Photometric Stereo

Computer Vision I

CSE252A

Lecture 8

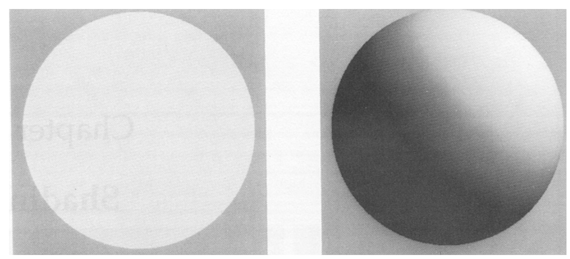
Announcement

- Read Chapter 2 of Forsyth & Ponce
- Office hours tomorrow:
  - 3-5, CSE 4127
  - 5-6, CSE B260A
- Piazza
- Next lecture

Shape-from-X

- Where X is
  - Shading
  - Photometric stereo
  - Stereo
  - Motion
  - Texture
  - Blur
  - Focus
  - Structured light
  - ….

Shading reveals 3-D surface geometry

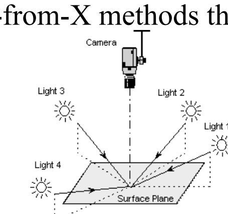


Two shape-from-X methods that use shading

- Shape-from-shading: Use just one image to recover shape. Requires knowledge of light source direction and BRDF everywhere. Too restrictive to be useful without priors/learning.

- Photometric stereo: Single viewpoint, multiple images under different lighting.
  1. Arbitrary known BRDF
  2. Lambertian BRDF, known lighting
  3. Lambertian BRDF, unknown lighting.

Two shape-from-X methods that use shading



- Photometric stereo: Single viewpoint, multiple images under different lighting.
  1. Arbitrary known BRDF
  2. Lambertian BRDF, known lighting
  3. Lambertian BRDF, unknown lighting.

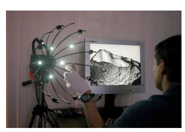
Photometric Stereo Rigs: One viewpoint, changing lighting

Because of single viewpoint, a pixel location sees the same point of the scene across all images

An example of photometric stereo

Multi-view stereo vs. Photometric Stereo: Assumptions

- Multi-view (binocular) Stereo
  - Multiple images
  - Dynamic scene
  - Multiple viewpoints
  - Fixed lighting
- Photometric Stereo
  - Multiple images
  - Static scene
  - Fixed viewpoint
  - Multiple lighting conditions



Photometric Stereo:

1. General BRDF and Reflectance Map
2. Lambertian Surface, known lighting
3. Lambertian surface

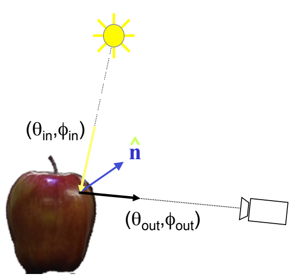
BRDF

- Bi-directional Reflectance Distribution Function

  $\rho(\theta_{in}, \phi_{in} ; \theta_{out}, \phi_{out})$
- Function of
  - Incoming light direction:

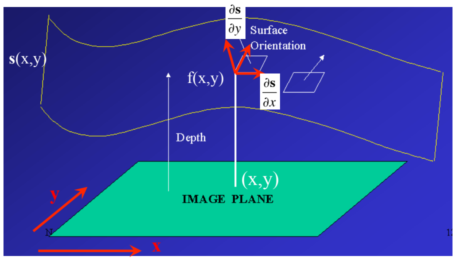
    θ_{in} , ϕ_{in}
  - Outgoing light direction:

    θ_{out} , ϕ_{out}
- Ratio of incident irradiance to emitted radiance



1. General BRDF and Reflectance Maps

Coordinate system

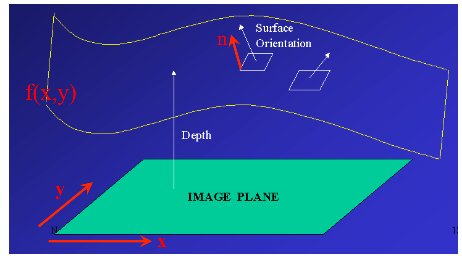


Surface: $\mathbf{s}(x,y) = (x, y, f(x,y))$

Tangent vectors:
$$\frac{\partial \mathbf{s}(x,y)}{\partial x} = \left(1, 0, \frac{\partial f}{\partial x}\right)$$
$$\frac{\partial \mathbf{s}(x,y)}{\partial y} = \left(0, 1, \frac{\partial f}{\partial y}\right)$$

Normal vector
$$\mathbf{n} = \boxed{\frac{\partial \mathbf{s}}{\partial x} \times \frac{\partial \mathbf{s}}{\partial y}} = \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right)$$

Gradient Space (p,q)

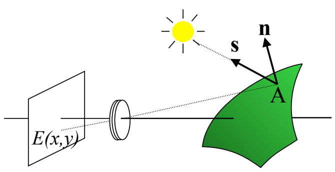


Gradient Space : (p,q)
$$p = \frac{\partial f}{\partial x}, \quad q = \frac{\partial f}{\partial y}$$

Tilt and slant

Normal vector
$$\mathbf{n} = \frac{\partial \mathbf{s}}{\partial x} \times \frac{\partial \mathbf{s}}{\partial y} = \left(-\frac{\partial f}{\partial x}, -\frac{\partial f}{\partial y}, 1\right)^T$$
$$\hat{\mathbf{n}} = \frac{1}{\sqrt{p^2 + q^2 + 1}}(p, q, -1)^T$$

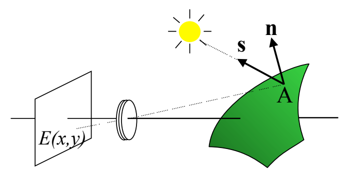
Image Formation



For a surface point A is illuminated by a unit strength distant light source, the image irradiance E(x,y) is a function of

1. The BRDF at A
2. The surface normal **n** at A
3. The direction **s** of the light source
4. The direction of A to the camera lens

Reflectance Map R(p q)



Now, if we

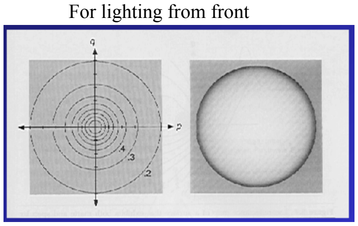
1. fix the lighting direction S
2. assume orthographic projection
3. consider constant BRDF

Then image irradiance is only a function of the surface normal, which we can write as R(p, q).

We can measure R(p,q) for a material from a sphere of that material.

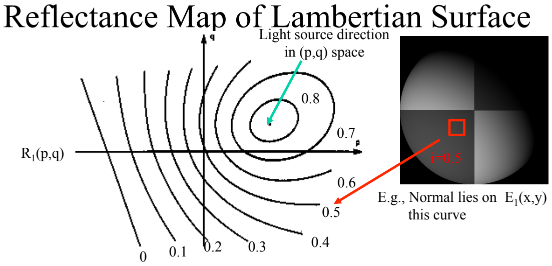
Example Reflectance Map: Lambertian surface

For lighting from front



R(p,q) Image of Lambertian sphere

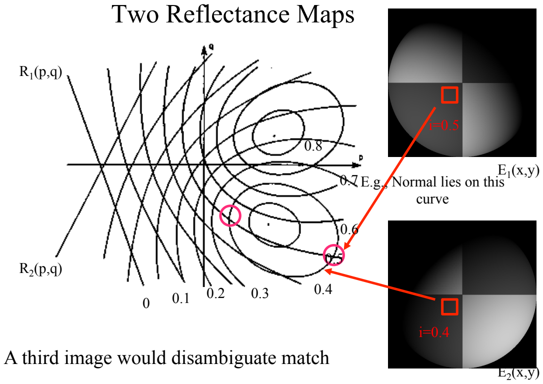
Reflectance Map of Lambertian Surface



What does the intensity (Irradiance) $E_1(x,y)$ of one pixel in one image tell us?

It constrains the surface normal projecting to a curve $R_1(p,q)=E_1(x,y)$

Second image under second light source
Two Reflectance Maps



A third image would disambiguate match

Three Source Photometric stereo: Step1

Offline:

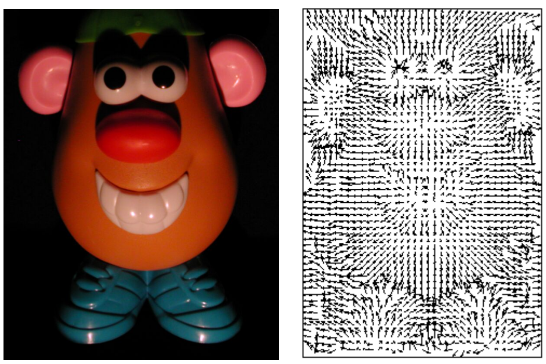
Using source directions & BRDF, construct reflectance map for each light source direction. $R_1(p,q)$, $R_2(p,q)$, $R_3(p,q)$

This can be done analytically or by taking images of spheres

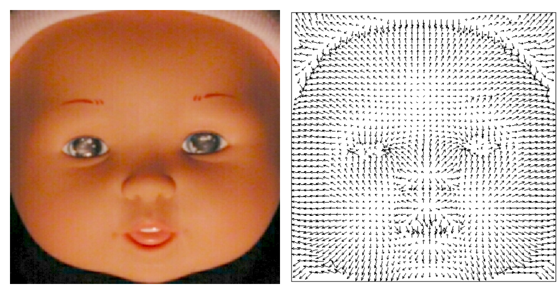
Online:

1. Acquire three images with the known light source directions. $E_1(x,y)$, $E_2(x,y)$, $E_3(x,y)$
2. For each pixel location (x,y), find (p,q) as the intersection of the three curves
   $R_1(p,q)=E_1(x,y)$
   $R_2(p,q)=E_2(x,y)$
   $R_3(p,q)=E_3(x,y)$
3. The estimated (p,q) gives the surface normal at pixel (x,y). Over image, the normal field [p(x,y), q(x,y)] is estimated

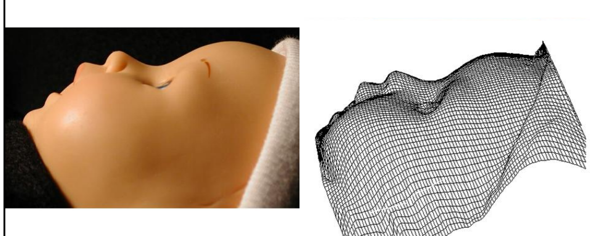
Normal Field



Plastic Baby Doll: Normal Field



Next step: Go from normal field to surface

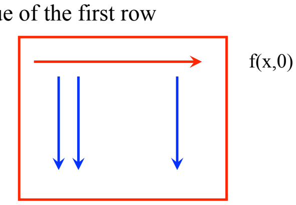


Step 2: Recovering the surface f(x,y)

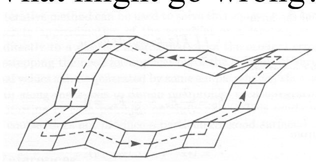
Many methods: Simplest approach

Assumption: Surface and derivative are continuous

1. From estimated $\mathbf{n}=(n_x,n_y,n_z)$, $p=n_x/n_z$, $q=n_y/n_z$
2. Integrate df/dx=p along a row (x,0) to get f(x,0)
3. Then integrate df/dy=q along each column starting with value of the first row



What might go wrong?



- Height z(x,y) is obtained by integration along a curve from (x_0, y_0).

$$z(x,y) = z(x_0,y_0) + \int_{(x_0,y_0)}^{(x,y)} (p\,dx + q\,dy)$$

- If one integrates the derivative field along any closed curve, on expects to get back to the starting value.
- Might not happen because of noisy estimates of (p,q)

What might go wrong?

Integrability. If f(x,y) is the height function, we expect that

$$\frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial}{\partial x}\frac{\partial f}{\partial y}$$

In terms of estimated gradient (p,q), this means:

$$\frac{\partial p}{\partial y} = \frac{\partial q}{\partial x}$$

where $p=n_x/n_z$, $q=n_y/n_z$ with $\mathbf{n}=[n_x\ n_y\ n_z]$

But N (and in turn p,q) are estimated independently at each, equality is not going to exactly hold

Horn's Method

["Robot Vision, B.K.P. Horn, 1986]

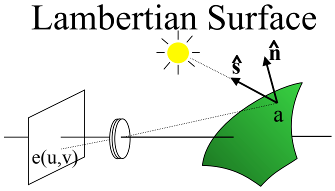
- Formulate estimation of surface height z(x,y) from gradient field by minimizing cost functional:

$$\iint_{\text{Image}} (z_x - p)^2 + (z_y - q)^2\, dx dy$$

where (p,q) are estimated components of the gradient while z_x and z_y are partial derivatives of best fit surface
- Solved using calculus of variations – iterative updating
- z(x,y) can be discrete or represented in terms of basis functions.
- Integrability is naturally satisfied.

II. Photometeric Stereo: Lambertian Surface, Known Lighting

Lambertian Surface



At image location (u,v), the intensity of a pixel x(u,v) is:

$$e(u,v) = [a(u,v)\,\hat{\mathbf{n}}(u,v)] \cdot [s_0\hat{\mathbf{s}}]$$
$$= \mathbf{b}(u,v) \cdot \mathbf{s}$$

where

- a(u,v) is the albedo of the surface projecting to (u,v).
- **n**(u,v) is the direction of the surface normal.
- s_0 is the light source intensity.
- **s** is the direction to the light source.

Lambertian Photometric stereo

- If the light sources $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$ are **known**, then we **can** recover b from as few as three images. (Photometric Stereo: Silver 80, Woodham81).

$$[e_1\ e_2\ e_3] = \mathbf{b}^T[\mathbf{s}_1\ \mathbf{s}_2\ \mathbf{s}_3]$$

- i.e., we measure e_1, e_2, and e_3 and we know $\mathbf{s}_1$, $\mathbf{s}_2$, and $\mathbf{s}_3$. We can then solve for **b** by solving a linear system.

$$\mathbf{b}^T = [e_1\ \ e_2\ \ e_3][\mathbf{s}_1\ \ \mathbf{s}_2\ \ \mathbf{s}_3]^{-1}$$

- Normal is: $\mathbf{n} = \mathbf{b}/|\mathbf{b}|$, albedo is: $|\mathbf{b}|$

What if we have more than 3 Images? Linear Least Squares

$[e_1\ e_2\ e_3 \ldots e_n] = \mathbf{b}^T[\mathbf{s}_1\ \mathbf{s}_2\ \mathbf{s}_3 \ldots \mathbf{s}_n]$

Rewrite as

$\mathbf{e} = \mathbf{Sb}$

where

- e is n by 1
- **b** is 3 by 1
- S is n by 3

Let the residual be

$r = e - Sb$

Squaring this:

$r^2 = r^T r = (e-Sb)^T (e-Sb)$
$= e^T e - 2b^T S^T e + b^T S^T S b$

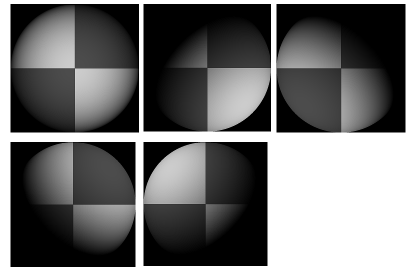
$(r^2)_b = 0$ - zero derivative is a necessary condition for a minimum, or

$-2S^T e + 2S^T S b = 0;$

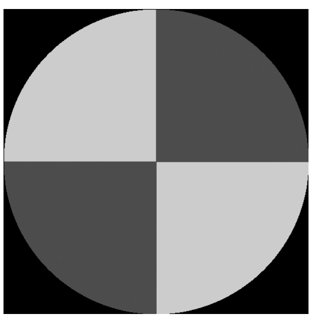
Solving for b gives

$$\boxed{\mathbf{b} = (\mathbf{S}^T\mathbf{S})^{-1}\mathbf{S}^T\mathbf{e}}$$

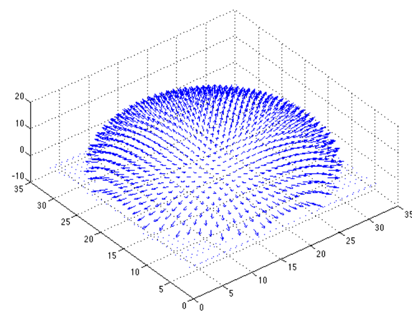
Input Images



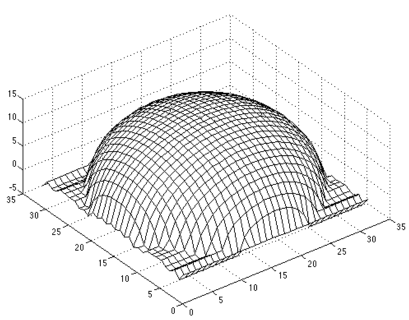
Recovered albedo



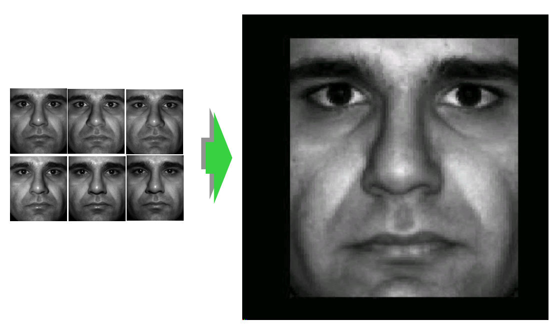
Recovered normal field



Surface recovered by integration



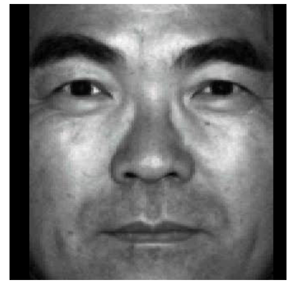
Lambertian Photometric Stereo



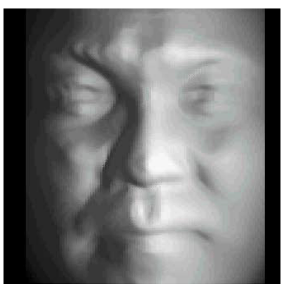
Reconsstruction with Albedo Map

Without the albedo map

Another person



No Albedo map



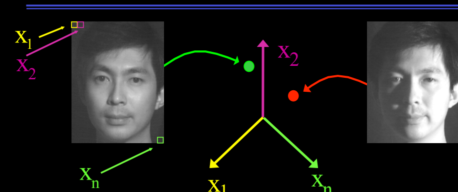
III. Photometric Stereo with unknown lighting and Lambertian surfaces

Covered in Illumination cone slides

What is the set of images of an object under all possible lighting conditions?

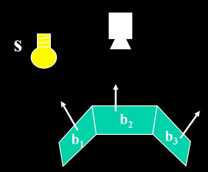
In answering this question, we'll arrive at a photometric setereo method for reconstructing surface shape w/ unknown lighting.

The Space of Images



- Consider an n-pixel image to be a point in an n-dimensional space, $\mathbf{x} \in \mathbb{R}^n$.
- Each pixel value is a coordinate of **x**.
- Many results will apply to linear transformations of image space (e.g. filtered images)
- Other image representations (e.g. Cayley-Klein spaces, See Koenderink's "pixel f#@king paper")

Model for Image Formation



Lambertian Assumption with shadowing:

$$\mathbf{x} = \max(\mathbf{B}\,\mathbf{s}, 0) \qquad \mathbf{B} = \begin{bmatrix} -\ \mathbf{b}^T_1\ - \\ -\ \mathbf{b}^T_2\ - \\ -\ \ldots\ldots \\ -\ \mathbf{b}^T_n\ - \end{bmatrix}_{n \times 3}$$

where

- $\mathbf{x}$ is an n-pixel image vector
- $\mathbf{B}$ is a matrix whose rows are unit normals scaled by the albedos
- $\mathbf{s} \in \mathbf{R}^3$ is vector of the light source direction scaled by intensity

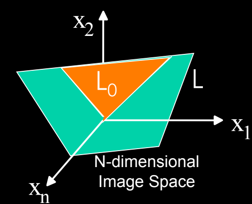
3-D Linear subspace

The set of images of a Lambertian surface with no shadowing is a subset of 3-D linear subspace.

[Moses 93], [Nayar, Murase 96], [Shashua 97]

$L = \{\mathbf{x} \mid \mathbf{x} = \mathrm{B}\mathbf{s}, \forall \mathbf{s} \in \mathrm{R}^3\}$

where B is a n by 3 matrix whose rows are product of the surface normal and Lambertian albedo

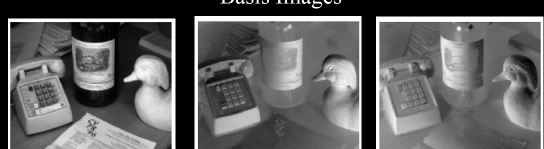


CSE252A

Still Life

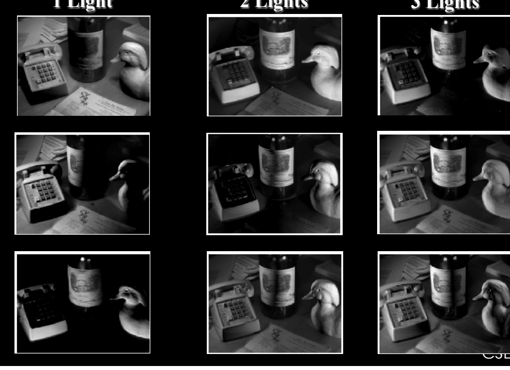
Original Images

Basis Images



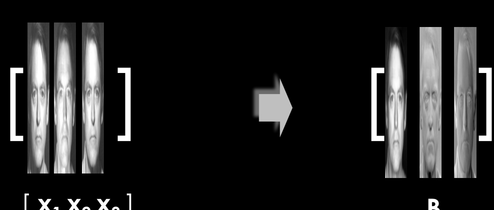
CSE252A

Rendering Images: $\Sigma_i \max(\mathbf{B}\mathbf{s}_i, 0)$



CSE252A

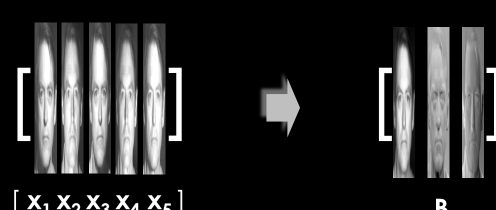
How do you construct subspace?



- Any three images w/o shadows taken under different lighting span L
- Not orthogonal
- Orthogonalize with Grahm-Schmidt

CSE252A

How do you construct subspace?



With more than three images, perform least squares estimation of **B** using Singular Value Decomposition (SVD)

CSE252A

Matrix Decompositions

- Definition: The factorization of a matrix $\boldsymbol{M}$ into two or more matrices $\boldsymbol{M}_1, \boldsymbol{M}_2, \ldots, \boldsymbol{M}_n$, such that $\boldsymbol{M} = \boldsymbol{M}_1\boldsymbol{M}_2\ldots\boldsymbol{M}_n$.
- Many decompositions exist…
  - ***QR*** Decomposition
  - ***LU*** Decomposition
  - ***LDU*** Decomposition
  - Etc.

Singular Value Decomposition

Excellent ref: 'Matrix Computations," Golub, Van Loan

- Any m by n matrix $\boldsymbol{A}$ may be factored such that

$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^T$$

$$[m \times n] = [m \times m][m \times n][n \times n]$$

- $\boldsymbol{U}$: m by m, orthogonal matrix
  - Columns of $\boldsymbol{U}$ are the eigenvectors of $\boldsymbol{AA}^T$
- $\boldsymbol{V}$: n by n, orthogonal matrix,
  - columns are the eigenvectors of $\boldsymbol{A}^T\boldsymbol{A}$
- $\boldsymbol{\Sigma}$: m by n, diagonal with non-negative entries ($\sigma_1, \sigma_2, \ldots, \sigma_s$) with s=min(m,n) are called the called the singular values
  - Singular values are the square roots of eigenvalues of both $\boldsymbol{AA}^T$ and $\boldsymbol{A}^T\boldsymbol{A}$
  - *Result of SVD algorithm:* $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_s$

SVD Properties

- In Matlab [u s v] = svd(A), and you can verify that: A=u*s*v'
- r=Rank($\boldsymbol{A}$) = # of non-zero singular values.
- $\boldsymbol{U}$, $\boldsymbol{V}$ give us orthonormal bases for the subspaces of $\boldsymbol{A}$:
  - 1st r columns of $\boldsymbol{U}$: Column space of $\boldsymbol{A}$
  - Last $m - r$ columns of $\boldsymbol{U}$: Left nullspace of $\boldsymbol{A}$
  - 1st r columns of $\boldsymbol{V}$: Row space of $\boldsymbol{A}$
  - last $n - r$ columns of $\boldsymbol{V}$: Nullspace of $\boldsymbol{A}$
- *For d<= r, the first d column of U provide the best d-dimensional basis for columns of A in least squares sense.*

Thin SVD

- Any m by n matrix $\boldsymbol{A}$ may be factored such that

$$\boldsymbol{A} = \boldsymbol{U\Sigma V}^T$$

$$[m \times n] = [m \times n][n \times n][n \times n]$$

- If m>n, then one can view $\boldsymbol{\Sigma}$ as:

$$\boxed{\begin{bmatrix} \Sigma' \\ 0 \end{bmatrix}}$$

- Where $\boldsymbol{\Sigma}'$ =diag($\sigma_1, \sigma_2, \ldots, \sigma_s$) with s=min(m,n), and lower matrix is (n-m by m) of zeros.
- Alternatively, you can write:

$$\boldsymbol{A} = \boldsymbol{U}'\boldsymbol{\Sigma}'\boldsymbol{V}^T$$

- In Matlab, thin SVD is:[U S V] = svds(A)

Estimating B with SVD

1. Construct data matrix

$$\mathbf{D} = [\, \mathbf{x}_1\, \mathbf{x}_2\, \mathbf{x}_3 \ldots \mathbf{x}_n \,]$$

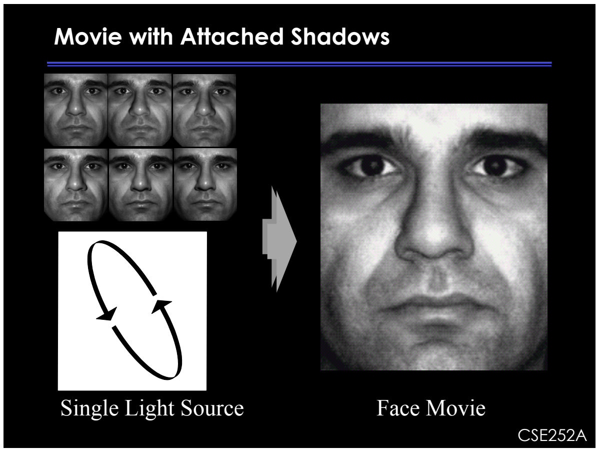
2. [**U S V**] = svds(D)

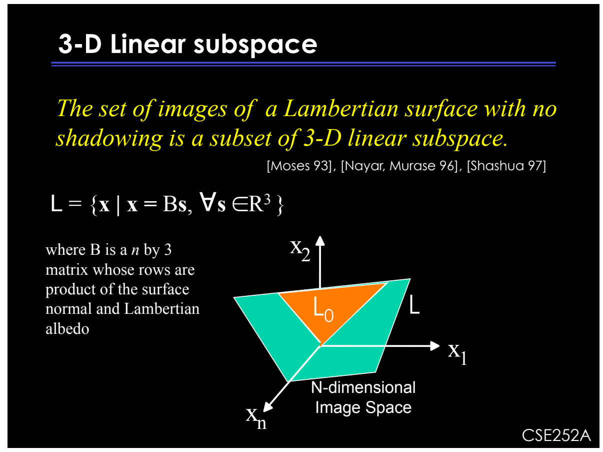
- If data had no noise, then rank(D)=3, and the first three singular values (S) would be positive and rest would be zero.
- Take first three column of **u** as **B.**

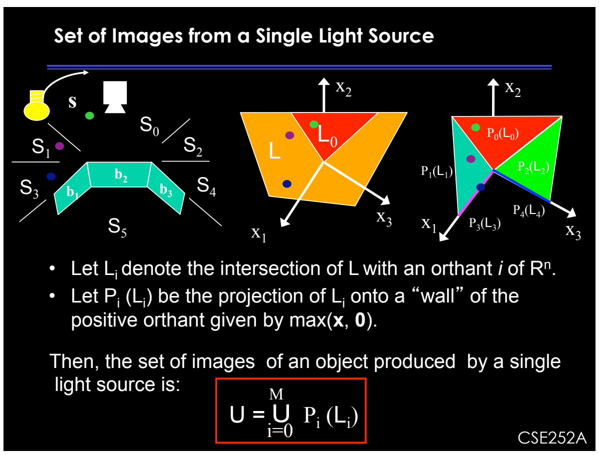
Face Basis

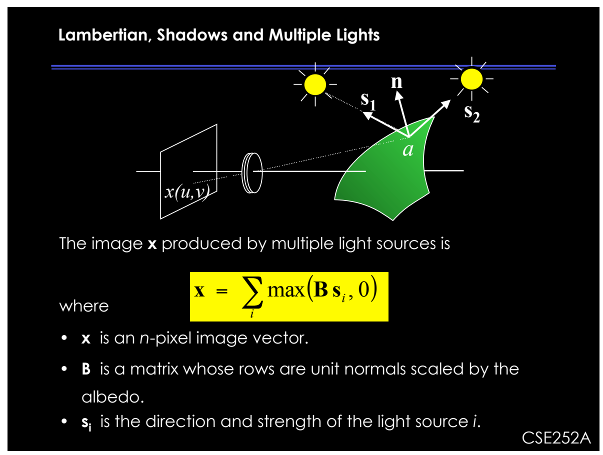
Original Images

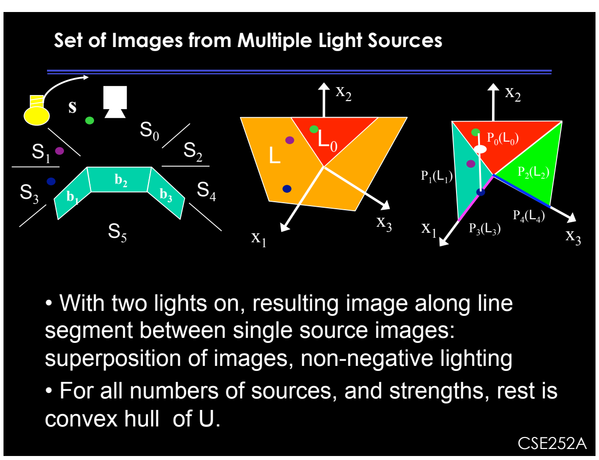
Basis Images

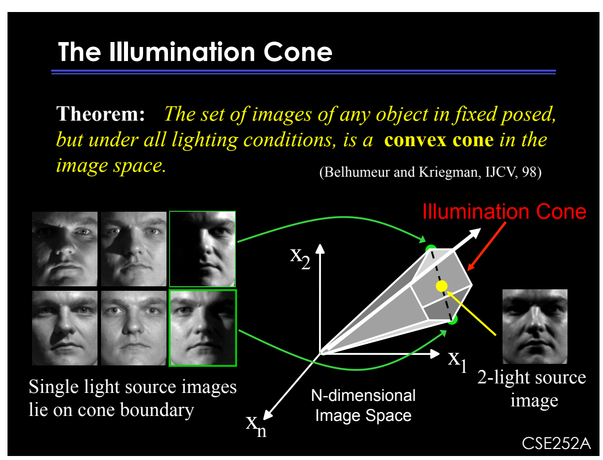

Movie with Attached Shadows
Single Light Source
Face Movie
CSE252A


3-D Linear subspace
The set of images of a Lambertian surface with no shadowing is a subset of 3-D linear subspace.
[Moses 93], [Nayar, Murase 96], [Shashua 97]
$L = \{\mathbf{x} \mid \mathbf{x} = B\mathbf{s}, \forall \mathbf{s} \in R^3\}$
where B is a n by 3 matrix whose rows are product of the surface normal and Lambertian albedo
x_2
L_0
L
x_1
x_n
N-dimensional Image Space
CSE252A

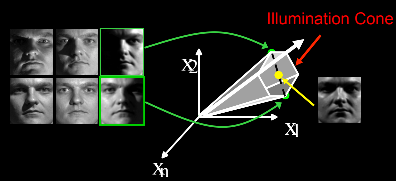

Set of Images from a Single Light Source
s
S_0
S_1
S_2
S_3
S_4
S_5
b_1
b_2
b_3
L
L_0
x_1
x_2
x_3
$P_0(L_0)$
$P_1(L_1)$
$P_2(L_2)$
$P_3(L_3)$
$P_4(L_4)$
• Let L_i denote the intersection of L with an orthant i of R^n.
• Let $P_i(L_i)$ be the projection of L_i onto a "wall" of the positive orthant given by max($\mathbf{x}$, $\mathbf{0}$).
Then, the set of images of an object produced by a single light source is:
$U = \bigcup_{i=0}^{M} P_i(L_i)$
CSE252A


Lambertian, Shadows and Multiple Lights
$\mathbf{s}_1$
$\mathbf{n}$
$\mathbf{s}_2$
a
$x(u,v)$
The image $\mathbf{x}$ produced by multiple light sources is
where
$\mathbf{x} = \sum_i \max(\mathbf{B}\,\mathbf{s}_i, 0)$
• $\mathbf{x}$ is an n-pixel image vector.
• $\mathbf{B}$ is a matrix whose rows are unit normals scaled by the albedo.
• $\mathbf{s}_i$ is the direction and strength of the light source i.
CSE252A


Set of Images from Multiple Light Sources
s
S_0
S_1
S_2
S_3
S_4
S_5
b_1
b_2
b_3
L
L_0
x_1
x_2
x_3
$P_0(L_0)$
$P_1(L_1)$
$P_2(L_2)$
$P_3(L_3)$
$P_4(L_4)$
• With two lights on, resulting image along line segment between single source images: superposition of images, non-negative lighting
• For all numbers of sources, and strengths, rest is convex hull of U.
CSE252A


The Illumination Cone
Theorem: The set of images of any object in fixed posed, but under all lighting conditions, is a convex cone in the image space.
(Belhumeur and Kriegman, IJCV, 98)
Illumination Cone
x_2
x_1
x_n
N-dimensional Image Space
Single light source images lie on cone boundary
2-light source image
CSE252A

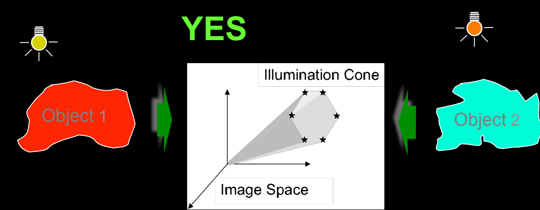
Some natural ideas & questions



- Can the cones of two different objects intersect?
- Can two different objects have the same cone?
- How big is the cone?
- How can cone be used for recognition?

Do Ambiguities Exist?

Can two objects of differing shapes produce the same illumination cone?



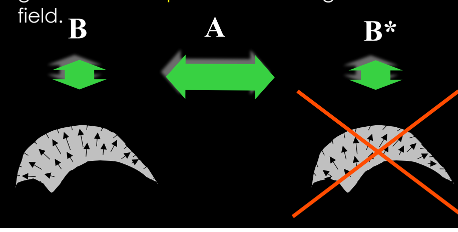
Do Ambiguities Exist? Yes

- Cone is determined by linear subspace **L**
- The columns of **B** span **L**
- For any **A** $\in$ GL(3), **B* = BA** also spans **L.**
- For any image of **B** produced with light source **S**, the same image can be produced by lighting **B*** with **S*=A**$^{-1}$**S** because
  X = B*S* = B AA^{-1}**S = BS**
- When we estimate B using SVD, the rows are NOT generally normal * albedo.
-

Surface Integrability

In general, **B*** does not have a corresponding surface.

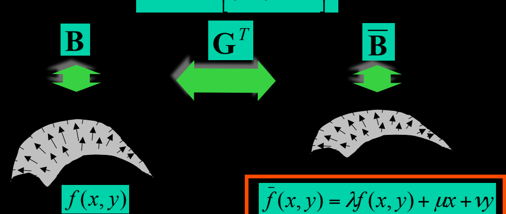
Linear transformations of the surface normals in general do not produce an integrable normal field.



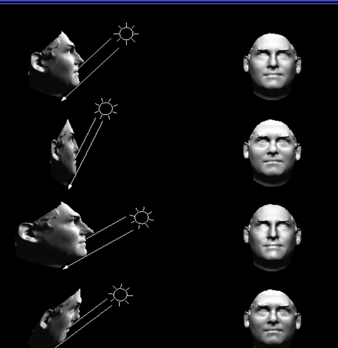
GBR Transformation

Only Generalized Bas-Relief transformations satisfy the integrability constraint:

$$\mathbf{A} = \mathbf{G}^T = \begin{bmatrix} \lambda & 0 & -\mu \\ 0 & \lambda & -\nu \\ 0 & 0 & 1 \end{bmatrix}^T$$



Generalized Bas-Relief Transformations



Objects differing by a GBR have the same illumination cone.

Without knowledge of light source location, one can only recover surfaces up to GBR transformations.

Uncalibrated photometric stereo

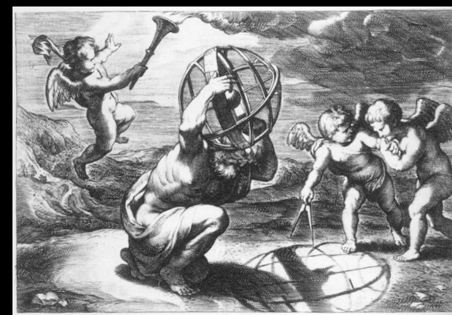
1. Take n images as input, perform SVD to compute B*.
2. Find some A such that B*A is close to integrable.
3. Integrate resulting gradient field to obtain height function f*(x,y).

Comments:

- f*(x,y) differs from f(x,y) by a GBR.
- Can use specularities to resolve GBR for non-Lambertian surface.

CSE252A

What about cast shadows for nonconvex objects?



P.P. Reubens in Opticorum Libri Sex, 1613

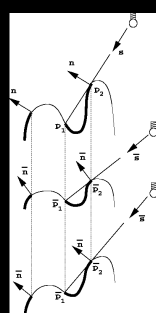
CSE252A

GBR Preserves Shadows

Given a surface f and a GBR transformed surface f' then for every light source **s** which illuminates f there exists a light source **s'** which illuminates f' such that the **attached** and **cast shadows** are identical.

GBR is the only transform that preserves shadows.

[Kriegman, Belhumeur 2001]



CSE252A

Bas-Relief Sculpture

CSE252A

Codex Urbinas

As far as light and shade are concerned low relief fails both as sculpture and as painting, because the shadows correspond to the low nature of the relief, as for example in the shadows of foreshortened objects, which will not exhibit the depth of those in painting or in sculpture in the round.

Leonardo da Vinci
Treatise on Painting
(Kemp)

CSE252A