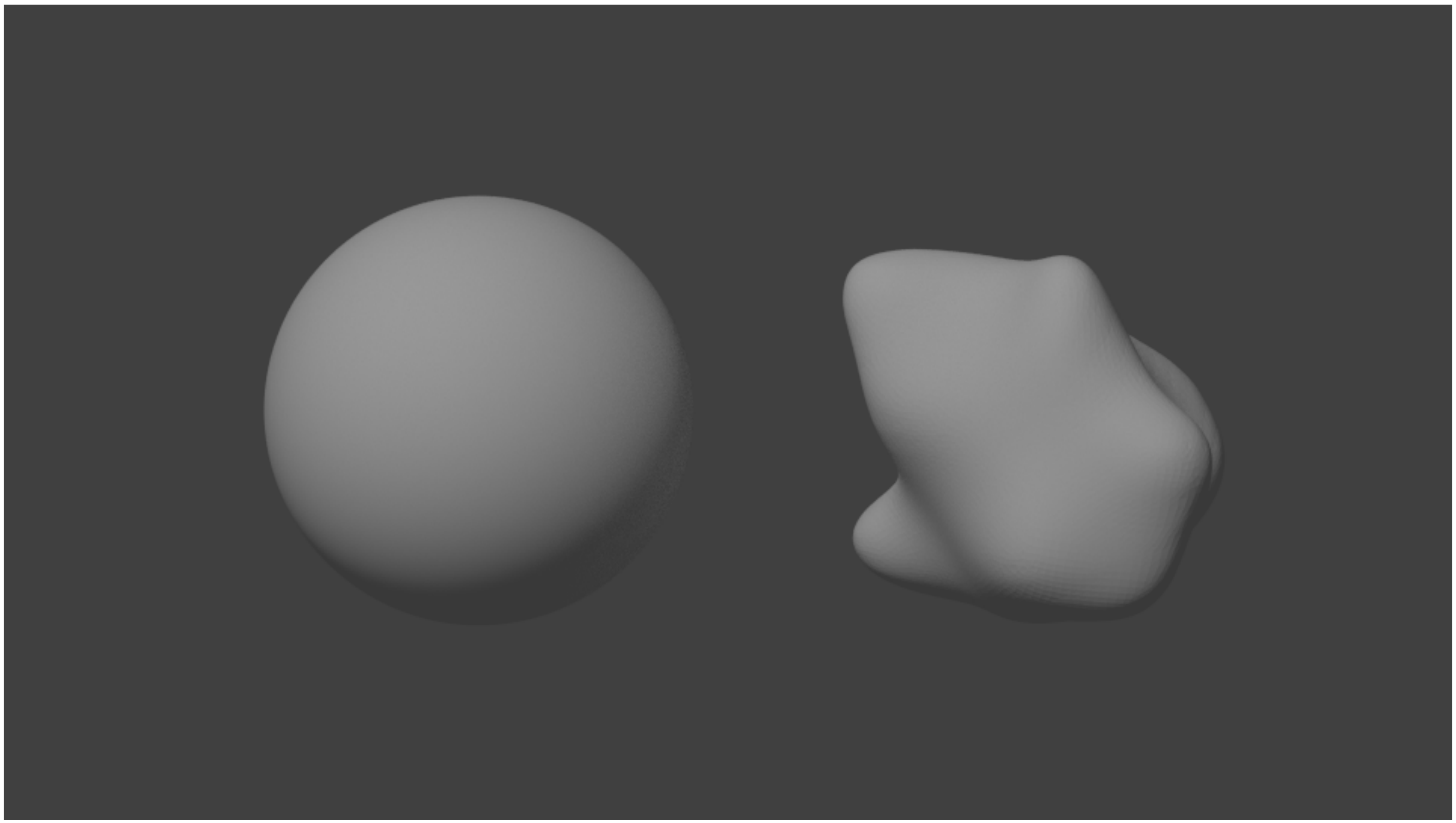
contains information about

the **surface normal** at that point.

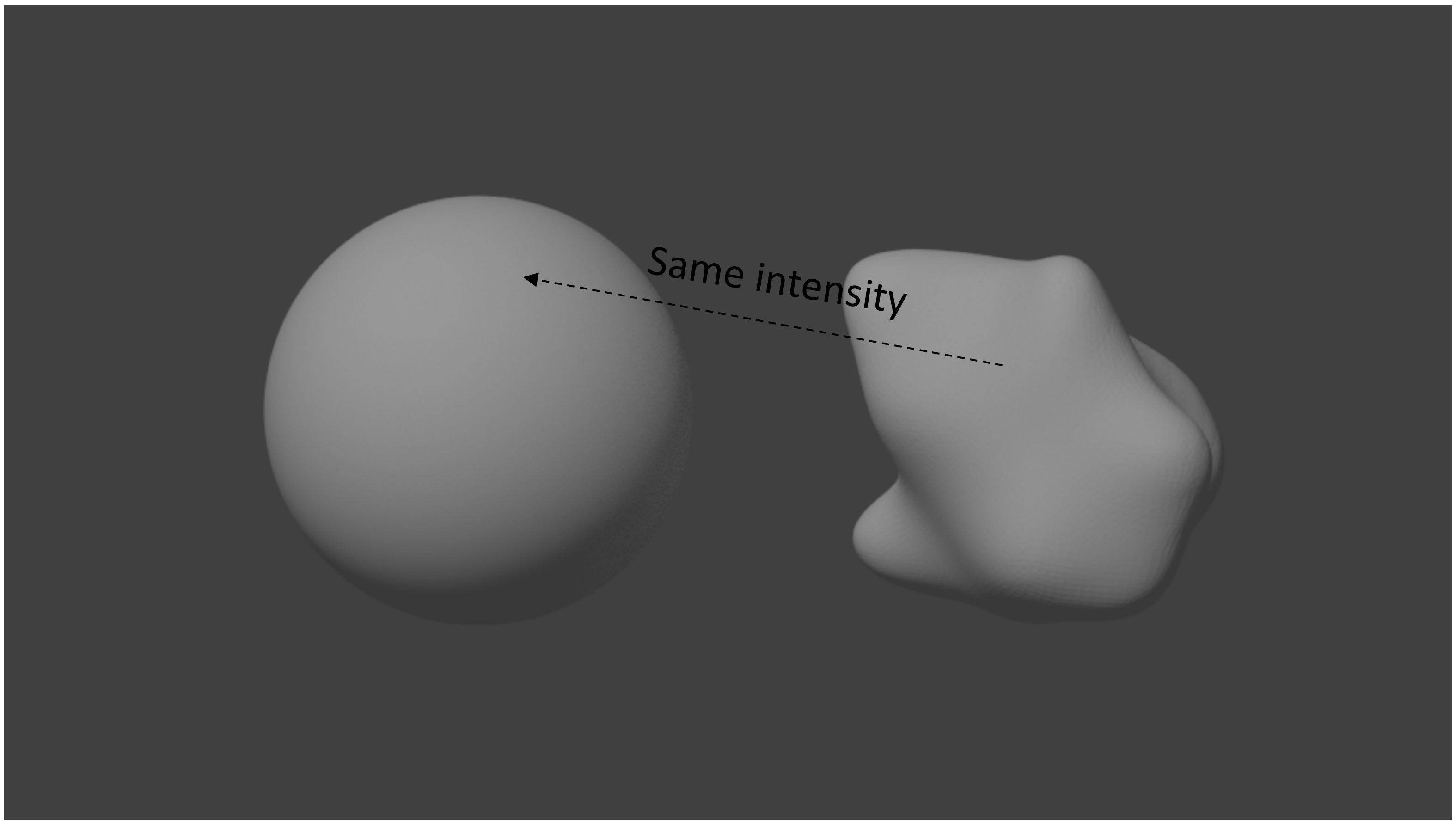


(Assume lighting direction is known.)

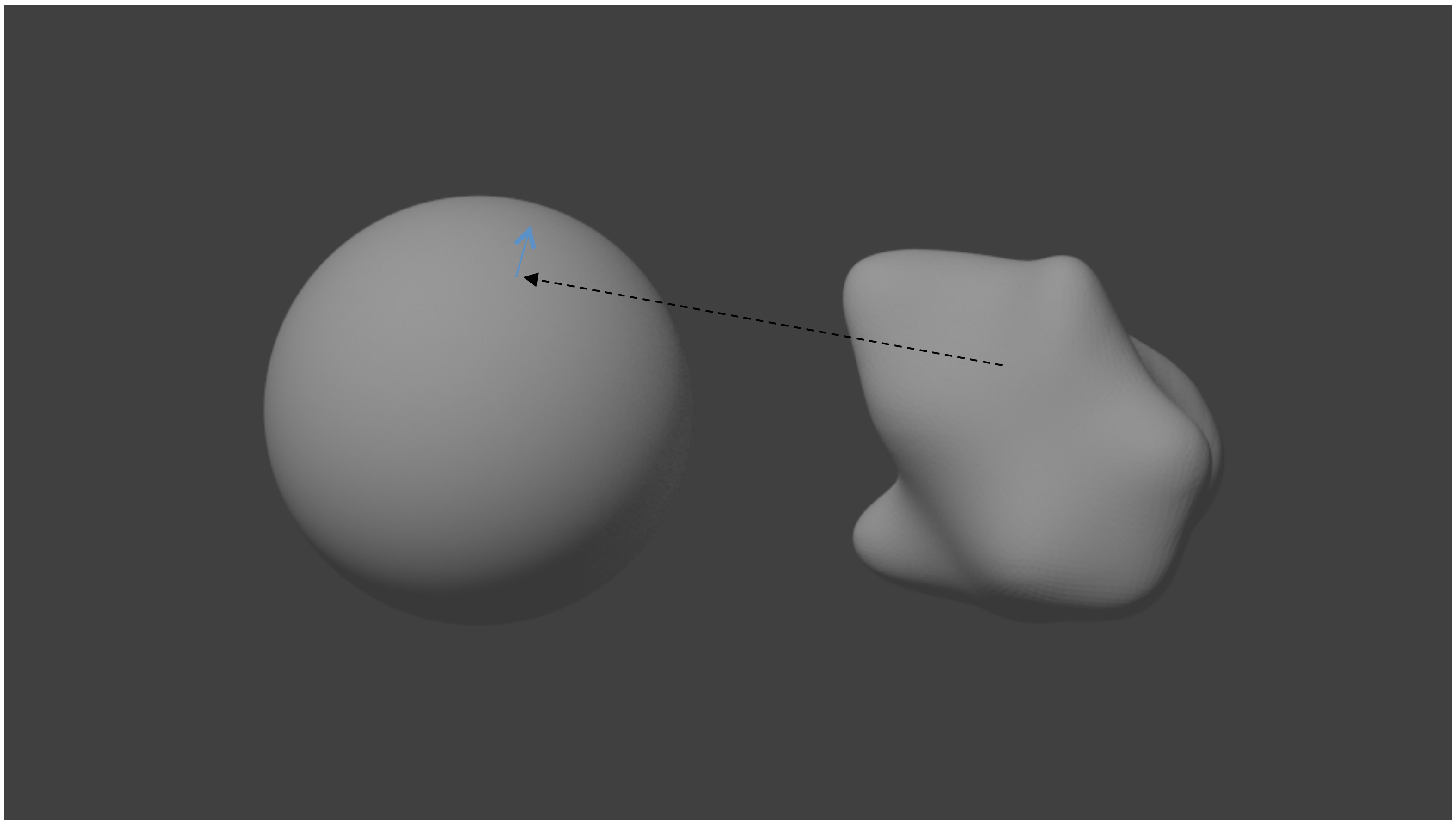
Ideally, every point could be mapped to one on the sphere



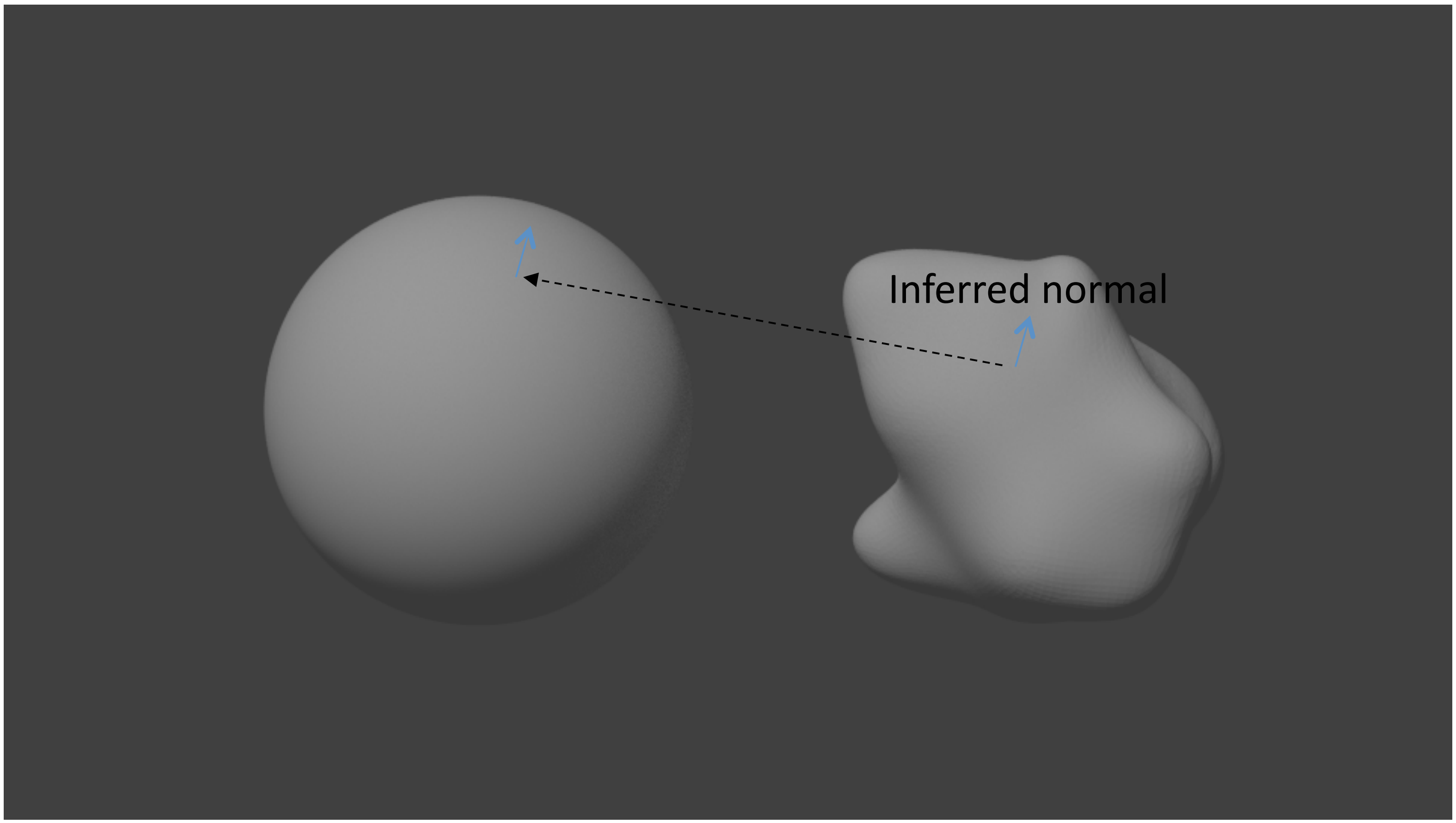
Ideally, every point could be mapped to one on the sphere



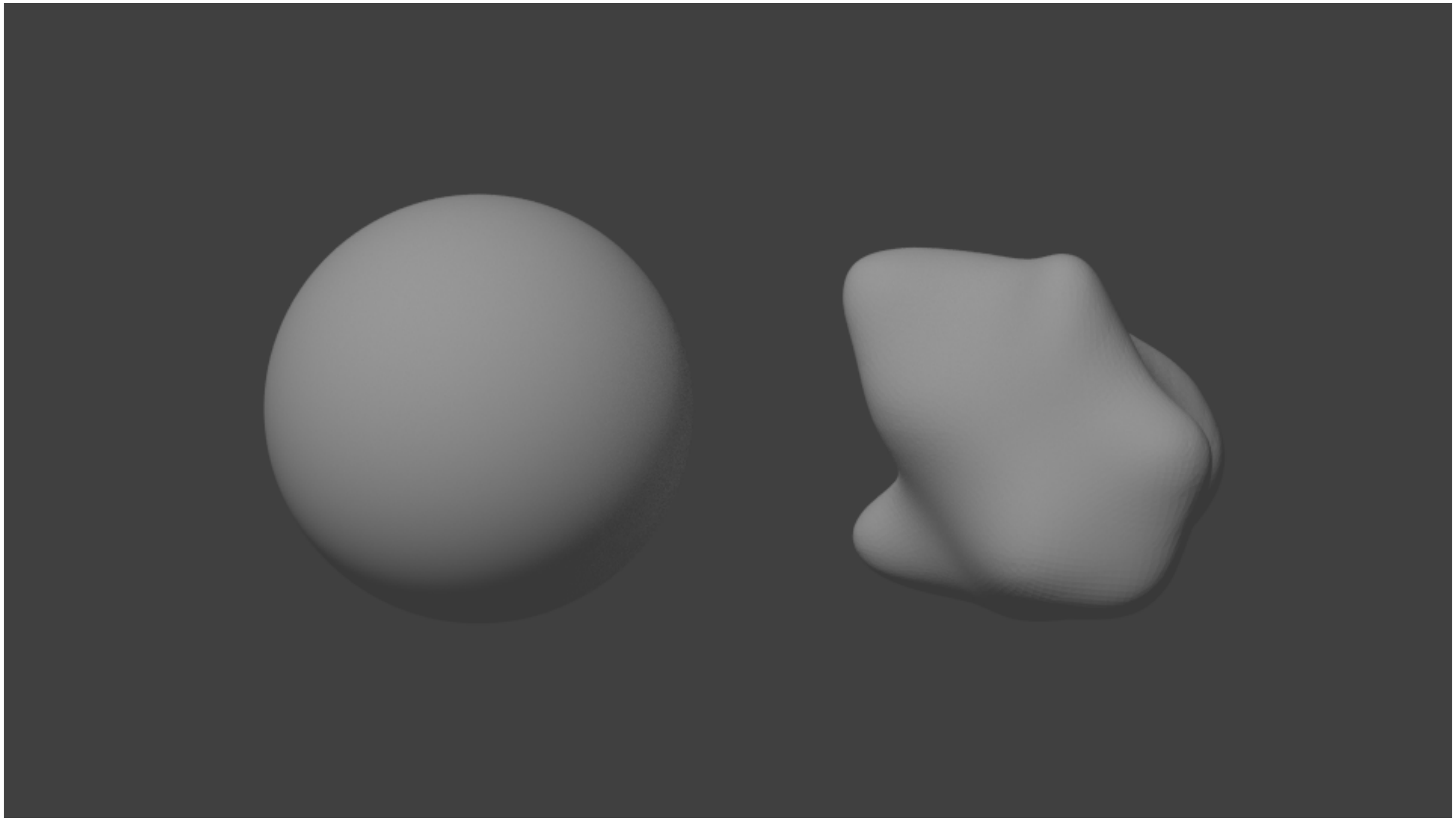
Ideally, every point could be mapped to one on the sphere



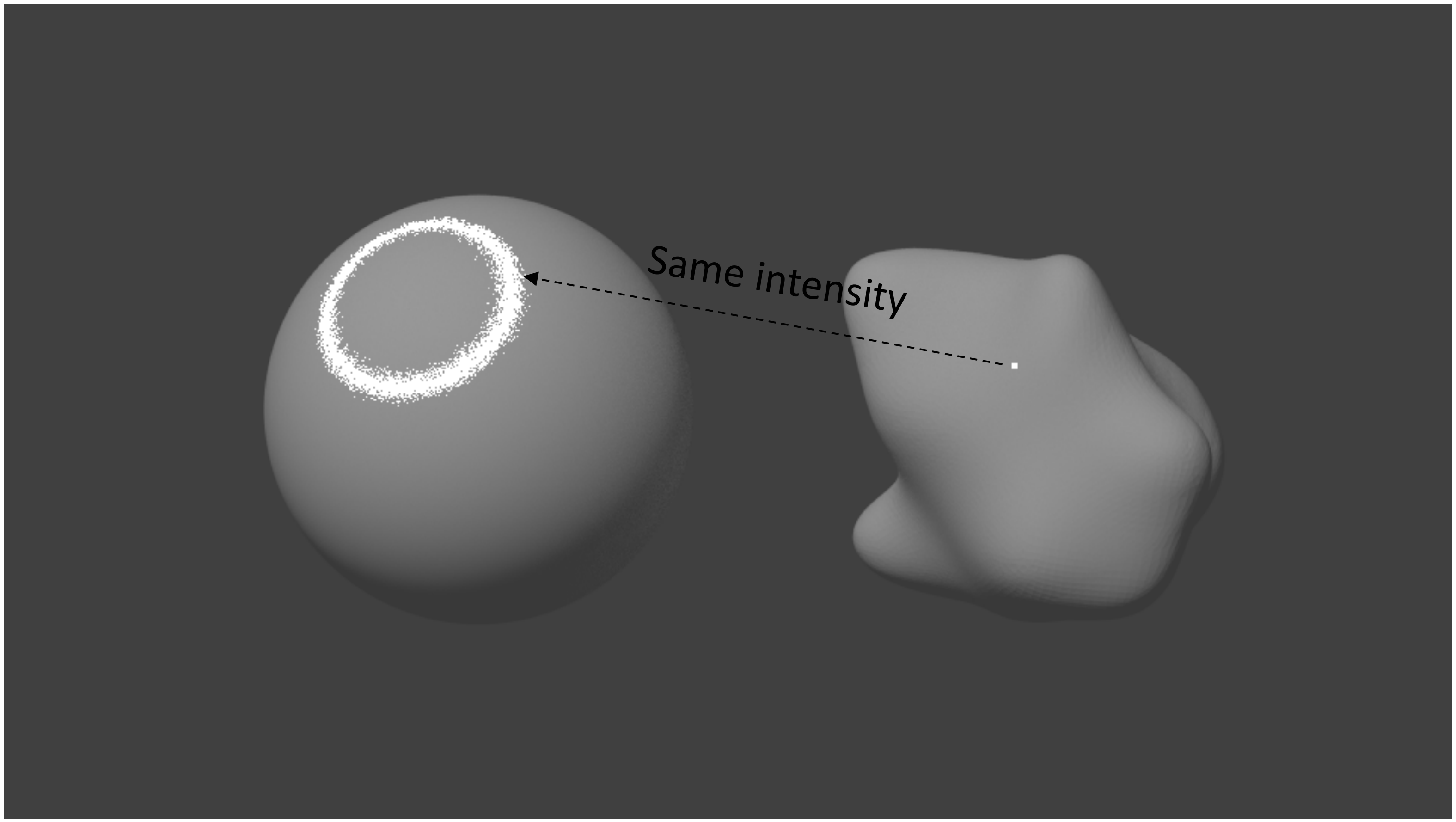
Ideally, every point could be mapped to one on the sphere



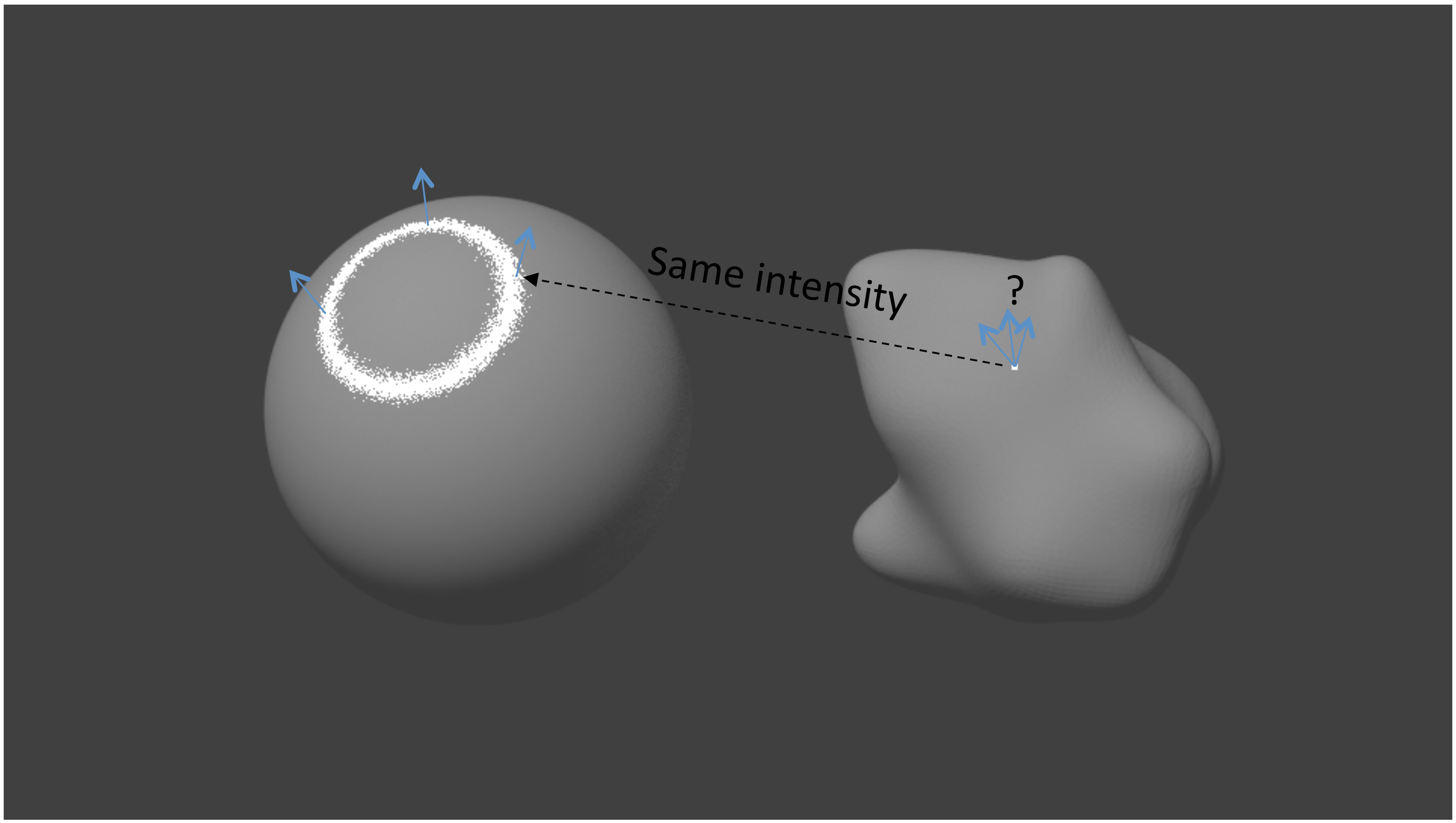
Unfortunately, there is ambiguity.



Unfortunately, there is ambiguity.



Unfortunately, there is ambiguity.



Technical Report 232

Shape From Shading: A Method for Obtaining the Shape of a Smooth Opaque Object From One View

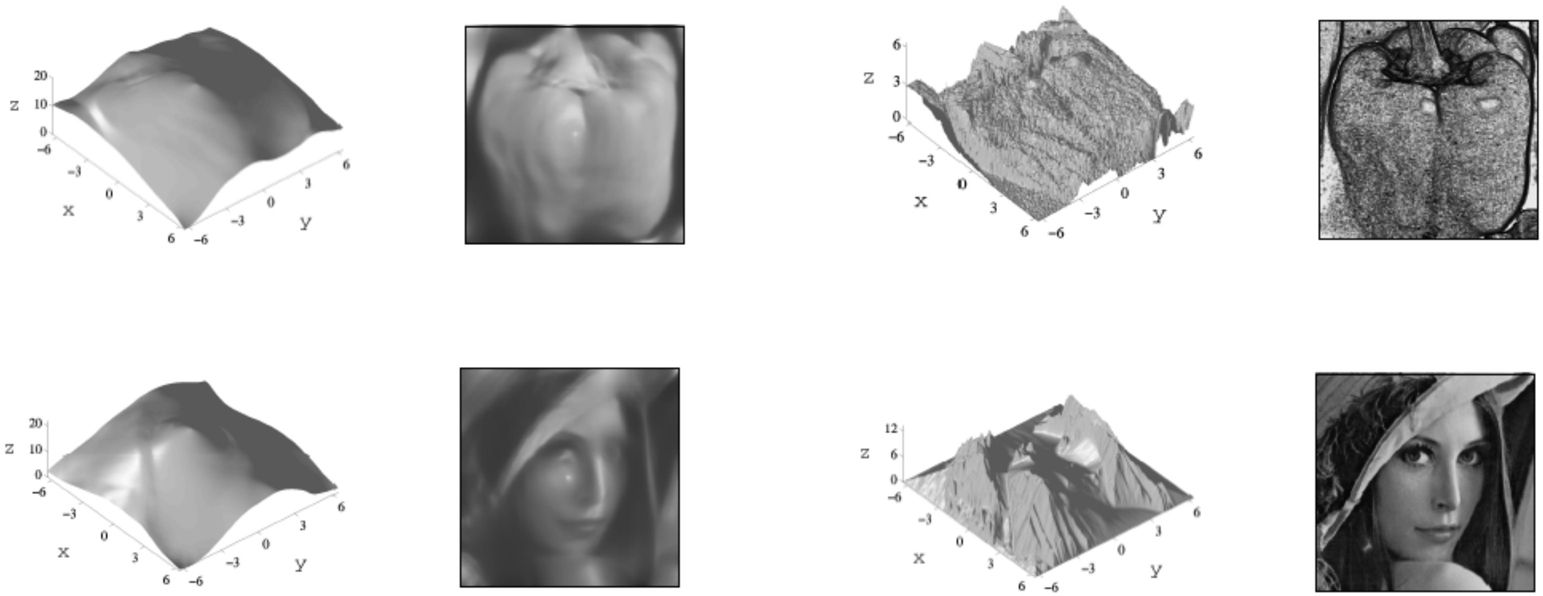
Berthold K. P. Horn

MIT Artificial Intelligence Laboratory

Horn showed:

Despite local ambiguity, if you know some normals,
you can infer the whole normal map by integrating across space

Results of classical shape-from-shading



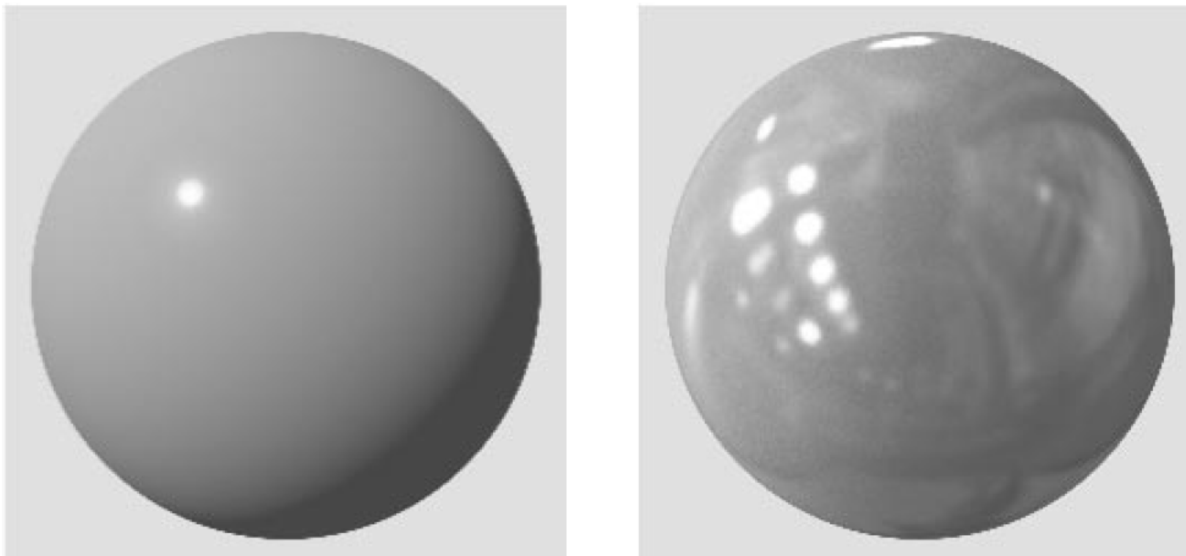
J.-D. Durou, M. Falcone, and M. Sagona.
Numerical methods for shape-from-shading: A new survey with benchmarks.
Computer Vision and Image Understanding, 109(1), 2008.

Shape Estimation in Natural Illumination

Micah K. Johnson and Edward H. Adelson
Massachusetts Institute of Technology

Idea:

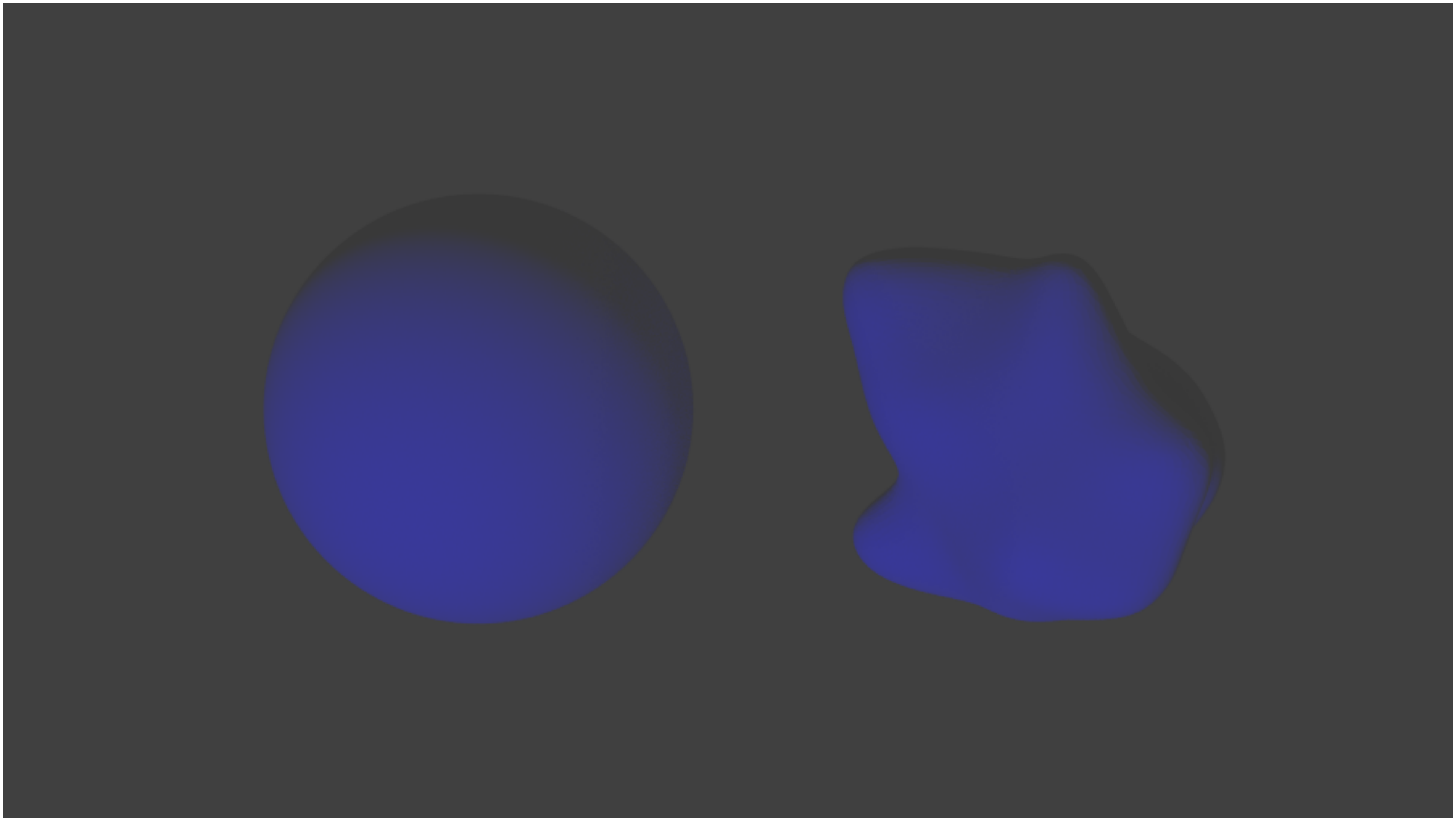
Natural scenes have complex lighting, which makes shape-from-shading **easier**



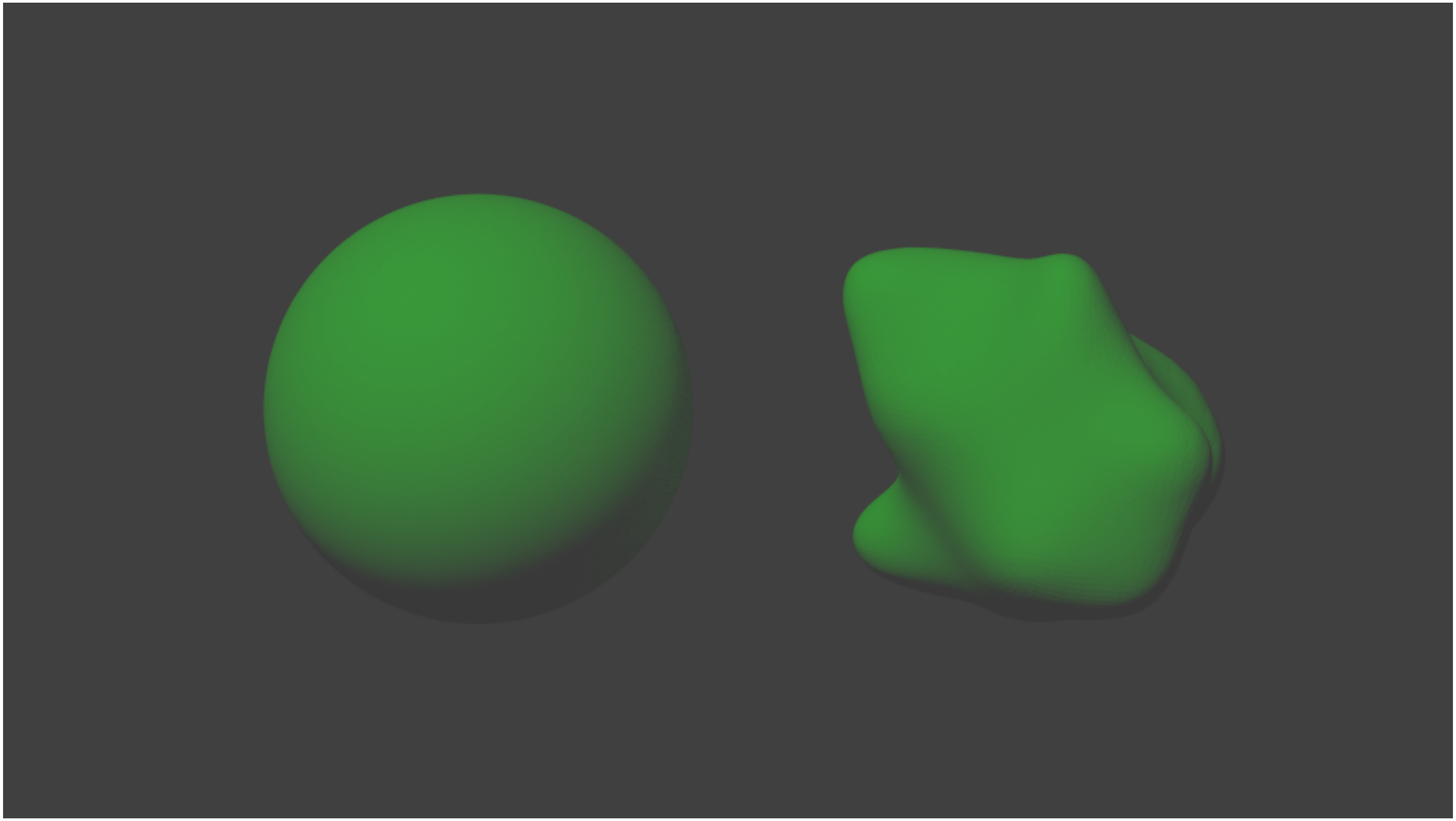
Statistics of Real-World Illumination

Dror, Leung, Adelson, and Willsky, *CVPR '01*

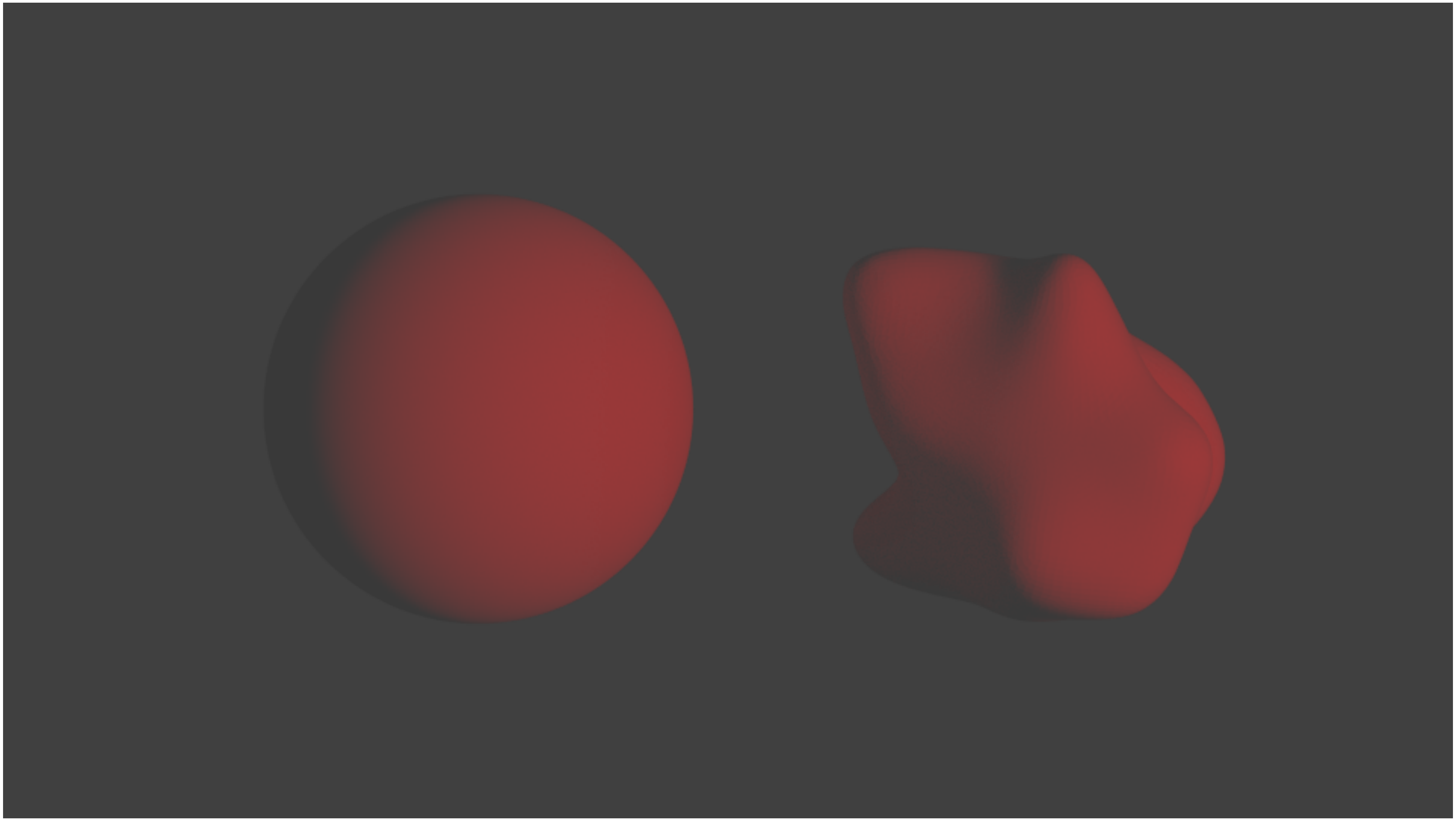
Photometric Stereo



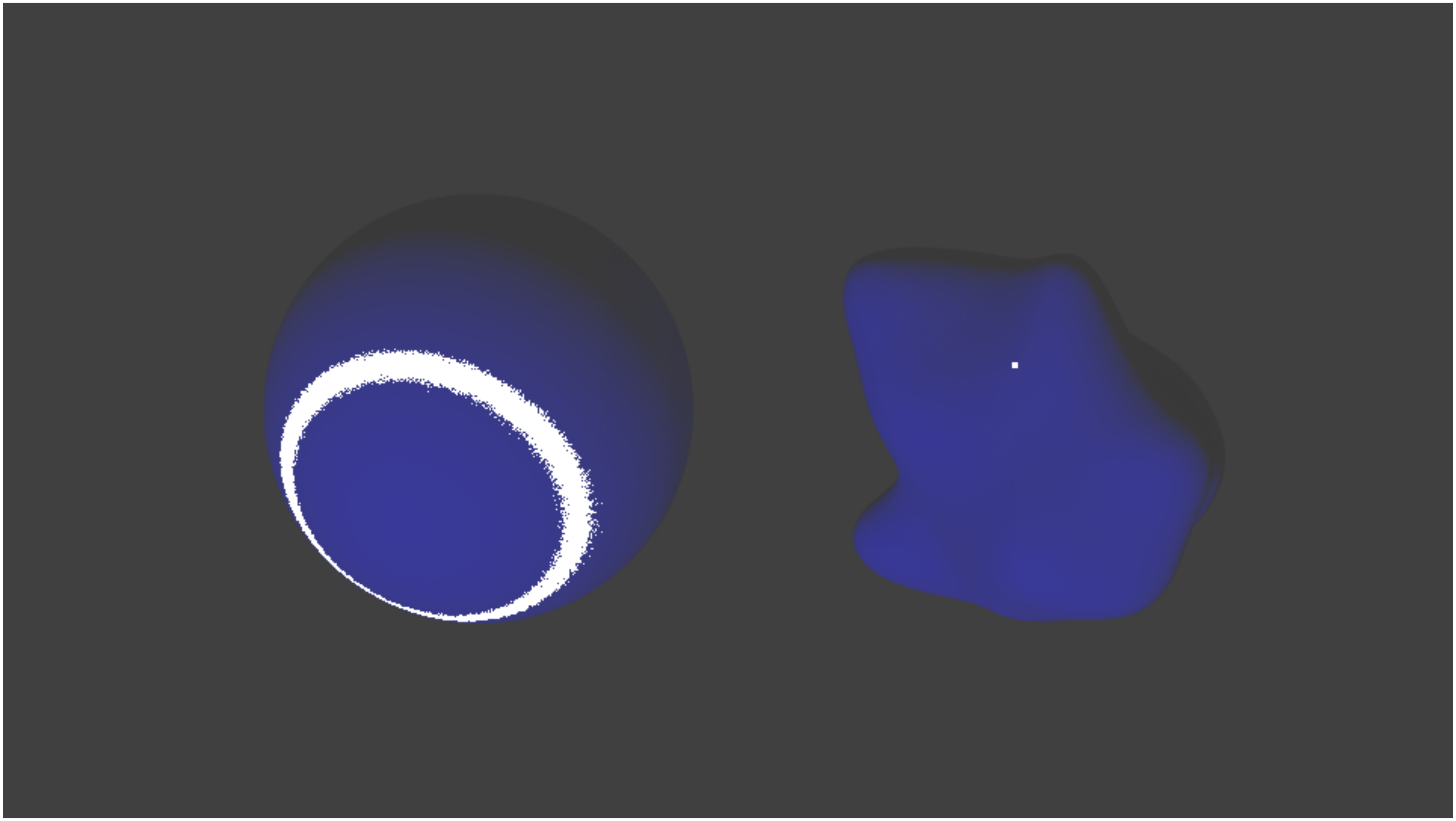
Photometric Stereo



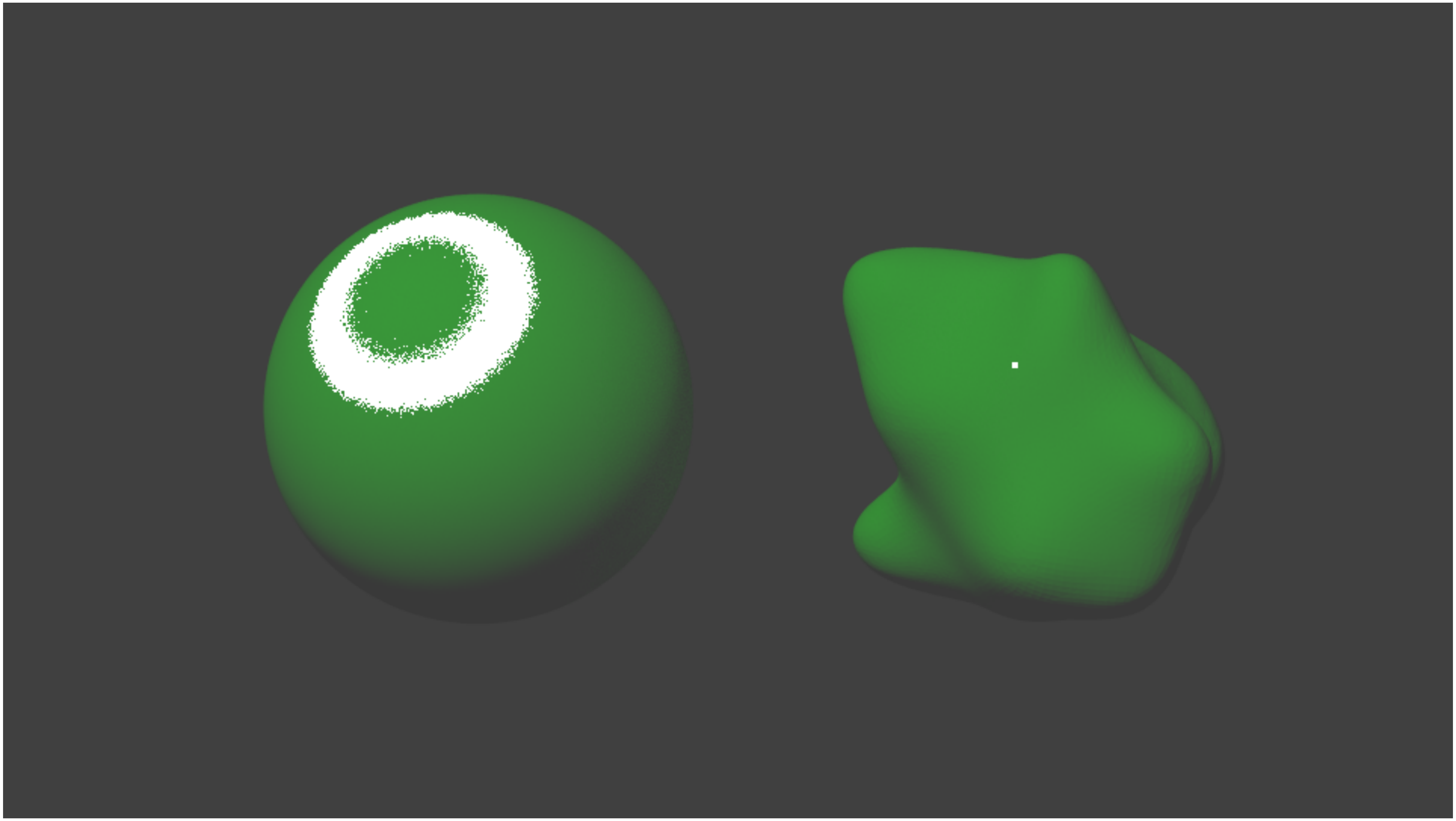
Photometric Stereo



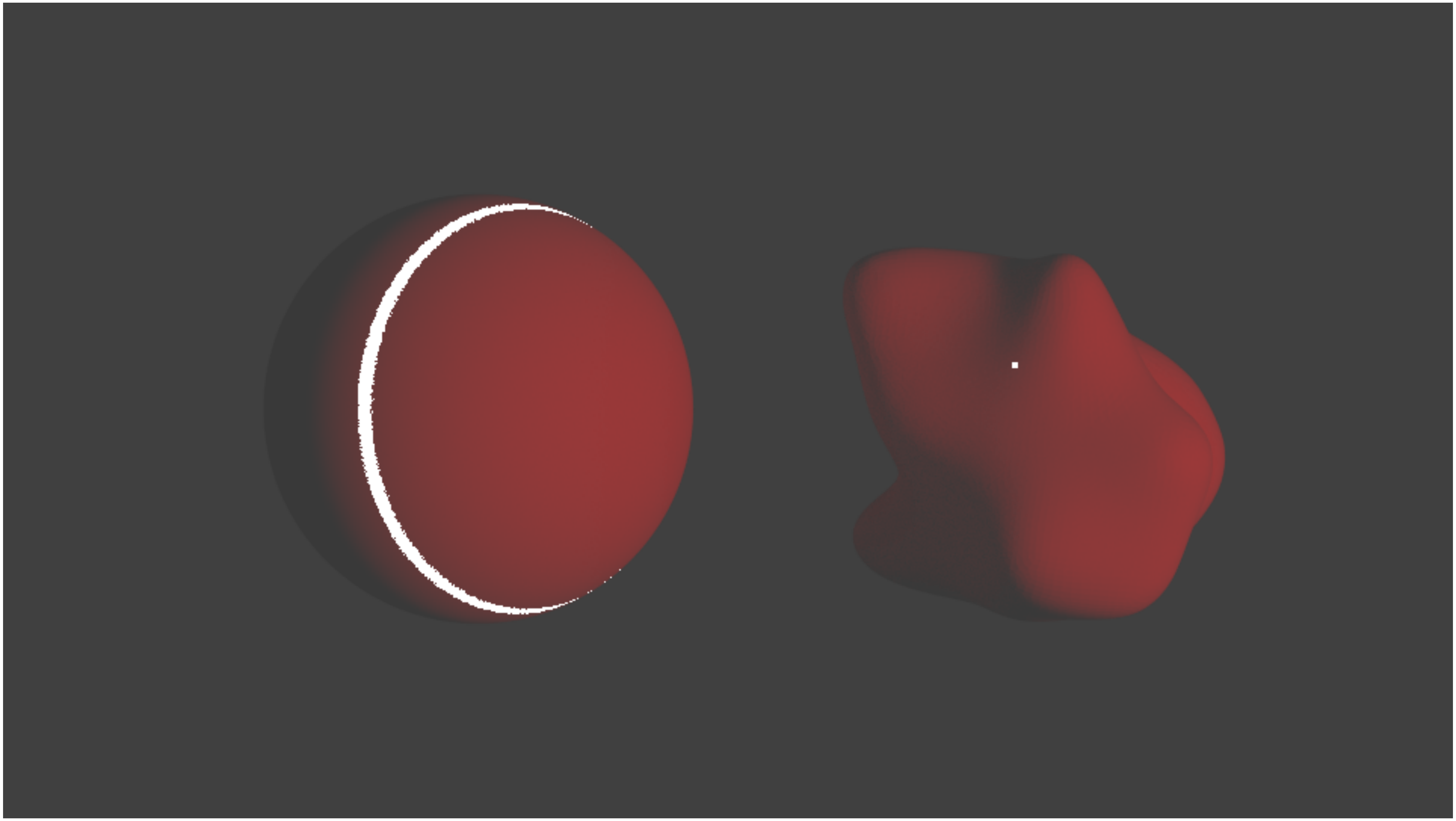
Photometric Stereo



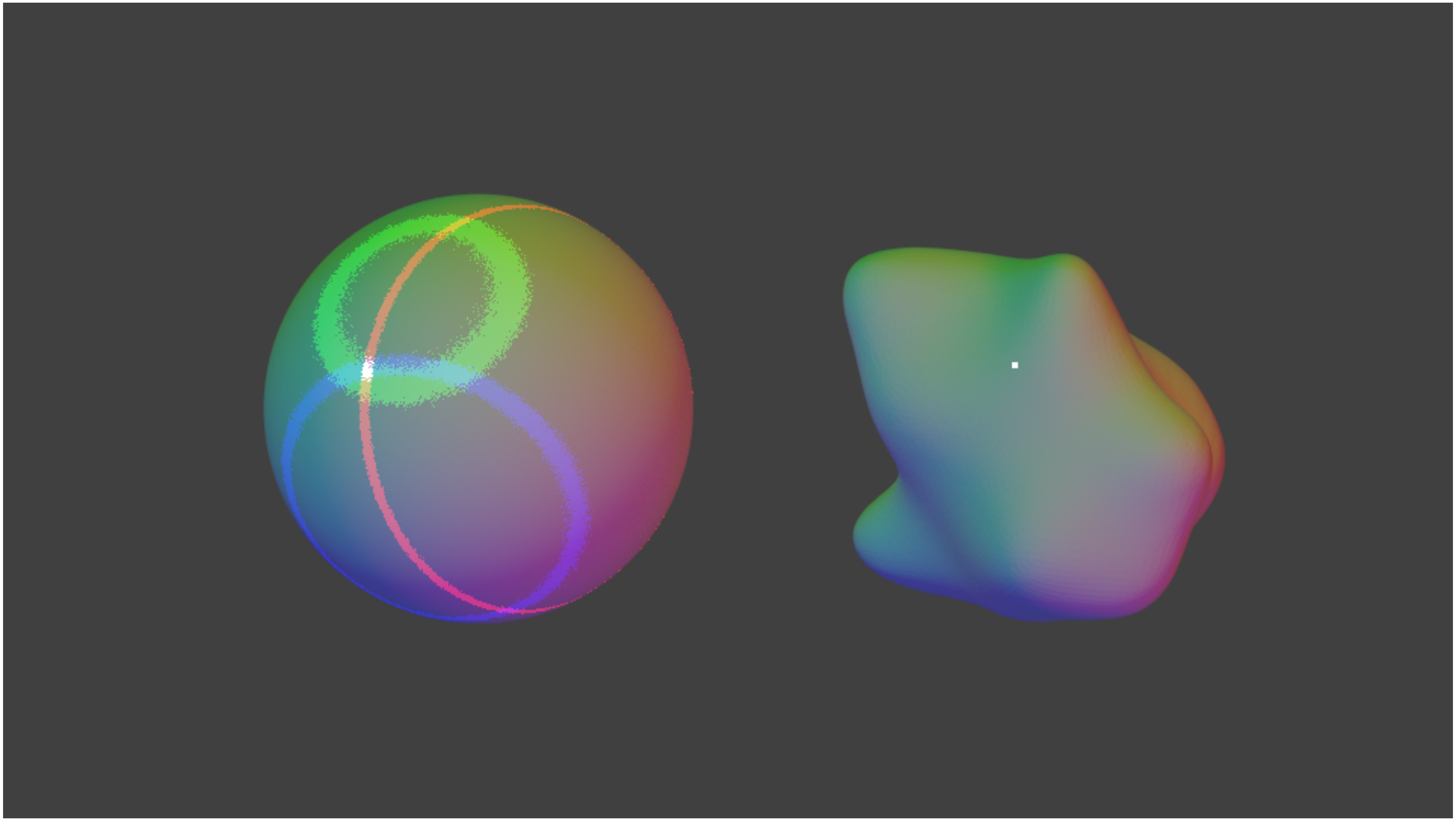
Photometric Stereo



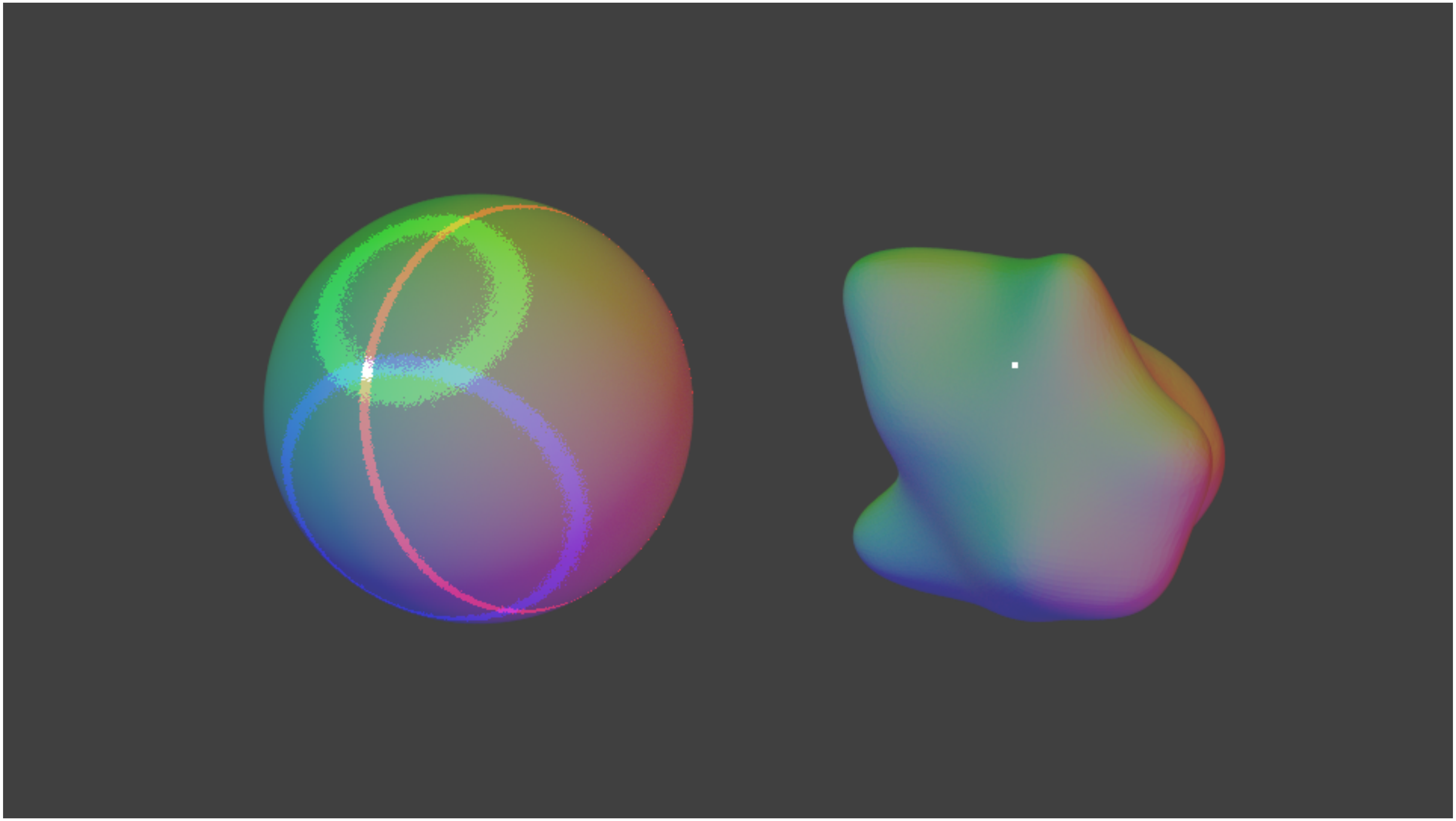
Photometric Stereo



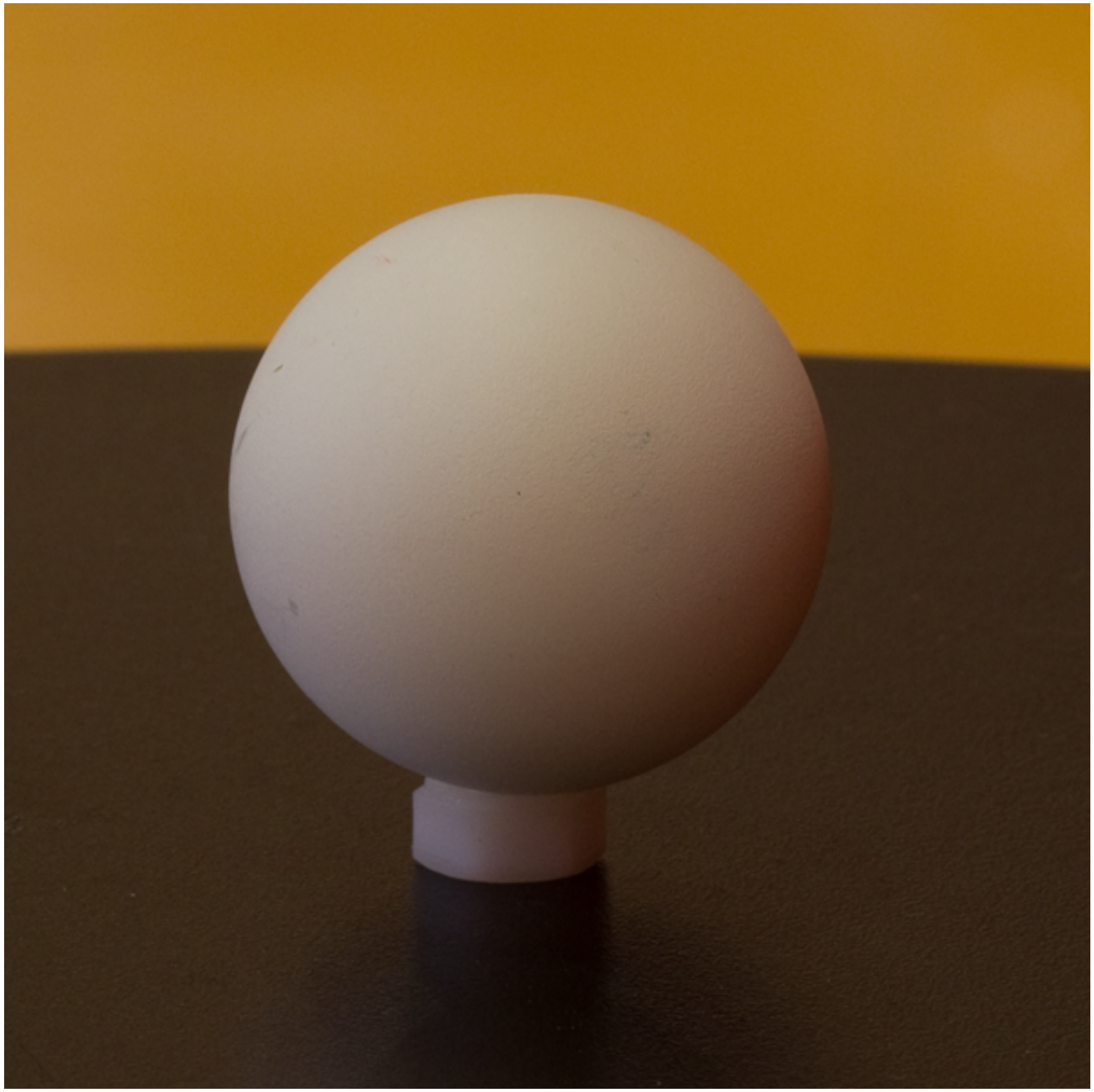
Photometric Stereo



Photometric Stereo

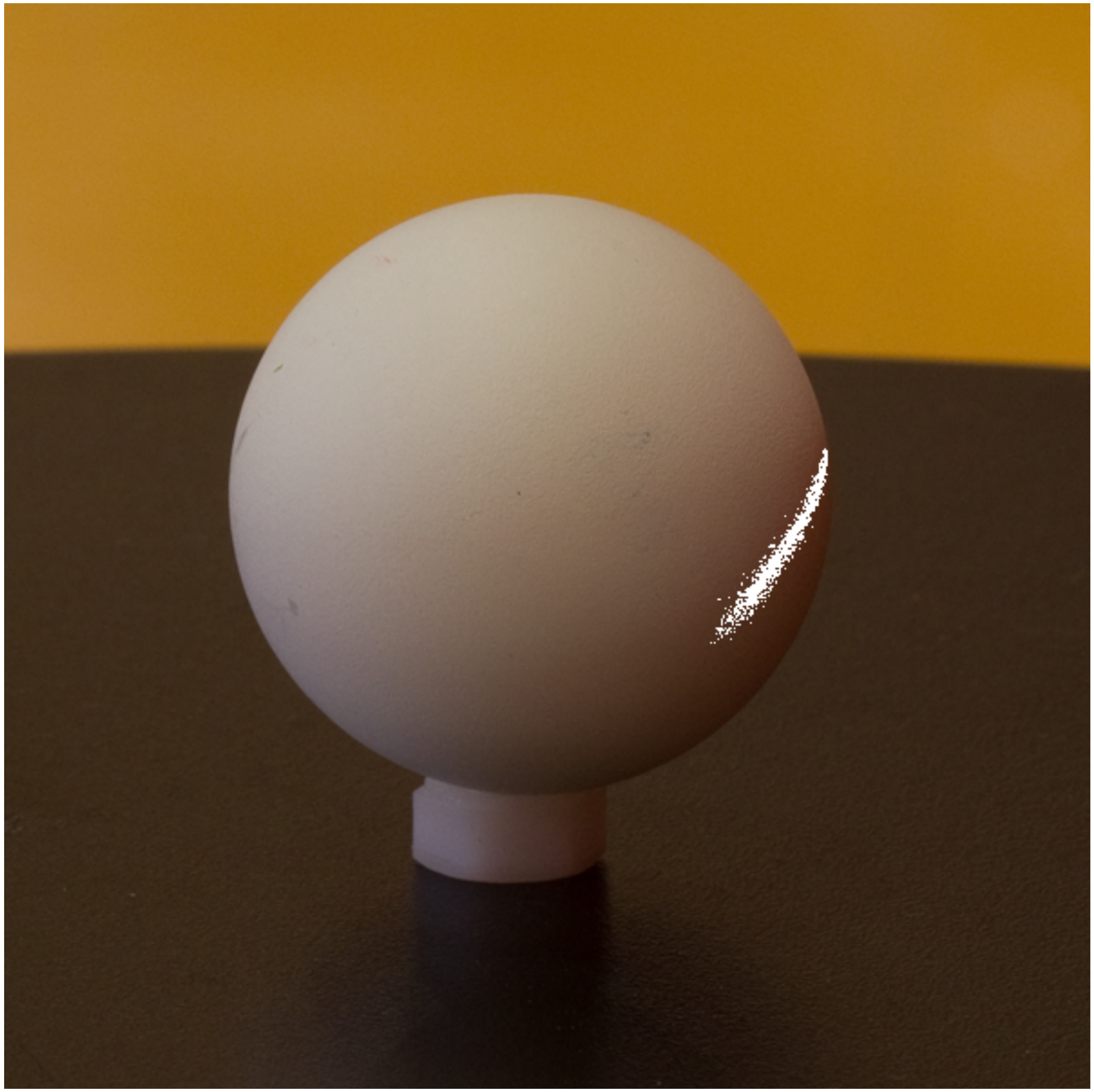


No local ambiguity → no complicated global algorithm!

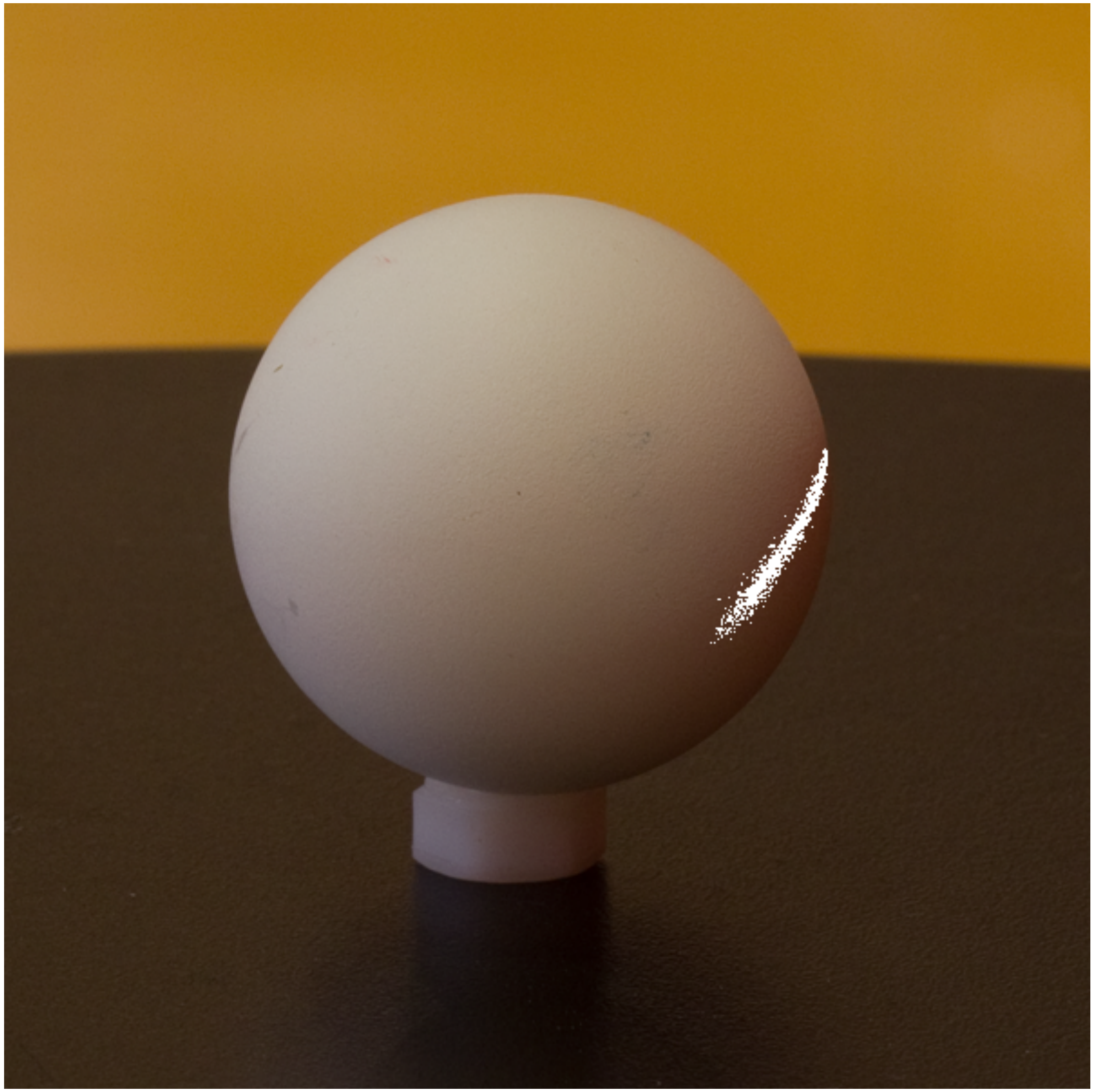


Not quite, photometric stereo, but not quite single-channel:

Not quite, photometric stereo, but not quite single-channel:



Not quite, photometric stereo, but not quite single-channel:



Still ambiguous, but global optimization should be easier

The Objective Function

Three terms:

- Pixel intensities implied by normals should match the observations
- Curl of the normal map should be close to 0 (normal map is the gradient of z(x,y))
- Along directions where the image does not change, neither should the normals (lighting and gradients are not coincidentally aligned)

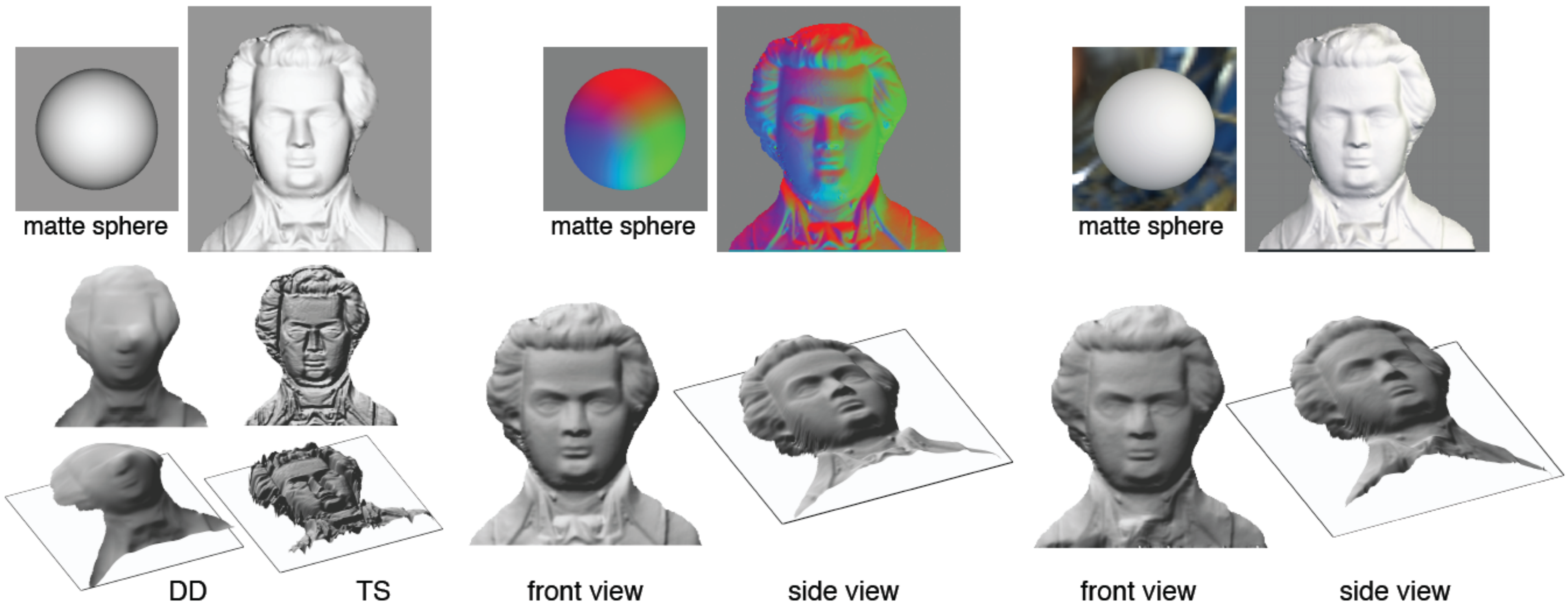
The Optimization Procedure

At each pixel, in raster scan order:

- Try optimizing 3x3 patch with three different initializations (patch, patch left, patch above)
- Pick the "best" outcome and keep it

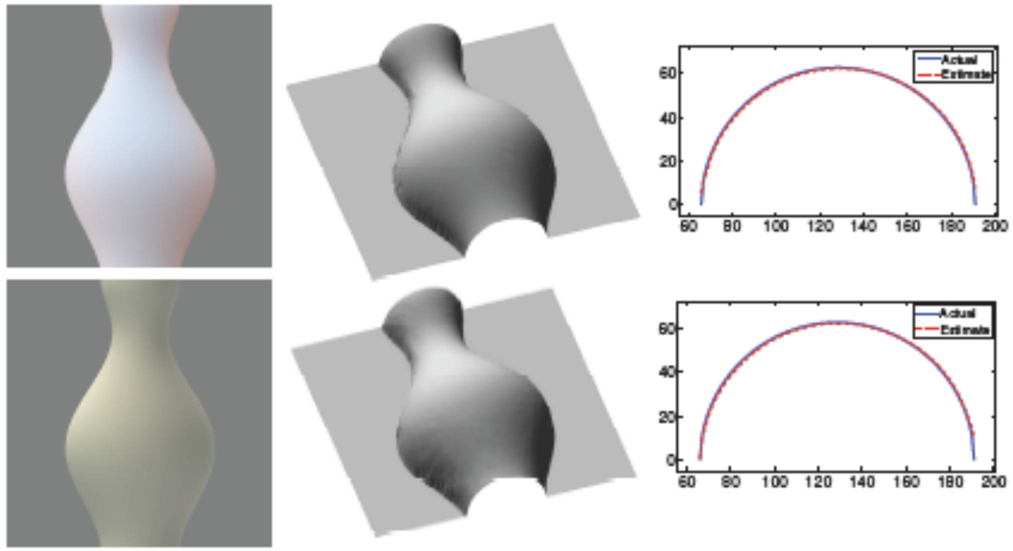
- Do this at coarse scales first, then upsample to initialize finer scales

Some results:

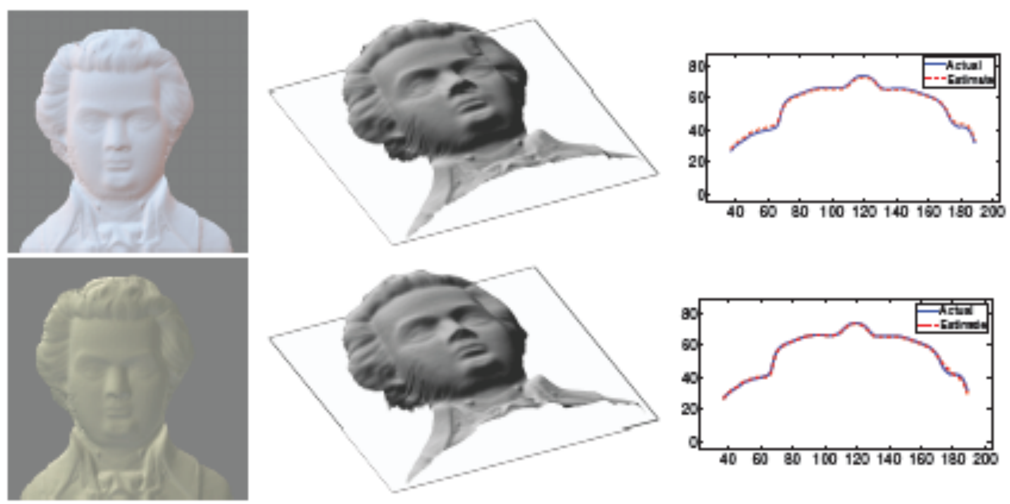


(a) Standard SFS **(b) RGB Photometric Stereo** **(c) Natural Illumination**

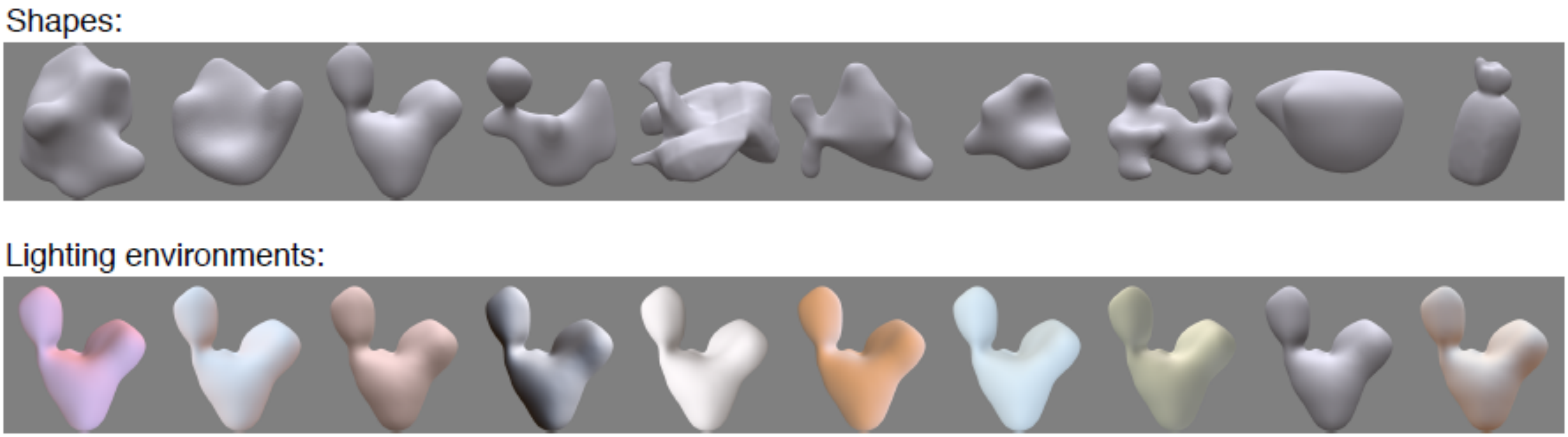
More results:



(a) image (b) estimated surface (c) profile



(a) image (b) estimated surface (c) profile



Across all surface normals from all 100 images,
90% have an angular error lower than 10 degrees

My thoughts:

Basic insight is very cool

Important step along the way to robust shape-from-shading

Constant albedo assumption must eventually be addressed

Cast shadows and occlusion must eventually be addressed

My questions:

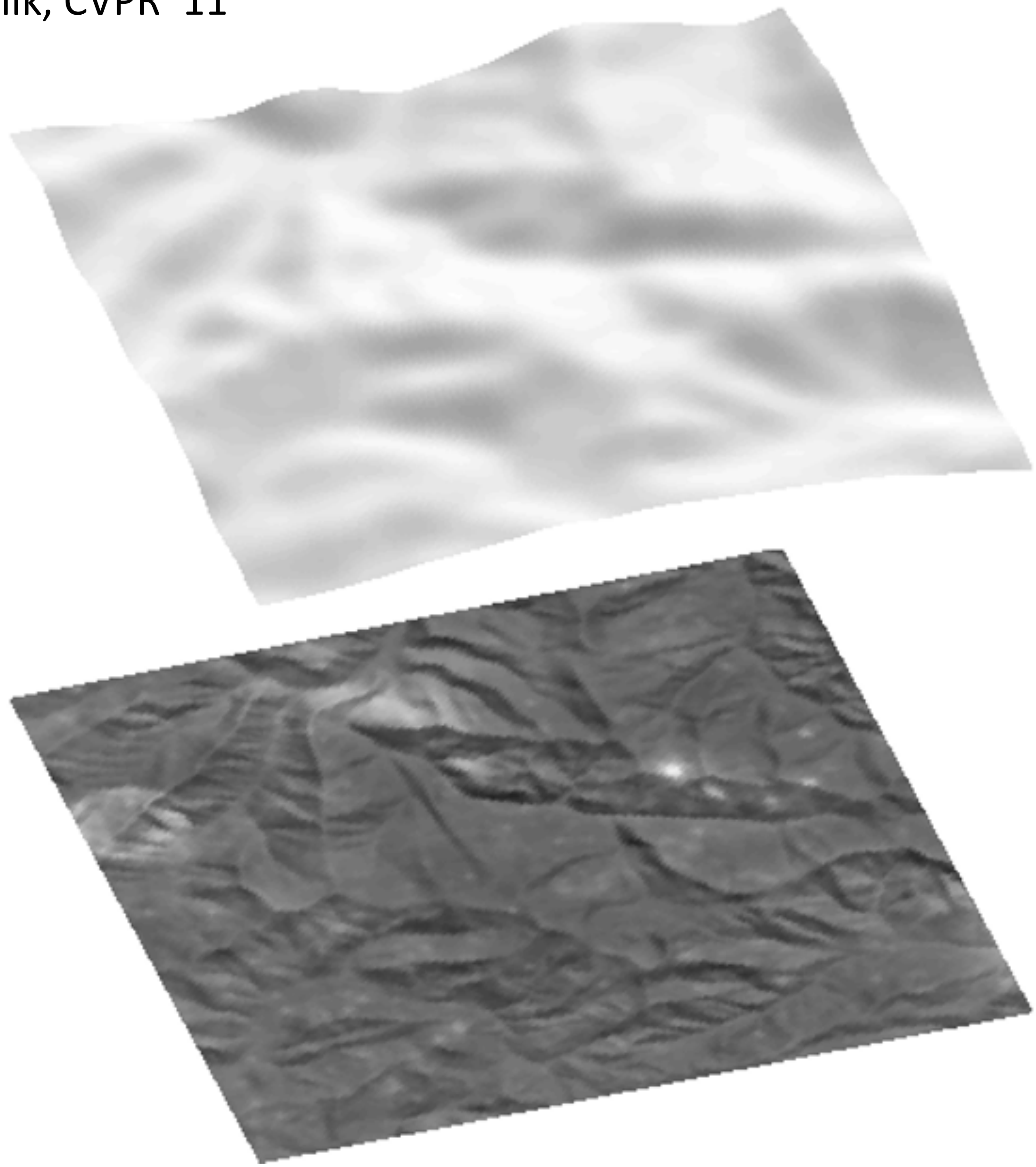
How good is their algorithm under uniform lighting?

Where do the environment maps come from? Natural?

There are self-occlusions in the blobs data set … how badly does the algorithm fail on them?

So far we have assumed that the surface is uniform in its photometric properties. Any non-uniformity will cause this algorithm to determine an incorrect shape. This is one of the uses of facial make-up; by darkening certain slopes they can be made to appear steeper for example.

Barron and Malik, CVPR '11



Sources

Flickr users Shawn Clover, Justin Frisch

http://www.flickr.com/photos/shawnclover/5382492085/

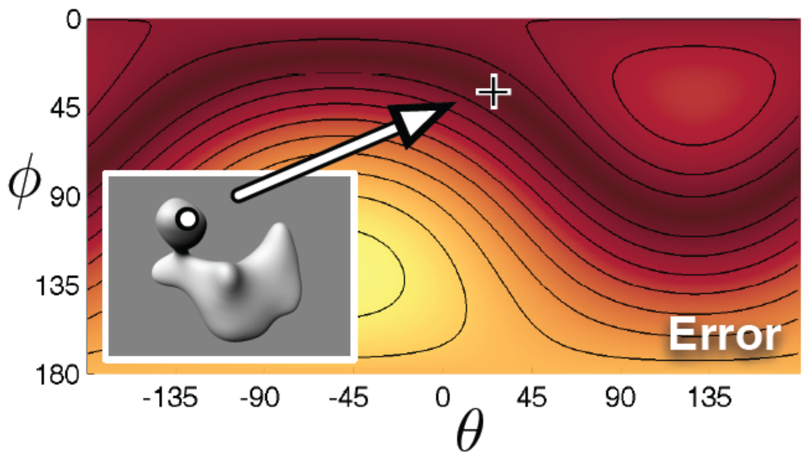
http://www.flickr.com/photos/shawnclover/5386000348

http://www.flickr.com/photos/purplegecko/2158408860/

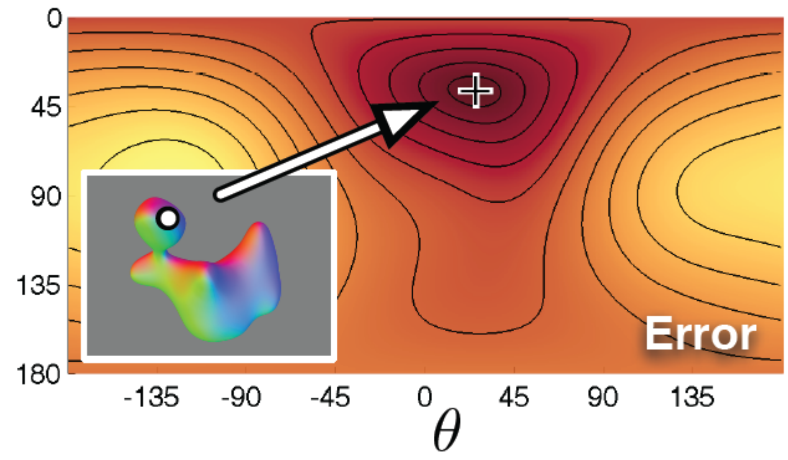
NASA/JPL care of Wikipedia

http://en.wikipedia.org/wiki/
File:Opportunity_photo_of_Mars_outcrop_rock.jpg

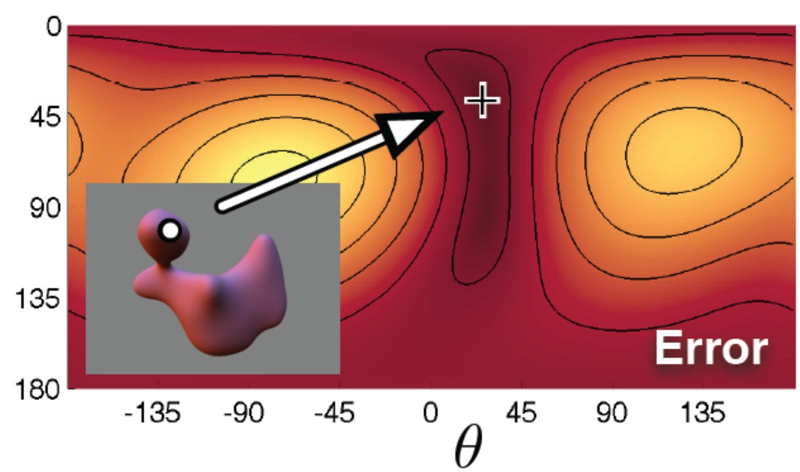
Thank you



(a) Point light source



(b) RGB photometric stereo



(c) Natural illumination

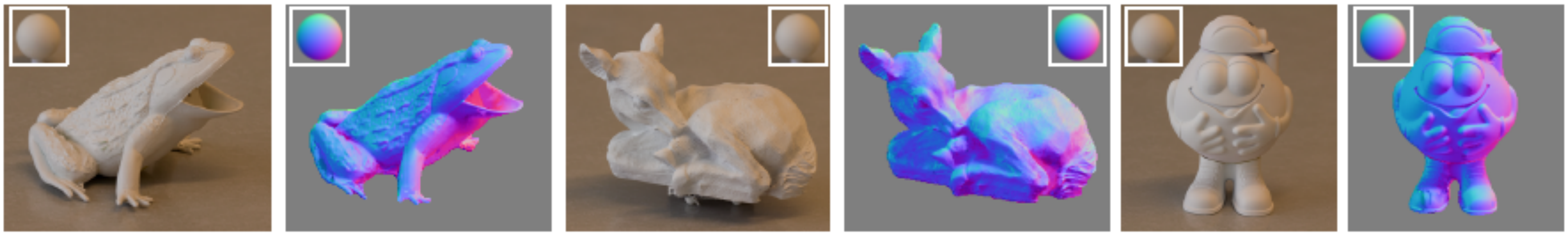
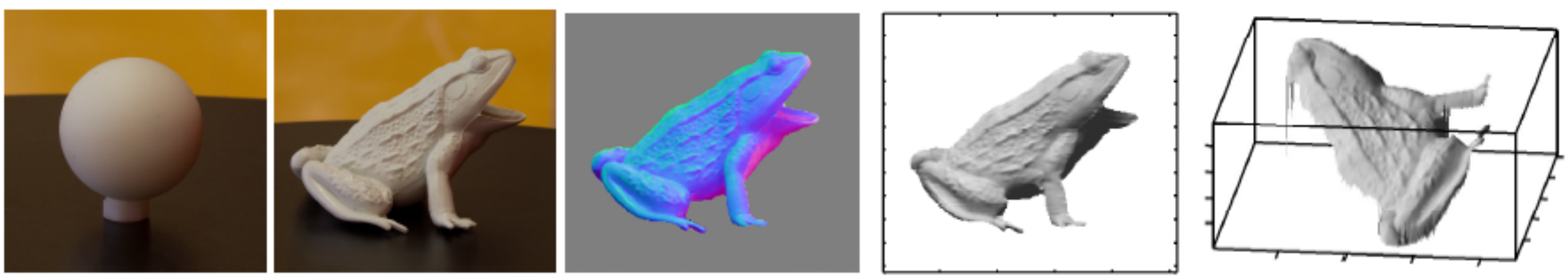


image normal map image normal map image normal map



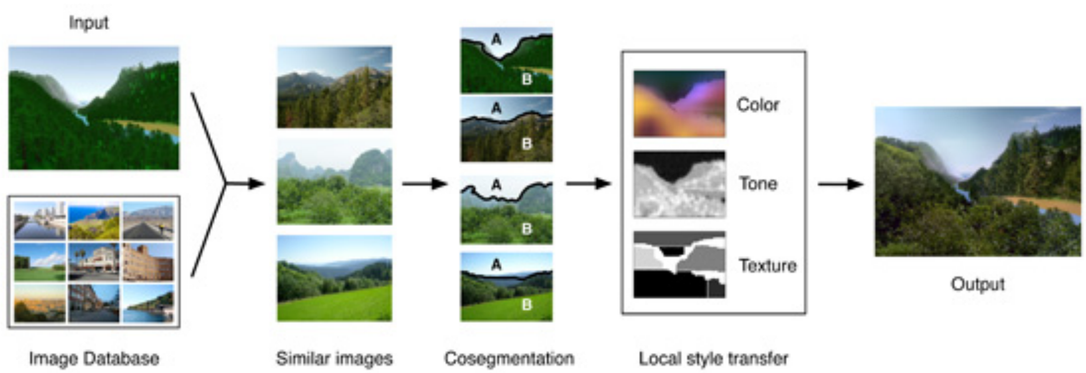
calibration sphere image normal map view 1 view 2

Micah Kimo Johnson
Computer Science and Artificial
Intelligence Laboratory
Massachusetts Institute of Technology

MIT 2008–2010
Postdoc, Brain & Cognitive Sciences
Advisor: Edward H. Adelson
Dartmouth 2002–2007
Ph.D. Computer Science
Lighting and Optical Tools for Image Forensics
Advisor: Hany Farid

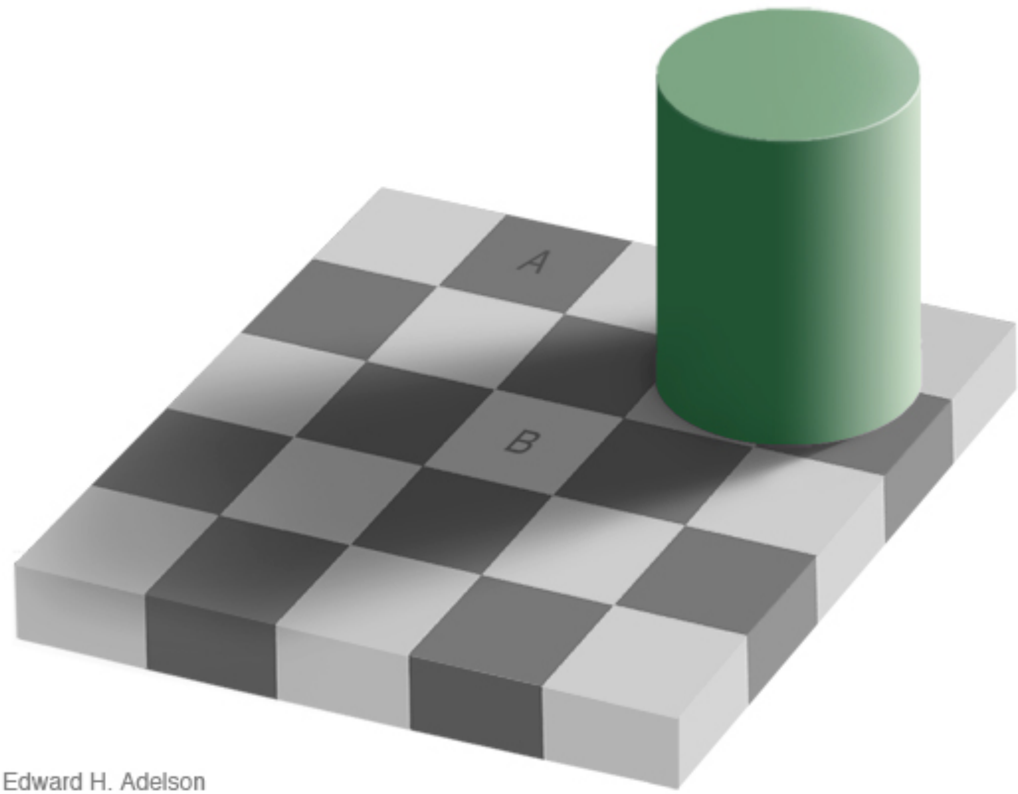
GelSight

CG2Real



Edward “Ted” Adelson
John and Dorothy Wilson Professor of Vision Science
Dept. of Brain and Cognitive Sciences
Massachusetts Institute of Technology

Education
1974-1979. Ph. D. in Experimental Psychology, University of Michigan
Dissertation: The response of the rod system to bright flashes of light



Edward H. Adelson

Figure 3: Frames 0, 15 and 30, of MPEG flower garden sequence.

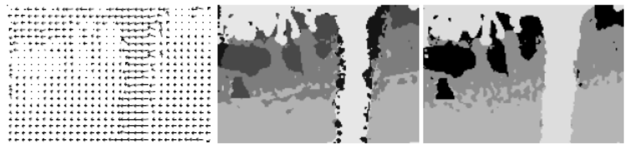


Figure 4: Affine motion segmentation of optic flow.

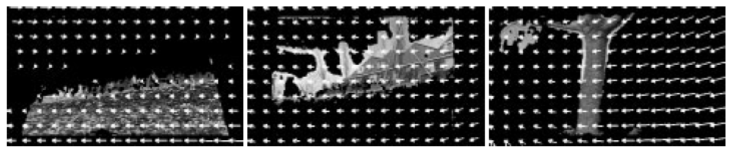


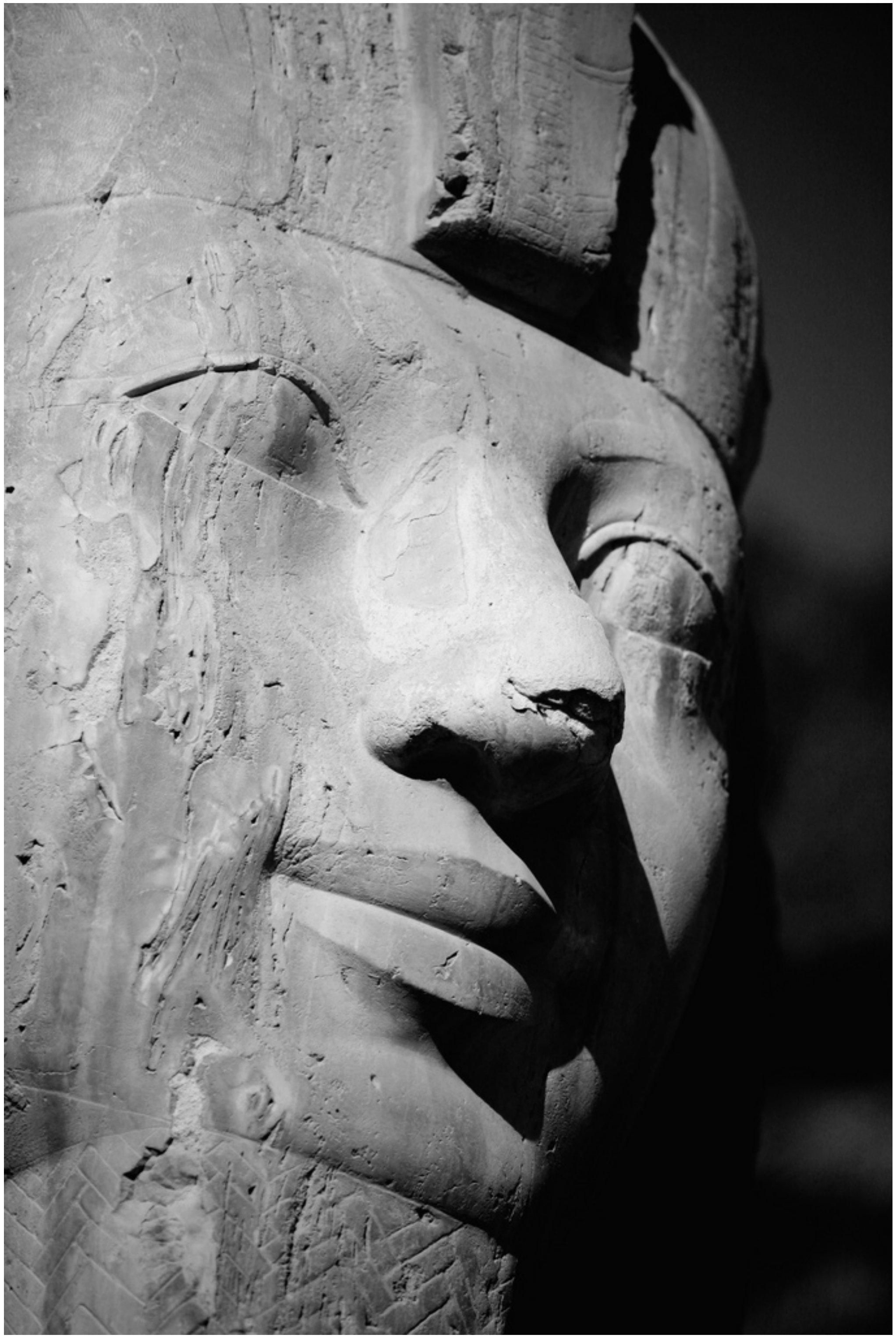
Figure 5: Images of the flower bed, houses, and tree. Affine motion fields are also shown here.

Figure 6: Corresponding frames of Figure 3 synthesized from layer images in Figure 5.

Figure 7: Corresponding frames of Figure 3 synthesized without the tree layer.

"Layered Representation for Motion Analysis"
Wang and Adelson, *CVPR* '93


ANNO DOMINI MDCCXXXV PONTIF VI
PERFECIT
BENEDICTVS XIV
ON·MAX
POSITIS SIGNIS ET ANA
CLEMENTIS XIII PONT MAX




new
When it doesn't fit in your trunk, we have delivery options that fit your budget.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Massachusetts Institute of Technology Project MAC	2a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED 2b. GROUP None
3. REPORT TITLE Shape from Shading; A Method for Obtaining the Shape of a Smooth Opaque Object From One View	
4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)* Ph.D. Thesis, Department of Electrical Engineering, June 1970	
5. AUTHOR(S) *(Last name, first name, initial)* Horn, Berthold K. P.	
6. REPORT DATE November 1970	7a. TOTAL NO. OF PAGES 198 7b. NO. OF REFS 14
8a. CONTRACT OR GRANT NO. Nonr-4102(02) b. PROJECT NO. c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) MAC TR-79 (THESIS) 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*
10. AVAILABILITY/LIMITATION NOTICES Distribution of this document is unlimited.	
11. SUPPLEMENTARY NOTES None	12. SPONSORING MILITARY ACTIVITY Advanced Research Projects Agency 3D-200 Pentagon

A number of interesting applications of this method can be mentioned. The first of these concerns the scanning electron microscope (SEM) which produces images which are particularly easy to interpret, since the intensity recorded is a function of the slope of the object at that point and is thus a form of shading (as opposed to optical and transmission electron

Another important application lies in the determination of lunar topography. Here the special reflectivity function of the material in the maria of the moon allows a very great simplification of the equations used in the shape-from-shading algorithm. The equations in fact reduce to one integral which has to be evaluated along each of a family of predetermined straight lines in the image, making for high accuracy. This problem was first tackled for areas near the

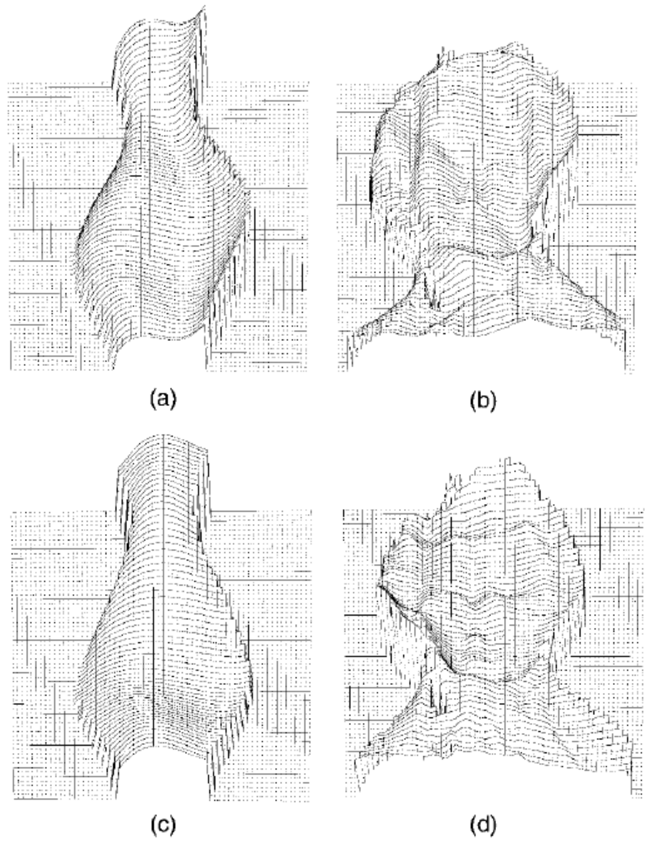


Fig. 11. Results for Pentland's method on synthetic images: (a) vase, (b) Mozart. (a1) and (b1) show the results for test images with light source $(0, 0, 1)$. (a2) and (b2) show the results for test images with light source $(1, 0, 1)$.

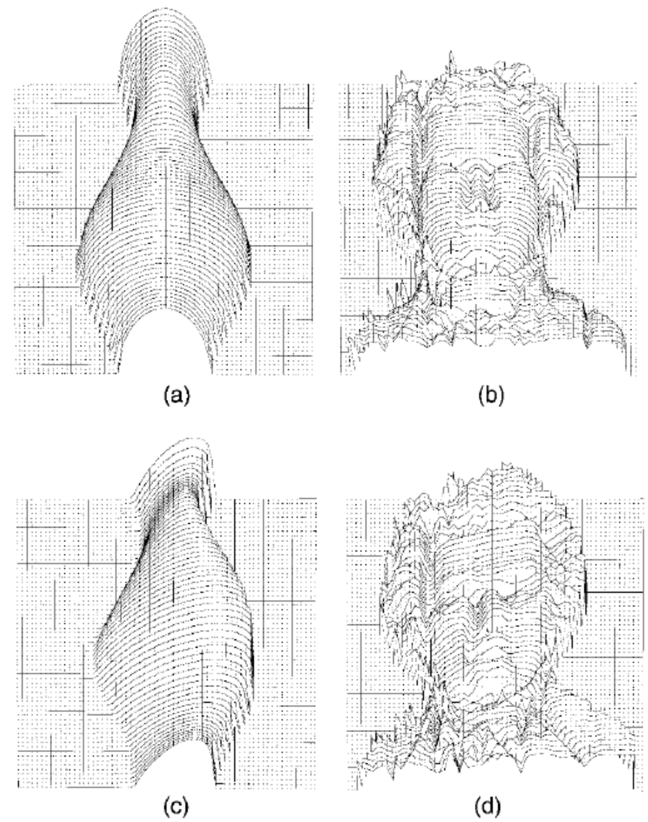
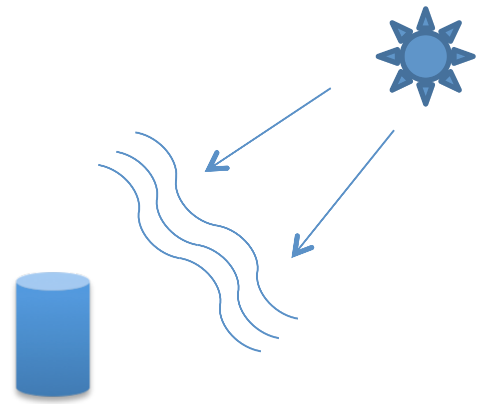


Fig. 12. Results for Tsai and Shah's method on synthetic images: (a) vase, (b) Mozart. (a1) and (b1) show the results for test images with light source $(0, 0, 1)$. (a2) and (b2) show the results for test images with light source $(1, 0, 1)$.

Zhang, Tsai, Cryer, and Shah, *PAMI* '99

Classical Shape-from-Shading

Horn's assumptions:



- light source is far away
- reflectance and albedo are constant at all points on the surface
- no shadows or interreflections

Classical Shape-from-Shading

- Under these assumptions, there is a simple model of image intensity:

$$I(x) = \mathbf{b} \cdot \mathbf{n}(x)$$

$\mathbf{b}$: lighting direction

$\mathbf{n}(x)$: surface normal at position x

Shading function modeled as quadratic:

$$s(\mathbf{n}) = \mathbf{n} \cdot A\mathbf{n} + \mathbf{b} \cdot \mathbf{n} + c$$

Given an intensity measurement $I(\mathrm{x})$ at position x, a SFS algorithm needs to recover the surface normal (or height) at position x. With only local information, the problem can be modeled as the minimization of an error function:

$$E(\mathbf{n}) = \|f(\mathbf{n})\|^2 = \|s(\mathbf{n}) - I(\mathrm{x})\|^2 . \tag{3}$$

In general, photometric stereo techniques use L lighting conditions rather than one. Mathematically, the shading function s can be represented as a vector-valued function of the surface normal:

$$\mathrm{s}(\mathbf{n}) = \begin{bmatrix} s_1(\mathbf{n}) \\ \vdots \\ s_L(\mathbf{n}) \end{bmatrix} = \begin{bmatrix} \mathbf{n}^T A_1 \mathbf{n} + \mathrm{b}_1^T \mathbf{n} + c_1 \\ \vdots \\ \mathbf{n}^T A_L \mathbf{n} + \mathrm{b}_L^T \mathbf{n} + c_L \end{bmatrix} . \tag{4}$$