Cameras & Camera Geometry

Forsyth&Ponce: Chap. 1,2,3
Szeliski: Chap. 2.1
Cameras & Camera Geometry

Ideal pinhole model
  perspective projection
  weak perspective and orthography
Ideal thin lens model
  thin lens equation
  finite field of view
  finite aperture and depth of field
General lens model
  spherical aberration
  geometric distortion
  vignetting
Geometric models of cameras
  intrinsic and extrinsic parameters
  perspective projection equations
  calibration
Pinhole Cameras
Equivalent Model with Virtual Image Plane
Basic Geometric Properties

- Distant objects are smaller
- Lines project to lines
- The projection of parallel lines meet at a single vanishing point
- Vanishing points of coplanar sets of lines are collinear, form the vanishing line of the plane (horizon)
Perspective Projection

\[ y' = f \frac{y}{z} \]
Special Case: Weak Perspective (Affine Projection)

\[ x' \approx -mx \]
\[ y' \approx -my \]

\[ m = -\frac{f'}{z_o} \]

If scene points are in a plane, projections are simply magnified by \( m \).
Special Case: Weak Perspective (Affine Projection)

If $\Delta z \ll -\bar{z}$:

$$x' \approx -mx$$
$$y' \approx -my$$

$$m = -\frac{f}{\bar{z}}$$

Justified if scene depth is small relative to average distance from camera
Orthographic Projection

Justified if scene depth is small compared to distance from camera and camera remains at approximately constant distance
Strong perspective:
Angles are not preserved
The projections of parallel lines intersect at one point
Strong perspective:
Angles are not preserved
The projections of parallel lines intersect at one point

Weak perspective:
Angles are better preserved
The projections of parallel lines are (almost) parallel
Limitations of the Pinhole Model

Ideal pinhole:
Single scene point generates single image
but:
Diffraction
Low light level

Finite-size pinhole:
Single scene point generates extended image.
Resulting image is blurry
All rays emanating from \( P \) converge to a single point \( P' \)

\[
\frac{1}{z'} + \frac{1}{z} = \frac{1}{f}
\]

Points at infinity are focused on plane \( z' = f \)

Ideal because: infinite aperture
infinite field of view
infinately small distance between surfaces
Finite Aperture
Finite Aperture

Ideal case: Only the points on one plane are in perfect focus.
Finite aperture: points within a region of depth $D$ (depth of field) are in focus.
For a given $f$, the larger the aperture, the smaller $D$
Depth of field controlled by $f/a$
• Previous approximation is incorrect
  → Aberrations and distortions
  → Blurring and incorrect shape in the image
Snell’s law of refraction:
\[ n_1 \sin(\alpha_1) = n_2 \sin(\alpha_2) \]

First-order optics: appropriate for ideal model of thin lens
\[ n_1 \alpha_1 \approx n_2 \alpha_2 \]

Higher order optics: necessary for real lenses
\[ \sin \alpha \approx \alpha - \frac{\alpha^3}{6} + \cdots \]

Aberrations:
  Blurring: e.g., spherical aberrations
  Geometric distortion
Spherical Aberrations
Rays further from the optical axis are focused closer to the lens.
Geometric Distortion

pincushion

barrel
From Mark Fiala, Univ. Alberta
Radial Distortion Model

Ideal:

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

Distorted:

\[ x'' = \frac{1}{\lambda} x' \]
\[ y'' = \frac{1}{\lambda} y' \]
\[ \lambda = 1 + k_1 r^2 + k_2 r^4 + \ldots \]
Vignetting

- Only part of the light reaches the sensor
- Periphery of the image is dimmer
- Problem in practice
| Perspective Projection | \( x' = f \frac{x}{z} \)  
| | \( y' = f \frac{y}{z} \)  
| | \( x, y: \) World coordinates  
| | \( x', y': \) Image coordinates  
| | \( f: \) pinhole-to-retina distance  
| Weak-Perspective Projection (Affine) | \( x' \approx -mx \)  
| | \( y' \approx -my \)  
| | \( m = -\frac{f}{z} \) \( x, y: \) World coordinates  
| | \( x', y': \) Image coordinates  
| | \( m: \) magnification  
| Orthographic Projection (Affine) | \( x' \approx x \)  
| | \( y' \approx y \) \( x, y: \) World coordinates  
| | \( x', y': \) Image coordinates  
| Common distortion model | \( x'' = \frac{1}{\lambda} x' \)  
| | \( y'' = \frac{1}{\lambda} y' \) \( \lambda = 1 + k_1 r^2 + k_2 r^4 + \cdots \)  
| | \( x', y': \) Ideal image coordinates  
| | \( x'', y'': \) Actual image coordinates  

Camera Geometry and Calibration

Chap. 2
Homogeneous coordinates in 2-D

- Physical point \( p = \begin{bmatrix} x \\ y \end{bmatrix} \) represented by three coordinates
  \[
  x = \frac{u}{w} \\
  y = \frac{v}{w}
  \]

- Two sets of homogeneous coordinate vectors are equivalent if they are proportional to each other:
  \[
  \begin{bmatrix} u \\ v \\ w \end{bmatrix} \equiv \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} \iff \begin{bmatrix} u \\ v \\ w \end{bmatrix} \equiv \lambda \begin{bmatrix} u' \\ v' \\ w' \end{bmatrix} \quad \lambda \neq 0
  \]

What if \( w = 0 \)???
Lines in 2-D

• General equation of a line in 2-D:

\[ ax + by + c = 0 \]

• In homogeneous coordinates:

\[
l^T p = l \cdot p = 0 \quad l = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad p = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}
\]
Using the cross-product

• The cross product of two 3-coordinate vectors \( p \) and \( q \) is: \( p \times q \)

• Properties:
  \[ p \cdot (p \times q) = q \cdot (p \times q) = p \times p = 0 \]

• Questions:

• Write \( p \equiv q \) using the cross-product

• What is the intersection of two lines in homogeneous coordinates \( l_1 \) and \( l_2 \)?

• What is the line \( l \) going through 2 points in homogeneous coordinates \( p_1 \) and \( p_2 \)?
Homogeneous coordinates in 3-D

- Physical point \( p = \begin{bmatrix} x \\ y \\ z \end{bmatrix} \) represented by four coordinates:
  \[
  \begin{bmatrix}
  u \\
  v \\
  w \\
  t
  \end{bmatrix}
  \]
  \[
  x = \frac{u}{t} \\
  y = \frac{v}{t} \\
  z = \frac{w}{t}
  \]

- Two sets of homogeneous coordinate vectors are equivalent if they are proportional to each other:
  \[
  \begin{bmatrix}
  u \\
  v \\
  w \\
  t
  \end{bmatrix}
  \equiv
  \begin{bmatrix}
  u' \\
  v' \\
  w' \\
  t'
  \end{bmatrix}
  \leftrightarrow
  \begin{bmatrix}
  u \\
  v \\
  w \\
  t
  \end{bmatrix}
  = \lambda
  \begin{bmatrix}
  u' \\
  v' \\
  w' \\
  t'
  \end{bmatrix}
  \quad \lambda \neq 0
  \]

What if \( t = 0 \)???
Planes in 3-D

- General equation of a line in 2-D:
  \[ ax + by + cz + d = 0 \]
- In homogeneous coordinates:
  \[ \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} = 0 \]
# Basic transformations (reminder)....

<table>
<thead>
<tr>
<th>Transformation</th>
<th>Vector Coordinates</th>
<th>Homogeneous Coordinates</th>
<th>Degrees of Freedom</th>
<th>Invariants</th>
</tr>
</thead>
</table>
| Translation     | \( y = x + t \)    | \[
\begin{bmatrix}
I & t \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 3                  | lengths, angles     |
| Rotation        | \( y = Rx \)       | \[
\begin{bmatrix}
R & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 3                  | lengths, angles     |
|                 | \( R^T R = RR^T = I \) | \[
Det(R) = +1
\] |                    |                             |
| Rigid           | \( y = Rx + t \)   | \[
\begin{bmatrix}
R & t \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 6                  | lengths, angles     |
| Affine          | \( y = Ax + t \)   | \[
\begin{bmatrix}
A & t \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 12                 | ratios of lengths, parallelism |
| Projective      | \( 4 \times 4 \) matrix \( M \) | \[
\] | 15                 | colinearity, incidence   |
<table>
<thead>
<tr>
<th>Transformation</th>
<th>Vector Coordinates</th>
<th>Homogeneous Coordinates</th>
<th>Degrees of Freedom</th>
<th>Invariants</th>
</tr>
</thead>
</table>
| Translation   | \( y = x + t \)   | \[
\begin{bmatrix} I & t \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 3                  | lengths, angles   |
| Rotation      | \( y = Rx \)      | \[
\begin{bmatrix} R & 0 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 3                  | lengths, angles   |
|               | \( R^T R = RR^T = I \) |                        |                    |            |
| Rigid         | \( y = Rx + t \)  | \[
\begin{bmatrix} R & t \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 6                  | lengths, angles   |
| Affine        | \( y = Ax + t \)  | \[
\begin{bmatrix} A & t \\
0 & 0 & 0 & 1
\end{bmatrix}
\] | 12                 | ratios of lengths, parallelism |
| Projective    | 4 × 4 matrix \( M \) |                        | 15                 | colinearity, incidence |
Retinal plane

Normalized image plane

\[ f \]

\[ 1 \]

\[ \hat{v}, \hat{p}, \hat{C}, \hat{u} \]

\[ \theta \]

\[ P \]

\[ x, y, z, \Omega \]
Standard Perspective Camera Model

Scale in x direction between world coordinates and image coordinates

\[
M = \begin{bmatrix}
\alpha & -\alpha \cot \theta & u_o \\
0 & \beta & v_o \\
0 & \sin \theta & 1
\end{bmatrix}
\]

Skew of camera axes. \( \theta = 90^\circ \)
if the axes are perpendicular

Principal point =
Image coordinates of the projection of camera origin on the retina

Translation between world coordinate system and camera

Rotation between world coordinate system and camera
Q: Is a given 3x4 matrix \( M \) the projection matrix of some camera?
A: Yes, if and only if \( \det(A) \) is not zero

Q: Is the decomposition unique?
A: There are multiple equivalent solutions
Applying the Projection Matrix

Homogeneous coordinates of point in image

\[ p = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \]

Homogeneous coordinates of point in world

\[ P = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \]

Homogeneous vector transformation:

\[ p \equiv MP \]

Computation of individual coordinates:

\[ u = \frac{m_1^T P}{m_3^T P} = \frac{m_1 \cdot P}{m_3 \cdot P} \]

\[ v = \frac{m_2^T P}{m_3^T P} = \frac{m_2 \cdot P}{m_3 \cdot P} \]
Observations:

- The equation
\[
\begin{align*}
    a_1^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + b_1 = 0
\end{align*}
\]

is the equation of a plane of normal direction \(a_1\).

- From the projection equation, it is also
\[
\begin{align*}
u = \frac{m_1^T P}{m_3^T P} = \frac{a_1^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + b_1}{a_3^T \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + b_3}
\end{align*}
\]
the plane defined by: \(u = 0\).

- Similarly:
  - \((a_2,b_2)\) describes the plane defined by: \(v = 0\)
  - \((a_3,b_3)\) describes the plane defined by:
    \[
    u = \infty \quad v = \infty
    \]
  \(\Rightarrow\) That is the plane passing through the pinhole \((z = 0)\).
Geometric Interpretation: The rows of the projection matrix describe the three planes defined by the image coordinate system.
Other useful geometric properties

Q: Given an image point $p$, what is the direction of the corresponding ray in space?

A: $w = A^{-1}p$

Q: Can we compute the position of the camera center $\Omega$?

A: $\Omega = -A^{-1}b$
Affine Cameras

Note: If the last row is

\[ m_3^T = \begin{bmatrix} 1 & 0 & 0 & 1 \end{bmatrix} \]

the coordinates equations degenerate to:

\[ u = m_1^T P = m_1 \cdot P \]
\[ v = m_2^T P = m_2 \cdot P \]

The mapping between world and image coordinates becomes linear. This is an affine camera.

- Example: Weak-perspective projection model
- Projection defined by 8 parameters
- Parallel lines are projected to parallel lines
- The transformation can be written as a direct linear transformation
perspective

weak perspective

increasing focal length

increasing distance from camera
Calibration: Recover $M$ from scene points $P_1,...,P_N$ and the corresponding projections in the image plane $p_1,...,p_N$.

In other words: Find $M$ that minimizes the distance between the actual points in the image, $p_i$, and their predicted projections $MP_i$.

Problems:
- The projection is (in general) non-linear
- $M$ is defined up to an arbitrary scale factor
Calibration: Recover $M$ from scene points $P_1,...,P_N$ and the corresponding projections in the image plane $p_1,...,p_N$.

Note:
1. This is the simplest possible way of thinking about calibration
2. This is one approach to the general problem called **PnP**: Relate $n$ points in image with $n$ 3D points
   - 3D model vs. image of object
   - Data from 3D sensor vs. corresponding image

Problems:
- The projection is (in general) non-linear
- $M$ is defined up to an arbitrary scale factor
Examples

3D + image data

Object recognition and localization (Hsiao&Collet)
The math for the calibration procedure follows a recipe that is used in many (most?) problems involving camera geometry, so it’s worth remembering:

Write relation between image point, projection matrix, and point in space:

\[ p_i = MP_i \]

Write non-linear relations between coordinates:

\[ u_i = \frac{m_1^T P_i}{m_3^T P_i} \quad v_i = \frac{m_2^T P_i}{m_3^T P_i} \]

Make them linear:

\[ m_1^T P_i - (m_3^T P_i) u_i = 0 \]
\[ m_2^T P_i - (m_3^T P_i) v_i = 0 \]
Put all the relations for all the points into a single matrix:

\[
\begin{bmatrix}
P_i^T & 0 & -u_i P_i^T \\
0 & P_i^T & -\nu_i P_i^T \\
\vdots & \vdots & \vdots \\
P_N^T & 0 & -u_N P_N^T \\
0 & P_N^T & -\nu_N P_N^T
\end{bmatrix} m = 0
\]

Solve by minimizing:

\[
|Lm|^2 = m^T L^T Lm
\]

Subject to:

\[
|m| = 1
\]
Slight digression: Homogeneous Least Squares

Suppose that we want to estimate the best vector of parameters $X$ from matrices $V_i$ computed from input data such that $VX_i = 0$

We can do this by minimizing:

$$
\sum_i |V_i X|^2 = \sum_i X^T V_i^T V_i \quad X = X^T VX
$$

Since $V=0$ is a trivial solution, we need to constrain the magnitude of $V$ to be non-zero, for example: $|V| = 1$. The problem becomes:

$$
\text{Min } X^T VX
\quad |X| = 1
$$

The key result (which we will use 50 times in this class) is:

*For any symmetric matrix $V$, the minimum of $X^T VX$ is reached at $X = \text{eigenvector of } V$ corresponding to the smallest eigenvalue.*
Calibration

\[(u_1, v_1), (u_i, v_i)\]

\[\left( u_i - \frac{m_1 \cdot P_i}{m_3 \cdot P_i} \right)^2 + \left( v_i - \frac{m_2 \cdot P_i}{m_3 \cdot P_i} \right)^2\]
Radial Distortion Model

Ideal:

\[ x' = f \frac{x}{z} \quad y' = f \frac{y}{z} \]

Distorted:

\[ x'' = \frac{1}{\lambda} x', \quad y'' = \frac{1}{\lambda} y', \quad \lambda = 1 + k_1 r^2 + k_2 r^4 + \cdots \]
We can follow exactly the same recipe with non-linear distortion:

Write non-linear relations between coordinates:

\[ u_i = \frac{1}{\lambda} \frac{m_1^T P_i}{m_3^T P_i} \quad v_i = \frac{1}{\lambda} \frac{m_2^T P_i}{m_3^T P_i} \]

Make them linear:

\[ v_i \begin{pmatrix} n_1^T P_i \end{pmatrix} - u_i \begin{pmatrix} n_2^T P_i \end{pmatrix} = 0 \]
Write them in matrix form:

\[
\begin{bmatrix}
  i P_i^T & -u_i P_i^T \\
\end{bmatrix} \begin{bmatrix} m \end{bmatrix} = 0 \quad m = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}
\]

Put all the relations for all the points into a single matrix:

\[
\begin{bmatrix}
  v_1 P_1^T & -u_1 P_1^T \\
  \vdots & \vdots \\
  \vdots & \vdots \\
  v_N P_N^T & -u_N P_N^T \\
\end{bmatrix} \begin{bmatrix} m \end{bmatrix} = 0
\]

Solve by minimizing:

\[
|Lm|^2 = m^T L^T Lm \\
L = \begin{bmatrix}
  v_1 P_1^T & -u_1 P_1^T \\
  \vdots & \vdots \\
  \vdots & \vdots \\
  v_N P_N^T & -u_N P_N^T \\
\end{bmatrix}
\]

Subject to:

\[
|m| = 1
\]
Key Results

Projection matrix: \( p = MP \)

\[
M = \begin{bmatrix}
\alpha & -\alpha \cot \theta & u_o \\
0 & \frac{\beta}{\sin \theta} & v_o \\
0 & 0 & 1
\end{bmatrix}
\]

\[
[R \quad t] = [A \quad b]
\]

5 intrinsic parameters 6 extrinsic parameters

Existence/unicity: \( \det(A) \neq 0 \)

Zero skew: \((a_1 \times a_3) \cdot (a_2 \times a_3) = 0\)

Aspect ratio 1: \( |a_1 \times a_3| = |a_2 \times a_3| \)
Planes: \( m_i \cdot P = 0 \)

Optical center: \( \Omega = -A^{-1}b \)

Viewing ray: \( w = A^{-1}p \)

Calibration: minimum 6 non-coplanar points

Linear: \( \min_{|m|=1} m^T L^T L m = \) Eigenvector of smallest eigenvalue of \( L^T L \)

Non-linear: \( \min_m \sum_i \left( u_i - \frac{m_1 \cdot P_i}{m_3 \cdot P_i} \right)^2 + \left( v_i - \frac{m_2 \cdot P_i}{m_3 \cdot P_i} \right)^2 \)
• More general calibration approaches (see “Geometric-Based Methods in Computer Vision”)

• MATLAB calibration toolbox

• OpenCV calibration package