Camera calibration & radiometry

- Reading:
  - Chapter 2, and section 5.4, Forsyth & Ponce
  - Chapter 10, Horn
- Optional reading:
  - Chapter 4, Forsyth & Ponce

Sept. 12, 2002
MIT 6.801/6.866
Profs. Freeman and Darrell

6.801/6.866 Machine Vision

Syllabus

#	Date	Description	Readings	Assignments	Materials
1	9/5	Course Introduction		Pset #0 (not collected)	Freeman Slides Darrell Slides Matlab Tutorial Diary
2	9/10	Cameras, Lenses, and Sensors	Req: FP 1 Opt: H 2.1, 2.3		Slides (ppt, html)
3	9/12	Radiometry and Shading Models	Req: FP 2, 5.4, H 10 Opt: FP 4	Pset #1 Assigned	
4	9/17	Color	Req: FP 6		

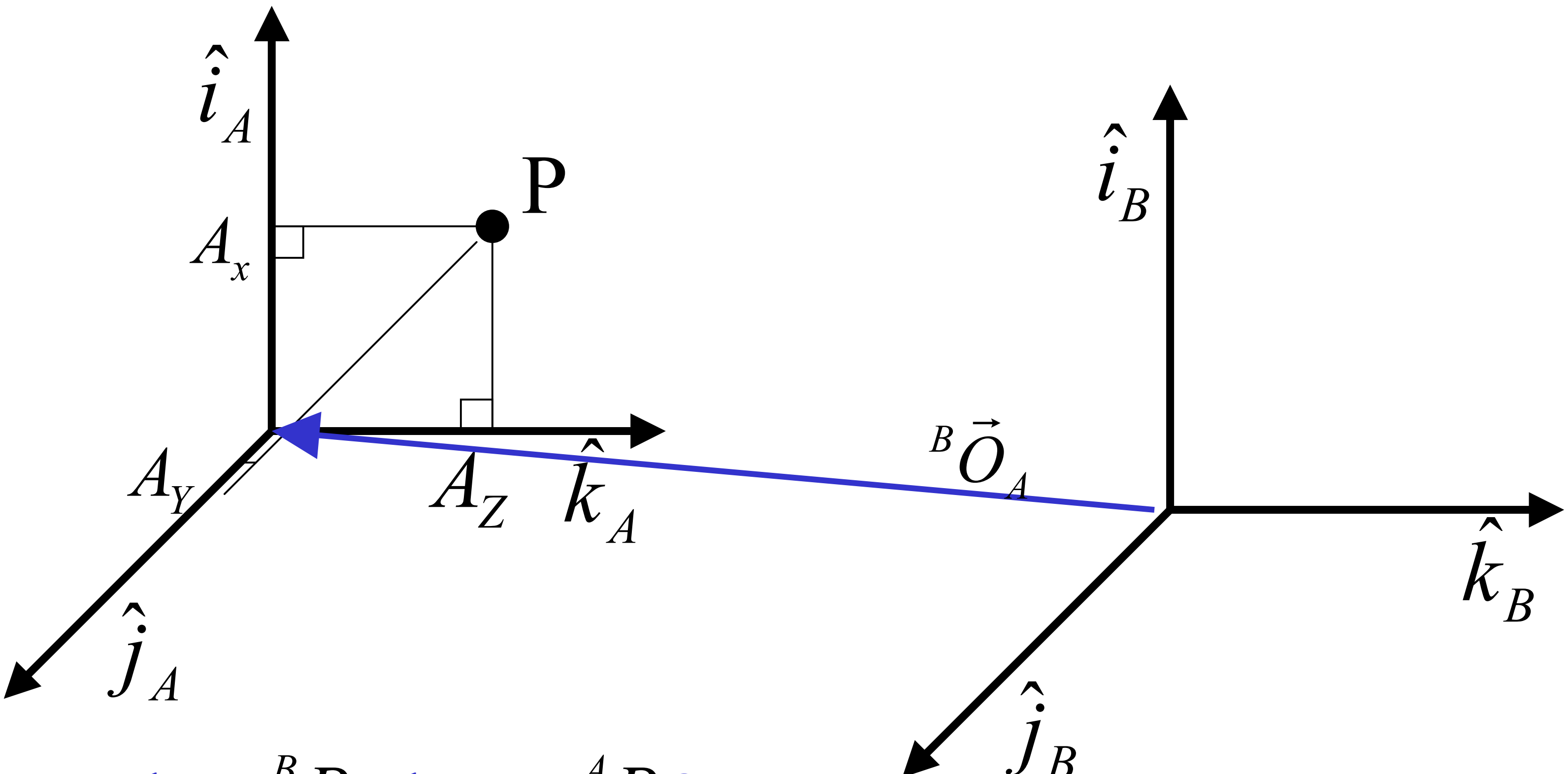
Photo to learn names

Today's class

- First part: how *positions* in the image relate to 3-d positions in the world.
- Second part: how the *intensities* in the image relate surface and lighting properties in the world.

Translation

$$^{A}P = \begin{pmatrix} A_x \\ A_Y \\ A_Z \end{pmatrix} \qquad ^{B}P = \begin{pmatrix} B_x \\ B_Y \\ B_Z \end{pmatrix}$$

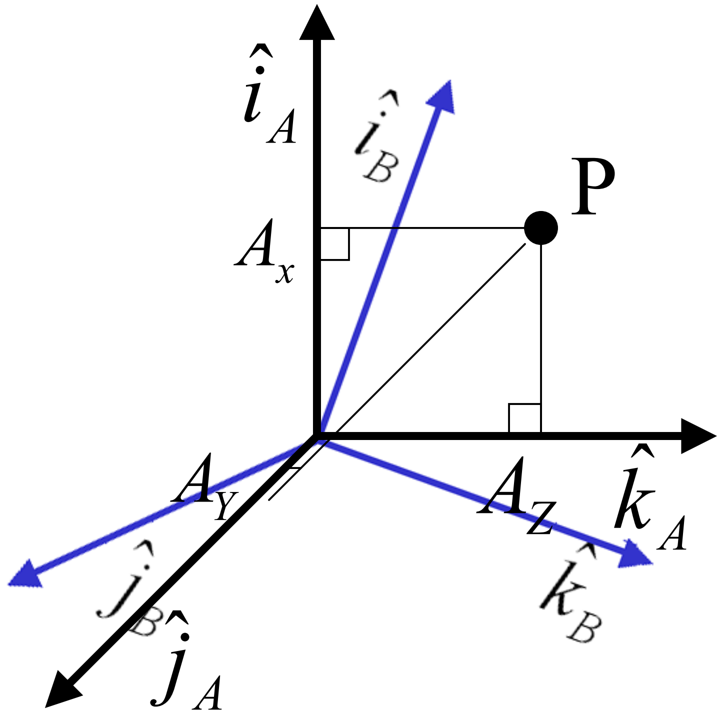


How does ^{B}P relate to ^{A}P ?

$$^{B}P = {}^{A}P + {}^{B}O_A$$

Rotation

$$^{A}P = \begin{pmatrix} A_x \\ A_Y \\ A_Z \end{pmatrix} \qquad ^{B}P = \begin{pmatrix} B_x \\ B_Y \\ B_Z \end{pmatrix}$$

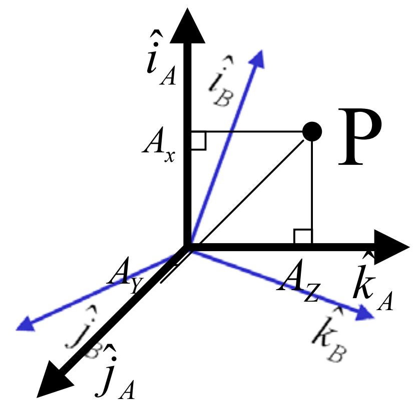


How does ^{B}P relate to ^{A}P ?

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

Find the rotation matrix

Project $\overrightarrow{OP} = \begin{pmatrix} \hat{i}_A & \hat{j}_A & \hat{k}_A \end{pmatrix} \begin{pmatrix} A_X \\ A_Y \\ A_Z \end{pmatrix}$



onto the B frame's coordinate axes.

$$\begin{pmatrix} B_X \\ B_Y \\ B_Z \end{pmatrix} = \begin{pmatrix} \hat{i}_B \bullet \hat{i}_A A_X & \hat{i}_B \bullet \hat{j}_A A_Y & \hat{i}_B \bullet \hat{k}_A A_Z \\ \hat{j}_B \bullet \hat{i}_A A_X & \hat{j}_B \bullet \hat{j}_A A_Y & \hat{j}_B \bullet \hat{k}_A A_Z \\ \hat{k}_B \bullet \hat{i}_A A_X & \hat{k}_B \bullet \hat{j}_A A_Y & \hat{k}_B \bullet \hat{k}_A A_Z \end{pmatrix}$$

Rotation matrix

this

$$\begin{pmatrix} B_X \\ B_Y \\ B_Z \end{pmatrix} = \begin{pmatrix} \hat{i}_B \bullet \hat{i}_A A_X & \hat{i}_B \bullet \hat{j}_A A_Y & \hat{i}_B \bullet \hat{k}_A A_Z \\ \hat{j}_B \bullet \hat{i}_A A_X & \hat{j}_B \bullet \hat{j}_A A_Y & \hat{j}_B \bullet \hat{k}_A A_Z \\ \hat{k}_B \bullet \hat{i}_A A_X & \hat{k}_B \bullet \hat{j}_A A_Y & \hat{k}_B \bullet \hat{k}_A A_Z \end{pmatrix}$$

implies

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

where

$$ {}^{B}_{A}R = \begin{pmatrix} \hat{i}_B \bullet \hat{i}_A & \hat{i}_B \bullet \hat{j}_A & \hat{i}_B \bullet \hat{k}_A \\ \hat{j}_B \bullet \hat{i}_A & \hat{j}_B \bullet \hat{j}_A & \hat{j}_B \bullet \hat{k}_A \\ \hat{k}_B \bullet \hat{i}_A & \hat{k}_B \bullet \hat{j}_A & \hat{k}_B \bullet \hat{k}_A \end{pmatrix}$$

Translation and rotation

Let's write $^{B}P = {}^{B}_{A}R\ {}^{A}P + {}^{B}O_{A}$

as a single matrix equation:

$$\begin{pmatrix} B_X \\ B_Y \\ B_Z \\ 1 \end{pmatrix} = \begin{pmatrix} \boxed{\begin{matrix} - & - & - \\ - & {}^{B}_{A}R & - \\ - & - & - \end{matrix}} & \boxed{\begin{matrix} | \\ {}^{B}O_{A} \\ | \end{matrix}} \\ \begin{matrix} 0 & 0 & 0 \end{matrix} & 1 \end{pmatrix} \begin{pmatrix} A_X \\ A_Y \\ A_Z \\ 1 \end{pmatrix}$$

Homogenous coordinates

- Add an extra coordinate and use an equivalence relation
- for 3D
  - equivalence relation k*(X,Y,Z,T) is the same as (X,Y,Z,T)
- Motivation
  - Possible to write the action of a perspective camera as a matrix

Homogenous/non-homogenous transformations for a 3-d point

- From non-homogenous to homogenous coordinates: add 1 as the 4th coordinate, ie

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

- From homogenous to non-homogenous coordinates: divide 1st 3 coordinates by the 4th, ie

$$\begin{pmatrix} x \\ y \\ z \\ T \end{pmatrix} \rightarrow \frac{1}{T} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Homogenous/non-homogenous transformations for a 2-d point

- From non-homogenous to homogenous coordinates: add 1 as the 3rd coordinate, ie

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

- From homogenous to non-homogenous coordinates: divide 1st 2 coordinates by the 3rd, ie

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \frac{1}{z} \begin{pmatrix} x \\ y \end{pmatrix}$$

The camera matrix, in homogenous coordinates

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

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

$$\begin{pmatrix} X \\ Y \\ \frac{Z}{f} \end{pmatrix} \rightarrow \frac{f}{Z} \begin{pmatrix} X \\ Y \end{pmatrix}$$

HC Non-HC

The projection matrix for orthographic projection, homogenous coordinates

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

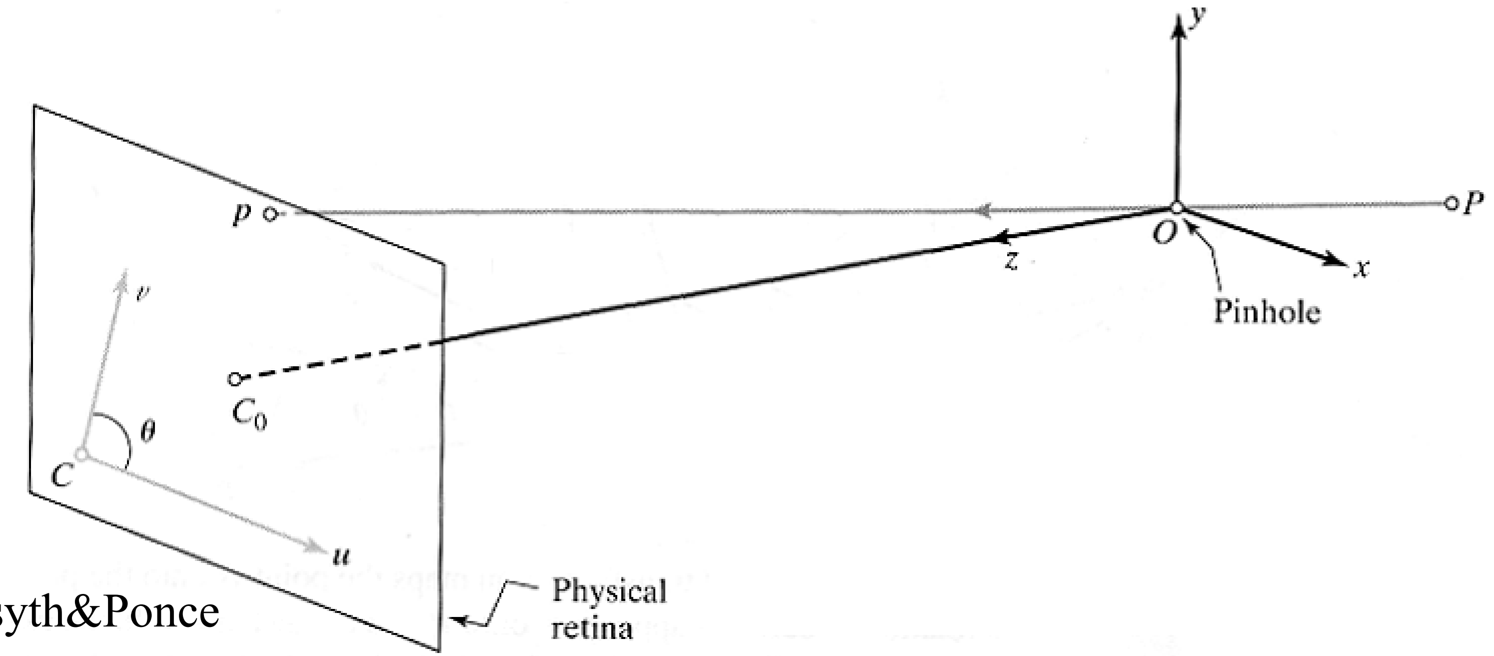
$$= \begin{pmatrix} X \\ Y \\ T \end{pmatrix} \rightarrow \frac{1}{T} \begin{pmatrix} X \\ Y \end{pmatrix}$$

HC   Non-HC

Camera calibration

- Use the camera to tell you things about the world.
  - Relationship between coordinates in the world and coordinates in the image: geometric camera calibration.
  - (Later we'll discuss relationship between intensities in the world and intensities in the image: photometric camera calibration.)

Intrinsic parameters



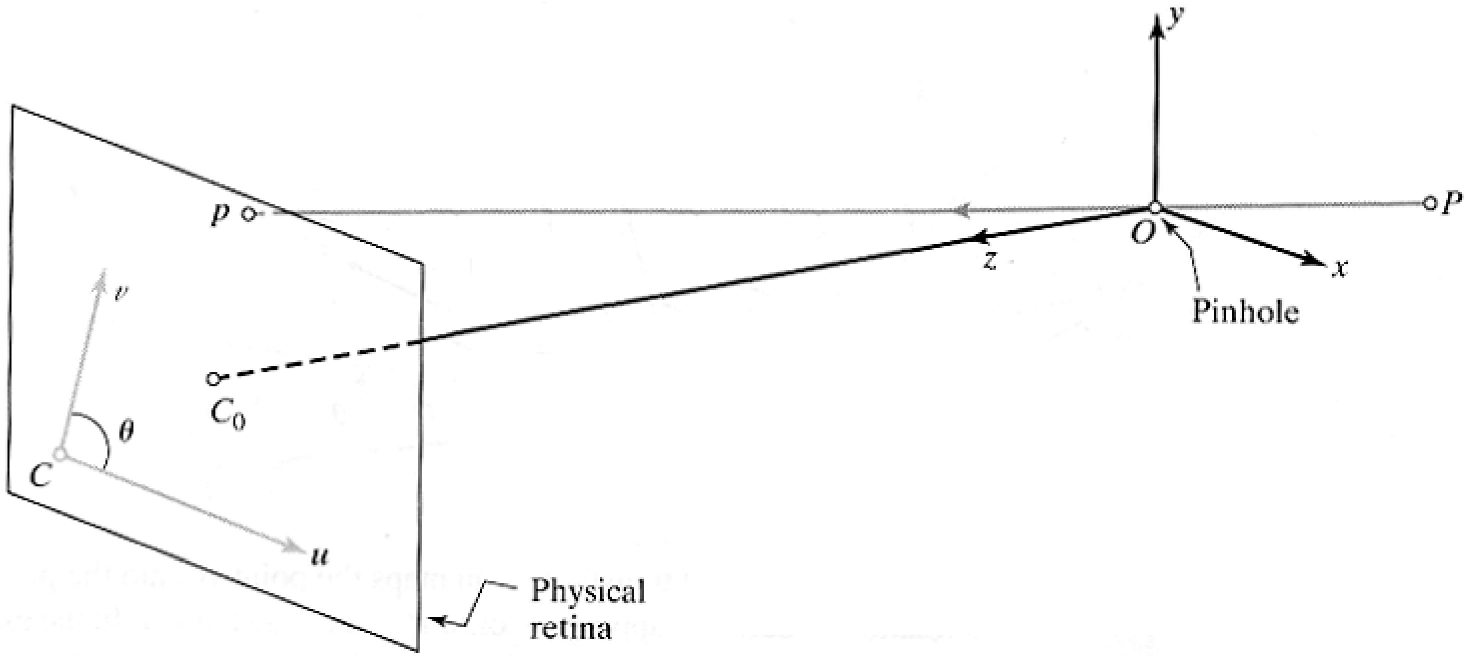
Forsyth&Ponce

Perspective projection

$$u = f\frac{x}{z}$$

$$v = f\frac{y}{z}$$

Intrinsic parameters

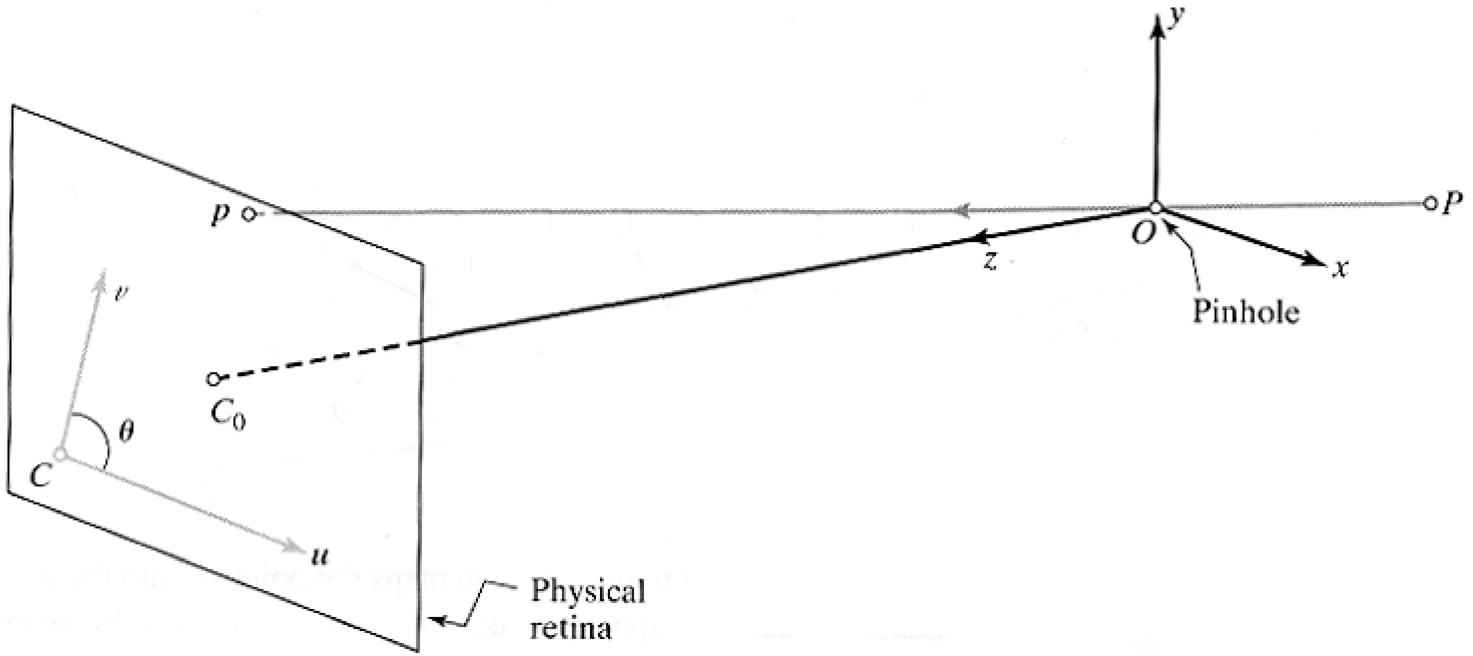


But “pixels” are in some arbitrary spatial units

$$u = \alpha \frac{x}{z}$$

$$v = \alpha \frac{y}{z}$$

Intrinsic parameters

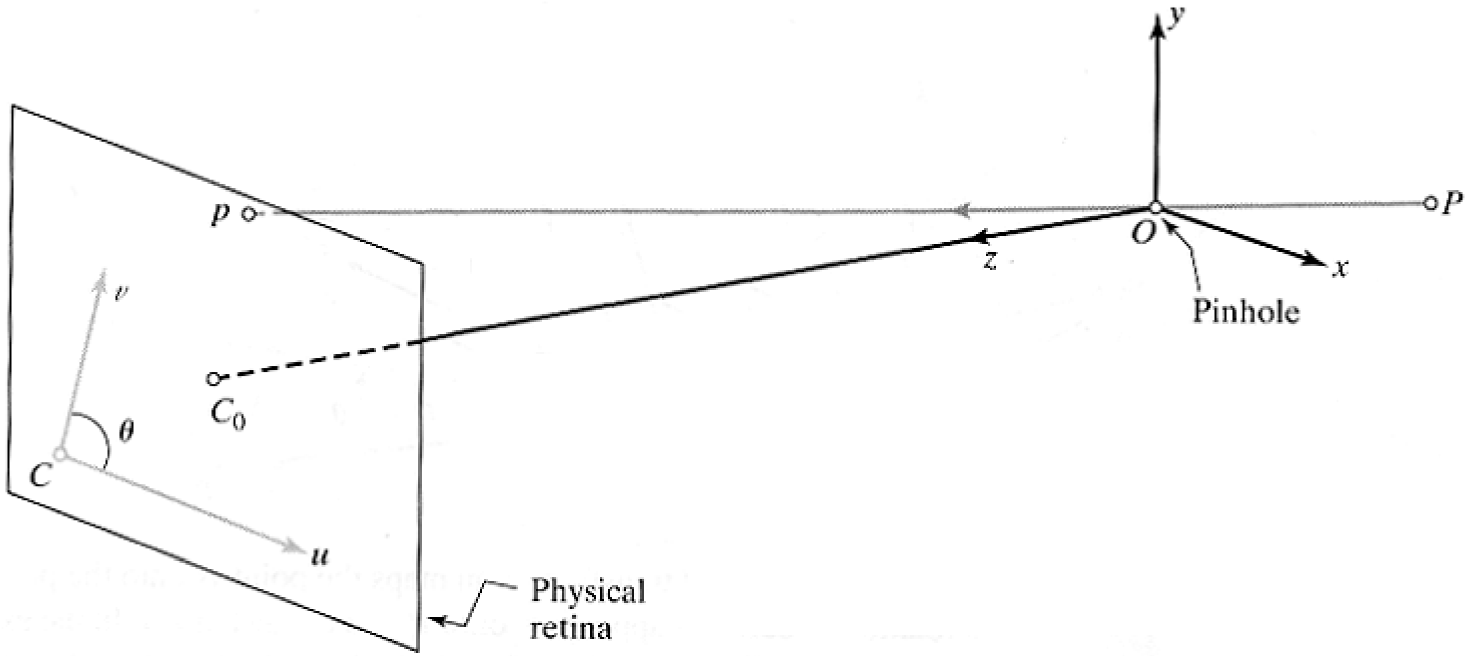


Maybe pixels are not square

$$u = \alpha \frac{x}{z}$$

$$v = \beta \frac{y}{z}$$

Intrinsic parameters

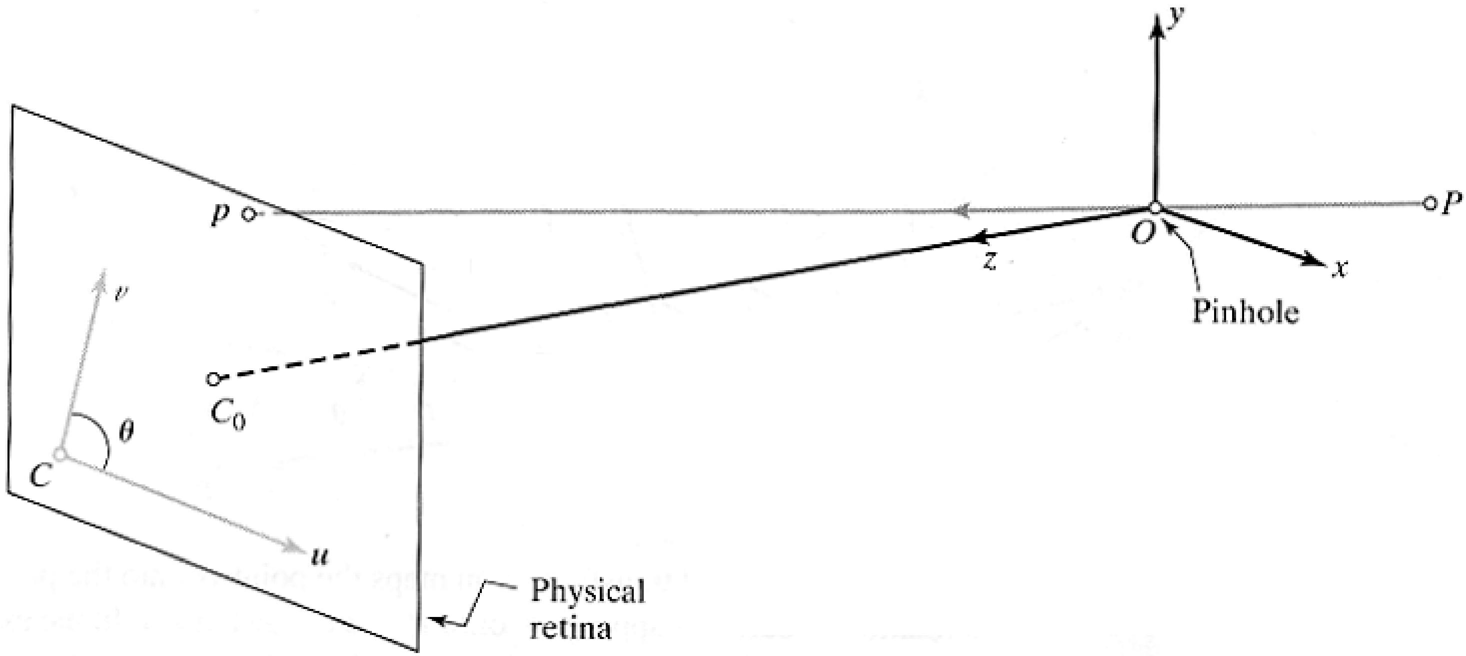


We don't know the origin of our camera pixel coordinates

$$u = \alpha \frac{x}{z} + u_0$$

$$v = \beta \frac{y}{z} + v_0$$

Intrinsic parameters

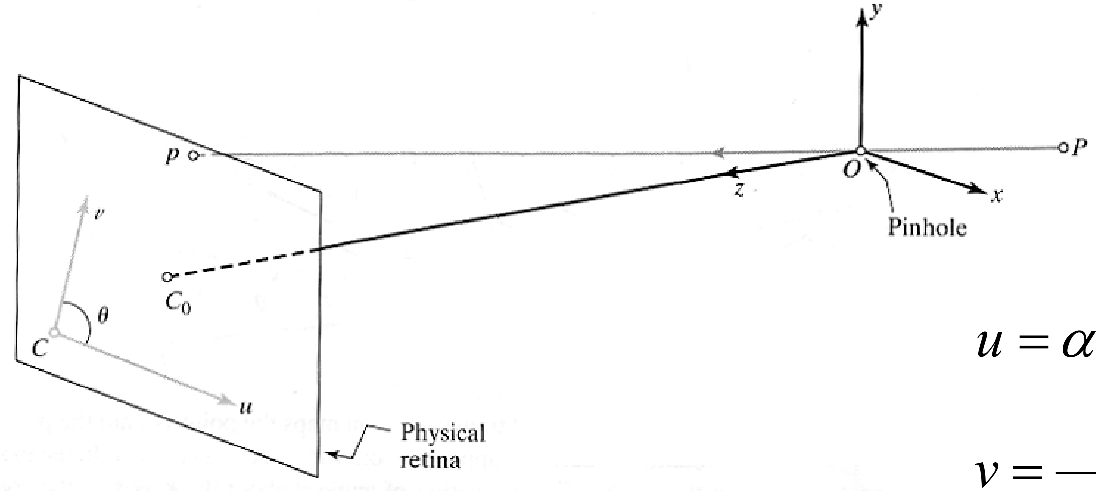


May be skew between camera pixel axes

$$u = \alpha \frac{x}{z} - \alpha \cot(\theta) \frac{y}{z} + u_0$$

$$v = \frac{\beta}{\sin(\theta)} \frac{y}{z} + v_0$$

Intrinsic parameters



$$u = \alpha \frac{x}{z} - \alpha \cot(\theta) \frac{y}{z} + u_0$$

$$v = \frac{\beta}{\sin(\theta)} \frac{y}{z} + v_0$$

Using homogenous coordinates,
we can write this as:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} \alpha & -\alpha\cot(\theta) & u_0 & 0 \\ 0 & \dfrac{\beta}{\sin(\theta)} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

or:

$$\vec{p} = \frac{1}{z} \begin{pmatrix} K & \vec{0} \end{pmatrix} \vec{P}$$

Extrinsic parameters: translation and rotation of camera frame

$$^{C}P = {}^{C}_{W}R\,{}^{W}P + {}^{C}O_{W}$$

Non-homogeneous coordinates

$$\begin{pmatrix} C_X \\ C_Y \\ C_Z \\ 1 \end{pmatrix} = \begin{pmatrix} - & - & - & | \\ - & {}^{C}_{W}R & - & {}^{C}O_{W} \\ - & - & - & | \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} W_X \\ W_Y \\ W_Z \\ 1 \end{pmatrix}$$

Homogeneous coordinates

$$\begin{pmatrix} {}^{C}P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{C}_{W}\mathcal{R} & {}^{C}O_{W} \\ \mathbf{0}^{T} & 1 \end{pmatrix} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

Block matrix form

Combining extrinsic and intrinsic calibration parameters

$$\boldsymbol{p} = \frac{1}{z}\mathcal{M}\boldsymbol{P}, \quad \text{where} \quad \mathcal{M} = \mathcal{K}(\mathcal{R} \quad \boldsymbol{t}), \tag{2.15}$$

$\mathcal{R} = {}^{C}_{W}\mathcal{R}$ is a rotation matrix, $\boldsymbol{t} = {}^{C}O_W$ is a translation vector, and $\boldsymbol{P} = ({}^{W}x, {}^{W}y, {}^{W}z, 1)^T$ denotes the *homogeneous* coordinate vector of P in the frame (W).

A projection matrix can be written explicitly as a function of its five intrinsic parameters (α, β, u_0, v_0, and θ) and its six extrinsic ones (the three angles defining $\mathcal{R}$ and the three coordinates of $\boldsymbol{t}$), namely,

$$\mathcal{M} = \begin{pmatrix} \alpha \boldsymbol{r}_1^T - \alpha \cot\theta \boldsymbol{r}_2^T + u_0 \boldsymbol{r}_3^T & \alpha t_x - \alpha \cot\theta t_y + u_0 t_z \\ \dfrac{\beta}{\sin\theta}\boldsymbol{r}_2^T + v_0 \boldsymbol{r}_3^T & \dfrac{\beta}{\sin\theta} t_y + v_0 t_z \\ \boldsymbol{r}_3^T & t_z \end{pmatrix}, \tag{2.17}$$

where $\boldsymbol{r}_1^T$, $\boldsymbol{r}_2^T$, and $\boldsymbol{r}_3^T$ denote the three rows of the matrix $\mathcal{R}$ and t_x, t_y, and t_z are the coordinates of the vector $\boldsymbol{t}$.

Forsyth&Ponce

Other ways to write the same equation

pixel coordinates

world coordinates

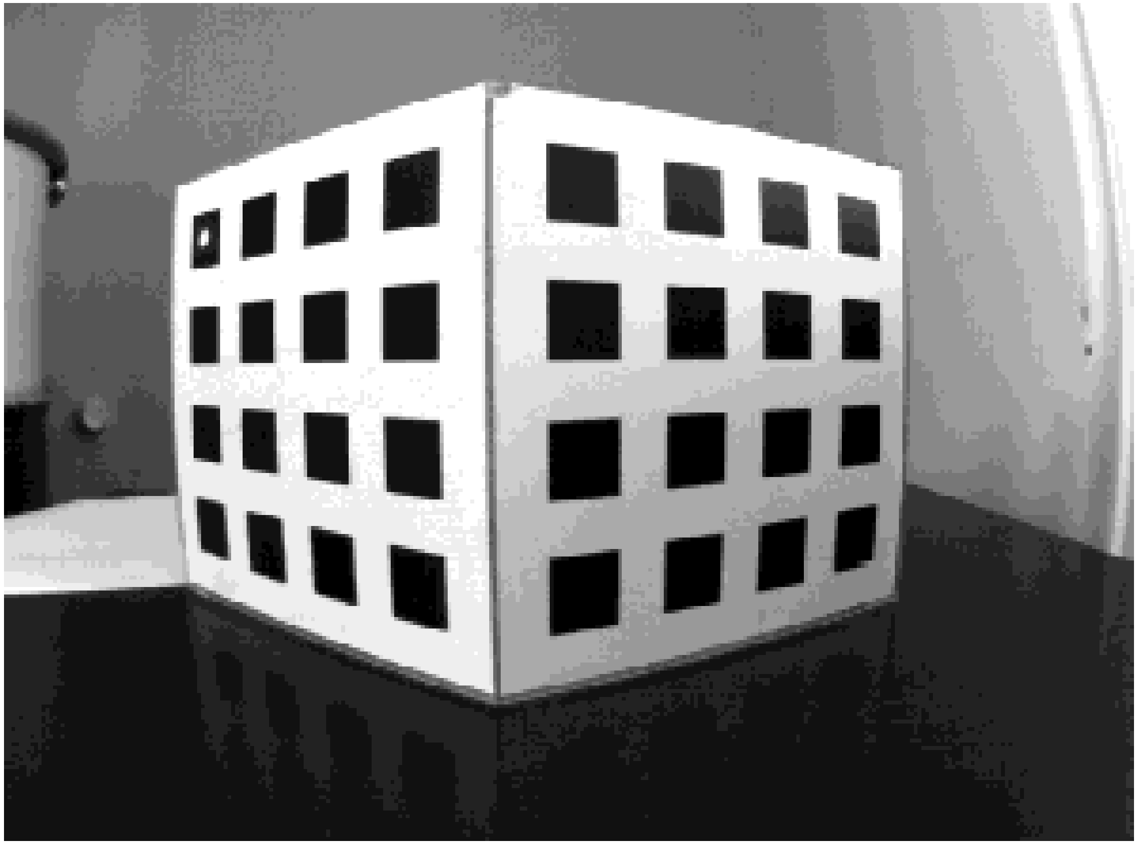
$$\vec{p} = \frac{1}{z} M\vec{P}$$

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} . & m_1^T & . & . \\ . & m_2^T & . & . \\ . & m_3^T & . & . \end{pmatrix} \begin{pmatrix} W_x \\ W_y \\ W_z \\ 1 \end{pmatrix}$$

$$\begin{cases} u = \dfrac{m_1 \cdot \vec{P}}{m_3 \cdot \vec{P}} \\[2ex] v = \dfrac{m_2 \cdot \vec{P}}{m_3 \cdot \vec{P}} \end{cases}$$

z is in the *camera* coordinate system, but we can solve for that, since $1 = \frac{m_3 \cdot \vec{P}}{z}$, leading to:

Calibration target



The Opti-CAL Calibration Target Image

http://www.kinetic.bc.ca/CompVision/opti-CAL.html

Camera calibration

From before, we had these equations relating image positions, u,v, to points at 3-d positions P (in homogeneous coordinates):

$$u = \frac{m_1 \cdot \vec{P}}{m_3 \cdot \vec{P}}$$

$$v = \frac{m_2 \cdot \vec{P}}{m_3 \cdot \vec{P}}$$

So for each feature point, i, we have:

$$(m_1 - u_i m_3) \cdot \vec{P}_i = 0$$

$$(m_2 - v_i m_3) \cdot \vec{P}_i = 0$$

Camera calibration

Stack all these measurements of i=1…n points

$$(m_1 - u_i m_3) \cdot \vec{P}_i = 0$$

$$(m_2 - v_i m_3) \cdot \vec{P}_i = 0$$

into a big matrix:

$$\begin{pmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ \cdots & \cdots & \cdots \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Camera calibration

In vector form:

$$\begin{pmatrix} P_1^T & 0^T & -u_1 P_1^T \\ 0^T & P_1^T & -v_1 P_1^T \\ \cdots & \cdots & \cdots \\ P_n^T & 0^T & -u_n P_n^T \\ 0^T & P_n^T & -v_n P_n^T \end{pmatrix} \begin{pmatrix} m_1 \\ m_2 \\ m_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Showing all the elements:

$$\begin{pmatrix} P_{1x} & P_{1y} & P_{1z} & 1 & 0 & 0 & 0 & 0 & -u_1 P_{1x} & -u_1 P_{1y} & -u_1 P_{1z} & -u_1 \\ 0 & 0 & 0 & 0 & P_{1x} & P_{1y} & P_{1z} & 1 & -v_1 P_{1x} & -v_1 P_{1y} & -v_1 P_{1z} & -v_1 \\ & & & & & & \cdots & \cdots & \cdots & & & \\ P_{nx} & P_{ny} & P_{nz} & 1 & 0 & 0 & 0 & 0 & -u_n P_{nx} & -u_n P_{ny} & -u_n P_{nz} & -u_n \\ 0 & 0 & 0 & 0 & P_{nx} & P_{ny} & P_{nz} & 1 & -v_n P_{nx} & -v_n P_{ny} & -v_n P_{nz} & -v_n \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

Camera calibration

$$\begin{pmatrix} P_{1x} & P_{1y} & P_{1z} & 1 & 0 & 0 & 0 & 0 & -u_1P_{1x} & -u_1P_{1y} & -u_1P_{1z} & -u_1 \\ 0 & 0 & 0 & 0 & P_{1x} & P_{1y} & P_{1z} & 1 & -v_1P_{1x} & -v_1P_{1y} & -v_1P_{1z} & -v_1 \\ & & & & & \cdots & \cdots & \cdots & & & & \\ P_{nx} & P_{ny} & P_{nz} & 1 & 0 & 0 & 0 & 0 & -u_nP_{nx} & -u_nP_{ny} & -u_nP_{nz} & -u_n \\ 0 & 0 & 0 & 0 & P_{nx} & P_{ny} & P_{nz} & 1 & -v_nP_{nx} & -v_nP_{ny} & -v_nP_{nz} & -v_n \end{pmatrix} \begin{pmatrix} m_{11} \\ m_{12} \\ m_{13} \\ m_{14} \\ m_{21} \\ m_{22} \\ m_{23} \\ m_{24} \\ m_{31} \\ m_{32} \\ m_{33} \\ m_{34} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}$$

P m = 0

We want to solve for the unit vector m (the stacked one) that minimizes $|Pm|^2$

The minimum eigenvector of the matrix P^TP gives us that (see Forsyth&Ponce, 3.1)

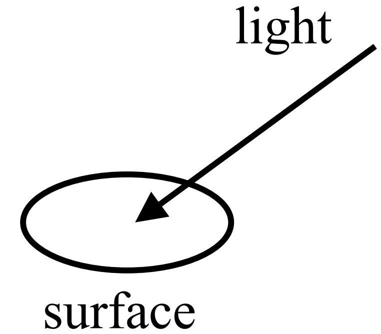
Once you have the M matrix, can recover the intrinsic and extrinsic parameters as in Forsyth&Ponce, sect. 3.2.2.

$$\mathcal{M} = \begin{pmatrix} \alpha \boldsymbol{r}_1^T - \alpha \cot\theta \boldsymbol{r}_2^T + u_0 \boldsymbol{r}_3^T & \alpha t_x - \alpha \cot\theta t_y + u_0 t_z \\ \dfrac{\beta}{\sin\theta} \boldsymbol{r}_2^T + v_0 \boldsymbol{r}_3^T & \dfrac{\beta}{\sin\theta} t_y + v_0 t_z \\ \boldsymbol{r}_3^T & t_z \end{pmatrix}$$

Today's class

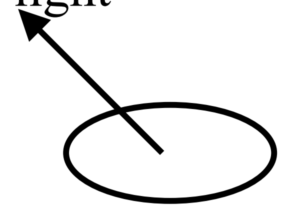
- First part: how *positions* in the image relate to 3-d positions in the world.
- Second part: how the *intensities* in the image relate surface and lighting properties in the world.

Irradiance, E



- Light power per unit area (watts per square meter) incident on a surface.
- The units tell you what to integrate over to find the energy impinging on a given area.
- E times pixel area, times exposure time gives the pixel intensity out (for linear sensor response)

Radiance, L

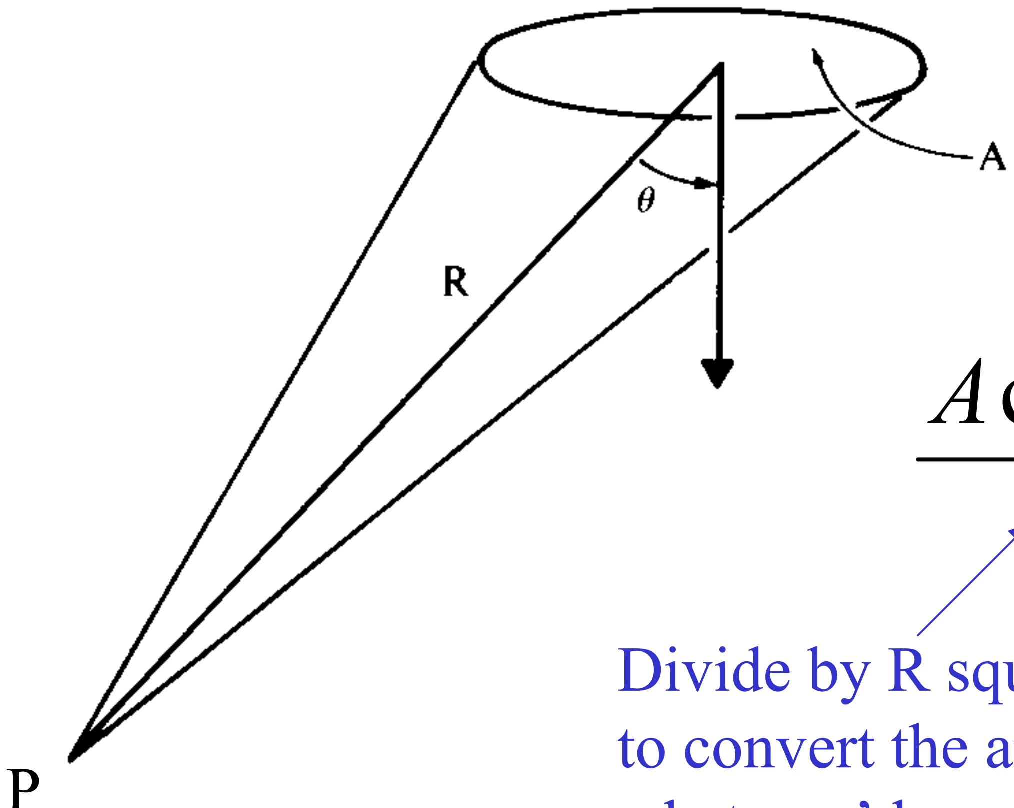


- Amount of light radiated from a surface into a given solid angle per unit area (watts per square meter per steradian).
- Note: the area is the foreshortened area, as seen from the direction that the light is being emitted.
- Informally, radiance tells you the “brightness”.

Solid angle

- The solid angle subtended by a cone of rays is the area of a unit sphere (centered at the cone origin) intersected by the cone.
- All possible angles from a point covers 4π steradians.
- A hemisphere covers 2π steradians, etc.

What's the solid angle subtended by this patch, area A, seen from P?



Multiply by cos(θ) to account for foreshortening

$$\frac{A\cos(\theta)}{R^2}$$

Divide by R squared to convert the area to what you'd see on a unit sphere

Image irradiance/scene radiance relationship

- The definition of scene radiance is constructed so that image irradiance is proportional to scene radiance.

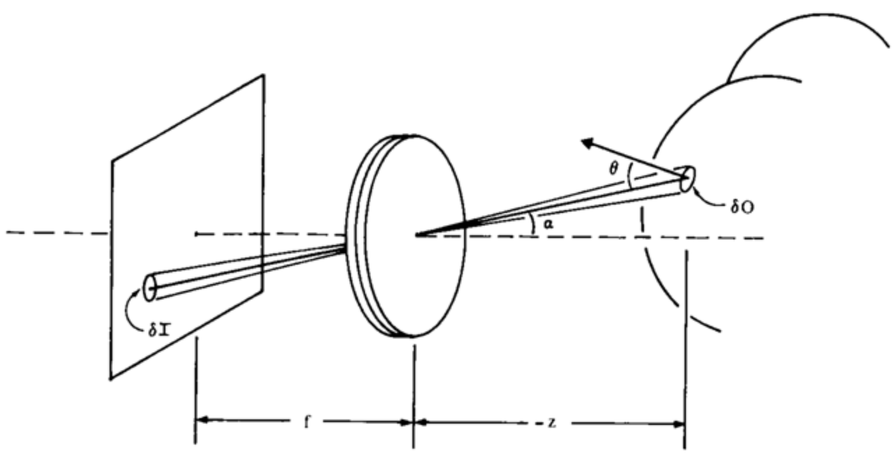


Figure 10-4. To see how image irradiance is related to the radiance of the surface, we must determine the size of the region in the image that corresponds to the patch on the surface.

Scene radiance

$$E = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4(\alpha)$$

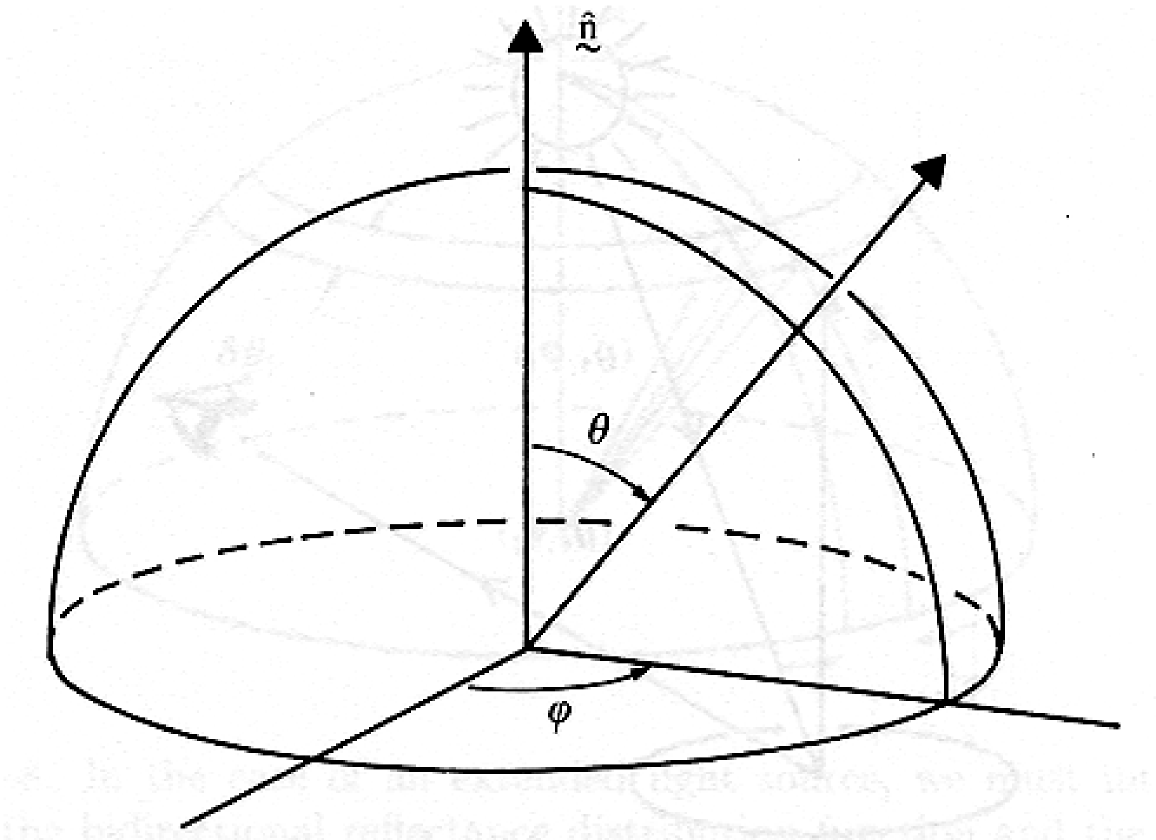
Image irradiance

Horn, sect. 10.3

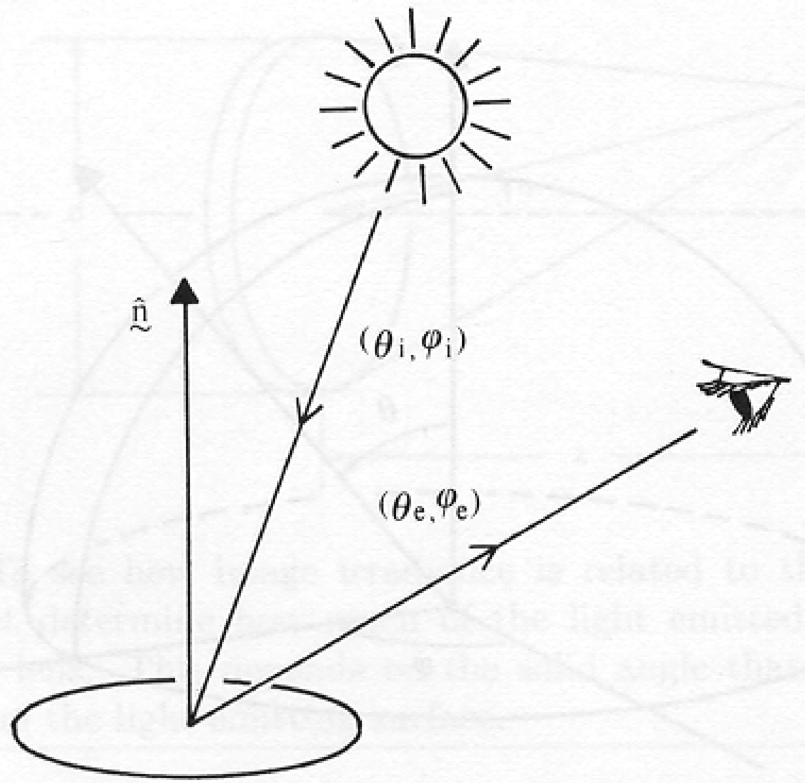
How the brightness depends on the surface properties: BRDF

- Bidirectional reflectance distribution function tells how bright a surface appears when viewed from one direction while light falls on it from another.

Coordinate system



Horn, 1986



Horn, 1986

Figure 10-7. The bidirectional reflectance distribution function is the ratio of the radiance of the surface patch as viewed from the direction (θ_e, ϕ_e) to the irradiance resulting from illumination from the direction (θ_i, ϕ_i).

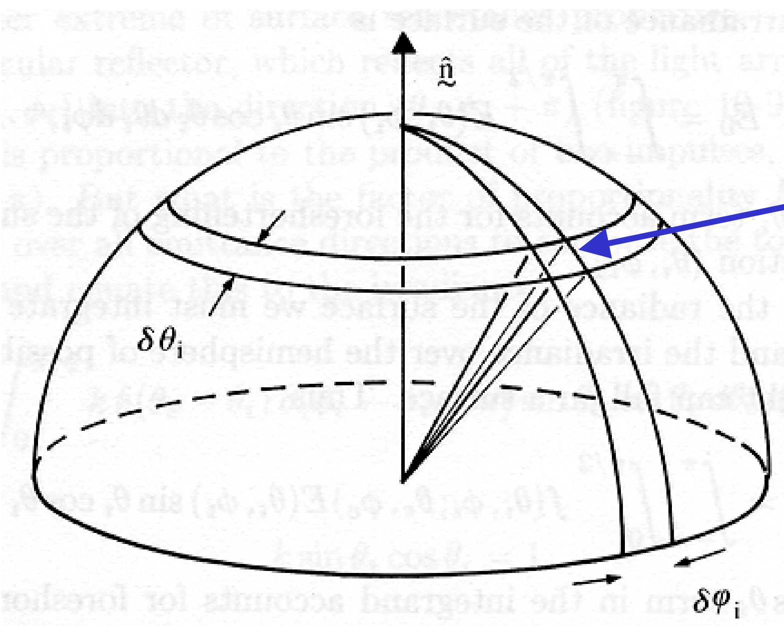
$$BRDF = f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{L(\theta_e, \phi_e)}{E(\theta_i, \phi_i)}$$

Helmholtz reciprocity condition

$$f(\theta_i, \phi_i, \theta_e, \phi_e) = f(\theta_e, \phi_e, \theta_i, \phi_i)$$

Otherwise, violate the 2^{nd} law of thermodynamics.

How does the world give us the brightness we observe at a point?



Solid angle of this patch:

$$\delta\omega = \sin(\theta_i)\,\delta\theta_i\,\delta\phi_i$$

Let radiance per solid angle be:

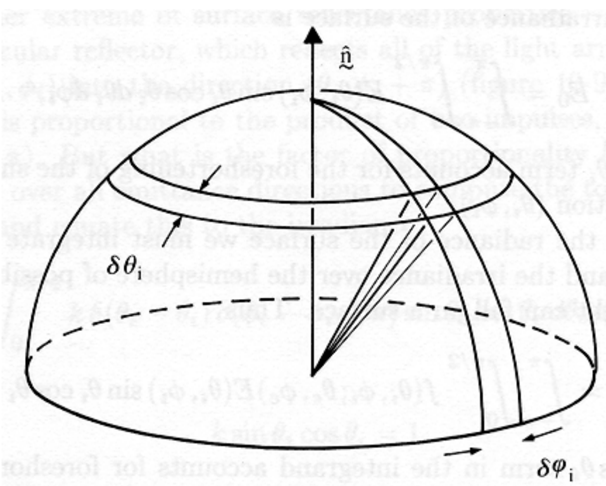
$$E(\theta_i, \phi_i)$$

The the radiance from this patch toward the origin is:

$$E(\theta_i, \phi_i)\sin(\theta_i)\,\delta\theta_i\,\delta\phi_i$$

Integrate all the source radiance impinging on the surface

Accounting for extended light sources



The total irradiance of the surface is:

$$E_0 = \int_{-\pi}^{\pi}\int_{0}^{\pi/2} E(\theta_i,\phi_i)\sin(\theta_i)\cos(\theta_i)\,d\theta_i\,d\phi_i$$

Accounting for the foreshortened area of center patch relative to illuminant.

The total radiance reflected from the surface patch is:

$$L(\theta_e,\phi_e) = \int_{-\pi}^{\pi}\int_{0}^{\pi/2} f(\theta_i,\phi_i,\theta_e,\phi_e)\,E(\theta_i,\phi_i)\sin(\theta_i)\cos(\theta_i)\,d\theta_i\,d\phi_i$$

Special case BRDF: Lambertian reflectance

BRDF is a constant. These surfaces look equally bright from all viewing directions.

$$f(\theta_i, \phi_i, \theta_e, \phi_e) = \frac{1}{\pi}$$

Radiance reflected from Lambertian surface illuminated by point source:

$$L(\theta_e, \phi_e) = \int_{-\pi}^{\pi} \int_{0}^{\pi/2} \frac{1}{\pi} \delta(\theta_i - \theta_0) \delta(\phi_i - \phi_0) \sin(\theta_i) \cos(\theta_i)\, d\theta_i\, d\phi_i$$

$$\propto \cos(\theta_0)$$

Show surfaces