BIL 719 - Computer Vision
April 06, 2012

Aykut Erdem
Dept. of Computer Engineering
Hacettepe University

Lecture #7
Cameras, projections and calibration
Today

- Pinhole cameras
- Cameras and lenses
- Geometric properties of pinhole cameras
- Other camera models
- Camera calibration
Camera and World Geometry

- How tall is this woman?
- How high is the camera?
- What is the camera rotation?
- What is the focal length of the camera?
- Which ball is closer?
Let’s design a camera

– Idea 1: put a piece of film in front of an object
– Do we get a reasonable image?
Idea 2: add a barrier to block off most of the rays

- This reduces blurring
- The opening known as the aperture
Pinhole camera model

- Pinhole model:
  - Captures **pencil of rays** – all rays through a single point
  - The point is called **Center of Projection (COP)**
  - The image is formed on the **Image Plane**
  - **Effective focal length** $f$ is distance from COP to Image Plane
Camera Obscura

Camera Obscura, Gemma Frisius, 1558

- The first camera
  - Known to Aristotle
  - Depth of the room is the effective focal length
Camera Obscura used for Tracing

Lens Based Camera Obscura, 1568
Vermeer and The Camera Obscura

http://www.essentialvermeer.com/camera_obscura

Officer and Laughing Girl, 1657
Camera Obscura

from BBC's Genius of Photography
Effect of pinhole size

(A) Source

(B) Source

Photo by Wandell, Foundations of Vision, Sinauer, 1995

Slide credit: B. Freeman and A. Torralba
Effect of pinhole size

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.
First Photograph

Oldest surviving photograph
- Took 8 hours on pewter plate

Joseph Niepce, 1826

Photograph of the first photograph

Stored at UT Austin

Niepce later teamed up with Daguerre, who eventually created Daguerrotypes

Slide credit: J. Hays
Cameras and Lenses

Shrinking aperture size

– Rays are mixed up

– Why the aperture cannot be too small?
  • Less light passes through
  • Diffraction effect
Cameras and Lenses

- A lens focuses light onto the film
  - Rays passing through the center are not deviated
  - All parallel rays converge to one point on a plane located at the **focal length** $f$
Dimensionality Reduction Machine
(3D to 2D)

3D world

2D image

Point of observation

Figures © Stephen E. Palmer, 2002
Projection can be tricky...
Projection can be tricky...
Projective Geometry

What is lost?

• Length

Who is taller?

Which is closer?

Slide credit: J. Hays
Length is not preserved
Projective Geometry

What is lost?

• Length
• Angles

Slide credit: J. Hays
Projective Geometry

What is preserved?

• Straight lines are still straight
Cartesian coordinates:
We have, by similar triangles, that
\((x, y, z) \rightarrow (f \frac{x}{z}, f \frac{y}{z}, -f)\)
Ignore the third coordinate, and get
\((x, y, z) \rightarrow (f \frac{x}{z}, f \frac{y}{z})\)
Geometric properties of projection

- Points go to points
- Lines go to lines
- Planes go to whole image or half-planes.
- Polygons go to polygons
- Degenerate cases
  - line through focal point to point
  - plane through focal point to line

Slide credit: B. Freeman and A. Torralba
Line in 3-space

\[ x(t) = x_0 + at \]
\[ y(t) = y_0 + bt \]
\[ z(t) = z_0 + ct \]

In the limit as \( t \to \pm \infty \)
we have (for \( c \neq 0 \)):

This tells us that any set of parallel lines (same \( a, b, c \) parameters)
project to the same point
(called the \textbf{vanishing point}).

Perspective projection
of that line

\[ x'(t) = \frac{fx}{z} = \frac{f(x_0 + at)}{z_0 + ct} \]
\[ y'(t) = \frac{fy}{z} = \frac{f(y_0 + bt)}{z_0 + ct} \]
Vanishing points and lines

Parallel lines in the world intersect in the image at a vanishing point
Vanishing points and lines

Vanishing point

Vertical vanishing point (at infinity)

Vanishing line

Vanishing point

Vanishing point

Photo by Criminisi

Slide credit: A. Efros
Vanishing points and lines

http://www.webexhibits.org/sciartperspective/tylerperspective.html
Vanishing points and lines
Vanishing points and lines
Vanishing points and lines

• Each set of parallel lines (=direction) meets at a different point
  – The *vanishing point* for this direction

• Sets of parallel lines on the same plane lead to *collinear* vanishing points.
  – The line is called the *horizon* for that plane

Slide credit: B. Freeman and A. Torralba
Reading Assignment

Automatic Photo Pop-up

Derek Hoiem  Alexei A. Efros  Martial Hebert
Carnegie Mellon University

Input Image  Automatic Photo Pop-up

• D. Hoiem, A.A. Efros and M. Hebert, Automatic Photo Pop-up, ACM SIGGRAPH 2005

http://www.cs.uiuc.edu/homes/dhoiem/projectspopup
What if you photograph a brick wall head-on?

Slide credit: B. Freeman and A. Torralba
Brick wall line in 3-space
\[ \begin{align*}
x(t) &= x_0 + at \\
y(t) &= y_0 \\
z(t) &= z_0
\end{align*} \]

Perspective projection of that line
\[ \begin{align*}
x'(t) &= \frac{f \cdot (x_0 + at)}{z_0} \\
y'(t) &= \frac{f \cdot y_0}{z_0}
\end{align*} \]

All bricks have same \( z_0 \). Those in same row have same \( y_0 \).

Thus, a brick wall, photographed head-on, gets rendered as set of parallel lines in the image plane.

Slide credit: B. Freeman and A. Torralba
A very large focal length camera approximates the orthographic projection.

$$(x, y, z) \rightarrow (x, y)$$
Other projection models:
Weak perspective

- Issue
  - perspective effects, but not over the scale of individual objects
  - collect points into a group at about the same depth, then divide each point by the depth of its group
  - Adv: easy
  - Disadv: only approximate

\[(x, y, z) \rightarrow \left( \frac{fx}{z_0}, \frac{fy}{z_0} \right)\]

Slide credit: B. Freeman and A. Torralba
Weak perspective projection

Qingming Festival by the Riverside  Zhang Zeduan ~900 AD
Weak perspective projection

An Ottoman miniature from the Surname-ı Vehbi  Abdulcelil Levni 1720
Three camera projections

(1) Perspective: \( (x, y, z) \mapsto \left( \frac{fx}{z}, \frac{fy}{z} \right) \)

(2) Weak perspective: \( (x, y, z) \mapsto \left( \frac{fx}{z_0}, \frac{fy}{z_0} \right) \)

(3) Orthographic: \( (x, y, z) \mapsto (x, y) \)

Slide credit: B. Freeman and A. Torralba
Homogeneous coordinates

Is this a linear transformation?
  • no—division by \( z \) is nonlinear

**Trick:** add one more coordinate:

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

homogeneous image coordinates  
homogeneous scene coordinates

Converting *from* homogeneous coordinates

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

Slide credit: S. Seitz
Perspective Projection

- Projection is a matrix multiply using homogeneous coordinates:

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

This is known as **perspective projection**
- The matrix is the projection matrix
Perspective Projection

• How does scaling the projection matrix change the transformation?

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

\[
\begin{bmatrix}
f & 0 & 0 & 0 \\
0 & f & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
y \\
z \\
1
\end{bmatrix} = \begin{bmatrix} fx \\ fy \\ z \end{bmatrix} \Rightarrow \left( f \frac{x}{z}, f \frac{y}{z} \right)
\]
Orthographic Projection

• Special case of perspective projection
  – Distance from the COP to the image plane is infinite

  – Also called *parallel projection*
  – What’s the projection matrix? \[
  \begin{bmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 0 & 1 \\
  0 & 0 & 0 & 1
  \end{bmatrix}
  \begin{bmatrix}
  x \\
  y \\
  z \\
  1
  \end{bmatrix}
  =
  \begin{bmatrix}
  x \\
  y \\
  1
  \end{bmatrix}
  \Rightarrow (x, y)\]
Scaled Orthographic Projection

- Special case of perspective projection
  - Object dimensions are small compared to distance to camera
  - Also called “weak perspective”
  - What’s the projection matrix?

\[
\begin{bmatrix}
  u \\
  v \\
  w \\
  1
\end{bmatrix} = \begin{bmatrix}
  f & 0 & 0 & 0 \\
  0 & f & 0 & 0 \\
  0 & 0 & 0 & s \\
  0 & 0 & 0 & 1
\end{bmatrix} \begin{bmatrix}
  x \\
  y \\
  z \\
  1
\end{bmatrix}
\]
2D Transformations

Translation

Euclidean

Similarity

Affine

Projective
**2D Transformations**

Example: translation

\[ x' = x + t \]
**2D Transformations**

Example: translation

\[ x' = x + t \]

\[ x' = \begin{bmatrix} 1 & t \end{bmatrix} \bar{x} \]

Slide credit: B. Freeman and A. Torralba
2D Transformations

Example: translation

\[ x' = x + t \]

\[ x' = \begin{bmatrix} I & t \end{bmatrix} \bar{x} \]

\[ \bar{x}' = \begin{bmatrix} I & t \\ 0^T & 1 \end{bmatrix} \bar{x} \]

Now we can chain transformations

Slide credit: B. Freeman and A. Torralba
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, see Szeliski, section 5.2, 5.3 for references

- (Relationship between intensities in the world and intensities in the image: photometric image formation, see Szeliski, sect. 2.2.)
One reason to calibrate a camera

Measuring height

vanishing line (horizon)

\[ v \cong (b \times b_0) \times (v_x \times v_y) \]

\[ t = (v \times t_0) \times (r \times b) \]

Slide credit: S. Seitz
Recovery of 3D structure

• Recovery of structure from one image is inherently ambiguous
Recovery of 3D structure

- Recovery of structure from one image is inherently ambiguous.
Intrinsic parameters: from idealized world coordinates to pixel values

Perspective projection

\[ u = f \frac{x}{z} \]
\[ v = f \frac{y}{z} \]

Figure by Forsyth & Ponce

Slide credit: B. Freeman and A. Torralba
Intrinsic parameters

But “pixels” are in some arbitrary spatial units

\[ u = \alpha \frac{x}{z} \]
\[ v = \alpha \frac{y}{z} \]
May be 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

\[ u = \alpha \frac{x}{z} + u_0 \]
\[ v = \beta \frac{y}{z} + v_0 \]

Figure by Forsyth & Ponce

Slide credit: B. Freeman and A. Torralba
Intrinsic parameters

May be skew between camera pixel axes

\[ u = \alpha \frac{x}{z} + u_0 \]
\[ v = \beta \frac{y}{z} + v_0 \]
Intrinsic parameters

May be skew between camera pixel axes

\[ u = \alpha \frac{x}{z} + u_0 \]

\[ v = \beta \frac{y}{z} + v_0 \]
Intrinsic parameters

May be skew between camera pixel axes

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

$\nu = \frac{\beta}{\sin(\theta)} \frac{y}{z} + \nu_0$

Figure by Forsyth & Ponce

Slide credit: B. Freeman and A. Torralba
Intrinsic parameters, homogeneous coordinates

Using homogenous coordinates, we can write this as:

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

In pixels \[\vec{p} = K \begin{pmatrix} x \\ y \end{pmatrix}\]
Camera parameters

- Intrinsic parameters
  - Principal point coordinates
  - Focal length
  - Pixel magnification factors
  - Skew (non-rectangular pixels)
  - Radial distortion

\[
K = \begin{bmatrix}
m_x & f & p_x \\
m_y & f & p_y \\
1 & 1 & 1 \\
\end{bmatrix} = \begin{bmatrix}
\alpha_x & \beta_x \\
\alpha_y & \beta_y \\
1 & 1 \\
\end{bmatrix}
\]
Camera parameters

• Intrinsic parameters
  – Principal point coordinates
  – Focal length
  – Pixel magnification factors
    – Skew (non-rectangular pixels)
    – Radial distortion

• Extrinsic parameters
  – Rotation and translation relative to world coordinate system
Extrinsic parameters: translation and rotation of camera frame

May be skew between camera pixel axes

\[
\begin{align*}
    u &= \alpha \frac{x}{z} - \alpha \cot(\theta) \frac{y}{z} + u_0 \\
    v &= \beta \frac{y}{\sin(\theta)} + v_0
\end{align*}
\]
Extrinsic parameters: translation and rotation of camera frame
Translated camera

\[ x = K[I \ t]X \]
3D Rotation of Points

Rotation around the coordinate axes, **counter-clockwise**:

\[
\begin{align*}
R_x(\alpha) &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \\
R_y(\beta) &= \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \\
R_z(\gamma) &= \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}
\end{align*}
\]

Slide credit: J. Hays, S. Saverese
Combining extrinsic and intrinsic calibration parameters, in homogeneous coordinates

\[
x = K[R \ t]X
\]

where

\[
\begin{bmatrix}
u \\ w
\end{bmatrix} = \begin{bmatrix}
\alpha & s & u_0 \\ 0 & \beta & v_0 \\ 1 & 0 & 0 & 1
\end{bmatrix}\begin{bmatrix}
r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1
\end{bmatrix}\begin{bmatrix}
x \\ y \\ z \\ 1
\end{bmatrix}
\]
Camera calibration

- Given n points with known 3D coordinates $X_i$ and known image projections $x_i$, estimate the camera parameters

You may use 1 or more images!
Camera calibration: Linear method

\[ \lambda x_i = PX_i \]

\[ x_i \times PX_i = 0 \]

\[ \left[ \begin{array}{ccc}
0 & -X_i^T & y_iX_i^T \\
X_i^T & 0 & -x_iX_i^T \\
-y_iX_i^T & x_iX_i^T & 0
\end{array} \right] \left[ \begin{array}{c}
P_1 \\
P_2 \\
P_3
\end{array} \right] = 0 \]

Two linearly independent equations

Slide credit: A. Efros
Camera calibration: Linear method

\[
\begin{bmatrix}
0^T & X_1^T & -y_1X_1^T \\
X_1^T & 0^T & -x_1X_1^T \\
\vdots & \vdots & \vdots \\
0^T & X_n^T & -y_nX_n^T \\
X_n^T & 0^T & -x_nX_n^T \\
\end{bmatrix}
\begin{pmatrix}
P_1 \\
P_2 \\
P_3 \\
\end{pmatrix}
= 0 \\
Ap = 0
\]

- \( P \) has 11 degrees of freedom (12 parameters, but scale is arbitrary)
- One 2D/3D correspondence gives us two linearly independent equations
- Homogeneous least squares
- 6 correspondences needed for a minimal solution

Slide credit: A. Efros
Camera calibration: Linear method

\[
\begin{bmatrix}
0^T & X_1^T & -y_1X_1^T \\
X_1^T & 0^T & -x_1X_1^T \\
\vdots & \vdots & \vdots \\
0^T & X_n^T & -y_nX_n^T \\
X_n^T & 0^T & -x_nX_n^T \\
\end{bmatrix}
\begin{pmatrix}
P_1 \\
P_2 \\
P_3 \\
\end{pmatrix} = 0 \quad \text{Ap} = 0
\]

- Note: for coplanar points that satisfy $\Pi^T X = 0$, we will get degenerate solutions $(\Pi,0,0)$, $(0,\Pi,0)$, or $(0,0,\Pi)$
Camera calibration: Linear method

- Advantages: easy to formulate and solve
- Disadvantages
  - Doesn’t directly tell you camera parameters
  - Doesn’t model radial distortion
  - Can’t impose constraints, such as known focal length and orthogonality

- Non-linear methods are preferred
  - Define error as difference between projected points and measured points
  - Minimize error using Newton’s method or other non-linear optimization

Slide credit: D. Hoiem
Calibration Demo

Camera Calibration Toolbox for Matlab

http://www.vision.caltech.edu/bouguetj/calib_doc/index.html
Calibration demo

Calibration images