Lecture 17: Photometric Stereo

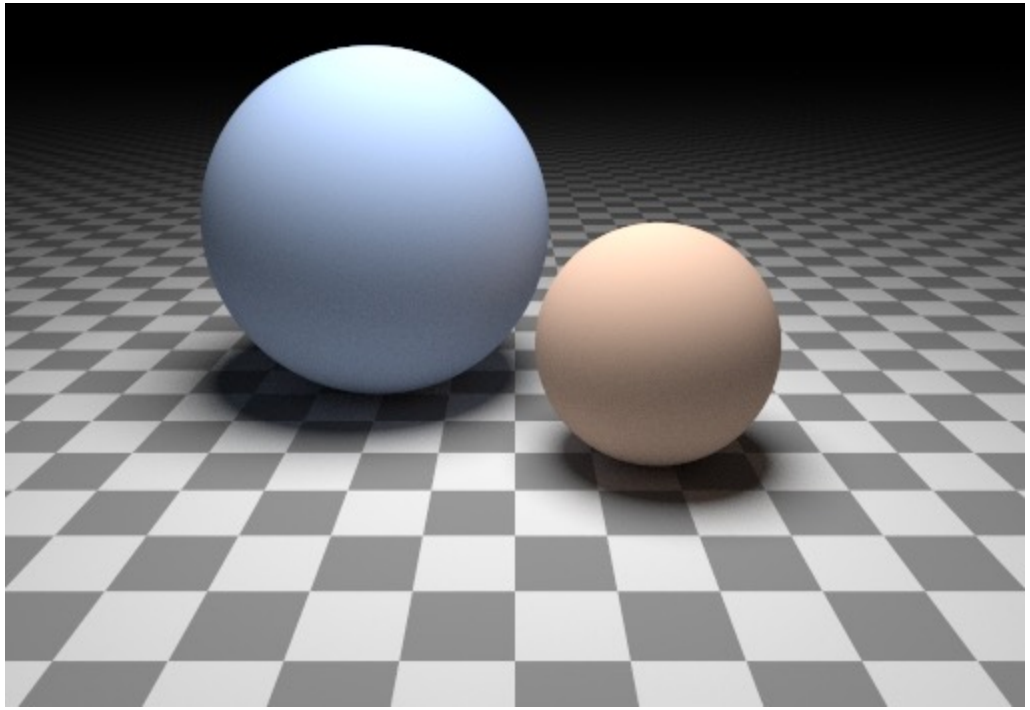
COMP 590/776: Computer Vision

Instructor: Soumyadip (Roni) Sengupta

TA: Mykhailo (Misha) Shvets

Course Website:
Scan Me!

Can we determine shape from lighting?



- Are these spheres?
  - Or just flat discs painted with varying color (albedo)?
  - There is ambiguity between *shading* and *reflectance*
  - But still, as humans we can understand the shapes of these objects

What we know: Stereo

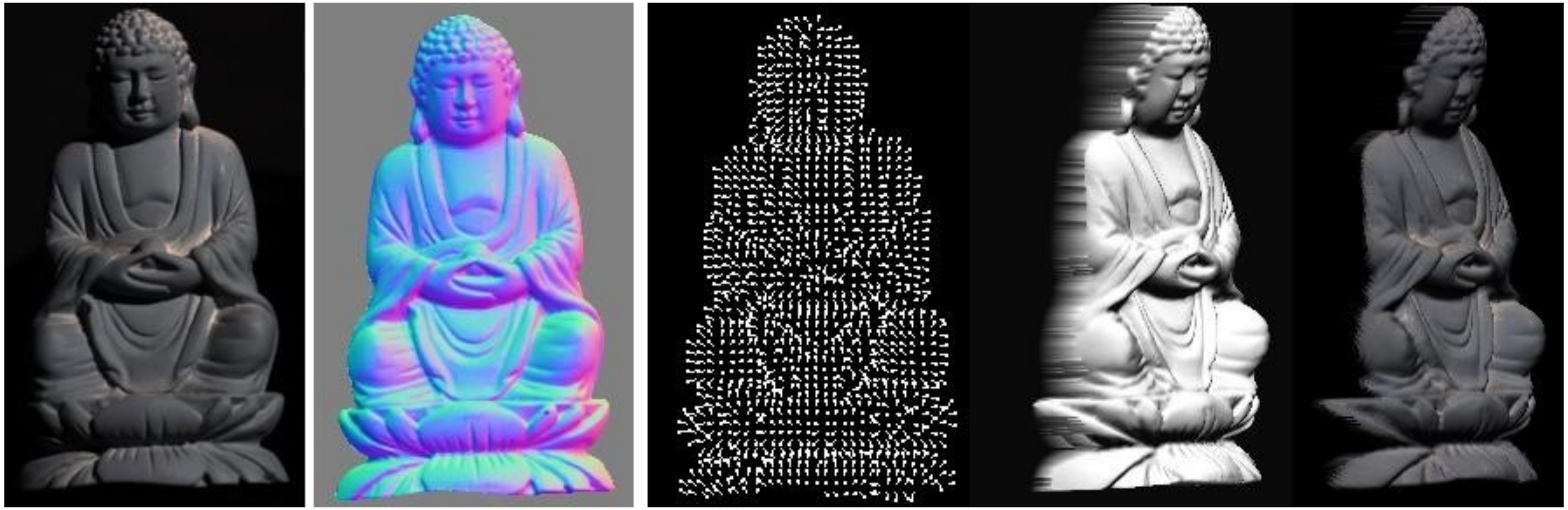
Key Idea: use camera motion to compute shape

Next: Photometric Stereo

Key Idea: use pixel brightness to understand shape

Photometric Stereo

What results can you get?



Input (1 of 12) | Normals (RGB colormap) | Normals (vectors) | Shaded 3D rendering | Textured 3D rendering

Today’s class

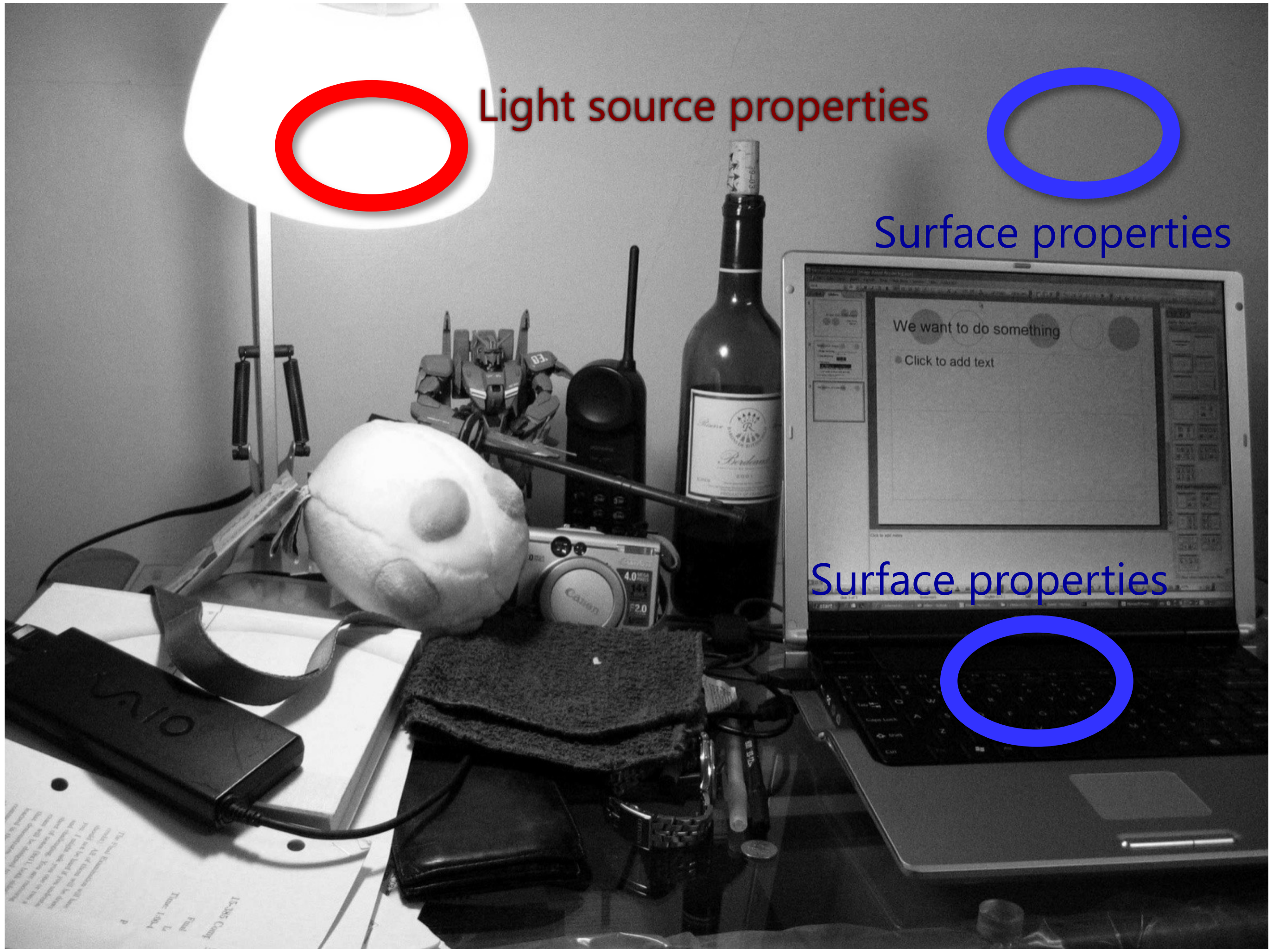
- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.

Today’s class

- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.

Radiometry

- What determines the brightness of a pixel?



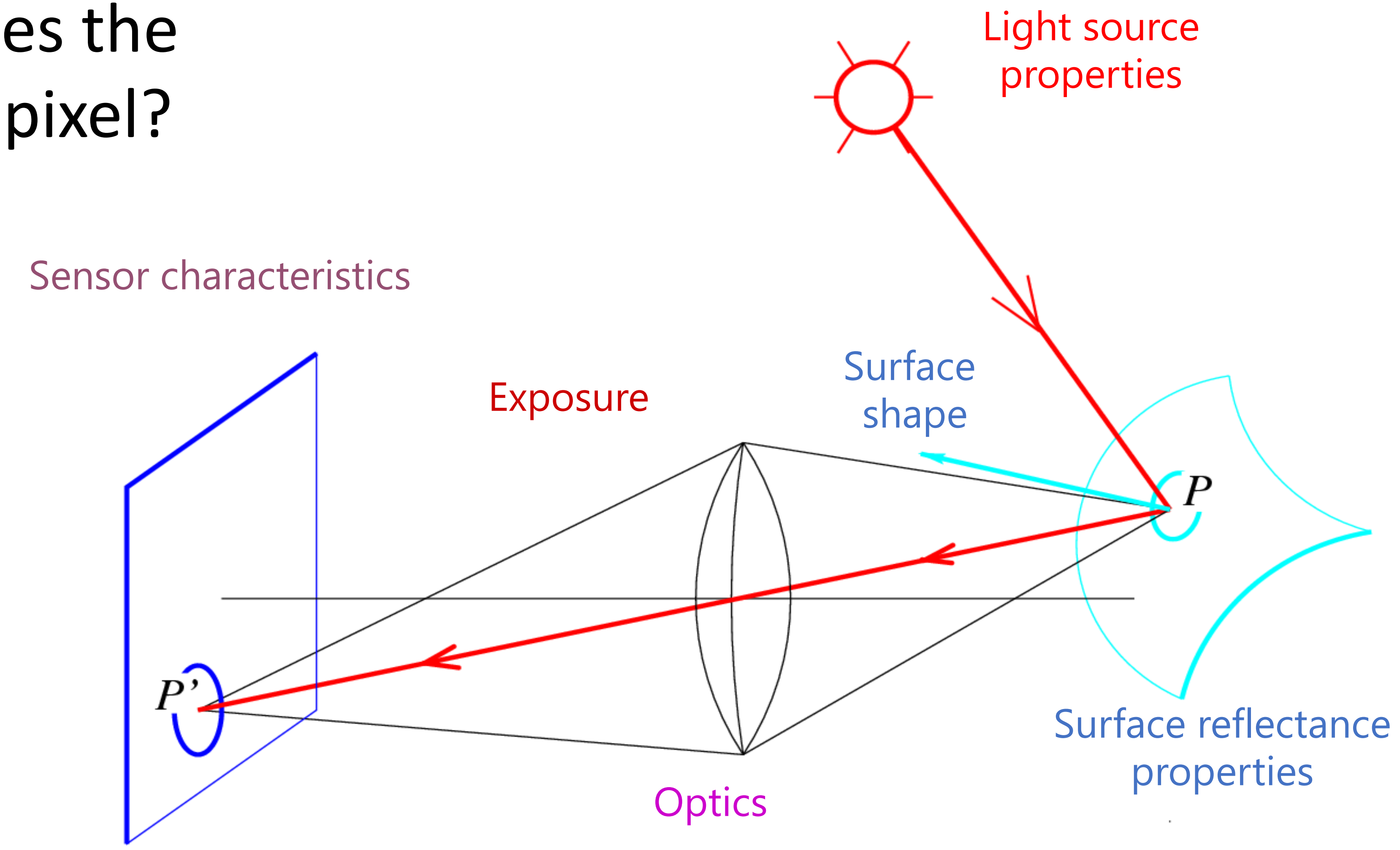
Radiometry

- What determines the brightness of a pixel?

[@robertwestonbreshears](https://www.instagram.com/p/BtgX55ZBhU-/)
https://www.instagram.com/p/BtgX55ZBhU-/

Radiometry

- What determines the brightness of a pixel?

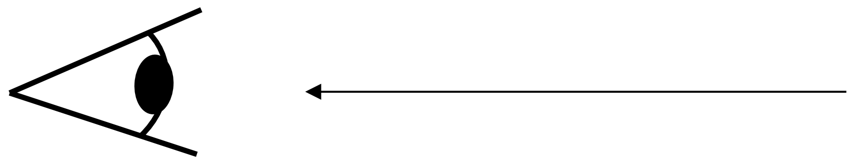


Slide by L. Fei-Fei

What is light?

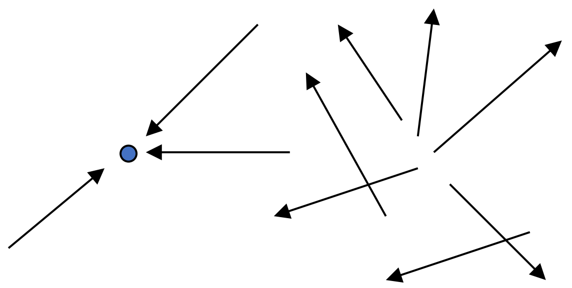
Electromagnetic radiation (EMR) moving along rays in space

- R(λ) is EMR, measured in units of power (watts)
  - λ is wavelength



Light field

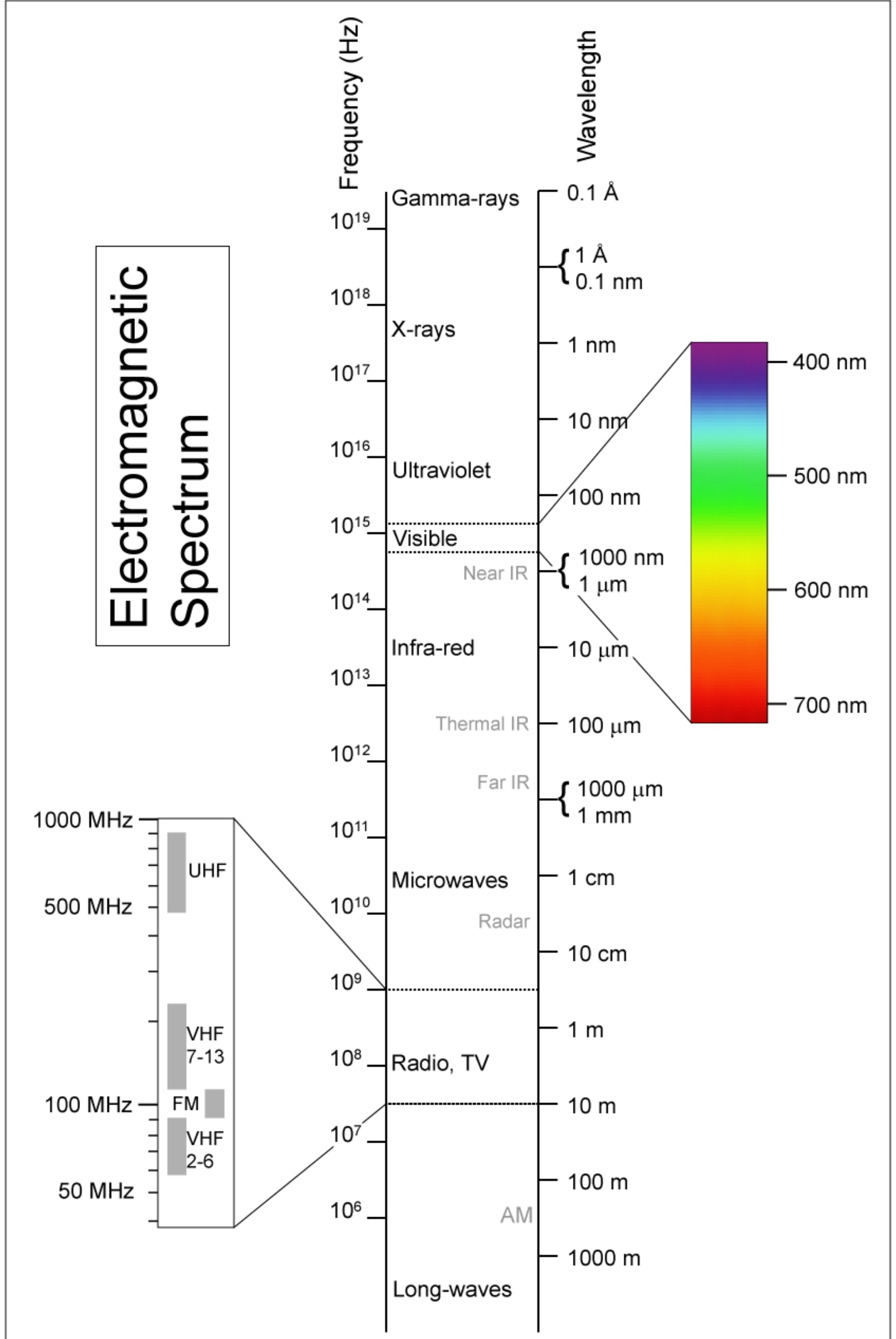
- We can describe all of the light in the scene by specifying the radiation (or **"radiance"** along all light rays) arriving at every point in space and from every direction



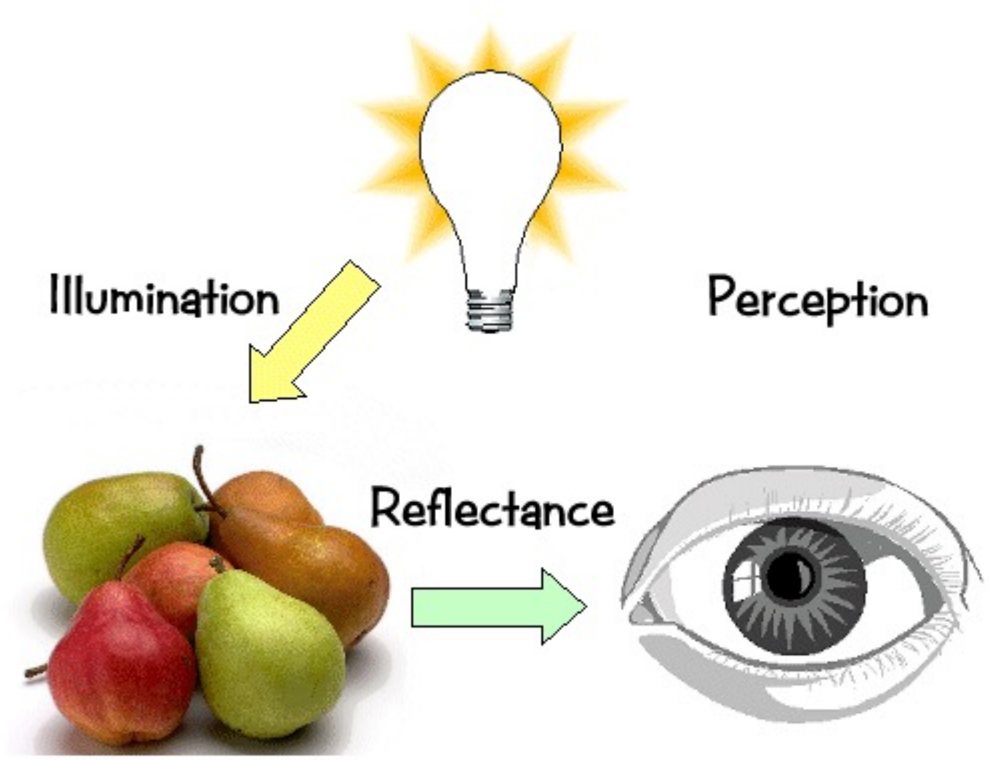
The ***plenoptic function*** describes all of this light: $R(X, Y, Z, \theta, \phi, \lambda, t)$

Visible light

We "see" electromagnetic radiation in a range of wavelengths



Light transport



Light sources

- Basic types
  - point source
  - directional source
    - a point source that is infinitely far away
  - area source
    - a union of point sources
- More generally
  - a light field can describe *any* distribution of light sources
  - Environment map
- What happens when light hits an object?

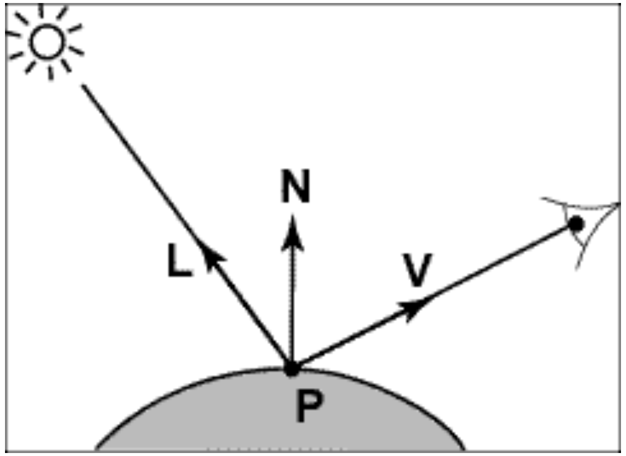
Today's class

- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in 'deep learning era'.

Modeling Image Formation

We need to reason about:

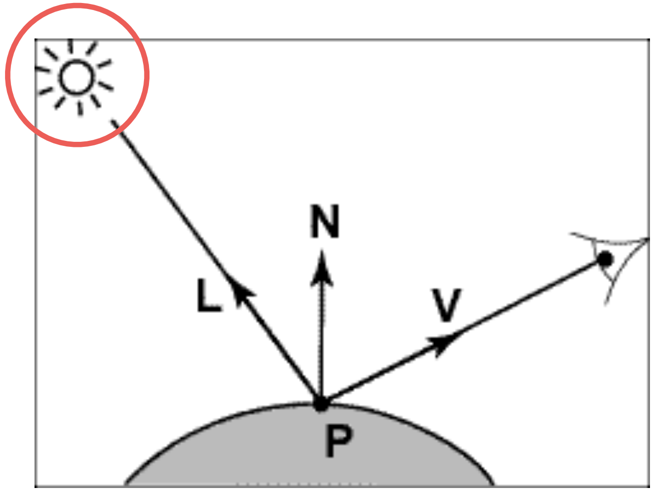
- How light interacts with the scene
- How a pixel value is related to light energy in the world



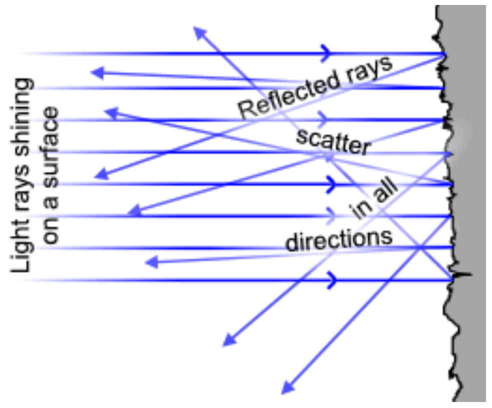
Track a “ray” of light all the way from light source to the sensor

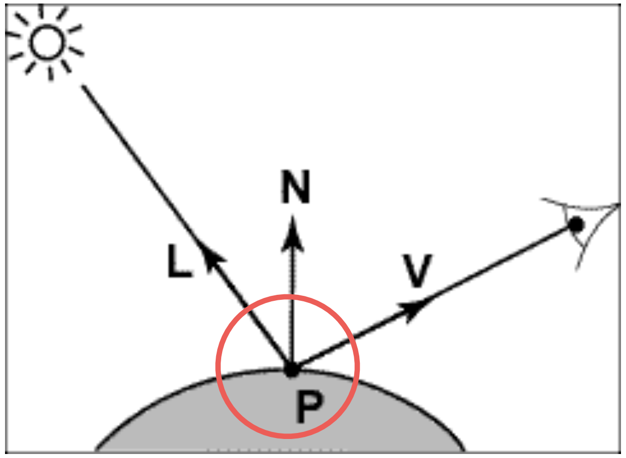
Directional Lighting

- Key property: all rays are parallel
- Equivalent to an infinitely distant point source



Lambertian Reflectance



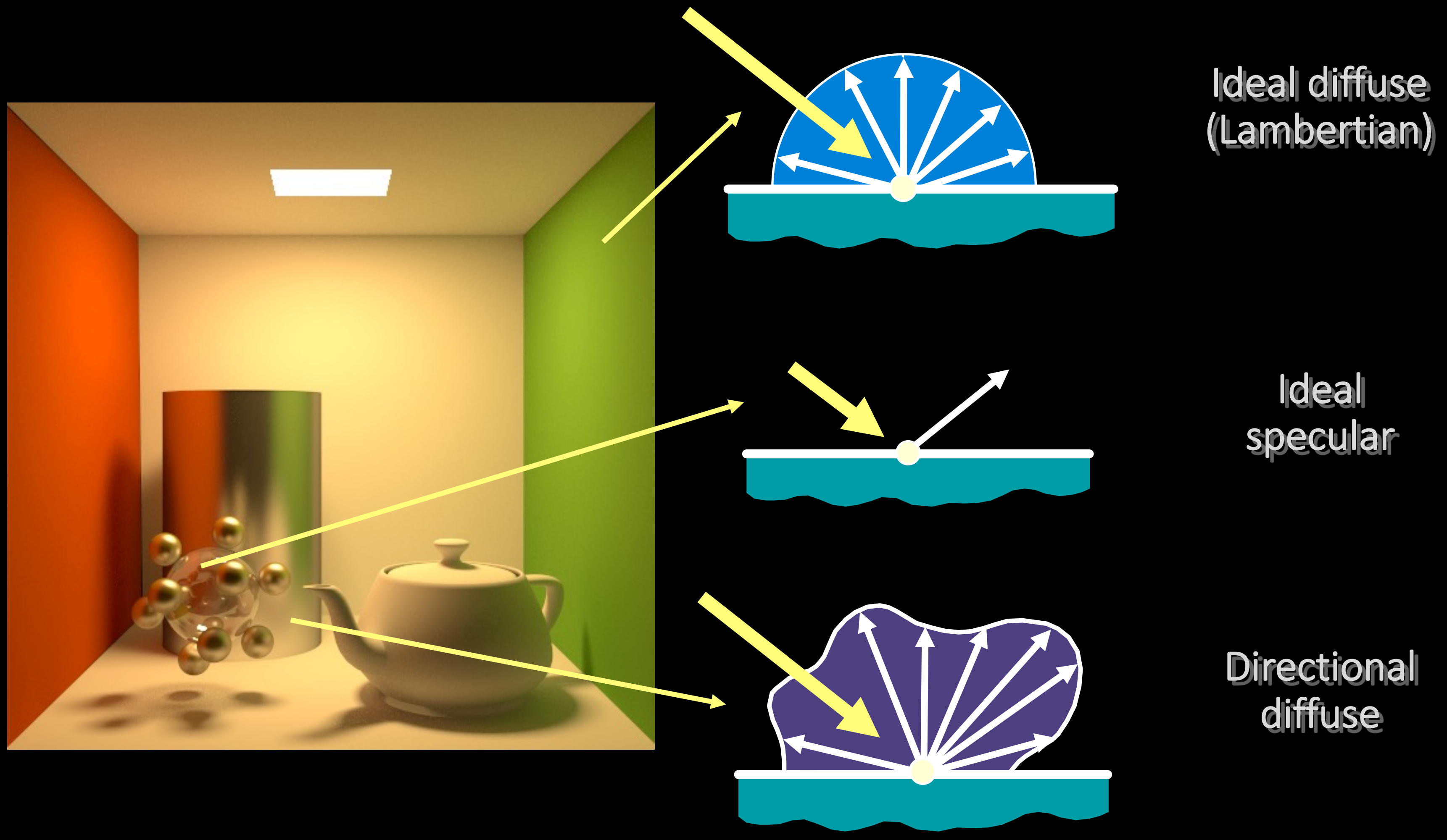


$$I = N \cdot L$$

Image intensity = Surface normal • Light direction

Image intensity $\propto$ cos(angle between N and L)

Materials - Three Forms

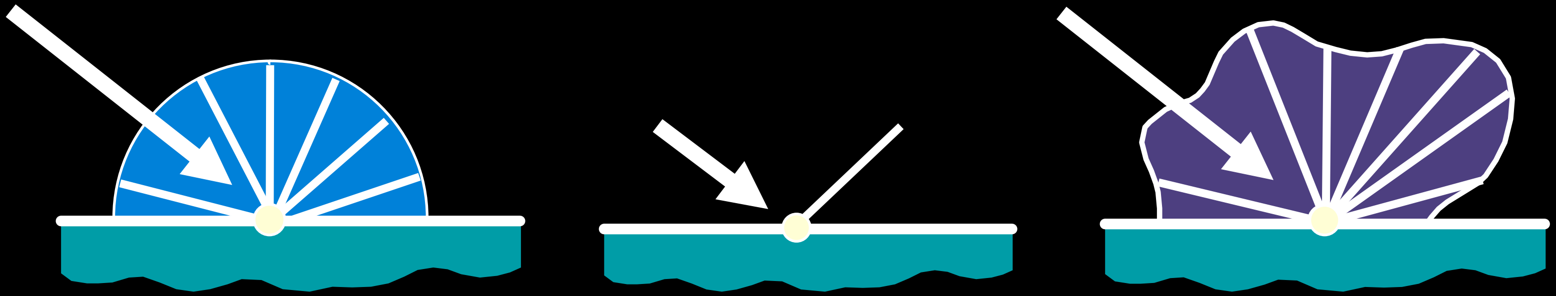


Reflec

Ideal diffuse (Lambertian)

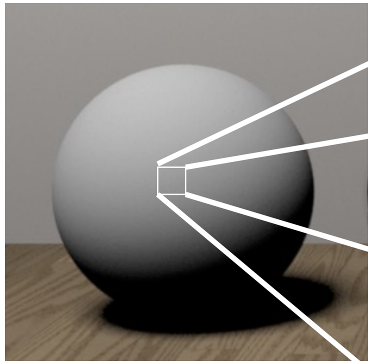
Ideal specular

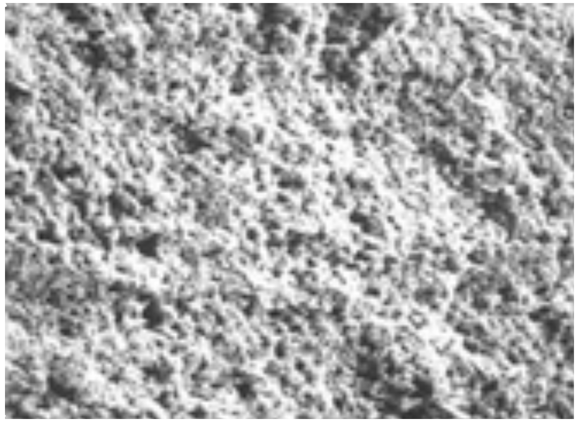
Directional diffuse



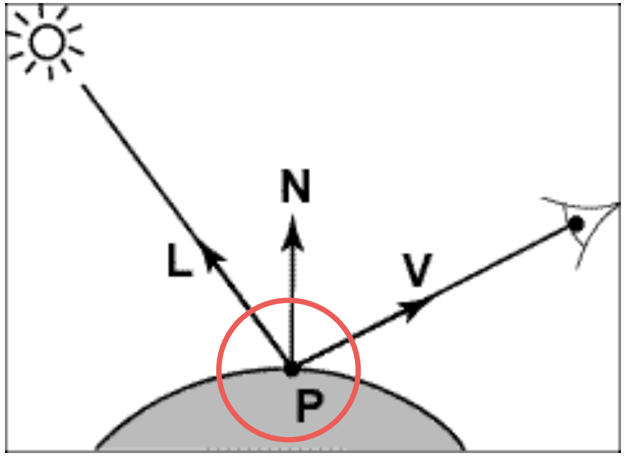
Ideal Diffuse Reflection

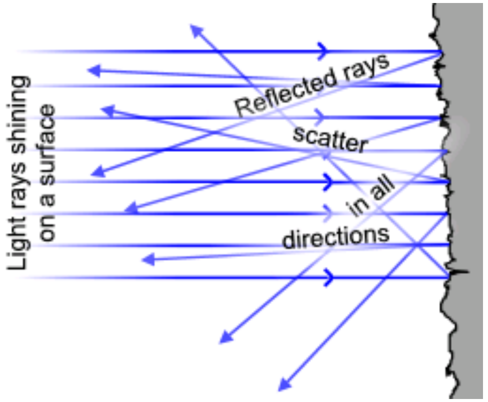
- Characteristic of multiple scattering materials
- An idealization but reasonable for matte surfaces





Lambertian Reflectance



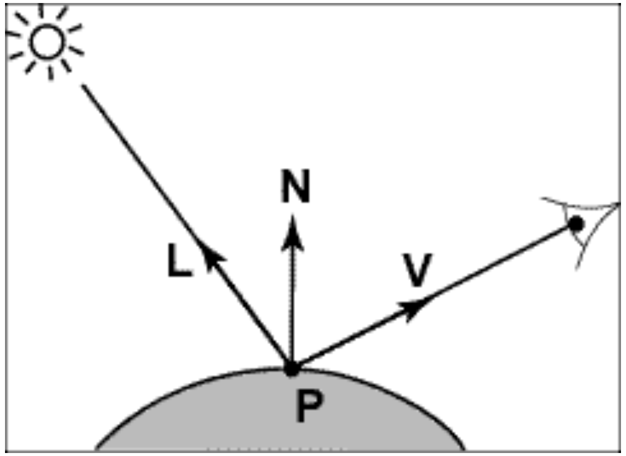


$$I = N \cdot L$$

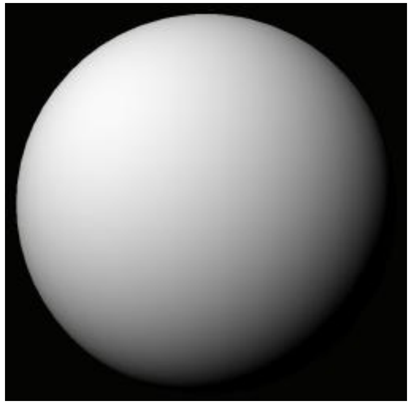
1. Reflected energy is proportional to cosine of angle between L and N **(incoming)**

2. Measured intensity is viewpoint-independent **(outgoing)**

Final Lambertian image formation model



$$I = k_d \mathbf{N} \cdot \mathbf{L}$$

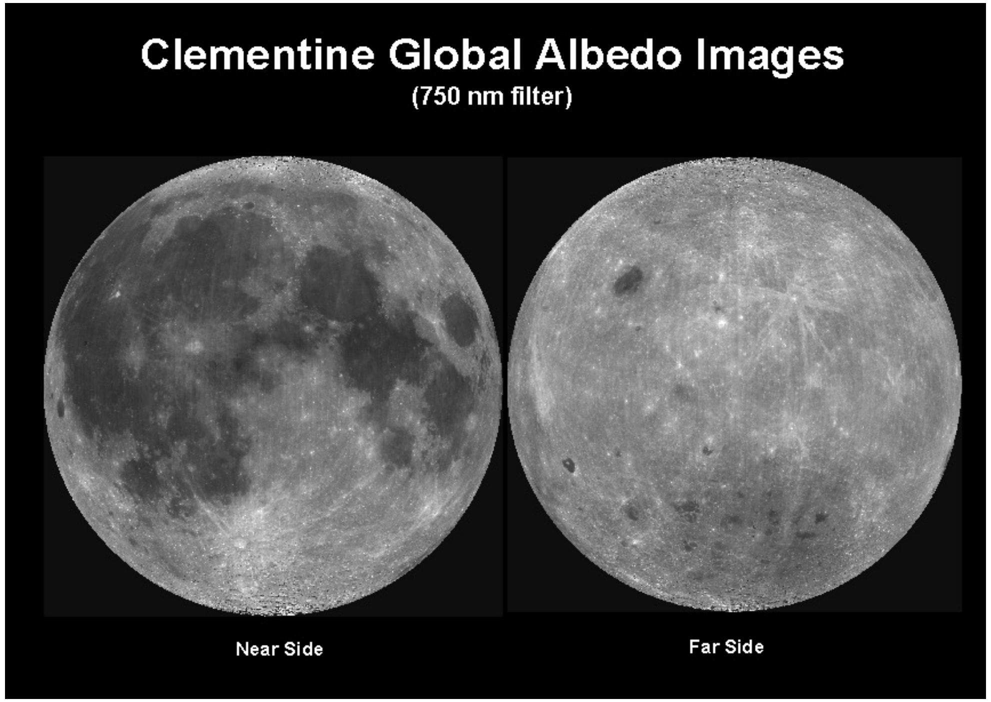


1. Diffuse **albedo**: what fraction of incoming light is reflected?
   - Introduce scale factor k_d
2. Light intensity: how much light is arriving?
   - Compensate with camera exposure (global scale factor)
3. Camera response function
   - Assume pixel value is linearly proportional to incoming energy (perform radiometric calibration if not)

Albedo

Sample albedos

Surface	Typical albedo
Fresh asphalt	0.04[4]
Open ocean	0.06[5]
Worn asphalt	0.12[4]
Conifer forest (Summer)	0.08,[6] 0.09 to 0.15[7]
Deciduous trees	0.15 to 0.18[7]
Bare soil	0.17[8]
Green grass	0.25[8]
Desert sand	0.40[9]
New concrete	0.55[8]
Ocean ice	0.5–0.7[8]
Fresh snow	0.80–0.90[8]



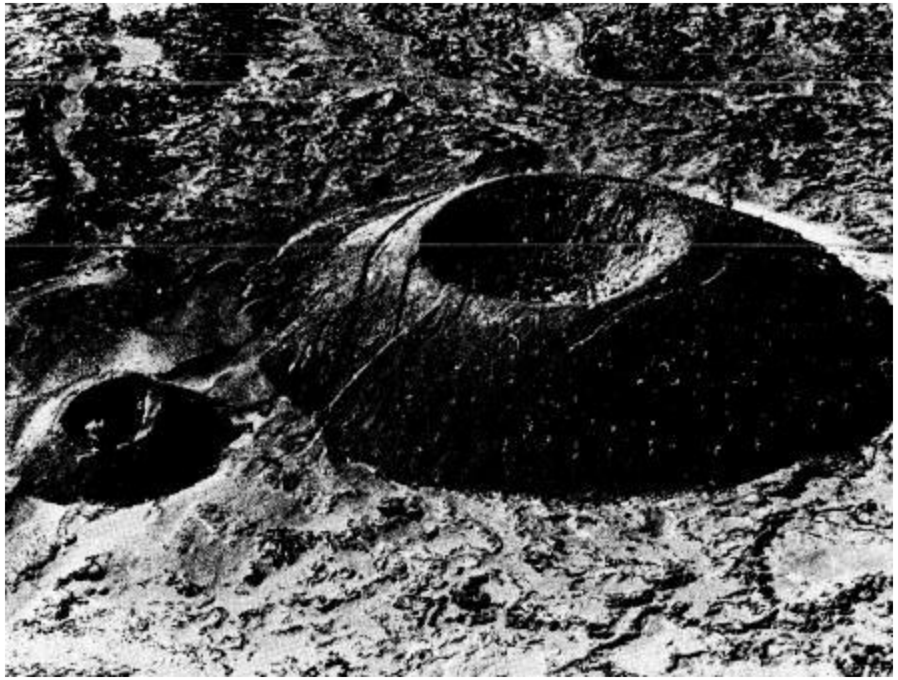
Objects can have varying albedo and albedo varies with wavelength

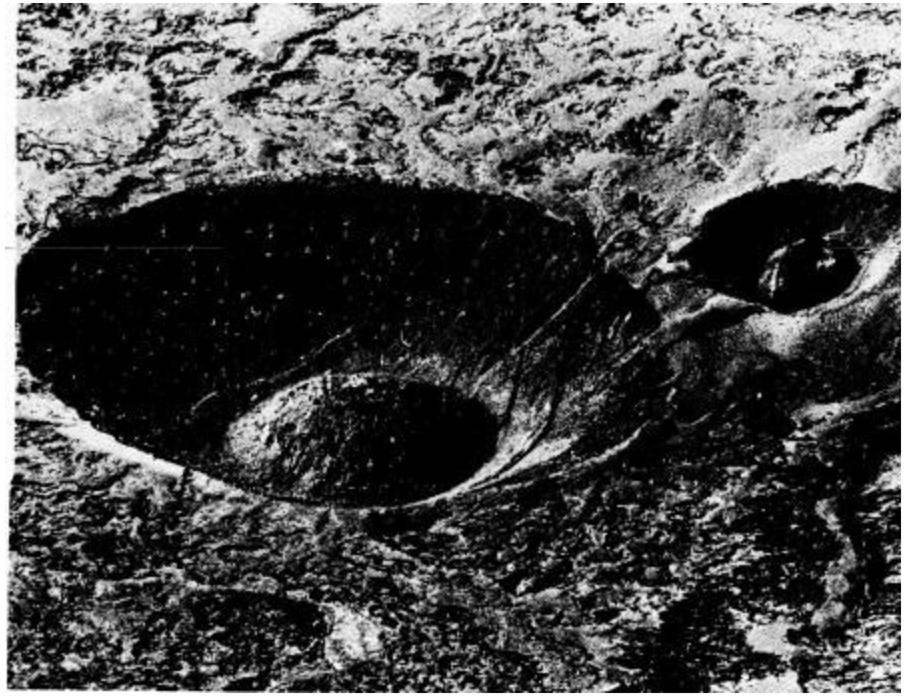
Source: https://en.wikipedia.org/wiki/Albedo

Today’s class

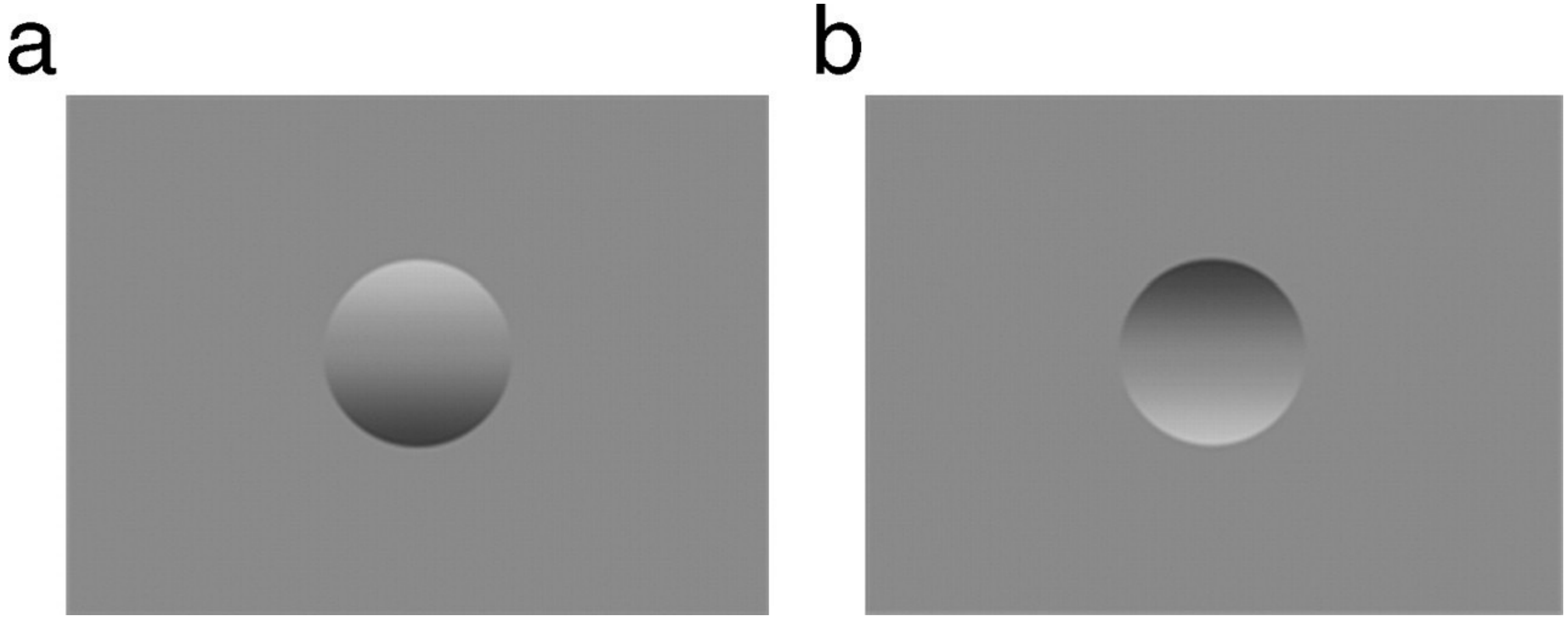
- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.

Human Perception





Examples of the classic bump/dent stimuli used to test lighting assumptions when judging shape from shading, with shading orientations (a) 0° and (b) 180° from the vertical.



Thomas R et al. J Vis 2010;10:6

journal of VISION

Human Perception

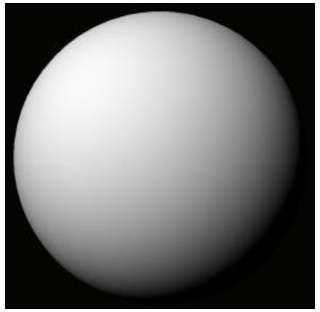
- Our brain often perceives shape from shading.
- Mostly, it makes many assumptions to do so.
- For example:

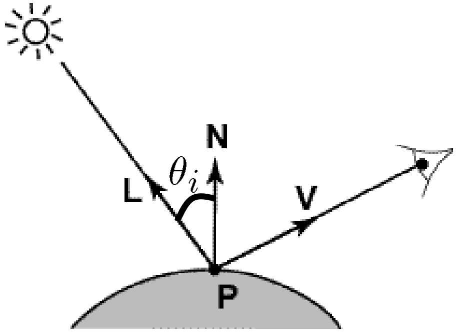
  Light is coming from above (sun).

  Biased by occluding contours.

by V. Ramachandran

A Single Image: Shape from shading





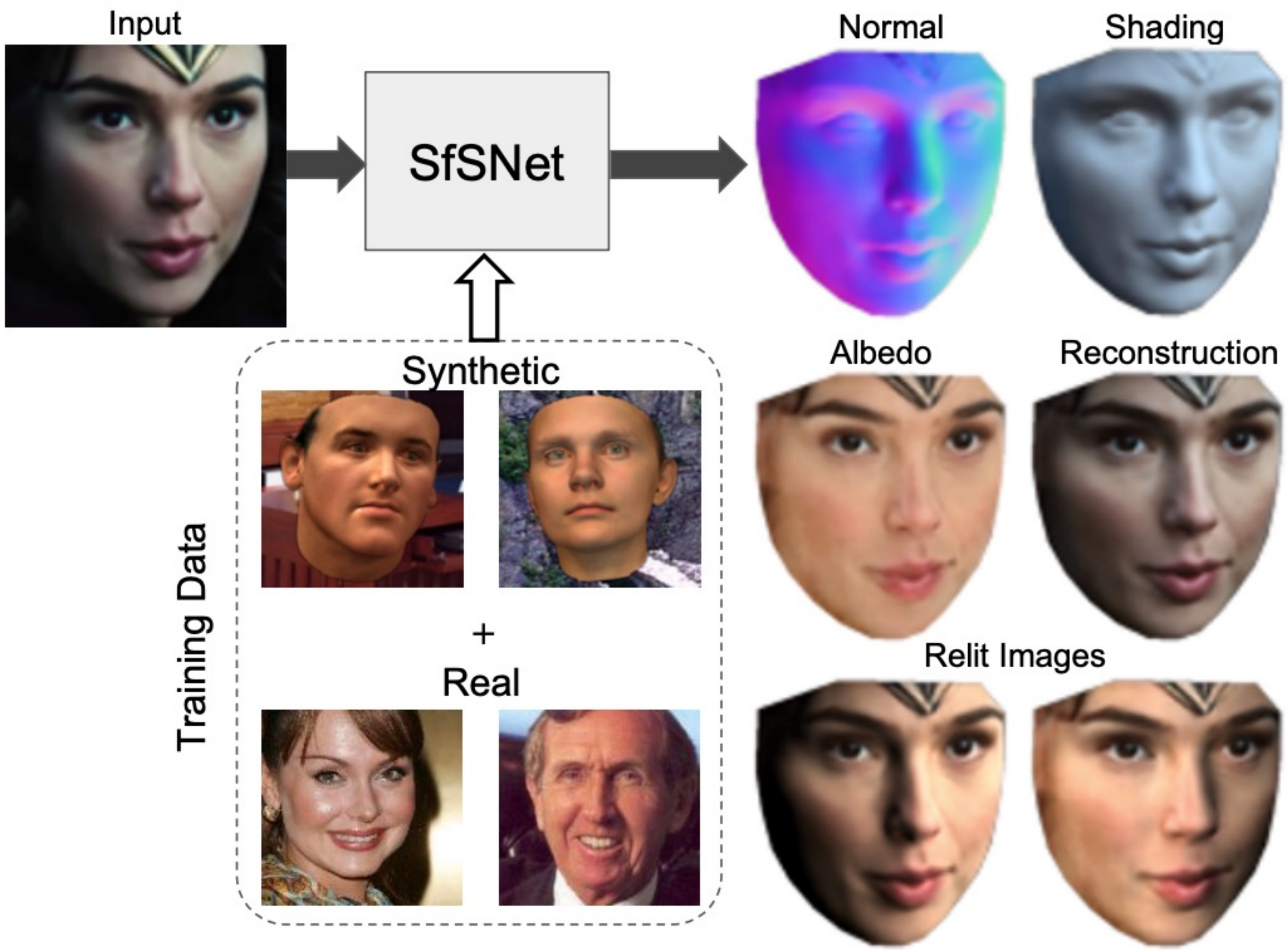
Suppose (for now) $k_d = 1$

$$\begin{aligned} I &= k_d \mathbf{N} \cdot \mathbf{L} \\ &= \mathbf{N} \cdot \mathbf{L} \\ &= cos\ \theta_i \end{aligned}$$

You can directly measure angle between normal and light source

- Not quite enough information to compute surface shape
- But can be if you add some additional info, for example
  - assume a few of the normals are known (e.g., along silhouette)
  - constraints on neighboring normals—"integrability"
  - smoothness
- Hard to get it to work well in practice
  - plus, how many real objects have constant albedo?
  - But, deep learning can help

Deep Learning for Shape from Shading



“SfSNet: Learning Shape, Reflectance and Illuminance of Faces in the Wild”, **Sengupta**, Kanazawa, Castillo, Jacobs, CVPR 2018.

InverseRenderNet: Learning single image inverse rendering

CVPR 2019

Ye Yu and William A. P. Smith
Department of Computer Science, University of York, UK
{yy1571,william.smith}@york.ac.uk

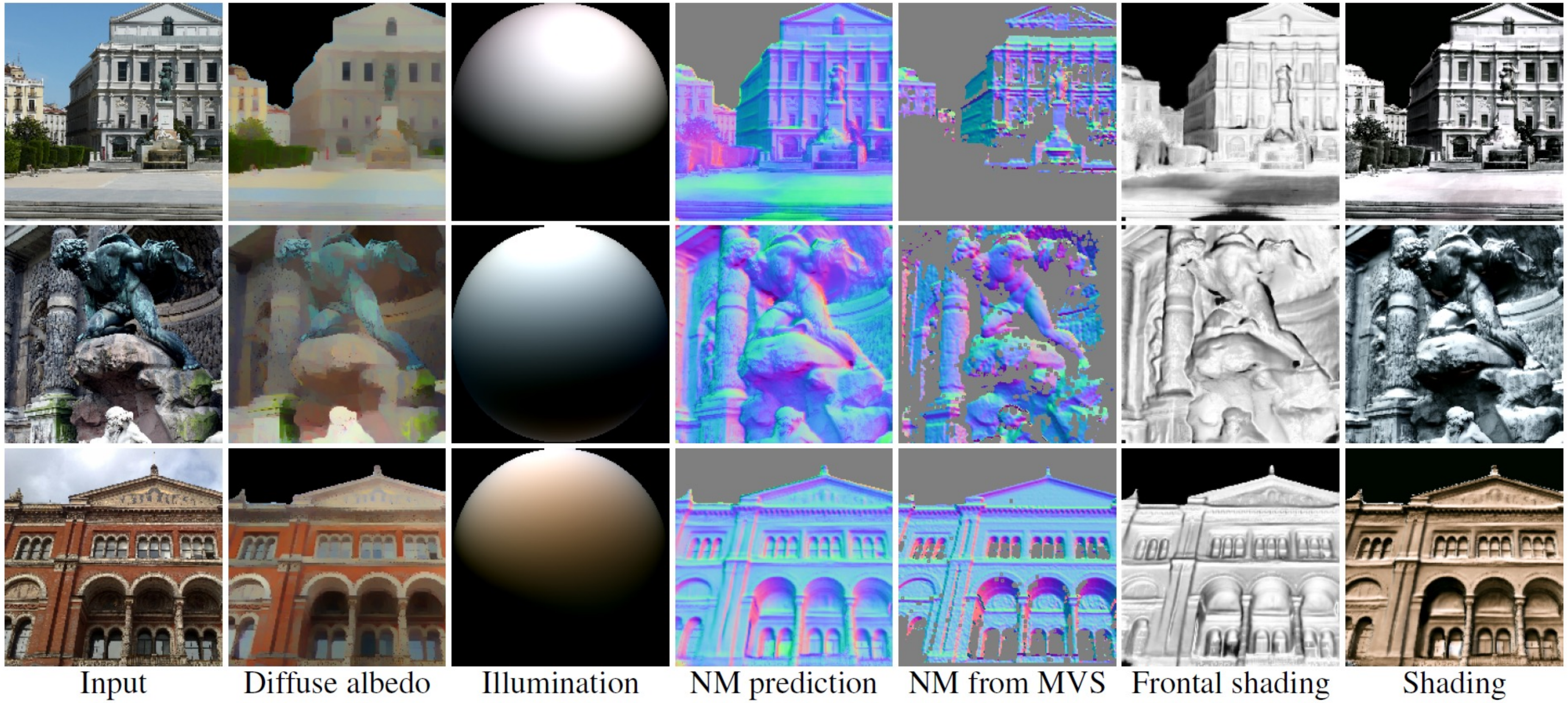


Figure 1: From a single image (col. 1), we estimate albedo and normal maps and illumination (col. 2-4); comparison multi-view stereo result from several hundred images (col. 5); re-rendering of our shape with frontal/estimated lighting (col. 6-7).

Application: Detecting composite photos

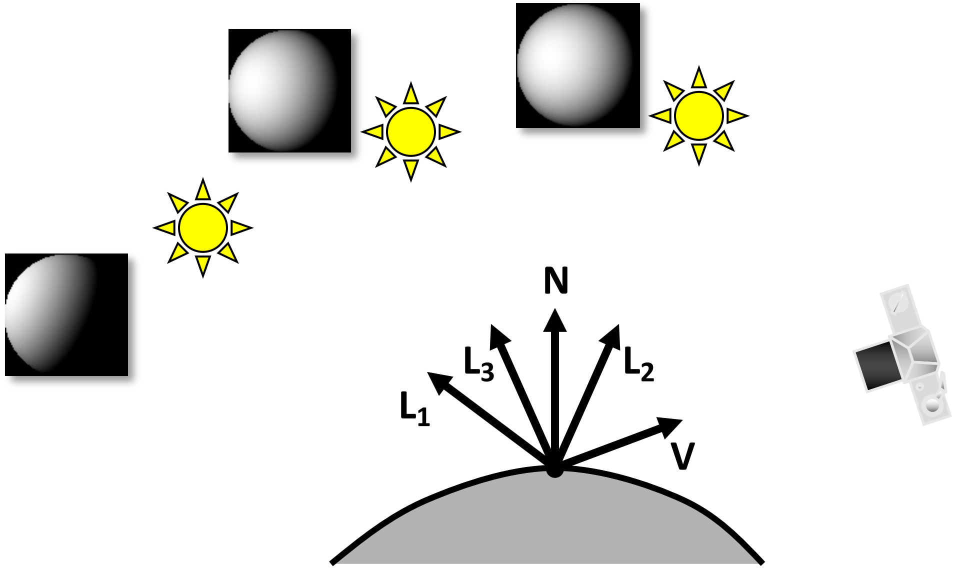
Fake photo

Real photo

Today’s class

- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.

Photometric stereo



$$I_1 = k_d \mathbf{N} \cdot \mathbf{L_1}$$
$$I_2 = k_d \mathbf{N} \cdot \mathbf{L_2}$$
$$I_3 = k_d \mathbf{N} \cdot \mathbf{L_3}$$

Can write this as a matrix equation:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = k_d \begin{bmatrix} \mathbf{L_1}^T \\ \mathbf{L_2}^T \\ \mathbf{L_3}^T \end{bmatrix} \mathbf{N}$$

Solving the equations

$$\underbrace{\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix}}_{\substack{\mathbf{I} \\ 3\times 1}} = \underbrace{\begin{bmatrix} \mathbf{L_1}^T \\ \mathbf{L_2}^T \\ \mathbf{L_3}^T \end{bmatrix}}_{\substack{\mathbf{L} \\ 3\times 3}} \underbrace{k_d\mathbf{N}}_{\substack{\mathbf{G} \\ 3\times 1}}$$

$$\mathbf{G} = \mathbf{L}^{-1}\mathbf{I}$$

$$k_d = \|\mathbf{G}\|$$

$$\mathbf{N} = \frac{1}{k_d}\mathbf{G}$$

Solve one such linear system **per pixel** to solve for that pixel's surface normal

More than three lights

Can get better results by using more than 3 lights

$$\begin{array}{ccc} \begin{bmatrix} I_1 \\ \vdots \\ I_n \end{bmatrix} = & \begin{bmatrix} \mathbf{L_1} \\ \vdots \\ \mathbf{L_n} \end{bmatrix} & k_d\mathbf{N} \\ \text{nx3} & \text{nx3} & \text{3x1} \end{array}$$

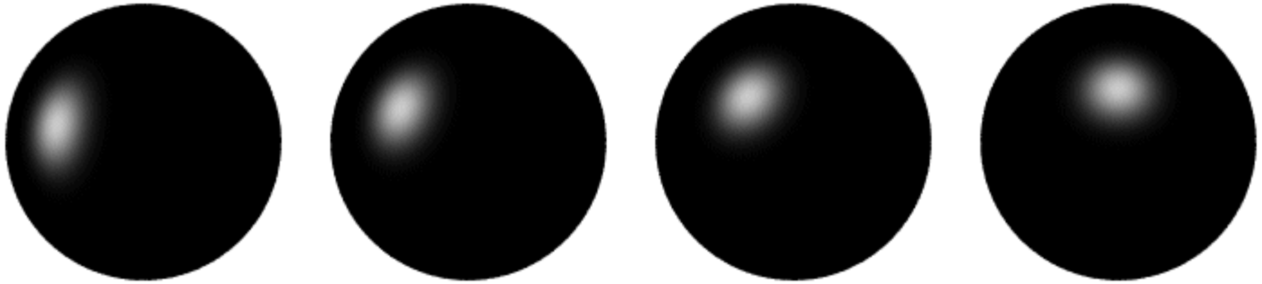
Least squares solution:

$$\begin{aligned} \mathbf{I} &= \mathbf{LG} \\ \mathbf{L^T I} &= \mathbf{L^T L G} \\ \mathbf{G} &= (\mathbf{L^T L})^{-1}(\mathbf{L^T I}) \end{aligned}$$

Solve for N, k_d as before

Calibrating Lighting Directions

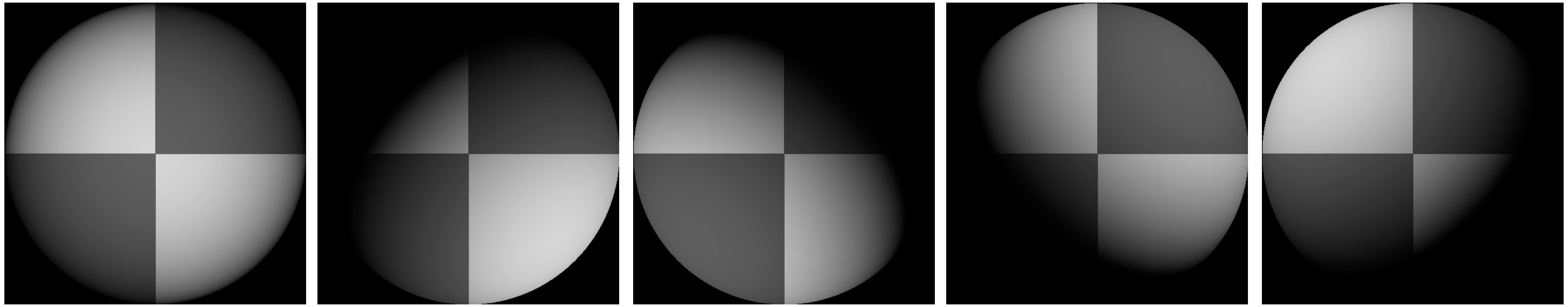
Trick: place a chrome sphere in the scene



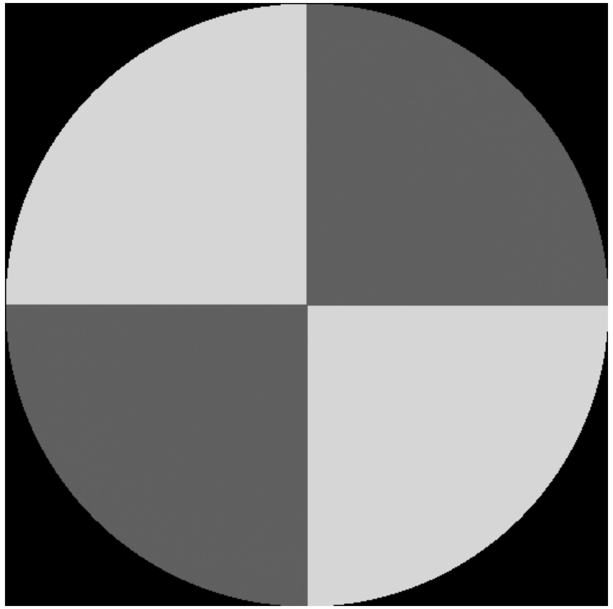
- the location of the highlight tells you where the light source is

Example

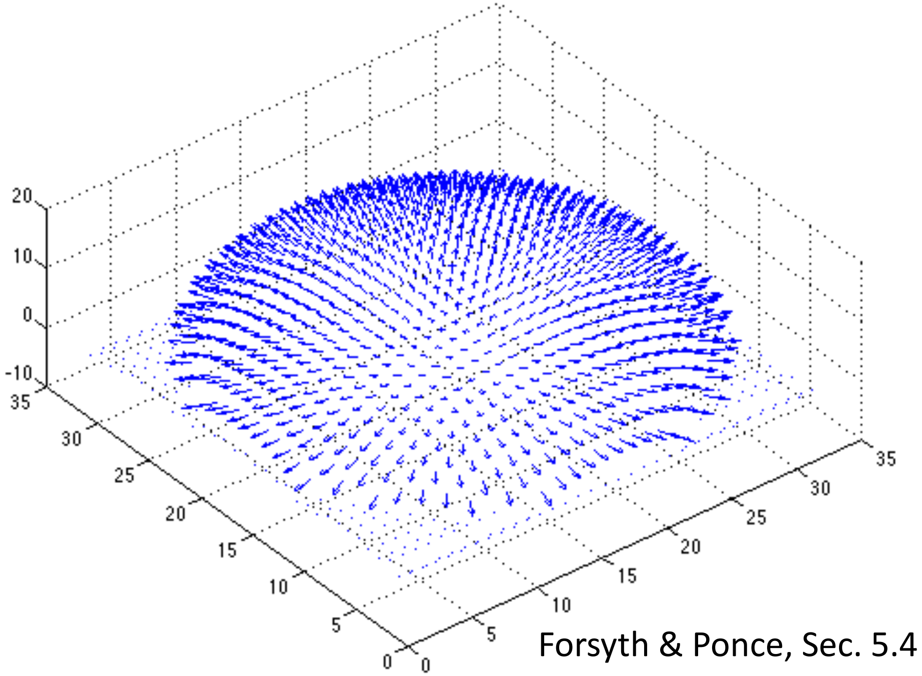
Input views



Recovered albedo



Recovered normal field



Forsyth & Ponce, Sec. 5.4

Depth from normals

- Solving the linear system per-pixel gives us an estimated surface normal for each pixel

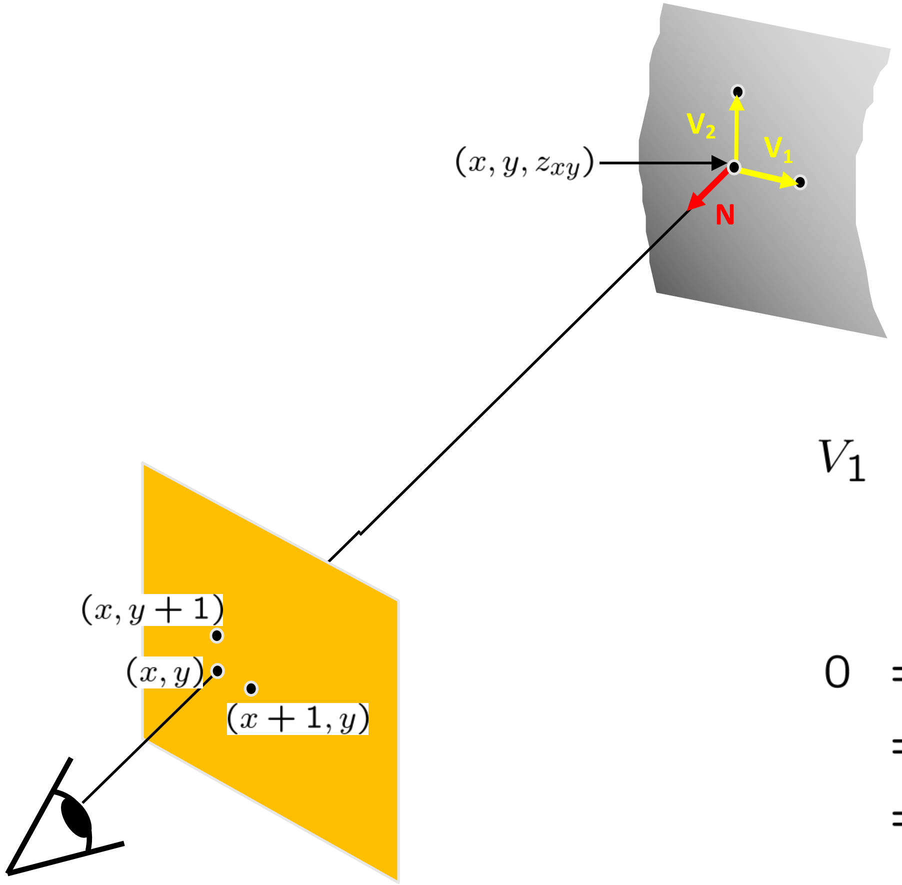
- How can we compute depth from normals?
    - Normals are like the “derivative” of the true depth

Input photo

Estimated normals

Estimated normals (needle diagram)

Depth from normals



$$
\begin{aligned}
V_1 &= (x+1, y, z_{x+1,y}) - (x, y, z_{xy}) \\
&= (1, 0, z_{x+1,y} - z_{xy})
\end{aligned}
$$

$$
\begin{aligned}
0 &= N \cdot V_1 \\
&= (n_x, n_y, n_z) \cdot (1, 0, z_{x+1,y} - z_{xy}) \\
&= n_x + n_z(z_{x+1,y} - z_{xy})
\end{aligned}
$$

Get a similar equation for $\mathbf{V_2}$

- Each normal gives us two linear constraints on z
- compute z values by solving a matrix equation

Normal Integration

$$\nabla z = [p, q]^\top$$

where: → Linear Partial Differential Equations

$$\begin{cases} p = -\dfrac{n_1}{n_3} \\ q = -\dfrac{n_2}{n_3} \end{cases}$$

Integrability Constraint:

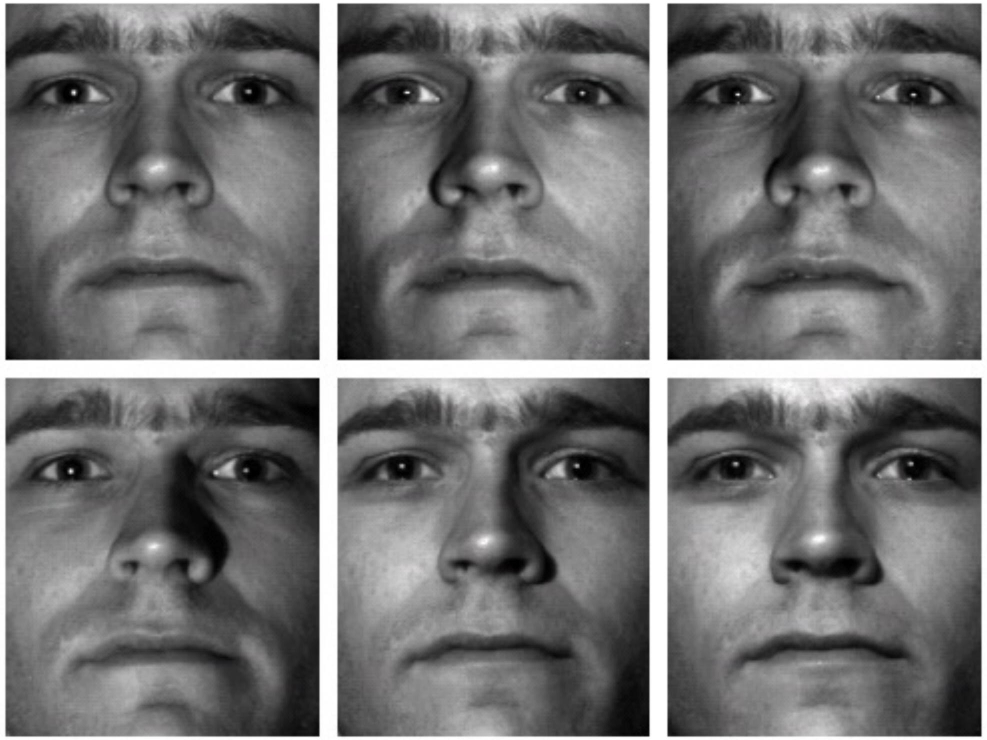
$$\partial_v p = \partial_u q$$

The order of taking 2nd order partial derivative with u & v (or x& y) shouldn't matter!

$$z(u, v) = z(u_0, v_0) + \int_{(r,s)=(u_0,v_0)}^{(u,v)} [p(r, s)\,\mathrm{d}r + q(r, s)\,\mathrm{d}s]$$

Read Normal Integration: A Survey (if interested)

Results

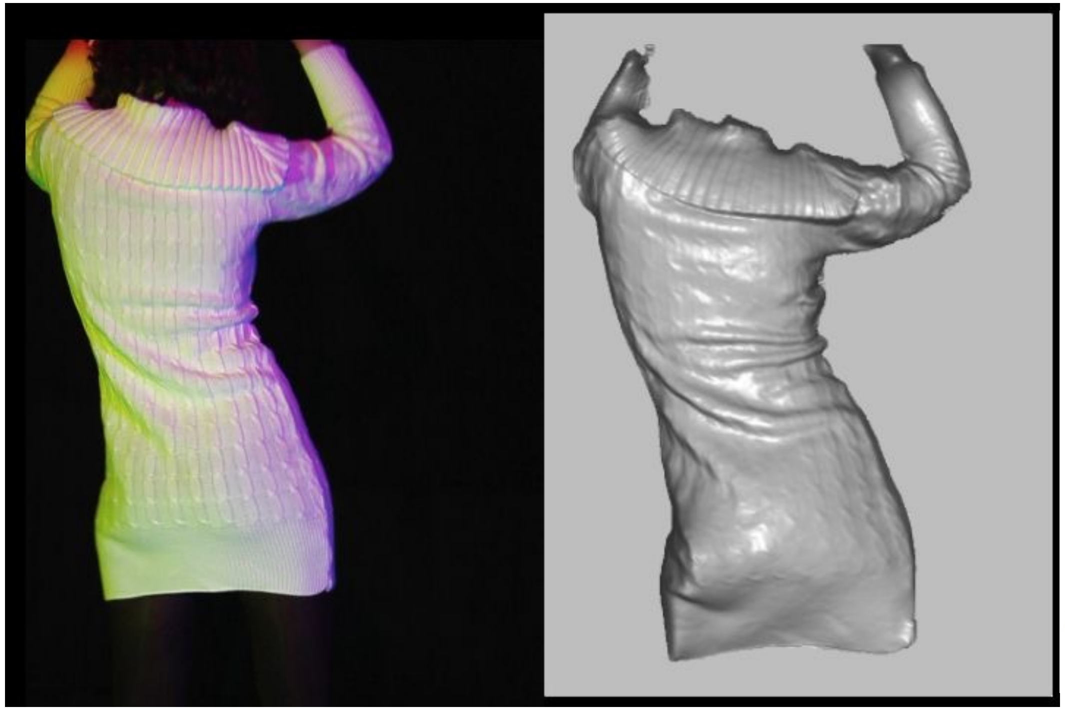


from Athos Georghiades

Results

Extension

- Photometric Stereo from Colored Lighting



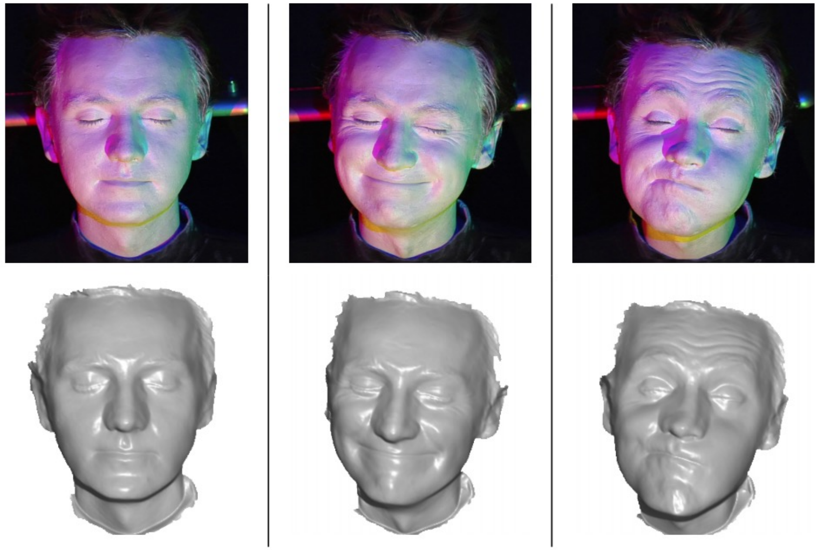


Fig. 2. **Applying the original algorithm to a face with white makeup.** Top: example input frames from video of an actor smiling and grimacing. Bottom: the resulting integrated surfaces.

Video Normals from Colored Lights
Gabriel J. Brostow, Carlos Hernández, George Vogiatzis, Björn Stenger, Roberto Cipolla
[IEEE TPAMI](), Vol. 33, No. 10, pages 2104-2114, October 2011.

Today’s class

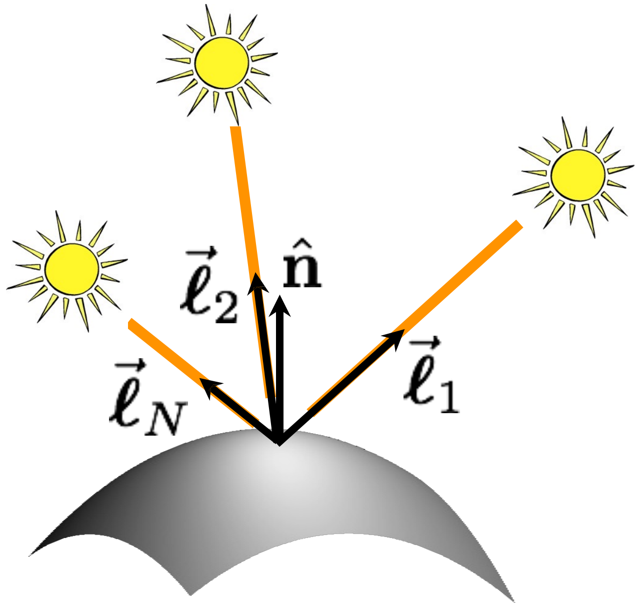
- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.

What if the light directions are unknown?

a = albedo.

Previously k_d was used for albedo.

$$I_1 = a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_1$$
$$I_2 = a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_2$$
$$\vdots$$
$$I_N = a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_N$$



define “pseudo-normal” $\vec{\boldsymbol{b}} \triangleq a\hat{\boldsymbol{n}}$

solve linear system for pseudo-normal

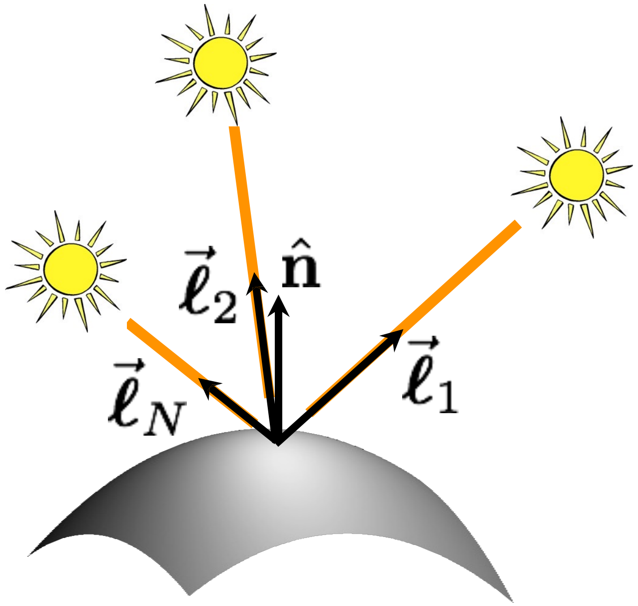
$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix}_{N\times 1} = \begin{bmatrix} \vec{\boldsymbol{\ell}}_1^\top \\ \vec{\boldsymbol{\ell}}_2^\top \\ \vdots \\ \vec{\boldsymbol{\ell}}_N^\top \end{bmatrix}_{N\times 3} \begin{bmatrix} \vec{\boldsymbol{b}} \end{bmatrix}_{3\times 1}$$

What if the light directions are unknown?

a = albedo.

Previously k_d was used for albedo.

$$
\begin{aligned}
I_1 &= a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_1 \\
I_2 &= a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_2 \\
&\vdots \\
I_N &= a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_N
\end{aligned}
$$



define "pseudo-normal" $\vec{\boldsymbol{b}} \triangleq a\hat{\boldsymbol{n}}$

solve linear system for pseudo-normal at each image pixel

$$
\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix}_{N\times M} = \begin{bmatrix} \vec{\boldsymbol{\ell}}_1^\top \\ \vec{\boldsymbol{\ell}}_2^\top \\ \vdots \\ \vec{\boldsymbol{\ell}}_N^\top \end{bmatrix}_{N\times 3} \begin{bmatrix} B \end{bmatrix}_{3\times M}
$$

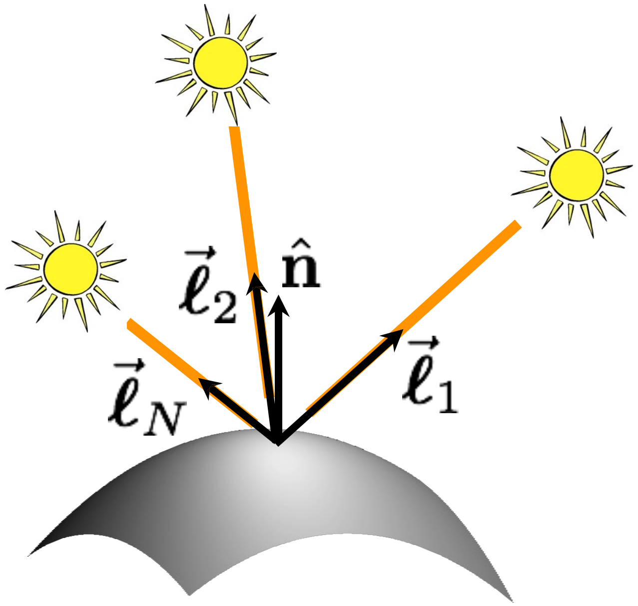
M: number of pixels

What if the light directions are unknown?

a = albedo.

Previously k_d was used for albedo.

$$\begin{aligned} I_1 &= a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_1 \\ I_2 &= a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_2 \\ &\vdots \\ I_N &= a\hat{\boldsymbol{n}}^\top \vec{\boldsymbol{\ell}}_N \end{aligned}$$



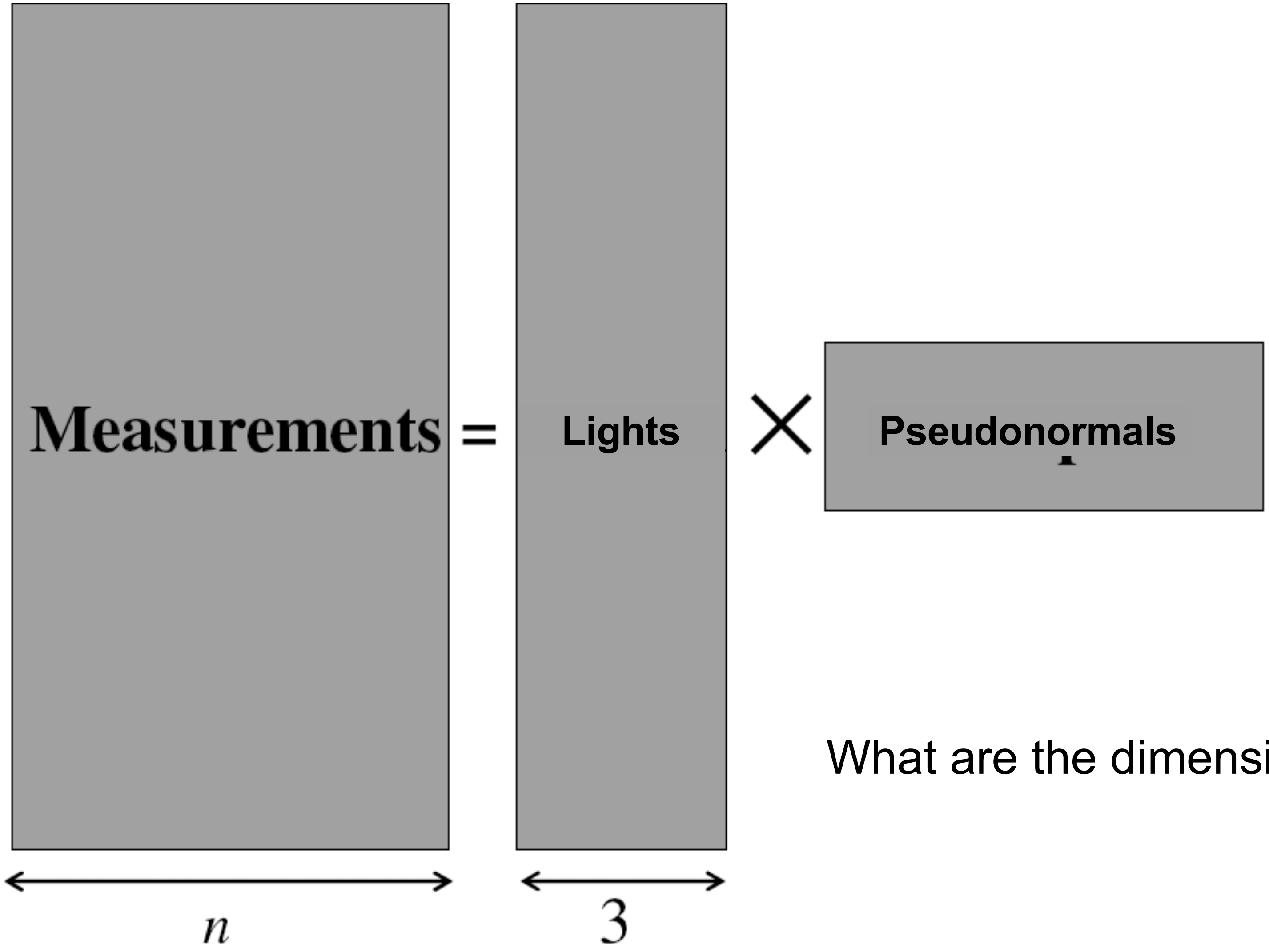
define "pseudo-normal" $\vec{\boldsymbol{b}} \triangleq a\hat{\boldsymbol{n}}$

solve linear system for pseudo-normal at each image pixel

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_N \end{bmatrix}_{N\times M} = \begin{bmatrix} \vec{\ell}_1^\top \\ \vec{\ell}_2^\top \\ \vdots \\ \vec{\ell}_N^\top \end{bmatrix}_{N\times 3} \begin{bmatrix} B \end{bmatrix}_{3\times M}$$

How do we solve this system without knowing light matrix L?

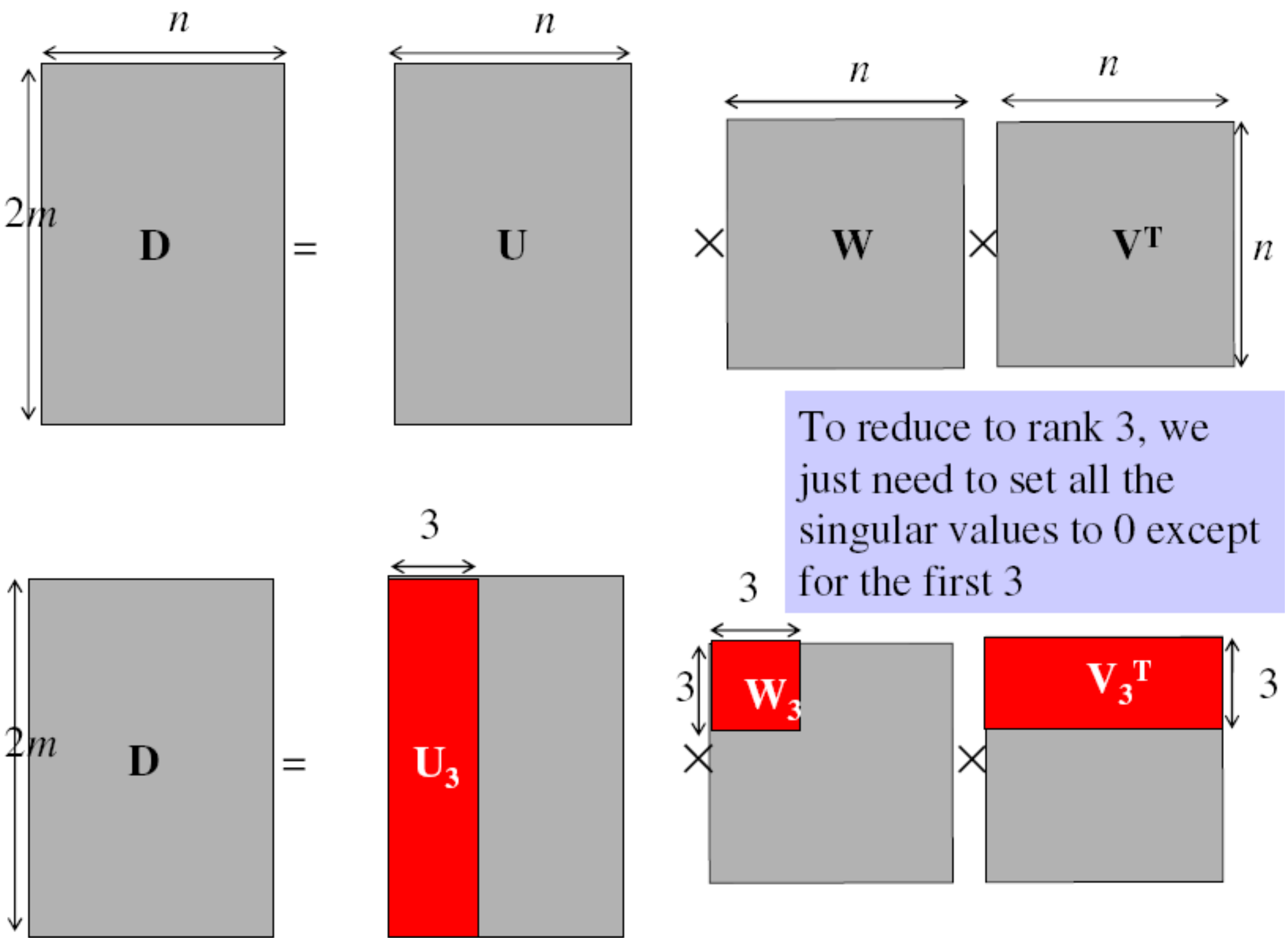
Factorizing the measurement matrix



What are the dimensions?

Factorizing the measurement matrix

- Singular value decomposition:



To reduce to rank 3, we just need to set all the singular values to 0 except for the first 3

This decomposition minimizes **$|I-LB|^2$**

Are the results unique?

We can insert any 3x3 matrix Q in the decomposition and get the same images:

$$\mathbf{I} = \mathbf{L}\,\mathbf{B} = (\mathbf{L}\,\mathbf{Q}^{-1})\,(\mathbf{Q}\,\mathbf{B})$$

Can we use any assumptions to remove some of these 9 degrees of freedom?

Today’s class

- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.

Generalized Bas-Relief ambiguity

We can insert any 3x3 matrix Q in the decomposition and get the same images:

$$\mathbf{I} = \mathbf{L}\,\mathbf{B} = (\mathbf{L}\,\mathbf{Q}^{-1})\,(\mathbf{Q}\,\mathbf{B})$$

Can we use any assumptions to remove some of these 9 degrees of freedom?

Generalized Bas-Relief ambiguity to rescue!

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}.$$

G has 3 degrees of freedom.

What does G mean?

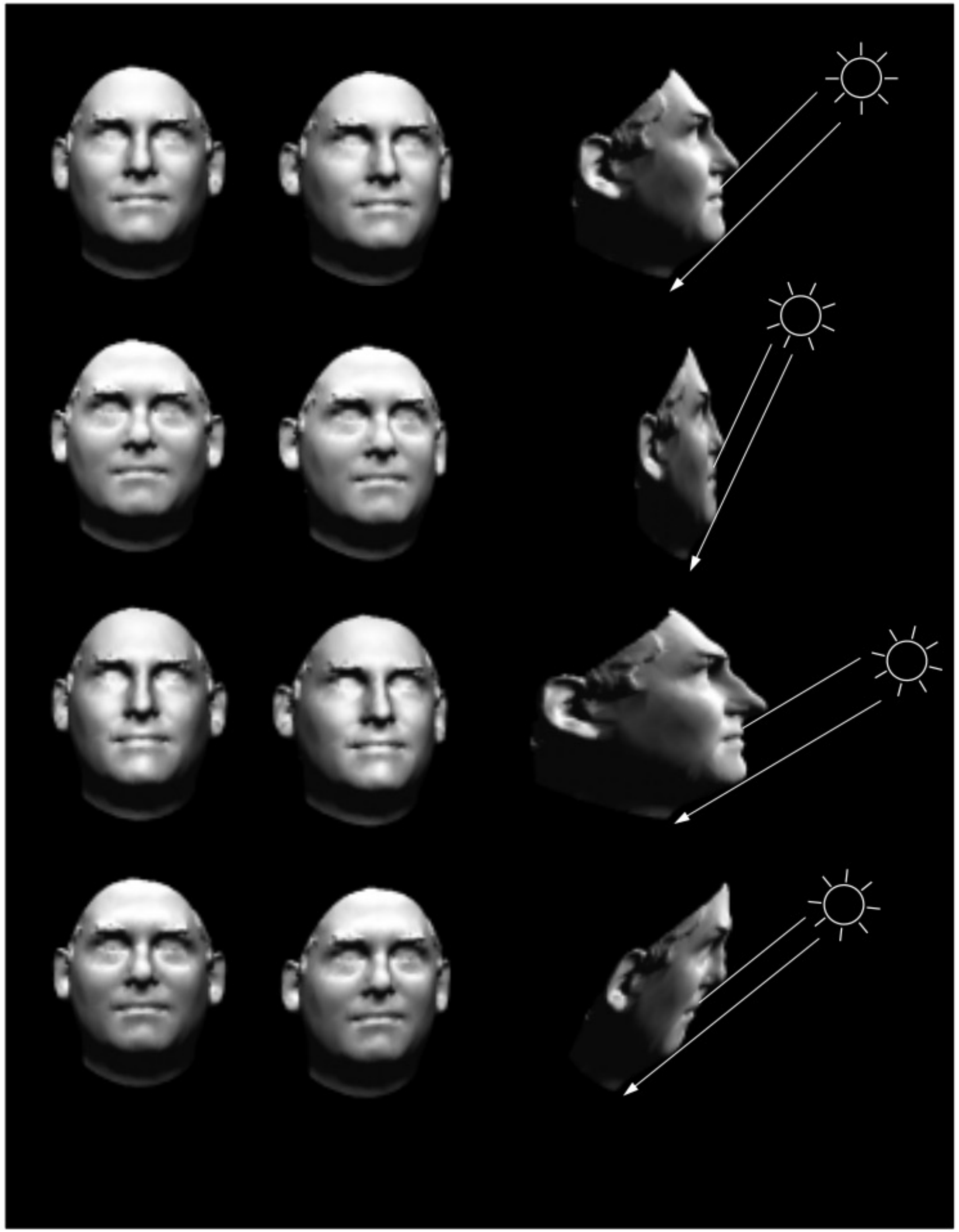
How do we obtain G? What constraints lead us to G?

Generalized Bas-Relief ambiguity

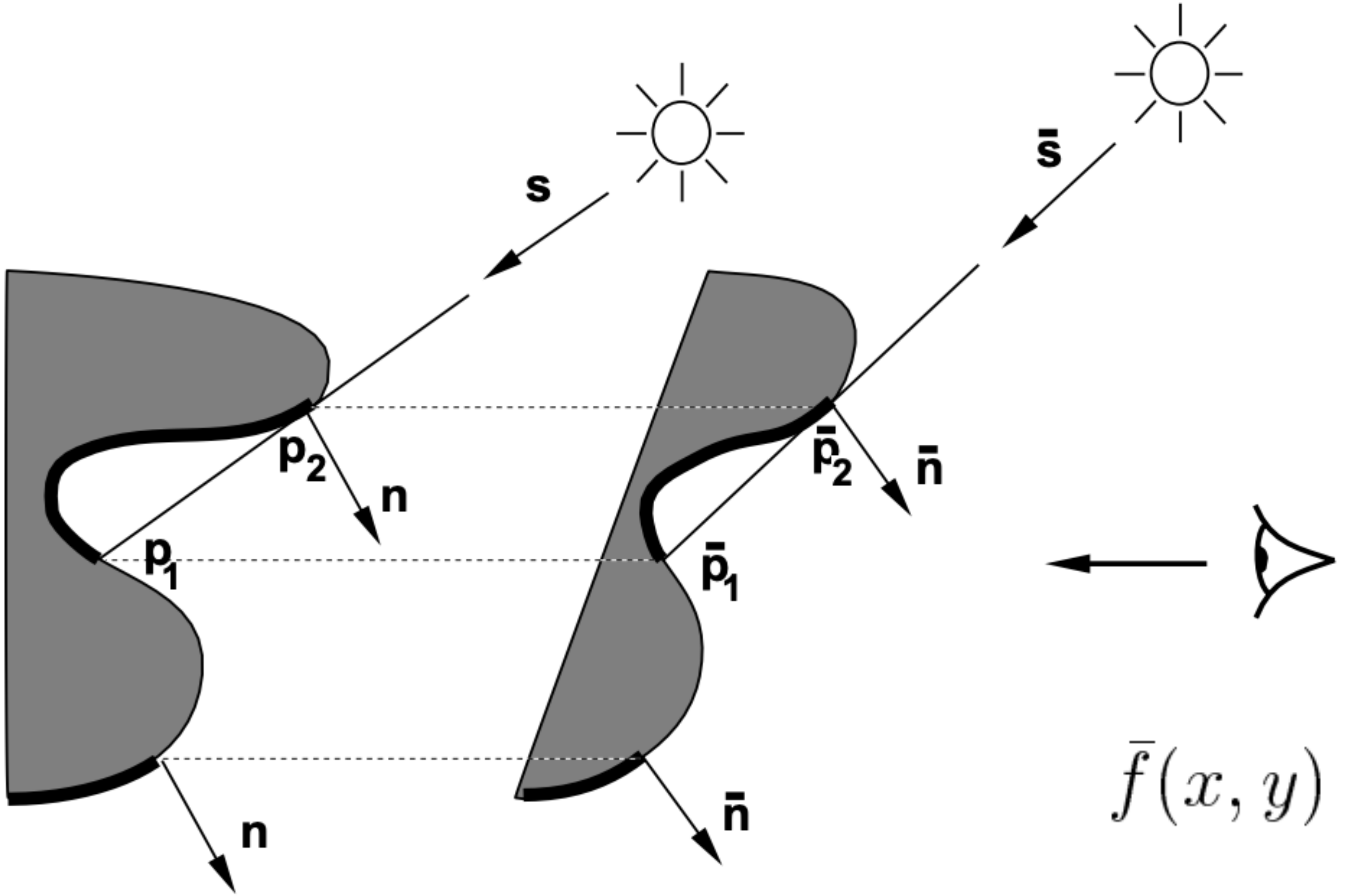
Artists have exploited GBR ambiguity in creating statues!

- On can flatten a surface and yet give an impression of full 3D to a viewer

“The Bas-Relief Ambiguity”, Peter N. Belhumeur, David J. Kriegman, Alan L. Yuille, IJCV 99



Generalized Bas-Relief ambiguity



$$z = f(x, y)$$

$$\mathbf{n}(x, y) = \begin{bmatrix} -f_x \\ -f_y \\ 1 \end{bmatrix}$$

$$\bar{f}(x, y) = \lambda f(x, y) + \mu x + \nu y$$

Generalized Bas-Relief ambiguity

Note that if $\mathbf{p} = (x, y, f(x, y))$ and $\bar{\mathbf{p}} = (x, y, \bar{f}(x, y))$, then $\bar{\mathbf{p}} = G\mathbf{p}$ where

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}.$$

$$\bar{\mathbf{n}} = G^{-T}\mathbf{n}$$

$$G^{-1} = \frac{1}{\lambda}\begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ -\mu & -\nu & 1 \end{bmatrix}.$$

Generalized Bas-Relief ambiguity

We can insert any 3x3 matrix Q in the decomposition and get the same images:

$$\mathbf{I} = \mathbf{L}\,\mathbf{B} = (\mathbf{L}\,\mathbf{Q}^{-1})\,(\mathbf{Q}\,\mathbf{B})$$

Can we use any assumptions to remove some of these 9 degrees of freedom?

Generalized Bas-Relief ambiguity to rescue!

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}.$$

G has 3 degrees of freedom.

G indicates integrable surface:
The order of taking 2nd order partial derivative with u & v (or x& y) shouldn't matter!

Enforcing integrability

What does the integrability constraint correspond to?

- Differentiation order should not matter:

$$\frac{d}{dy}\frac{df(x,y)}{dx} = \frac{d}{dx}\frac{df(x,y)}{dy}$$

$$\mathbf{I} = \mathbf{L}\,\mathbf{B} = (\mathbf{L}\,\mathbf{Q}^{-1})\,(\mathbf{Q}\,\mathbf{B})$$

If B is integrable, then:

- $B' = G^{-T} \cdot B$ is also integrable for all G of the form ($\lambda \neq 0$)

$$G = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \mu & \nu & \lambda \end{bmatrix}$$

For now, ignore specular reflection

Slides from Photometric Methods for 3D Modeling, Matsushita, Wilburn, Ben-Ezra

And Refraction…

Slides from Photometric Methods for 3D Modeling, Matsushita, Wilburn, Ben-Ezra

And Interreflections…

Slides from Photometric Methods for 3D Modeling, Matsushita, Wilburn, Ben-Ezra

And Subsurface Scattering…

Slides from Photometric Methods for 3D Modeling, Matsushita, Wilburn, Ben-Ezra

What assumptions have we made for all this?

- Lambertian BRDF
- Directional lighting
- Distant Lighting
- Orthographic camera
- No interreflections or scattering

Limitations

Bigger problems

- doesn’t work for shiny things, semi-translucent things
- shadows, inter-reflections

Smaller problems

- camera and lights have to be distant
- calibration requirements
  - measure light source directions, intensities
  - camera response function

Newer work addresses some of these issues

Some pointers for further reading:

- Zickler, Belhumeur, and Kriegman, "*Helmholtz Stereopsis: Exploiting Reciprocity for Surface Reconstruction*." IJCV, Vol. 49 No. 2/3, pp 215-227.
- Hertzmann & Seitz, “*Example-Based Photometric Stereo: Shape Reconstruction with General, Varying BRDFs*.” IEEE Trans. PAMI 2005

Today’s class

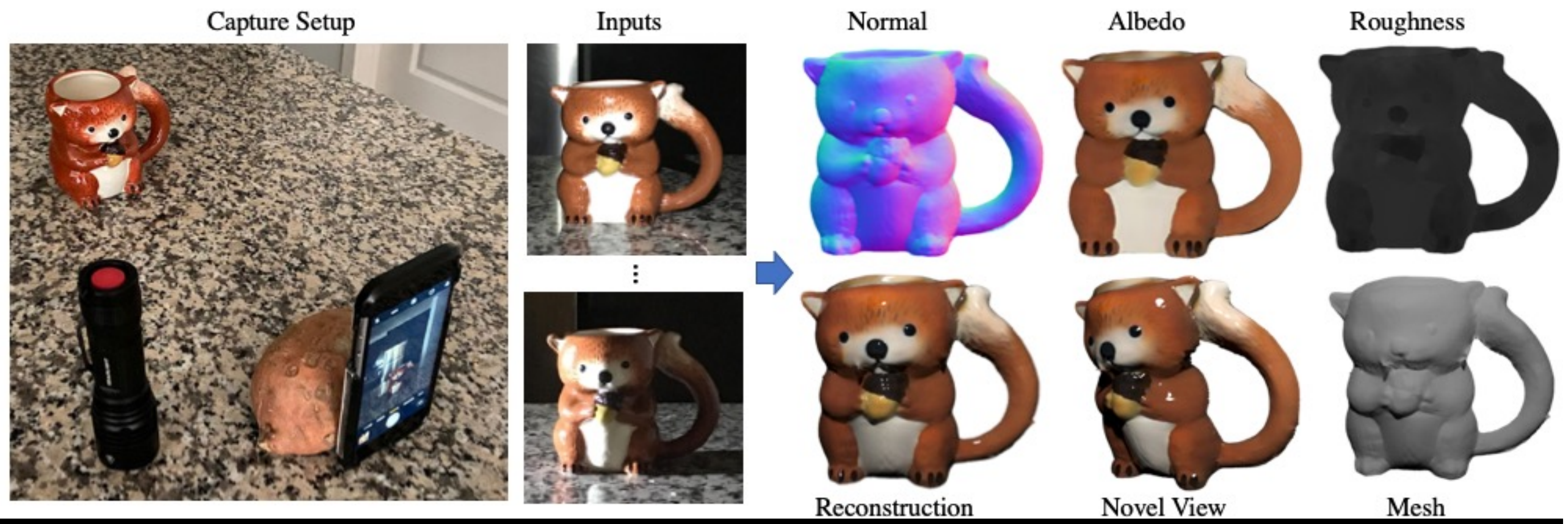
- Measuring Light (recap)
- Image formation with shape, reflectance, and illumination
- Shape from Shading
- Photometric Stereo
- Uncalibrated Photometric Stereo
- Generalized Bas-Relief Ambiguity
- Photometric Stereo in ‘deep learning era’.

Photometric Stereo now … in Deep Learning era!

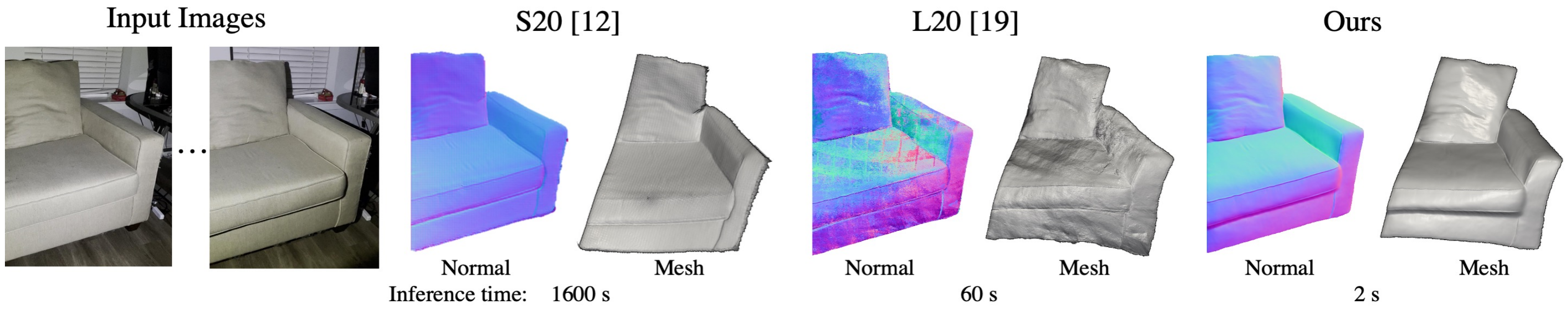
- Exploiting High-quality CG rendering for training data
- Designing deep neural network architectures
- Designing loss functions

- GBR ambiguity is still a problem! -> Flattened objects reconstructed.

Using lighting as a cue for 3D reconstruction (Photometric Stereo)



"Shape & Material Capture at Home", Lichy, Wu, Sengupta, Jacobs, CVPR 2021



"Real-Time Light-Weight Near-Field Photometric Stereo", Lichy, Sengupta, Jacobs, CVPR 2022

Captured Images: Right

Single iPhone Image with Built-In Flash

Image 1/1

Mesh

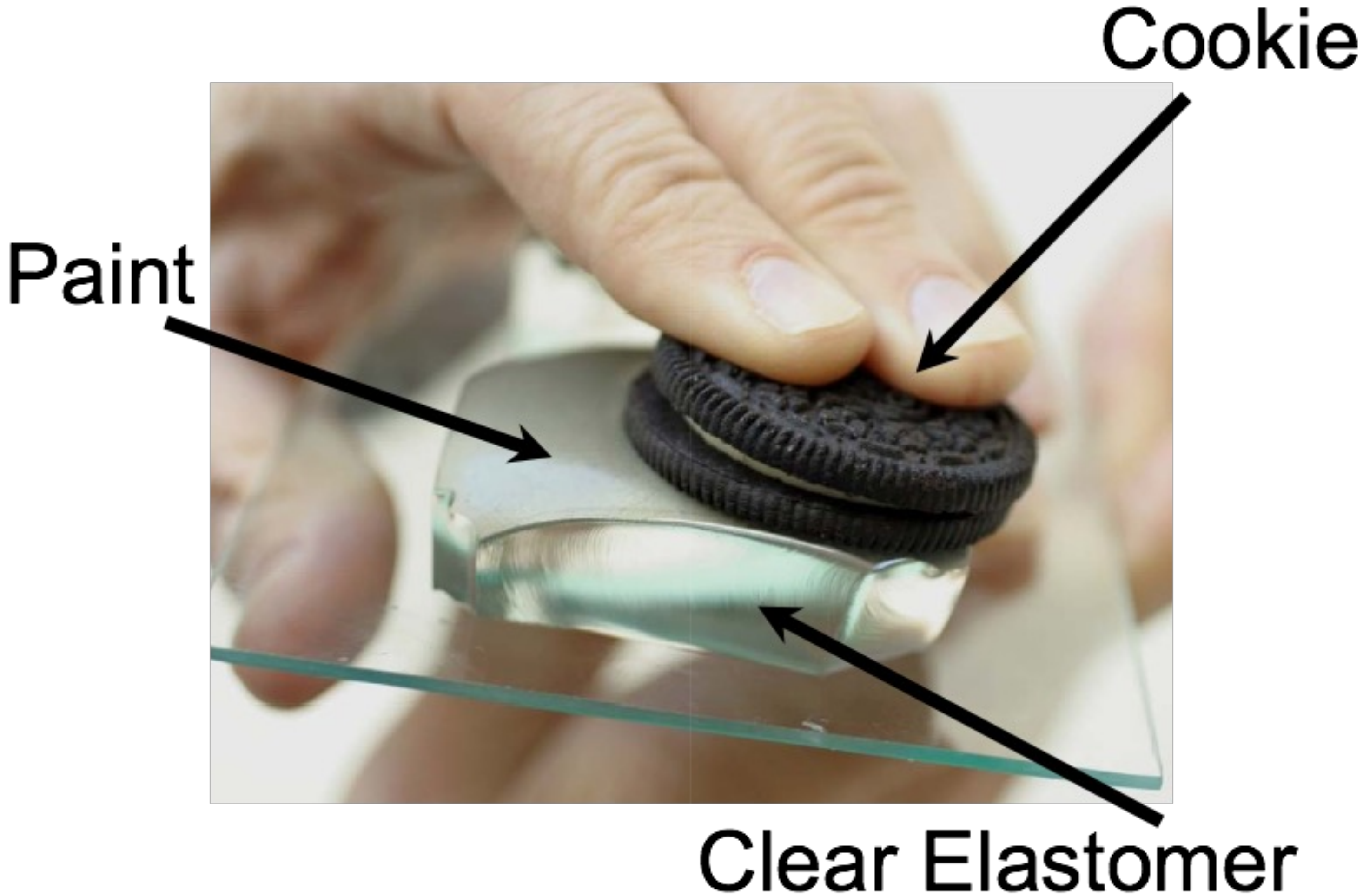
Photometric Stereo + SLAM for colon reconstruction in colonoscopy

SLAM only

Photometric Stereo + SLAM (Ours)

“A Surface-normal Based Neural Framework for Colonoscopy Reconstruction”, Sherry Wang, Yubo Zhang, Sarah McGill, Julian Rosenman, Jan-Michael Frahm, Soumyadip Sengupta, Steve Pizer, IPMI 2023.

Johnson and Adelson, 2009

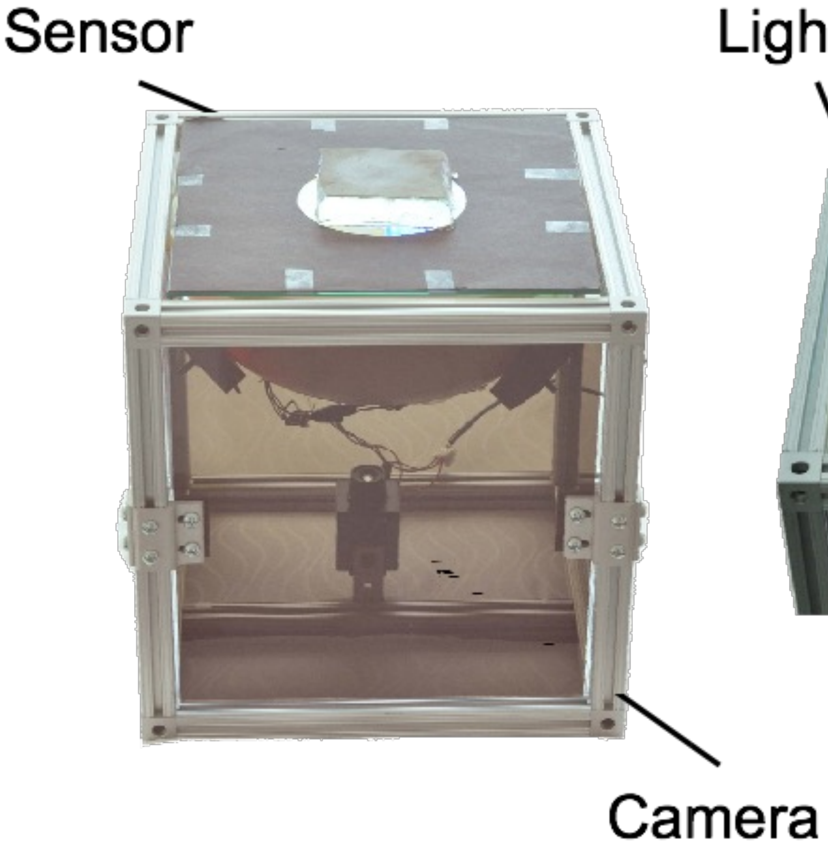


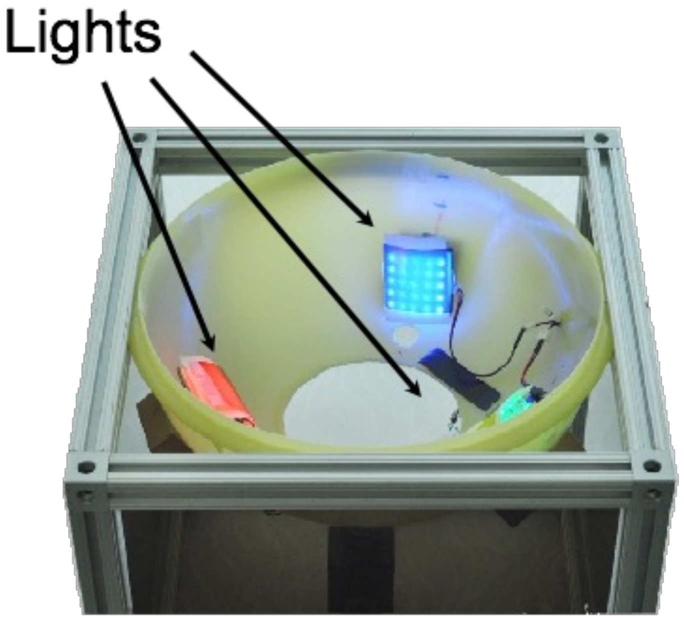
Johnson and Adelson, 2009

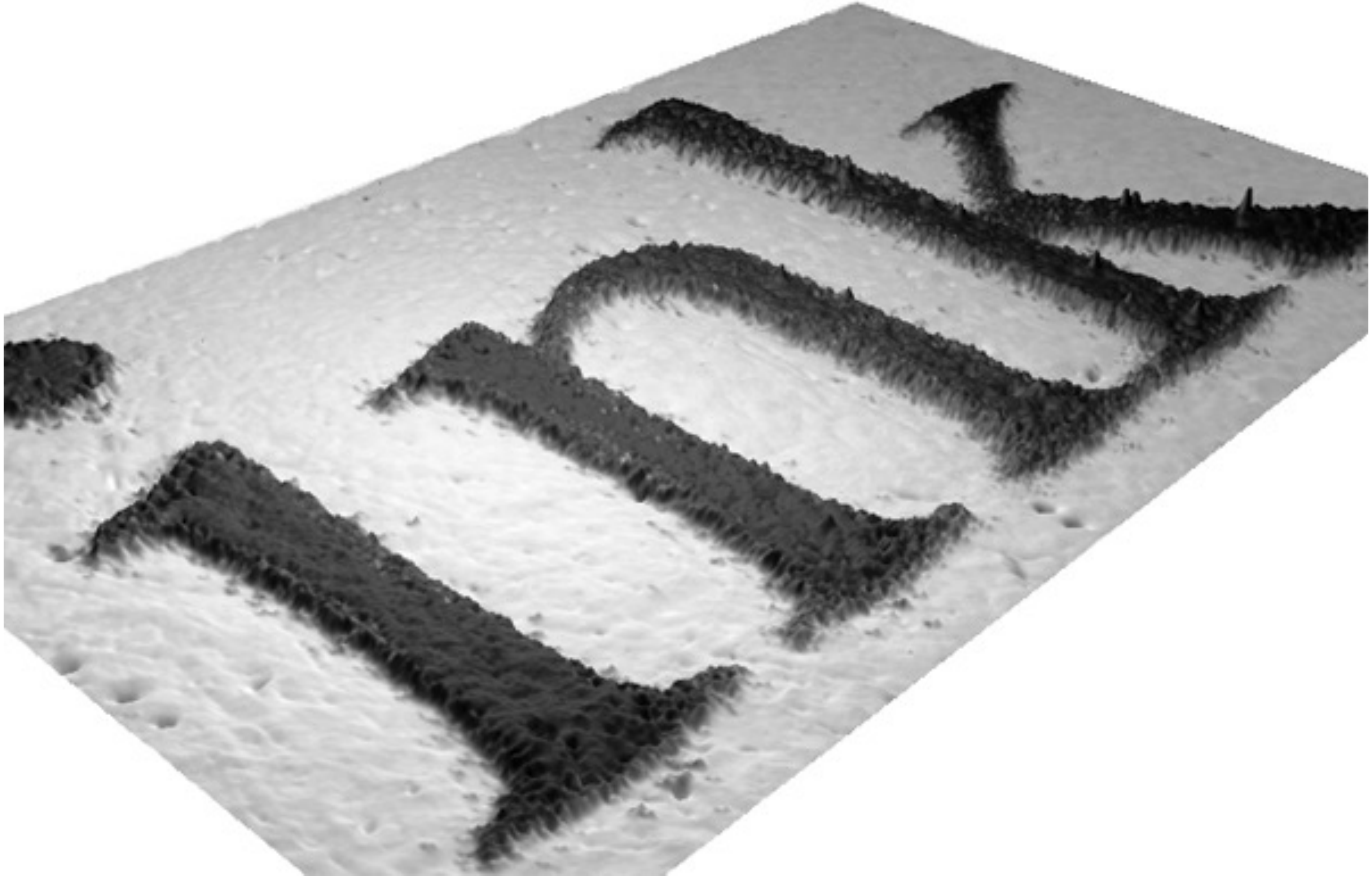

OREO

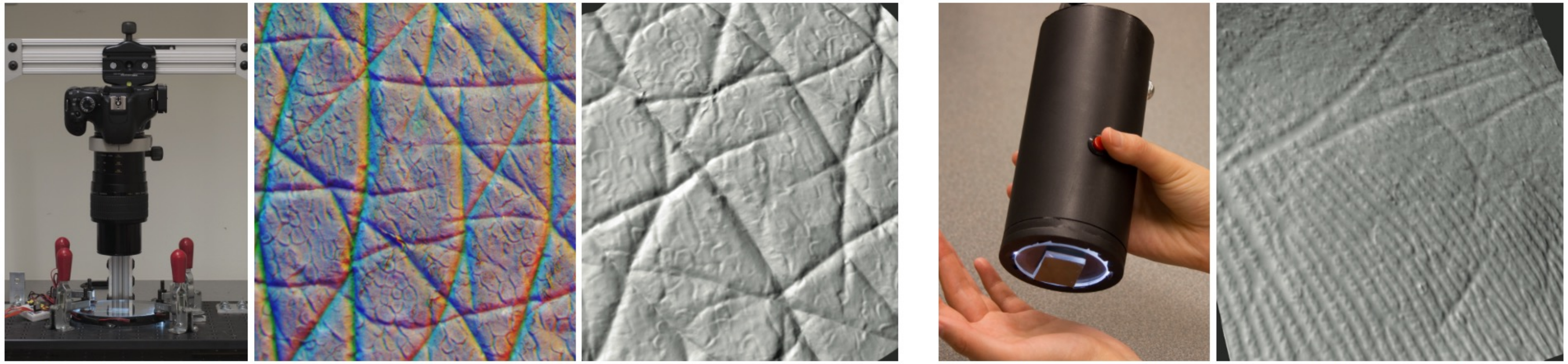

OREO

Lights, camera, action









(a) bench configuration (b) captured (c) reconstruction (d) portable configuration (e) reconstruction

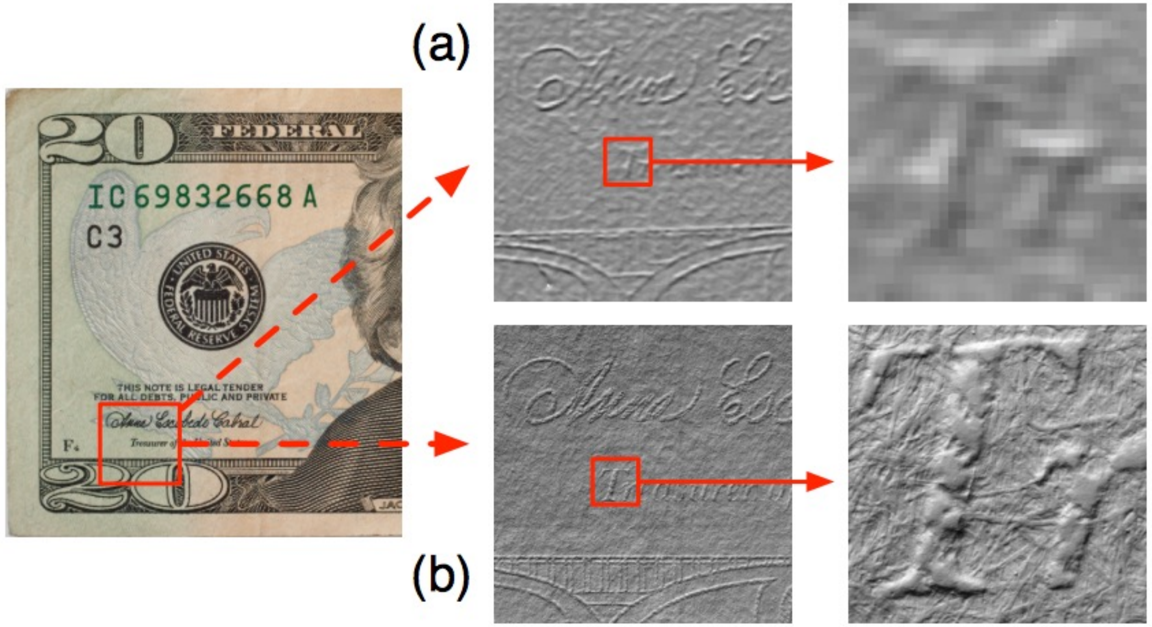


Figure 7: *Comparison with the high-resolution result from the original retrographic sensor. (a) Rendering of the high-resolution \$20 bill example from the original retrographic sensor with a close-up view. (b) Rendering of the captured geometry using our method.*

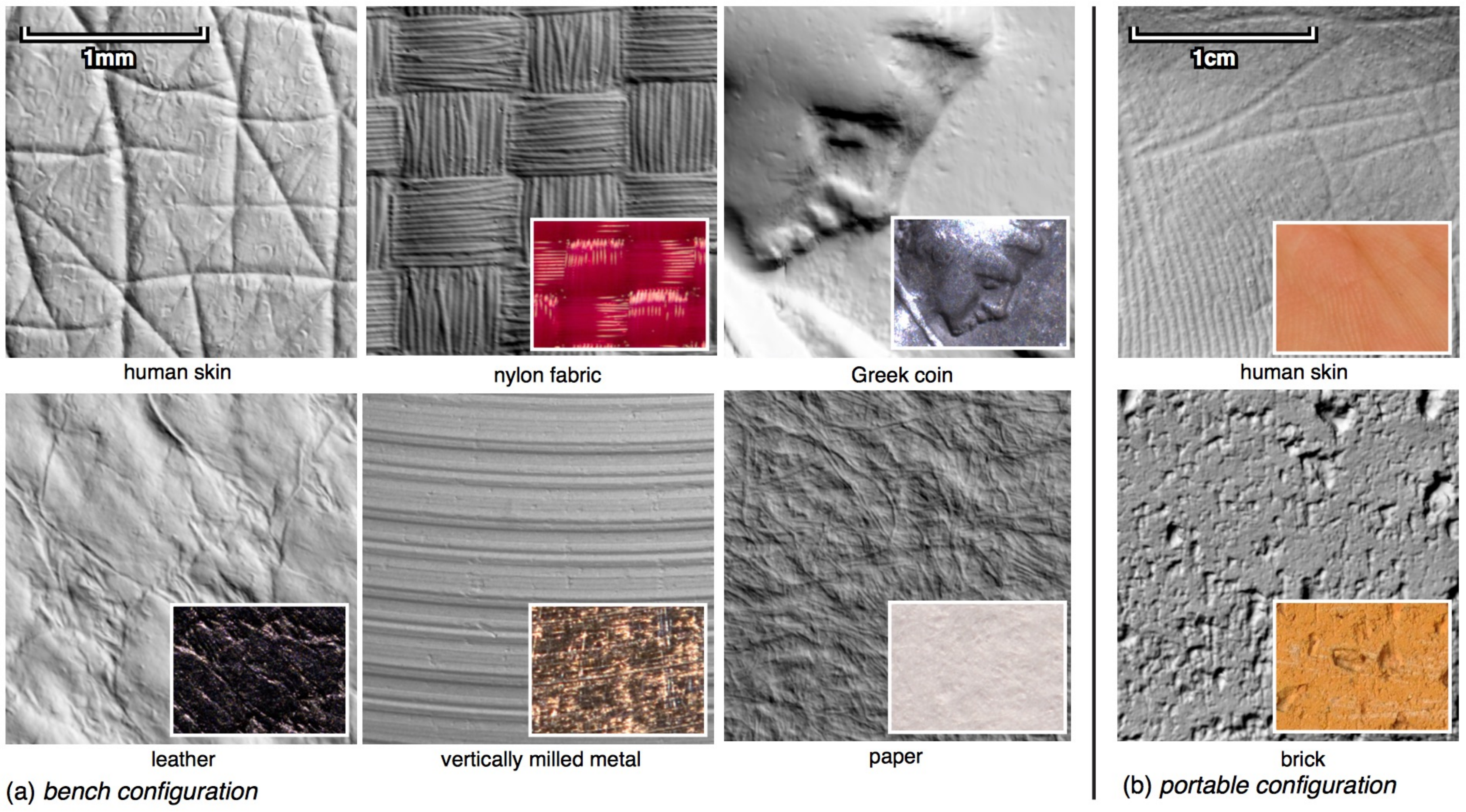
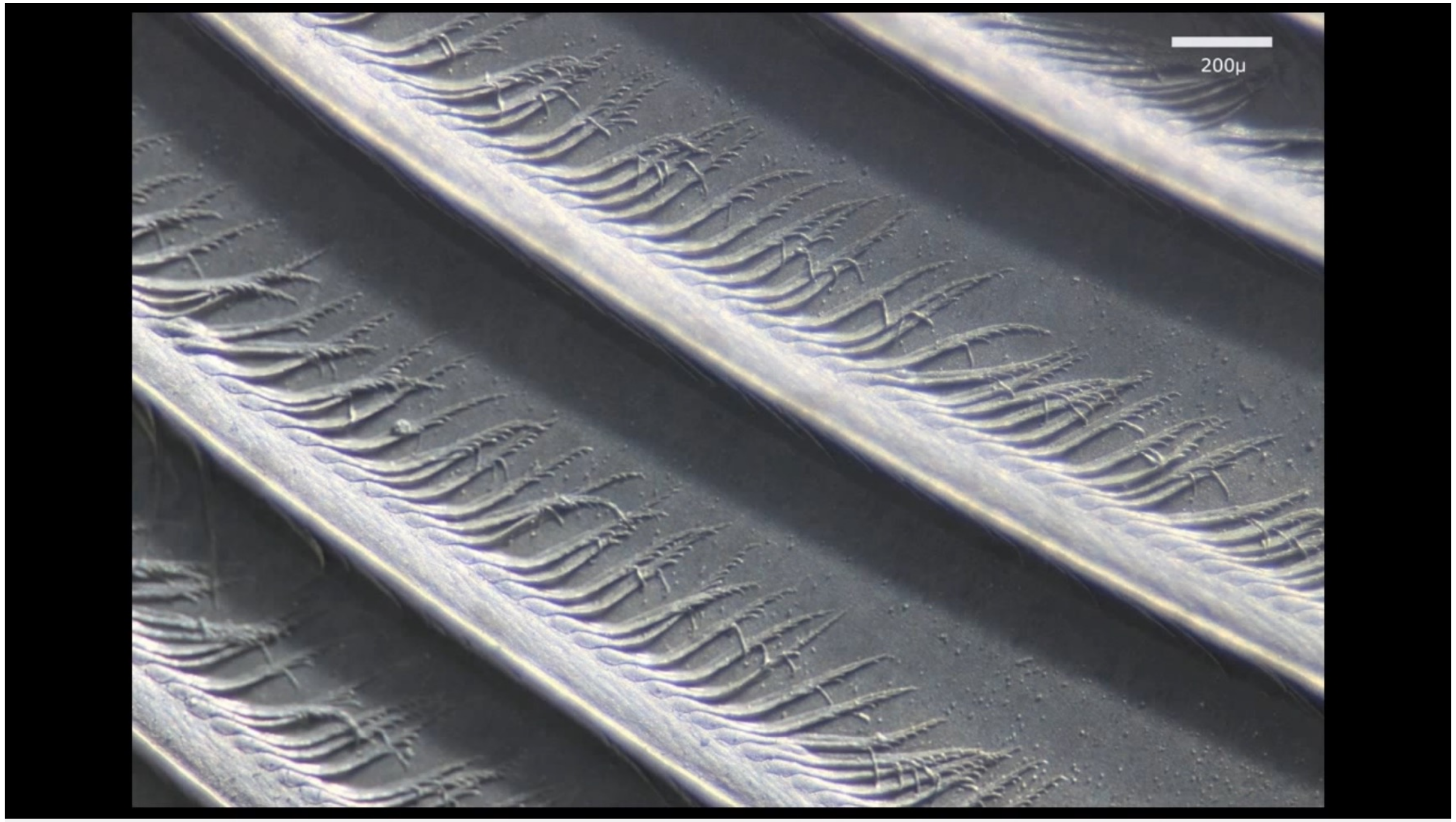


Figure 9: *Example geometry measured with the bench and portable configurations. Outer image: rendering under direct lighting. Inset: macro photograph of original sample. Scale shown in upper left. Color images are shown for context and are to similar, but not exact scale.*



Sensing Surfaces with GelSight

kimoatmit

Subscribe 64

138,850 views

https://www.youtube.com/watch?v=S7gXih4XS7A