Image Formation I
Chapter 1 (Forsyth&Ponce)
Cameras

Guido Gerig
CS 6320 Spring 2013

Acknowledgements:
• Slides used from Prof. Trevor Darrell, (http://www.eecs.berkeley.edu/~trevor/CS280.html)
• Some slides modified from Marc Pollefeys, UNC Chapel Hill. Other slides and illustrations from J. Ponce, addendum to course book.
GEOMETRIC CAMERA MODELS

- The Intrinsic Parameters of a Camera
- The Extrinsic Parameters of a Camera
- The General Form of the Perspective Projection Equation
- Line Geometry

Reading: Chapter 1.
Images are two-dimensional patterns of brightness values. They are formed by the projection of 3D objects.
Animal eye:
a looonng time ago.

Pinhole perspective projection: Brunelleschi, XVth Century.
Camera obscura: XVIth Century.

Photographic camera: Niepce, 1816.
Camera model

Relation between pixels and rays in space
The **camera obscura** (Latin for 'dark room') is an optical device that projects an **image** of its surroundings on a screen (source Wikipedia).
2.18 DIFFRACTION LIMITS THE QUALITY OF PINHOLE OPTICS. These three images of a bulb filament were made using pinholes with decreasing size. (A) When the pinhole is relatively large, the image rays are not properly converged, and the image is blurred. (B) Reducing the size of the pinhole improves the focus. (C) Reducing the size of the pinhole further worsens the focus, due to diffraction. From Ruechardt, 1958.
Physical parameters of image formation

- Geometric
  - Type of projection
  - Camera pose

- Photometric
  - Type, direction, intensity of light reaching sensor
  - Surfaces’ reflectance properties

- Optical
  - Sensor’s lens type
  - focal length, field of view, aperture

- Sensor
  - sampling, etc.

Figure 2.1 A few components of the image formation process: (a) perspective projection; (b) light scattering when hitting a surface; (c) lens optics; (d) Bayer color filter array.
Physical parameters of image formation

• Geometric
  – Type of projection
  – Camera pose

• Optical
  – Sensor’s lens type
  – focal length, field of view, aperture

• Photometric
  – Type, direction, intensity of light reaching sensor
  – Surfaces’ reflectance properties

• Sensor
  – sampling, etc.
Perspective and art

• Use of correct perspective projection indicated in 1st century B.C. frescoes
• Skill resurfaces in Renaissance: artists develop systematic methods to determine perspective projection (around 1480-1515)
Perspective projection equations

- 3d world mapped to 2d projection in image plane

\[(x, y, z) \rightarrow \left(\frac{fx}{z}, \frac{fy}{z}\right)\]

Scene point $\rightarrow$ Image coordinates
Affine projection models:
Weak perspective projection

where

\[ z' = \frac{m y}{z_0} \]

is the magnification.

When the scene relief is small compared to its distance from the Camera, \( m \) can be taken constant: weak perspective projection.
Affine projection models:
Orthographic projection

When the camera is at a (roughly constant) distance from the scene, take \( m=1 \).
Homogeneous coordinates

Is this a linear transformation?
- no—division by $z$ is nonlinear

Trick: add one more coordinate:

\[
(x, y) \Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad \text{homogeneous image coordinates}
\]

\[
(x, y, z) \Rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad \text{homogeneous scene coordinates}
\]

Converting \textit{from} homogeneous coordinates

\[
\begin{bmatrix} x \\ y \\ w \end{bmatrix} \Rightarrow (x/w, y/w)
\]

\[
\begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \Rightarrow (x/w, y/w, z/w)
\]
Perspective Projection Matrix

- Projection is a matrix multiplication using homogeneous coordinates:

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1/f' & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
z/f'
\end{bmatrix}
= \begin{bmatrix}
x \\
y \\
z/f' \\
z/f'
\end{bmatrix}
\Rightarrow (f'/x, f'/y) \Rightarrow (x', y')
\]

divide by the third coordinate to convert back to non-homogeneous coordinates

Complete mapping from world points to image pixel positions?
Points at infinity, vanishing points

Points from infinity represent rays into camera which are close to the optical axis.

Image source: wikipedia
Perspective projection & calibration

• Perspective equations so far in terms of camera’s reference frame....

• Camera’s intrinsic and extrinsic parameters needed to calibrate geometry.
The CCD camera

CCD camera
Perspective projection & calibration

Extrinsic:
Camera frame $\leftrightarrow$ World frame

Intrinsic:
Image coordinates relative to camera
$\leftrightarrow$ Pixel coordinates

Perspective projection matrix (3x4)
Camera to pixel coord. trans. matrix (3x3)

World to camera coord. trans. matrix (4x4)

K. Grauman
Intrinsic parameters: from idealized world coordinates to pixel values

Perspective projection: World point and pixels in camera coordinates

\[ x' = -f \frac{x}{z} \]
\[ y' = -f \frac{y}{z} \]
Intrinsic parameters

Pixel dimensions: $1/k^*1/k$

$k$: cells/mm, units $[\text{mm}^{-1}]$

But “pixels” are in some arbitrary spatial units, which can be described by $\text{#pixels per mm}$.

$$u = -\alpha \frac{x}{z}, \text{ with } \alpha = f \times k$$

$$v = -\alpha \frac{y}{z}, \text{ with } \alpha = f \times k$$

$\alpha$ represents magnification

W. Freeman
Intrinsic parameters

Maybe pixels are not square and have different horizontal and vertical dimensions. (u,v): pixel numbers, x,y,z): World point in camera coordinates.

\[ u = -\alpha \frac{x}{z} \text{, with } \alpha = f \cdot k \]
\[ v = -\beta \frac{y}{z} \text{, with } \beta = f \cdot l \]

\( \alpha, \beta \) represent magnifications

Pixel dimensions: \( \frac{1}{k} \times \frac{1}{l} \)

k,l: cells/mm, units [mm\(^{-1}\)]
Intrinsic parameters

We don’t know the origin of our camera pixel coordinates: $(u_0,v_0)$ represent intersection of optical axis with image plane: $(u_0, v_0)$: image center in 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 due to manufacturing errors and eventually line-by-line readouts.

\[ v \sin(\theta) = v' \rightarrow v = \frac{v'}{\sin(\theta)} \]

\[ u = u' - \cos(\theta) v = u' - \cot(\theta) v' \]

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

\[ v = \frac{\beta}{\sin(\theta)} \frac{y}{z} + v_0 \]
Intrinsic parameters, homogeneous coordinates

Using homogenous coordinates, we can write this as:

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

or:

\[
\vec{p} = \frac{1}{z} (K) \vec{c}_p
\]

World point in camera-based coordinates

W. Freeman
Perspective projection & calibration

Extrinsic:
Camera frame $\leftrightarrow$ World frame

Intrinsic:
Image coordinates relative to camera $\leftrightarrow$ Pixel coordinates

3D point (4x1)
Coordinate Changes: Pure Translations

\[ \overrightarrow{O_B P} = \overrightarrow{O_B O_A} + \overrightarrow{O_A P} \quad , \quad \overrightarrow{B P} = \overrightarrow{A P} + \overrightarrow{B O_A} \]
Coordinate Changes: Pure Rotations

\[
\begin{bmatrix}
    \mathbf{i}_A \cdot \mathbf{i}_B \\
    \mathbf{i}_A \cdot \mathbf{j}_B \\
    \mathbf{i}_A \cdot \mathbf{k}_B \\
\end{bmatrix}
\begin{bmatrix}
    \mathbf{j}_A \cdot \mathbf{i}_B \\
    \mathbf{j}_A \cdot \mathbf{j}_B \\
    \mathbf{j}_A \cdot \mathbf{k}_B \\
\end{bmatrix}
\begin{bmatrix}
    \mathbf{k}_A \cdot \mathbf{i}_B \\
    \mathbf{k}_A \cdot \mathbf{j}_B \\
    \mathbf{k}_A \cdot \mathbf{k}_B \\
\end{bmatrix}
= \begin{bmatrix}
    \mathbf{A}^T \\
    \mathbf{A}^T \\
    \mathbf{A}^T \\
\end{bmatrix}
\begin{bmatrix}
    \mathbf{i}_B \\
    \mathbf{j}_B \\
    \mathbf{k}_B \\
\end{bmatrix}
\]
Coordinate Changes: Rotations about the $k$ Axis

$$\mathbf{R} = \begin{bmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
A rotation matrix is characterized by the following properties:

• Its inverse is equal to its transpose, and

• its determinant is equal to 1.

Or equivalently:

• Its rows (or columns) form a right-handed orthonormal coordinate system.
Coordinate Changes:
Pure Rotations

\[ \overrightarrow{OP} = [i_A \ j_A \ k_A] \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix} = [i_B \ j_B \ k_B] \begin{bmatrix} B_x \\ B_y \\ B_z \end{bmatrix} \]

\[ \Rightarrow \quad ^B P = ^A R ^A P \]
Coordinate Changes: Rigid Transformations

\[ B \mathbf{P} = B R_A^A \mathbf{P} + B O_A \]
Block Matrix Multiplication

\[
A = \begin{bmatrix}
A_{11} & A_{12} \\
A_{21} & A_{22}
\end{bmatrix} \quad B = \begin{bmatrix}
B_{11} & B_{12} \\
B_{21} & B_{22}
\end{bmatrix}
\]

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

\[
\begin{bmatrix}
^B P \\
1
\end{bmatrix} = \begin{bmatrix}
^B R & ^B O_A \\
0^T & 1
\end{bmatrix} \begin{bmatrix}
^A P \\
1
\end{bmatrix} = \begin{bmatrix}
^B R^A P + ^B O_A \\
1
\end{bmatrix} = \begin{bmatrix}
^A P \\
1
\end{bmatrix}
\]
Extrinsic parameters: translation and rotation of camera frame

\[ \mathbf{c} \mathbf{p} = \mathbf{w} \mathbf{R} \mathbf{w} \mathbf{p} + \mathbf{w} \mathbf{t} \]

\[
\begin{pmatrix}
\mathbf{c} \mathbf{p}
\end{pmatrix} = \begin{pmatrix}
\mathbf{0} & \mathbf{c} \mathbf{R} & \mathbf{c} \mathbf{t}
\end{pmatrix}
\begin{pmatrix}
\mathbf{w} \mathbf{p}
\end{pmatrix}
\]

Non-homogeneous coordinates

Homogeneous coordinates

W. Freeman
Combining extrinsic and intrinsic calibration parameters, in homogeneous coordinates

\[
\mathbf{p} = \frac{1}{z} \mathbf{K} \mathbf{R} \mathbf{t} \mathbf{p}
\]

\[
\mathbf{p} = \frac{1}{z} \mathbf{M} \mathbf{w} \mathbf{p}
\]

**Intrinsic**

**Extrinsic**

**World coordinates**

**Camera coordinates**

**Pixels**
Other ways to write the same equation

\[
\vec{p} = \frac{1}{z} M \vec{w} \vec{p}
\]

Converting back from homogeneous coordinates leads to (note that \( z = m_{3}^{T} \cdot \vec{P} \)):

\[
\begin{align*}
    u &= \frac{m_{1} \cdot \vec{P}}{m_{3} \cdot \vec{P}} \\
    v &= \frac{m_{2} \cdot \vec{P}}{m_{3} \cdot \vec{P}}
\end{align*}
\]
Extrinsic Parameters

- When the camera frame \((C')\) is different from the world frame \((W)\),

\[
\begin{pmatrix}
  C P \\
  1
\end{pmatrix}
= \begin{pmatrix}
  C W R \\
  0^T
\end{pmatrix}
\begin{pmatrix}
  C O_W \\
  1
\end{pmatrix}
\begin{pmatrix}
  W P \\
  1
\end{pmatrix}.
\]

- Thus,

\[ p = \frac{1}{z} \mathcal{M} P, \]

where

\[
\mathcal{M} = \mathcal{K}(\mathcal{R} \quad t),
\]

\[ \mathcal{R} = C_W R, \]

\[ t = C O_W, \]

\[ P = \begin{pmatrix}
  W P \\
  1
\end{pmatrix}. \]

- Note: \( z \) is *not* independent of \( \mathcal{M} \) and \( P \):

\[
\mathcal{M} = \begin{pmatrix}
  m_1^T \\
  m_2^T \\
  m_3^T
\end{pmatrix} \Rightarrow z = m_3 \cdot P, \quad \text{or} \quad \begin{cases}
  u = \frac{m_1 \cdot P}{m_3 \cdot P}, \\
  v = \frac{m_2 \cdot P}{m_3 \cdot P}.
\end{cases}
\]
Explicit Form of the Projection Matrix

\[
\mathcal{M} = \begin{pmatrix}
\alpha r_1^T - \alpha \cot \theta r_2^T + u_0 r_3^T & \alpha t_x - \alpha \cot \theta t_y + u_0 t_z \\
\frac{\beta}{\sin \theta} r_2^T + v_0 r_3^T & \frac{\beta}{\sin \theta} t_y + v_0 t_z \\
r_3^T & t_z
\end{pmatrix}
\]

Note: If \( \mathcal{M} = (\mathcal{A} \ b) \) then \(|a_3| = 1\).

Replacing \( \mathcal{M} \) by \( \lambda \mathcal{M} \) in

\[
\begin{cases}
u = \frac{m_1 \cdot P}{m_3 \cdot P} \\
v = \frac{m_2 \cdot P}{m_3 \cdot P}
\end{cases}
\]

does not change \( u \) and \( v \).

\( M \) is only defined up to scale in this setting!!
Find the position, $u_i$ and $v_i$, in pixels, of each calibration object feature point.