Image Formation

- Image formation, Imaging process
- In general: “image” = Any 2D signal (e.g. range, Xray)
- Here: Images generated by light reflection

The connection between the world and its image:
- Geometry - Which image locations correspond to specific world locations?
- Photometry - What determine the intensity at every image point?

How cameras are built?
- to provide naturally looking pictures.

Familiar Observations about human imaging

- A projection process - lossy
  - Objects are occluded
  - Many world points are projected to the same image point
- Geometric properties
  - Points project to points, lines to lines, plans to planes
  - Special cases …
- Deformations
  - Far objects look smaller
  - Foreshortening - apparent size depends on object orientation
  - Lengths and angles are not preserved
  - Parallel lines look converging
Pinhole camera

The common simplified geometric imaging model

- **Geometrical optics** = Light travels in straight lines
- **The camera mimics the eye**
  - Dark box, A small hole in one side
  - Many light rays leave the object and only a small fraction enter the box through the pinhole
  - Image is upside down

Pinhole camera imaging geometry

- **Camera coordinate system**
- **Light travels in straight lines**
**Imaging geometry**

- Geometric transformation from world to camera is the Perspective transformation

\[ x' = (-) f \frac{X}{Z_c}; \quad y' = (-) f \frac{Y}{Z_c} \]

(in camera coordinates)

---

**Pinhole size**

**Focusing** - all ray that reach a point \( p \) in the image plane leave the same point \( P \) in 3D.

- What happens if we have two pinholes?
- What happens if the pinhole is large?

- Large pinhole - unfocused image
- Small pinhole - little light energy

Sharpness vs. light energy tradeoff
Camera with lens

- Can we have BOTH good focusing and High intensity?
- Answer: A lens can focus and capture more light

What lenses do?

- Usual lenses - complex and multi components

The simple “thin lens”:
- All parallel rays from one side of the lens meet at one focal point in the other side, and wise versa
- Some geometry → Lens formula: \( \frac{1}{z} + \frac{1}{Z} = \frac{1}{f} \)

- Far object’s points focus at approx. the same distance.
- Near points do not
- Lenses solve the energy-focus tradeoff but have limited depth of focus
Perspective projection properties

\[ x_{in} = (-f) \frac{X}{Z_c}; \quad y_{in} = (-f) \frac{Y}{Z_c} \]

- Depth, size, angles are not directly sensed
- Ambiguity

Properties of perspective projection

- Straight lines project to straight lines (board)
- Vanishing point - Each set of parallel lines in space project to a set of lines that meet at a point (board).
- Vanishing line (horizon) - Sets of parallel lines parallel to a plan correspond to collinear vanishing points (board).
Discovery of Perspective projection principles

Challenge: Drawing a building before it is built
Filippo Brunelleschi 1377-1446

Source: Jacobs
Imaging geometry - II

\[ x = (-1) \frac{X}{Z}; \quad y = (-1) \frac{Y}{Z} \]

- The perspective projection is non-linear (in world coordinates. Can we make it linear?)
- Points at infinity project to well behaved image points. Can we express this mathematically?
- What if the image plane is not normal to \( Z \)?

- Yes, we can! …
- Representing (2D and 3D) points in a different way.

---

Projective/homogeneous Coordinates

2D points,
- The projective plane
- Represent a 2D point by a triple of real numbers, (not all zeros.) \( \mathbf{p} = (x, y, z) \quad \mathbf{p} \in \mathbb{P}^2 \)
- Non-unique representation,
- Two points are equivalent ("congruent") iff
  \[ \mathbf{p}, \mathbf{p}' \in \mathbb{P}^2 \quad \mathbf{p} \sim \mathbf{p}' \iff (x, y, z) = \lambda (x', y', z') \]
- Geometric interpretation: projective points may be interpreted as lines (through origin) in 3D space.
Projective/homogeneous Coordinates

Coordinates change:

• Cartesian to homogeneous (one option)
  \[ q = (u, v) \in \mathbb{R}^2 \rightarrow p = (u, v, 1) \in \mathbb{P}^2 \]

• Homogeneous to Cartesian
  \[ p = (x, y, z) \in \mathbb{P}^2 \rightarrow q = (u, v) \in \mathbb{R}^2 \]
  s.t. \( u = x / z; v = y / z \)

Projective/homogeneous Coordinates

3D points

• The projective space
  • A point - a 4-tuple of real numbers, not all zeros.
  \[ P = (x, y, z, w) \quad P \in \mathbb{P}^3 \]

• Two points are equivalent iff
  \[ P, P' \in \mathbb{P}^3 \quad P \sim P' \iff (x, y, z, w) = \lambda(x', y', z', w') \]

• (1D coordinates .... \( p = (x, y) \quad p \in \mathbb{P}^1 \))
Perspective projection

- In Cartesian coordinates (simplified)
  \[ x_c = \frac{X_c}{Z_c}; \quad y_c = \frac{Y_c}{Z_c} \]

- In homogeneous coordinates
  \[
  \begin{pmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0
  \end{pmatrix}
  \begin{pmatrix}
  X \\
  Y \\
  Z
  \end{pmatrix} = \begin{pmatrix}
  X_c \\
  Y_c \\
  Z_c
  \end{pmatrix}
  \]

  \[ \mathbf{p} \sim \mathbf{p} = (I \mid 0) \mathbf{P} \]

Note: Every projective transformation is specified up to a multiplicative constant

---

Non camera coordinate system

- The Camera coordinate system be specified, in world coordinates, by
  - A center point \( O' \)
  - 3 axes along 3 unit vectors \( R_x, R_y, R_z \)

- Then, a point \( P_w = (X_w, Y_w, Z_w)^T \)

  is mapped to \( P_c = (X_c, Y_c, Z_c)^T = R^T (P_w - O') \)

  where \( R^T = \begin{pmatrix}
  R_x^T \\
  R_y^T \\
  R_z^T
  \end{pmatrix} \)

(Cartesian coordinates) (board)
Perspective projection

- In homogeneous coordinates
- For a camera in any position,

\[
p_c \sim (I | 0) p_c = (I | 0) \begin{pmatrix} R_x^T & -R_x^T O' \\ R_y^T & -R_y^T O' \\ R_z^T & -R_z^T O' \\ 0 & 1 \end{pmatrix} p_w = (I | 0) \begin{pmatrix} R^T \\ -R^T O' \\ 0 \\ 1 \end{pmatrix} p_w
\]

(3 1) columns

Inside the camera

- Real cameras invert, focal length = f
- The sensor is at the focal plan,
  Sensor’s center = intersection with optical axis
  center in \( o_x, o_y \) (in pixel coordinates)
- Pixel sizes \( s_x, s_y \)
- Image coordinates satisfy
  - in pixels, \( x_i, y_i \)
  - In meters, \( x^l, y^l \) (ideal camera, \( f=1 \)), \( x = x^l, y = y^l \)

\[
x = -(x_i - o_x) s_x
\]
\[
y = -(y_i - o_y) s_y
\]

Subject to: non rotated sensor, no lens deformations

\[
f x^l = -(x_i - o_x) s_x \quad ; \quad s_x x_i = -f x^l + o_x s_x
\]
Extrinsic+intrinsic transformation

\[
\begin{pmatrix}
  x_i \\
  y_i \\
  1
\end{pmatrix} = \begin{pmatrix}
  -f/s_x & 0 & o_x \\
  0 & -f/s_y & o_y \\
  0 & 0 & 1
\end{pmatrix} \begin{pmatrix}
  I & 0 \\
  R^T & -R^T O^T
\end{pmatrix} \begin{pmatrix}
  P
\end{pmatrix}
\]

\[= M_{\text{int}} M_{\text{ext}} P\]

- The projection matrix \( M = M_{\text{int}} M_{\text{ext}} \) is specified up to a constant.

- Notation: Sometimes \( M_{\text{int}} = K \)

- Calibration (Exterior/Interior)
  - Finding the projection matrices

General projection

- Often, the projection is considered as one process

\[p_c = MP\]

\[
M = \begin{pmatrix}
  p_{11} & p_{12} & p_{13} & p_{14} \\
  p_{21} & p_{22} & p_{23} & p_{24} \\
  p_{31} & p_{32} & p_{33} & p_{34}
\end{pmatrix}
\]

- 11 independent parameters

- Resulting in ...

\[
x_i = \frac{p_{11}X + p_{12} Y + p_{13} Z + p_{14}}{p_{31} X + p_{32} Y + p_{33} Z + p_{34}}; \quad y_i = \ldots
\]
Practical special cases-I

- Idealized perspective projection \( s_x = s_y = 1; o_x = o_y = 0 \)
- The projection matrix is specified by ... parameters

\[
\mathbf{M} = \begin{pmatrix}
-fr_{xx} & -fr_{xy} & -fr_{xz} & fR_x^T O' \\
-fr_{yx} & -fr_{yy} & -fr_{zy} & fR_y^T O' \\
\quad r_{zx} & \quad r_{zy} & \quad r_{zz} & -R_z^T O'
\end{pmatrix}
\]

Practical special cases-II

Weak perspective projection:
A practical approximation when the object/scene size is much smaller than its average depth

\[
\frac{Z_c - \bar{Z}_c}{Z_c} \ll 1
\]

\[
\rightarrow x_c \approx \frac{X_c}{Z_c}; \quad y_c \approx \frac{Y_c}{Z_c}
\]

Source: Moses
**Practical special cases-II**

- **Weak perspective projection:** The object/scene size is much smaller than its average depth
  \[
  \frac{Z_c - \bar{Z}_c}{Z_c} \ll 1 \quad \Rightarrow \quad x_c \equiv \frac{X_c}{Z_c}; \quad y_c \equiv \frac{Y_c}{Z_c}
  \]

- The projection is linear! (also in non-homogenous coor.)
  \[
  M_{wp} = \begin{pmatrix}
  -fr_{xx} & -fr_{xy} & -fr_{xz} & fR_x^T O' \\
  -fr_{yx} & -fr_{yy} & -fr_{yz} & fR_y^T O' \\
  0 & 0 & 0 & R_z^T (Z_c - O')
  \end{pmatrix}
  \]

- **Affine projection** - non-zero elements are arbitrary.

**Practical special cases-II**

- **Weak perspective = orthographic + uniform scaling**

  Scaled image

  \[
  p_{c} - \begin{pmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 0 & c
  \end{pmatrix} p_{c}
  \]

- **Orthographic projection** - idealization

  \[
  p_{c} - \begin{pmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 1 & 0
  \end{pmatrix} p_{c}
  \]

  (perspective)

  Image
Practical special cases-III

- **Projecting a plane**: Suppose the scene is planar.

\[
\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \sim MP = M \begin{pmatrix} X \\ Y \\ Z = aX + bY + c \end{pmatrix}
\]

\[
\begin{pmatrix} p_{11} + p_{12}a & p_{12} + p_{13}b & p_{14} + p_{15}c \\ p_{21} + p_{22}a & p_{22} + p_{23}b & p_{24} + p_{25}c \\ p_{31} + p_{32}a & p_{32} + p_{33}b & p_{34} + p_{35}c \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = H \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}
\]

- True for all 3D coordinate systems, including ....
- H - homography, collineation, projectivity,
- H - 2D transformation ! Invertible ! (in general position)

Practical special cases-IV

- **Less general 2D projections**
- **Affine homography**: 

\[
\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} \sim H \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix}
\]

- **Affine transformation (non-homogeneous Coordinates)**

\[
\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}
\]

Note: Here, unlike all previous transformation, parallel lines remain parallel, length ratios are preserved.
Practical special cases-IV+

- What if the plane is perpendicular to the viewing direction?
- Affine transformation
- Similarity
- Euclidean
- Scaling (uniform or not)
- Translation

All these special cases are very practical for image comparison, panoramas, warping,…

Imaging photometry

What determines the brightness of the image?

Illumination - reflection model

Plan:
1. Basic light units
2. How much light is reflected?
3. How much light the sensor gets?
How to measure (incident) light?

Irradiance, \( E \) = Light power per unit area (watts per square meter) incident on a surface. \( W \ m^{-2} \)

- If surface tilts away from light, same amount of light strikes bigger surface (lower irradiance).

Solid angle

- Solid angle - a measure to the size of an object as seen from a point
- For a patch seen from a point,
- The solid angle = observed area / (distance)
  \[ \delta \Omega = \frac{\delta A \cos \theta}{R^2} \]
- e.g. 1 steradian on a sphere

Source: Geiger
How to measure (emitted) light?

Radiance, $L$ - Amount of light radiated from a surface into a given direction per solid angle per unit area (watts per square meter per steradian). $W/m^2sr^{-1}$

Note: the area is the foreshortened area, as seen from the direction to which the light is being emitted*.

* there is a reason

Source: Geiger

From Surface Radiance to Image Irradiance

a. Finding area relations

Same solid angle

$$\frac{\delta A \cos \theta}{(z / \cos \alpha)^2} = \frac{\delta I \cos \alpha}{(f / \cos \alpha)^2} \quad \rightarrow \quad \frac{\delta A}{\delta l} = \frac{\cos \alpha}{\cos \theta} \left(\frac{z}{f}\right)^2$$

Source: Geiger
From Surface Radiance to Image Irradiance

b. Finding Solid angle subtended by the lens, as seen by the patch \( \delta A \)

\[
\Omega = \frac{\pi}{4} \frac{d^2 \cos \alpha}{(z / \cos \alpha)^2} = \frac{\pi}{4} \left( \frac{d}{z} \right)^2 \cos^3 \alpha
\]

c. Finding power from patch \( \delta A \) through the lens

\[
\delta P = L \Omega \delta A \cos \theta = L \delta A \frac{\pi}{4} \left( \frac{d}{z} \right)^2 \cos^3 \alpha \cos \theta
\]

Thus, we conclude

\[
E = \frac{\delta P}{\delta \Omega} = L \frac{\delta A}{\delta I} \frac{\pi}{4} \left( \frac{d}{z} \right)^2 \cos^3 \alpha \cos \theta = L \frac{\pi}{4} \left( \frac{d}{f} \right)^2 \cos^4 \alpha
\]

Source: Geiger

---

From Surface Radiance to Image Irradiance

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

- Scene radiance determines image irradiance
- Scene distance \( Z \) does not matter! Intuitively …
- The angle \( \alpha \) distance from optical axis reduces brightness.

- Irradiance and camera pixel area determines pixel brightness

Source: Geiger
What determines surface radiance

- Illumination + Surface properties

Bidirectional Reflectance Distribution Function (BRDF)

\[
BRDF(\theta, \phi, \theta_e, \phi_e) = \frac{\delta L(\theta, \phi)}{\delta E(\theta_e, \phi_e)}
\]

Source: Geiger

What determines surface radiance

- Body Reflection:
  - Diffuse Reflection
  - Matte Appearance
  - Clay, paper, etc

- Surface Reflection:
  - Specular Reflection
  - Glossy Appearance
  - Dominant for Metals

Image Intensity = Body Reflection + Surface Reflection
What determines surface radiance

Lambertian surface -
• Emits $\rho_d$ of the surface irradiance for all angles
  \[ BRDF = \frac{\rho_d}{\pi} \]
• When illuminated by a point source, the irradiance is
  \[ E = E_0 \cos \theta_i \]
  and the radiance is
  \[ L(\theta_e, \phi_e) = BRDF(\theta) \quad E = \frac{\rho_d}{\pi} E_0 \cos \theta_i \]

Radiance (and subjective brightness) do not depend on emission (viewing) direction!

Diffused surfaces (e.g. Lambertian) - The reflection is a result of light entering inside the surface, internal reflections, and the reflection is colored.

What determines surface radiance

Idealized mirror surface - reflects only in the direction

\[ R \sim (\theta_o, \phi_o) = (\theta_i, \phi_i + \pi) \quad ; \quad E \sim (\theta_i, \phi_i) \]

Brightness depends on viewing direction

The reflection is from surface, and may be colored or not.

Realistic surfaces are often rough, and reflection is spread (Models: Phong or Torrance-Sparrow)
What determines surface radiance

What if the illumination is not a point source?

Special case:
- Light is reflected from all the sky
- Diffused surfaces,

Result:
- The brightness/radiance is constant
- Shape is not perceived

Source: Narasimhan

What determines surface radiance

Mixed model - e.g. Phong (dichromatic reflectance) model
- Models diffused and specular effects

\[ I = I_{ambient} + I_{diffused}(E \cdot n) + I_{specular}(V \cdot R)^\alpha \]

Ambient + Diffuse + Specular = Phong Reflection
"30 seconds" on Real Camera structure

Other stages that influence the image
- Lens (Chromatic aberrations, Vingetting)
- Aperture
- Shutter
- Sensor (CCD or CMOS)
- Gain + AGC
- ADC
- Demosaicing
- Color balance
- Gamma correction
- Color transformation to YCbCr + JPEG coding

The space of image under different illuminations

How do images change under variable illumination?

For Lambertian model, the irradiance is

\[ I = \rho \overline{s} \cdot \overline{n} \]

The vector of intensities ...
Changing the illumination ... (board)
Reconstruction using Photometry - I

\[ I = \rho \bar{s} \cdot \bar{n} \]

- 1 camera
- 3 images using 3 Known illuminations
- Lambertian case
- One point at a time
- Solve for \( n \)

Photometric stereo

\[
\begin{bmatrix}
I_1 \\
I_2 \\
I_3
\end{bmatrix} = \rho
\begin{bmatrix}
s_1^T \\
s_2^T \\
s_3^T
\end{bmatrix} n
\]

Input: 3 images
Reconstruction result: 3D shape

Reconstruction result: albedo
Reconstruction using Photometry - II

Shape from shading

Lambertian case:

\[ I = \rho \bar{s} \cdot \bar{n} \]

Shape from Shading

- What if we only have one image?
- The task is underconstrained.
- Add smoothness assumption (regularization)

Minimize

\[ e = \iint_{\text{image}} \left( (n_x)^2 + (n_y)^2 \right) + \left( (n_x)_{,x}^2 + (n_y)_{,y}^2 \right) + \lambda \left( I_{\text{given}}(x,y) - I_{\text{synth}}(n) \right)^2 dx dy \]

Input & output

by Ikeuchi and Horn