Image Formation
and Cameras

Biometrics

CSE 190A

Lecture 5

Image Formation: Outline

- Factors in producing images
- Projection
- Perspective
- Vanishing points
- Orthographic
- Lenses
- Sensors
- Quantization/Resolution
- Illumination
- Reflectance

Earliest Surviving Photograph

- First photograph on record, “la table service” by Nicephore Niepce in 1822.
- Note: First photograph by Niepce was in 1816.

Images are two-dimensional patterns of brightness values.

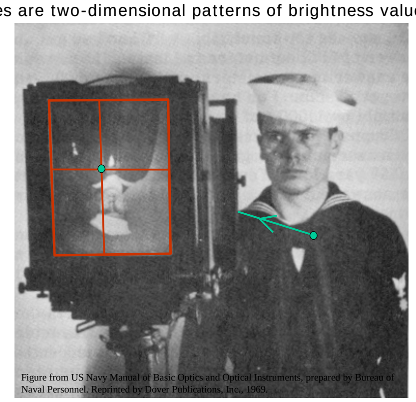


Figure from US Navy Manual of Basic Optics and Optical Instruments, prepared by Bureau of Naval Personnel. Reprinted by Dover Publications, Inc., 1969.

They are formed by the projection of 3D objects.

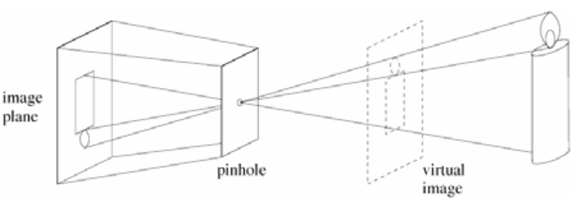
Effect of Lighting: Monet

Change of Viewpoint: Monet

Haystack at Chailly at Sunrise (1865)

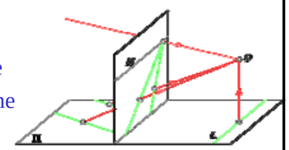
Pinhole Camera: **Perspective projection**

- Abstract camera model - box with a small hole in it



Geometric properties of projection

- Points go to points
- Lines go to lines
- Planes go to whole image or half-plane
- Polygons go to polygons
- Angles & distances not preserved



- Degenerate cases:
  - line through focal point yields point
  - plane through focal point yields line

Parallel lines meet in the image

- vanishing point

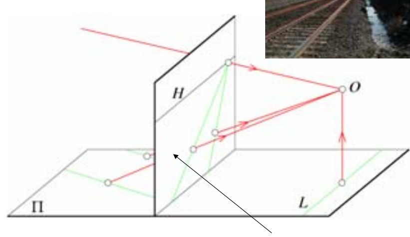
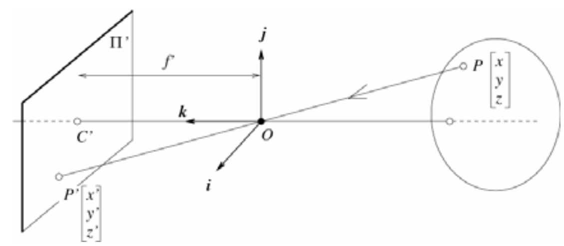


Image plane

The equation of projection



Cartesian coordinates:

- We have, by similar triangles, that (x, y, z) -> (f x/z, f y/z, -f)
- Ignore the third coordinate, and get

$$(x, y, z) \rightarrow (f\frac{x}{z}, f\frac{y}{z})$$

A Digression

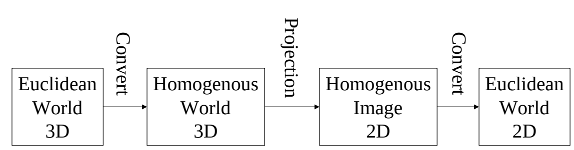
Homogenous Coordinates

and

Camera Matrices

Homogenous coordinates

- Our usual coordinate system is called a Euclidean or affine coordinate system
- Rotations, translations and projection in Homogenous coordinates can be expressed linearly as matrix multiplies



Homogenous coordinates
A way to represent points in a projective space

1. Add an extra coordinate
   e.g., (x,y) -> (x,y,1)=(u,v,w)
2. Impose equivalence relation such that (λ not 0)
   (u,v,w) ≈ λ*(u,v,w)
   i.e., (x,y,1) ≈ (λx, λy, λ)
3. "Point at infinity" – zero for last coordinate
   e.g., (x,y,0)

- Why do this?
  - Possible to represent points "at infinity"
    - Where parallel lines intersect
    - Where parallel planes intersect
  - Possible to write the action of a perspective camera as a matrix

Euclidean -> Homogenous-> Euclidean

In 2-D
- Euclidean -> Homogenous: (x, y) -> k (x,y,1)
- Homogenous -> Euclidean: (u,v,w) -> (u/w, v/w)

In 3-D
- Euclidean -> Homogenous: (x, y, z) -> k (x,y,z,1)
- Homogenous -> Euclidean: (x, y, z, w) -> (x/w, y/w, z/w)

The camera matrix

$$(x,y,z) \rightarrow \left(f\frac{x}{z}, f\frac{y}{z}\right)$$

Turn this expression into homogenous coordinates
- HC's for 3D point are (X,Y,Z,T)
- HC's for point in image are (U,V,W)

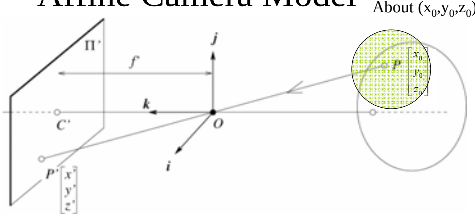
$$\begin{pmatrix}U\\V\\W\end{pmatrix} = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & \frac{1}{f} & 0\end{pmatrix}\begin{pmatrix}X\\Y\\Z\\T\end{pmatrix}$$

Perspective Camera Matrix
A 3x4 matrix

End of the Digression

Affine Camera Model

Appropriate in Neighborhood About (x_0,y_0,z_0)



- Take Perspective projection equation, and perform Taylor Series Expansion about (some point (x_0,y_0,z_0)).
- Drop terms of higher order than linear.
- Resulting expression is affine camera model

- Perspective

$$\begin{bmatrix}u\\v\end{bmatrix} = \frac{f}{z}\begin{bmatrix}x\\y\end{bmatrix}$$

- Assume that $f=1$, and perform a Taylor series expansion about (x_0, y_0, z_0)

$$\begin{bmatrix}u\\v\end{bmatrix} = \frac{1}{z_0}\begin{bmatrix}x_0\\y_0\end{bmatrix} - \frac{1}{z_0^2}\begin{bmatrix}x_0\\y_0\end{bmatrix}(z-z_0) + \frac{1}{z_0}\begin{bmatrix}1\\0\end{bmatrix}(x-x_0) + \frac{1}{z_0}\begin{bmatrix}0\\1\end{bmatrix}(y-y_0) + \frac{1}{2}\frac{2}{z_0^3}\begin{bmatrix}x_0\\y_0\end{bmatrix}(z-z_0)^2 + \cdots$$

- Dropping higher order terms and regrouping.

$$\begin{bmatrix}u\\v\end{bmatrix} \approx \frac{1}{z_0}\begin{bmatrix}x_0\\y_0\end{bmatrix} + \begin{bmatrix}1/z_0 & 0 & -x_0/z_0^2\\ 0 & 1/z_0 & -y_0/z_0^2\end{bmatrix}\begin{bmatrix}x\\y\\z\end{bmatrix} = \mathbf{Ap} + \mathbf{b}$$

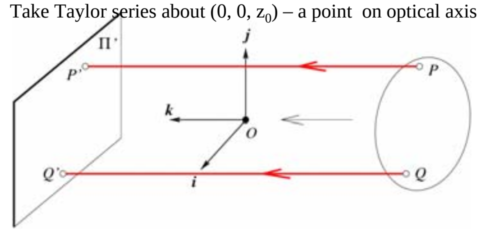
$$\begin{bmatrix} u \\ v \end{bmatrix} \approx \frac{1}{z_0}\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + \begin{bmatrix} 1/z_0 & 0 & -x_0/z_0^2 \\ 0 & 1/z_0 & -y_0/z_0^2 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{Ap} + \mathbf{b}$$

Rewrite Affine camera model
in terms of Homogenous Coordinates

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \approx \begin{bmatrix} 1/z_0 & 0 & -x_0/z_0^2 & x_0/z_0 \\ 0 & 1/z_0 & -y_0/z_0^2 & y_0/z_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Orthographic projection

Starting with Affine camera mode
Take Taylor series about (0, 0, z_0) – a point on optical axis



$$\begin{bmatrix} u \\ v \end{bmatrix} = \frac{1}{z_0}\begin{bmatrix} x \\ y \end{bmatrix}$$

The projection matrix for orthographic projection

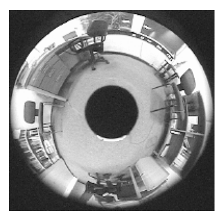
$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1/z_0 & 0 & 0 & 0 \\ 0 & 1/z_0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Parallel lines project to parallel lines
Ratios of distances are preserved under orthographic

Other camera models

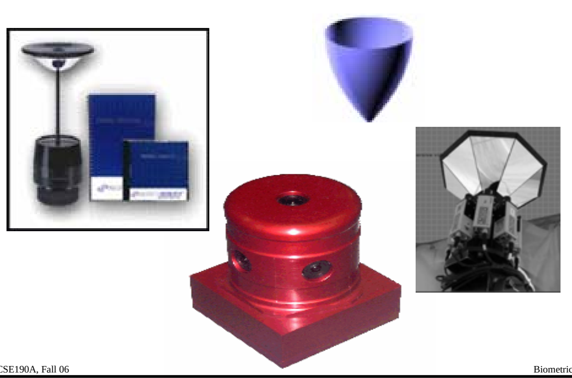
- Generalized camera – maps points lying on rays and maps them to points on the image plane.

Omnicam (hemispherical)

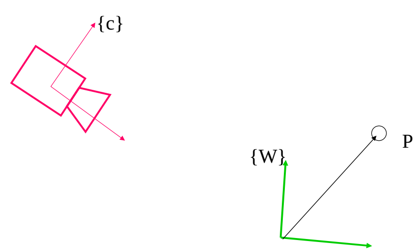


Light Probe (spherical)

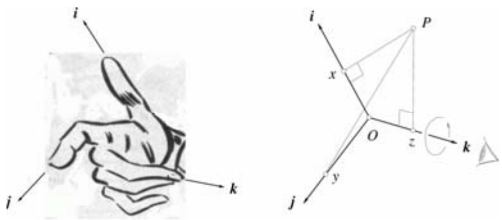
Some Alternative "Cameras"



What if camera coordinate system differs from object coordinate system

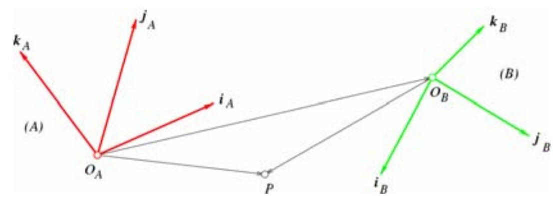


Euclidean Coordinate Systems



$$\begin{cases} x = \overrightarrow{OP}.\mathbf{i} \\ y = \overrightarrow{OP}.\mathbf{j} \\ z = \overrightarrow{OP}.\mathbf{k} \end{cases} \Leftrightarrow \quad \overrightarrow{OP} = x\mathbf{i} + y\mathbf{j} + z\mathbf{k} \quad \Leftrightarrow \quad \mathbf{P} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Coordinate Changes: Rigid Transformations



$${}^{B}P = {}^{B}_{A}R \; {}^{A}P + {}^{B}O_{A}$$

Rotation Matrix Translation vector

A rotation matrix ***R*** has the following properties:

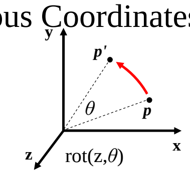
- Its inverse is equal to its transpose $R^{-1} = R^T$
- its determinant is equal to 1: ***det(R)=1.***

Or equivalently:

- Rows (or columns) of ***R*** form a right-handed orthonormal coordinate system.

Rotation: Homogenous Coordinates

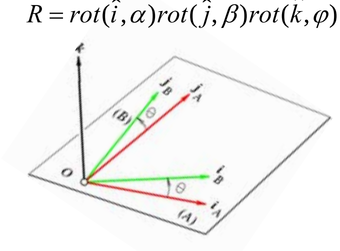
- About z axis



$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Roll-Pitch-Yaw

$$R = rot(\hat{i}, \alpha) rot(\hat{j}, \beta) rot(\hat{k}, \varphi)$$

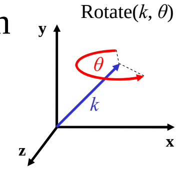


Euler Angles

$$R = rot(\hat{k}'', \alpha) rot(\hat{j}', \beta) rot(\hat{k}, \varphi)$$

Rotation

Rotate(k, θ)



- About (k_x, k_y, k_z), a unit vector on an arbitrary axis (Rodrigues Formula)

$$\begin{pmatrix} x' \\ y' \\ z' \\ 1 \end{pmatrix} = \begin{pmatrix} k_x k_x(1-c)+c & k_z k_x(1-c)-k_z s & k_x k_z(1-c)+k_y s & 0 \\ k_y k_x(1-c)+k_z s & k_z k_x(1-c)+c & k_y k_z(1-c)-k_x s & 0 \\ k_z k_x(1-c)-k_y s & k_z k_x(1-c)-k_x s & k_z k_z(1-c)+c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

where $c = \cos\theta$ & $s = \sin\theta$

Block Matrix Multiplication

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \qquad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

What is AB ?

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}$$

Homogeneous Representation of Rigid Transformations

$$\begin{bmatrix} {}^{B}P \\ 1 \end{bmatrix} = \begin{bmatrix} {}^{B}_{A}R\,{}^{A}P + {}^{B}O_A \\ 1 \end{bmatrix} = \begin{bmatrix} {}^{B}_{A}R & {}^{B}O_A \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} {}^{A}P \\ 1 \end{bmatrix} = {}^{B}_{A}T \begin{bmatrix} {}^{A}P \\ 1 \end{bmatrix}$$

Transformation represented by 4 by 4 Matrix

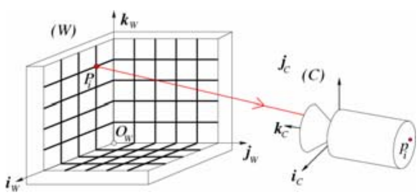
Camera parameters

- Issue
  - camera may not be at the origin, looking down the z-axis
    - extrinsic parameters (Rigid Transformation)
  - one unit in camera coordinates may not be the same as one unit in world coordinates
    - intrinsic parameters - focal length, principal point, aspect ratio, angle between axes, etc.

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{intrinsic parameters} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{extrinsic parameters} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

3 x 3 4 x 4

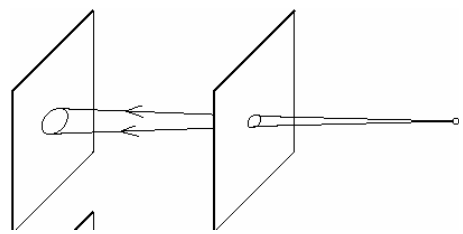
Camera Calibration



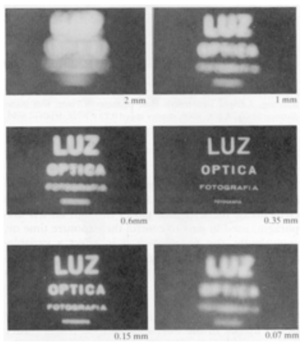
Given n points $P_1, \ldots, P_n$ with *known* positions and their images $p_1, \ldots, p_n$, estimate intrinsic and extrinsic camera parameters

- See Text book for how to do it.

Getting more light – Bigger Aperture



Pinhole Camera Images with Variable Aperture

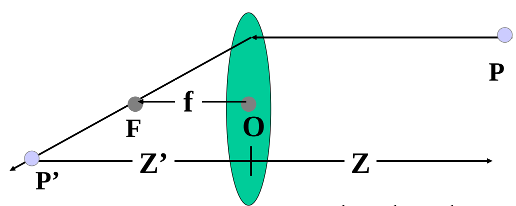


2 mm **1mm**

.6 mm **.35 mm**

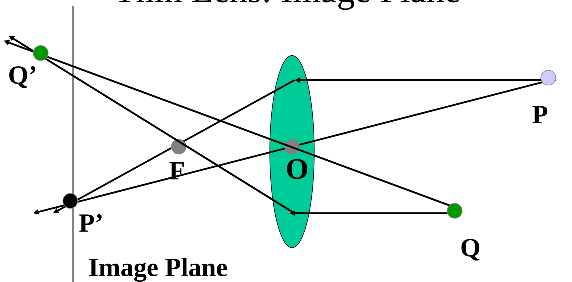
.15 mm **.07 mm**

Thin Lens: Image of Point



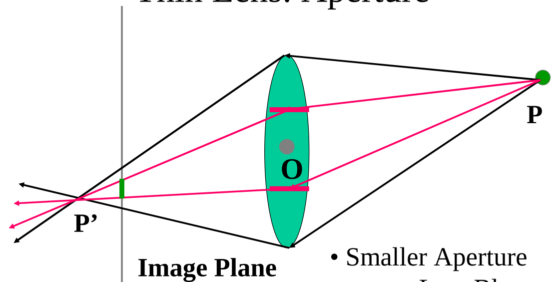
$$\frac{1}{z'} - \frac{1}{z} = \frac{1}{f}$$

Thin Lens: Image Plane



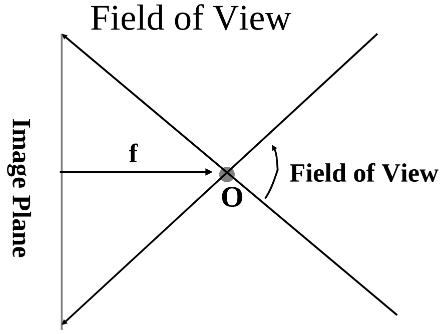
A price: Whereas the image of **P** is in focus, the image of **Q** isn't.

Thin Lens: Aperture



- Smaller Aperture -> Less Blur
- Pinhole -> No Blur

Field of View



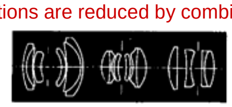
Deviations from the lens model

Deviations from this ideal are *aberrations*

Two types

1. geometrical
   - spherical aberration
   - astigmatism
   - distortion
   - coma
2. chromatic

Aberrations are reduced by combining lenses

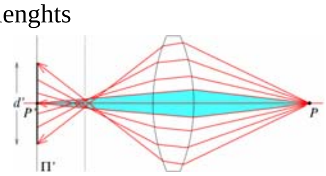


Compound lenses

Spherical aberration

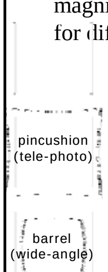
rays parallel to the axis do not converge

outer portions of the lens yield smaller focal lenghts



Distortion

magnification/focal length different for different angles of inclination



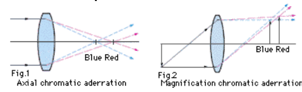
Can be corrected! (if parameters are know)

Chromatic aberration

Index of refraction of lens depends on wavelength of light

Chromatic aberration

rays of different wavelengths focused in different planes



The image is blurred and appears colored at the fringe.

cannot be removed completely

sometimes *achromatization* is achieved for more than 2 wavelengths

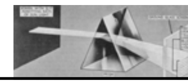
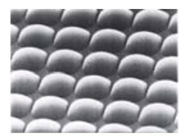
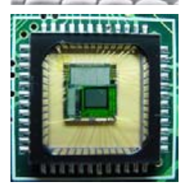


Image Brightness

How Cameras Produce Images

- Basic process:
  - photons hit a detector
  - the detector becomes charged
  - the charge is read out as brightness
- Sensor types:
  - CCD (charge-coupled device)
    - high sensitivity
    - high power
    - cannot be individually addressed
    - blooming
  - CMOS
    - most common
    - simple to fabricate (cheap)
    - lower sensitivity, lower power
    - can be individually addressed





Camera's sensor

- Measured pixel intensity is a function of irradiance integrated over
  - pixel's area
  - over a range of wavelengths
  - For some time

$$I = \int_t \int_\lambda \int_x \int_y E(x, y, \lambda, t) s(x, y) q(\lambda) \, dy\, dx\, d\lambda\, dt$$

Light at surfaces

Many effects when light strikes a surface -- could be:

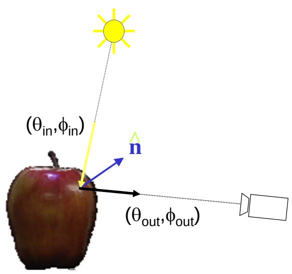
- transmitted
  - Skin, glass
- reflected
  - mirror
- scattered
  - milk
- travel along the surface and leave at some other point
- absorbed
  - sweaty skin

Assume that

- surfaces don't fluoresce
  - e.g. scorpions, detergents
- surfaces don't emit light (i.e. are cool)
- all the light leaving a point is due to that arriving at that point

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

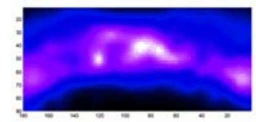


Surface Reflectance Models

Common Models

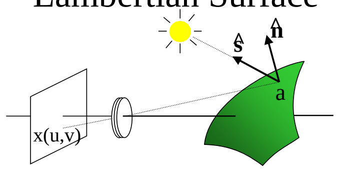
- Lambertian
- Phong
- Physics-based
  - Specular
    [Blinn 1977], [Cook-Torrance 1982], [Ward 1992]
  - Diffuse
    [Hanrahan, Kreuger 1993]
  - Generalized Lambertian
    [Oren, Nayar 1995]
  - Thoroughly Pitted Surfaces
    [Koenderink et al 1999]
- Phenomenological
  [Koenderink, Van Doorn 1996]

Arbitrary Reflectance



- Non-parametric model
- Anisotropic
- Non-uniform over surface
- BRDF Measurement *[Dana et al, 1999], [Marschner]*

Lambertian Surface



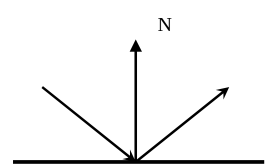
At image location (u,v), the intensity of a pixel x(u,v) is:

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

where

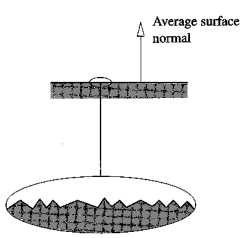
- a(u,v) is the albedo of the surface projecting to (u,v).
- **n**(u,v) is the direction of the surface normal.
- s_0 is the light source intensity.
- $\hat{s}$ is the direction to the light source.

Specular Reflection: Smooth Surface

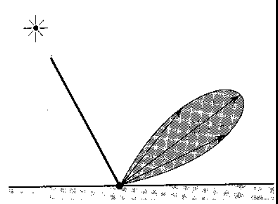


Phong – rough, specular

Rough Specular Surface



Symmetric V-shaped grooves – 'microfacets'



Phong Lobe