Announcements

• Piazza: https://piazza.com/cornell/fall2018/cs6670
• OH
  • BH: Tue, Thur, 3-4 pm, 311 Gates Hall
  • GY: Fri 4-5 pm, G17 Gates Hall
• Course webpage: http://www.cs.cornell.edu/courses/cs6670/2018fa/
• Instructor webpage: http://home.bharathh.info/
Geometry of Image Formation
Today

• Geometry of image formation: where a pixel projects in the world
• Deriving perspective effects
• Ways of using perspective effects in recognition
Camera obscura
The pinhole camera

Let’s get into the math
The pinhole camera
The pinhole camera
The pinhole camera

\[ P = (X, Y, Z) \]

\[ p = (x, y) \]

\[ Z = -1 \]
The pinhole camera

\[ P = (X, Y, Z) \]

\[ p = (x, y) \]
The pinhole camera

\[ \mathbf{P} = (X, Y, Z) \]

\[ \mathbf{p} = (x, y) \]

\[ Q(\lambda) = 0 + \lambda (P - O) \]

\[ Q(\lambda) = (0 + \lambda(X - 0), 0 + \lambda(Y - 0), 0 + \lambda(Z - 0)) \]

\[ = (\lambda X, \lambda Y, \lambda Z) \]

\[ \lambda = 0 \Rightarrow Q(\lambda) = O \]

\[ \lambda = 1 \Rightarrow Q(\lambda) = P \]

\[ Z = -1 \]
The pinhole camera

- Pinhole camera collapses *ray OP to point p*

- Any point on ray \( OP = O + \lambda(P - O) = (\lambda X, \lambda Y, \lambda Z) \)

- For this point to lie on \( Z=-1 \) plane:
  \[ \lambda^* Z = -1 \]
  \[ \Rightarrow \lambda^* = \frac{-1}{Z} \]

- Coordinates of point \( p \):
  \[ (\lambda^* X, \lambda^* Y, \lambda^* Z) = \left( \frac{-X}{Z}, \frac{-Y}{Z}, -1 \right) \]
The projection equation

- A point $P = (X, Y, Z)$ in 3D projects to a point $p = (x, y)$ in the image

$$
x = \frac{-X}{Z}
$$

$$
y = \frac{-Y}{Z}
$$

- But pinhole camera’s image is inverted, invert it back!

$$
x = \frac{X}{Z}
$$

$$
y = \frac{Y}{Z}
$$
Another derivation

\[ P = (X, Y, Z) \]

\[ \frac{Y}{Z} = \frac{y}{1} \]
A virtual image plane

- A pinhole camera produces an inverted image
- Imagine a "virtual image plane" in the front of the camera
The projection equation

\[ x = \frac{X}{Z} \]
\[ y = \frac{Y}{Z} \]
Consequence 1: Farther away objects are smaller

Image of foot: \((\frac{X}{Z}, \frac{Y}{Z})\)

Image of head: \((\frac{X}{Z}, \frac{Y+h}{Z})\)

\[
\frac{Y+h}{Z} - \frac{Y}{Z} = \frac{h}{Z}
\]
Consequence 2: Parallel lines converge at a point

- Point on a line passing through point $A$ with direction $D$:
  \[ Q(\lambda) = A + \lambda D \]
- Parallel lines have the same direction but pass through different points:
  \[ Q(\lambda) = A + \lambda D \]
  \[ R(\lambda) = B + \lambda D \]
Consequence 2: Parallel lines converge at a point

- Parallel lines have the same direction but pass through different points

\[
Q(\lambda) = A + \lambda D \\
R(\lambda) = B + \lambda D
\]

- \( A = (A_X, A_Y, A_Z) \)
- \( B = (B_X, B_Y, B_Z) \)
- \( D = (D_X, D_Y, D_Z) \)
Consequence 2: Parallel lines converge at a point

- \[ Q(\lambda) = (A_X + \lambda D_X, A_Y + \lambda D_Y, A_Z + \lambda D_Z) \]
- \[ R(\lambda) = (B_X + \lambda D_X, B_Y + \lambda D_Y, B_Z + \lambda D_Z) \]
- \[ q(\lambda) = \left( \frac{A_X + \lambda D_X}{A_Z + \lambda D_Z}, \frac{A_Y + \lambda D_Y}{A_Z + \lambda D_Z} \right) \]
- \[ r(\lambda) = \left( \frac{B_X + \lambda D_X}{B_Z + \lambda D_Z}, \frac{B_Y + \lambda D_Y}{B_Z + \lambda D_Z} \right) \]
- Need to look at these points as \( Z \) goes to infinity
- Same as \( \lambda \to \infty \)
Consequence 2: Parallel lines converge at a point

\[ q(\lambda) = \left( \frac{A_X + \lambda D_X}{A_Z + \lambda D_Z}, \frac{A_Y + \lambda D_Y}{A_Z + \lambda D_Z} \right) \]

\[ r(\lambda) = \left( \frac{B_X + \lambda D_X}{B_Z + \lambda D_Z}, \frac{B_Y + \lambda D_Y}{B_Z + \lambda D_Z} \right) \]

\[
\lim_{\lambda \to \infty} \frac{A_X + \lambda D_X}{A_Z + \lambda D_Z} = \lim_{\lambda \to \infty} \frac{A_Y + \lambda D_Y}{A_Z + \lambda D_Z} = \frac{D_X}{D_Z}
\]

\[
\lim_{\lambda \to \infty} q(\lambda) = \left( \frac{D_X}{D_Z}, \frac{D_Y}{D_Z} \right) \quad \lim_{\lambda \to \infty} r(\lambda) = \left( \frac{D_X}{D_Z}, \frac{D_Y}{D_Z} \right)
\]
Consequence 2: Parallel lines converge at a point

• Parallel lines have the same direction but pass through different points
  \[ Q(\lambda) = A + \lambda D \]
  \[ R(\lambda) = B + \lambda D \]

• Parallel lines converge at the same point \( \left( \frac{D_x}{D_z}, \frac{D_y}{D_z} \right) \)

• This point of convergence is called the vanishing point

• What happens if \( D_z = 0? \)
Consequence 2: Parallel lines converge at a point
What about planes?

\[ N_X X + N_Y Y + N_Z Z = d \]

\[ \Rightarrow N_X \frac{X}{Z} + N_Y \frac{Y}{Z} + N_Z = \frac{d}{Z} \]

\[ \Rightarrow N_X x + N_Y y + N_Z = \frac{d}{Z} \]

Take the limit as \( Z \) approaches infinity

\[ N_X x + N_Y y + N_Z = 0 \]

Vanishing line of a plane
What about planes?

Normal: \((N_X, N_Y, N_Z)\)

What do parallel planes look like?

\[
N_X X + N_Y Y + N_Z Z = d
\]

\[
N_X x + N_Y y + N_Z = 0
\]

\[
N_X X + N_Y Y + N_Z Z = c
\]

\[
N_X x + N_Y y + N_Z = 0
\]

Vanishing lines

Parallel planes converge!
Vanishing line

\[ N_X X + N_Y Y + N_Z Z = d \]

• What happens if \( N_X = N_Y = 0 \)?
• Equation of the plane: \( Z = c \)
• Vanishing line?
Accidental pinholes
Accidental pinholes

Accidental pinholes

Geometry for recognition

• Which of these is likely to be an adult human?
The math

What is the vanishing line of the ground plane?

$Y = -c$
The math

$(X, H - c, Z)$

$(X, -c, Z)$

$Y = -c$
The math

\[(x, y_1)\]

\[(x, y_2)\]
The math

\[ y_2 = \frac{-c}{Z_c} \]

\[ \Rightarrow Z = \frac{Z_c}{y_2} \]

\[ y_1 = \frac{H - c}{Z} \]

\[ \Rightarrow H = Z y_1 + c \]

\[ \Rightarrow H = \frac{c(y_2 - y_1)}{y_2} = \frac{ch}{|y_2|} \]
Geometry for recognition

(f) $P(\text{person} \mid \text{viewpoint})$

Changing coordinate systems

P = (X,Y,Z)
Changing coordinate systems
Changing coordinate systems
Changing coordinate systems
Changing coordinate systems
Changing coordinate systems
Rotations and translations

• How do you represent a rotation?
• A point in 3D: (X,Y,Z)
• Rotations can be represented as a matrix multiplication

\[ \mathbf{v}' = \mathbf{Rv} \]

• What are the properties of rotation matrices?
Properties of rotation matrices

• Rotation does not change the length of vectors

\[ \mathbf{v}' = R \mathbf{v} \]
\[ \| \mathbf{v}' \|^2 = \mathbf{v}'^T \mathbf{v}' \]
\[ = \mathbf{v}^T R^T R \mathbf{v} \]
\[ \| \mathbf{v} \|^2 = \mathbf{v}^T \mathbf{v} \]
\[ \Rightarrow R^T R = I \]
Properties of rotation matrices

\[ R^T R = I \]

\[ \Rightarrow \det(R)^2 = 1 \]

\[ \Rightarrow \det(R) = \pm 1 \]

\[
\begin{align*}
\det(R) = 1 & \quad \text{Rotation} \\
\det(R) = -1 & \quad \text{Reflection}
\end{align*}
\]
Rotation matrices

• Rotations in 3D have an axis and an angle
• Axis: vector that does not change when rotated
  \[ Rv = v \]
• Rotation matrix has eigenvector that has eigenvalue 1
Rotation matrices from axis and angle

• Rotation matrix for rotation about axis $\mathbf{v}$ and $\theta$
• First define the following matrix

$$
[\mathbf{v}] \times = \begin{bmatrix}
0 & -v_z & v_y \\
v_z & 0 & -v_x \\
-v_y & v_x & 0
\end{bmatrix}
$$

• Interesting fact: this matrix represents cross product

$$
[\mathbf{v}] \times \mathbf{x} = \mathbf{v} \times \mathbf{x}
$$
Rotation matrices from axis and angle

• Rotation matrix for rotation about axis $\mathbf{v}$ and $\theta$
• Rodrigues’ formula for rotation matrices

$$R = I + (\sin \theta)[\mathbf{v}] \times + (1 - \cos \theta)[\mathbf{v}]^2 \times$$
Translations

\[ x' = x + t \]

- Can this be written as a matrix multiplication?
Putting everything together

• Change coordinate system so that center of the coordinate system is at pinhole and Z axis is along viewing direction

\[ \mathbf{x}'_w = R \mathbf{x}_w + \mathbf{t} \]

• Perspective projection

\[ \mathbf{x}'_w \equiv (X, Y, Z) \]
\[ \mathbf{x}'_{img} \equiv (x, y) \]
\[ x = \frac{X}{Z} \]
\[ y = \frac{Y}{Z} \]
The projection equation

\[ x = \frac{X}{Z} \]
\[ y = \frac{Y}{Z} \]

• Is this equation linear?
• Can this equation be represented by a matrix multiplication?
Is projection linear?

\[
\begin{align*}
X' &= aX + b \\
Y' &= aY + b \\
Z' &= aZ + b
\end{align*}
\]

\[
\begin{align*}
x' &= \frac{aX + b}{aZ + b} \\
y' &= \frac{aY + b}{aZ + b}
\end{align*}
\]
Can projection be represented as a matrix multiplication?

Matrix multiplication:

\[
\begin{bmatrix}
  a & b & c \\
p & q & r
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix} =
\begin{bmatrix}
  aX + bY + cZ \\
pX + qY + rZ
\end{bmatrix}
\]

Perspective projection:

\[
x = \frac{X}{Z}
\]

\[
y = \frac{Y}{Z}
\]
The space of rays

- Every point on a ray maps it to a point on the image plane.
- Perspective projection maps rays to points.
- All points \((\lambda x, \lambda y, \lambda)\) map to the same image point \((x, y, 1)\).
Projective space

- Standard 2D space (plane) $\mathbb{R}^2$: Each point represented by 2 coordinates $(x,y)$

- Projective 2D space (plane) $\mathbb{P}^2$: Each “point” represented by 3 coordinates $(x,y,z)$, BUT:
  - $(\lambda x, \lambda y, \lambda z) \equiv (x, y, z)$

- Mapping $\mathbb{R}^2$ to $\mathbb{P}^2$ (points to rays):
  $$(x, y) \rightarrow (x, y, 1)$$

- Mapping $\mathbb{P}^2$ to $\mathbb{R}^2$ (rays to points):
  $$(x, y, z) \rightarrow \left(\frac{x}{z}, \frac{y}{z}\right)$$
Projective space and homogenous coordinates

• Mapping $\mathbb{R}^2$ to $\mathbb{P}^2$ (points to rays):
  $$(x, y) \rightarrow (x, y, 1)$$

• Mapping $\mathbb{P}^2$ to $\mathbb{R}^2$ (rays to points):
  $$(x, y, z) \rightarrow \left(\frac{x}{z}, \frac{y}{z}\right)$$

• A change of coordinates
• Also called *homogenous coordinates*
Homogenous coordinates

• In standard Euclidean coordinates
  • 2D points : (x,y)
  • 3D points : (x,y,z)

• In homogenous coordinates
  • 2D points : (x,y,1)
  • 3D points : (x,y,z,1)
Why homogenous coordinates?

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1 \\
\end{bmatrix}
= 
\begin{bmatrix}
X \\
Y \\
Z \\
1 \\
\end{bmatrix} 
\equiv 
\begin{bmatrix}
\frac{X}{Z} \\
\frac{Y}{Z} \\
1 \\
\end{bmatrix}
\]

Homogenous coordinates of world point

Homogenous coordinates of image point
Why homogenous coordinates?

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 \\
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1 \\
\end{bmatrix} =
\begin{bmatrix}
X \\
Y \\
Z \\
1 \\
\end{bmatrix} \equiv
\begin{bmatrix}
\frac{X}{Z} \\
\frac{Y}{Z} \\
1 \\
\end{bmatrix}
\]

\[P\vec{x}_w = \vec{x}_{img}\]

- Perspective projection is matrix multiplication in homogenous coordinates!
Why homogenous coordinates?

\[
\begin{bmatrix}
1 & 0 & 0 & t_x \\
0 & 1 & 0 & t_y \\
0 & 0 & 1 & t_z \\
0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]

• Translation is matrix multiplication in homogenous coordinates!
Homogenous coordinates

\[
\begin{bmatrix}
a & b & c & t_x \\
d & e & f & t_y \\
g & h & i & t_z \\
0 & 0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1 \\
\end{bmatrix}
= 
\begin{bmatrix}
\begin{array}{c}
aX + bY + cZ + t_x \\
dX + eY + fZ + t_y \\
gX + hY + iZ + t_z \\
1 \\
\end{array}
\end{bmatrix}
\]

\[
\begin{bmatrix}
M & t \\
0^T & 1 \\
\end{bmatrix}
\begin{bmatrix}
\mathbf{x}_w \\
1 \\
\end{bmatrix}
= 
\begin{bmatrix}
\begin{array}{c}
M\mathbf{x}_w + t \\
1 \\
\end{array}
\end{bmatrix}
\]
Homogenous coordinates

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
= 
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix} 
\equiv 
\begin{bmatrix}
\frac{X}{Z} \\
\frac{Y}{Z} \\
\frac{Z}{Z}
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}
\rightarrow 
\begin{bmatrix}
I & 0
\end{bmatrix}
\]
Perspective projection in homogenous coordinates

\[
\vec{x}_{img} = \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \vec{x}_w
\]

\[
\vec{x}_{img} = \begin{bmatrix} R & t \end{bmatrix} \vec{x}_w
\]
More about matrix transformations

\[
\begin{bmatrix}
I & 0 \\
I & t \\
0^T & 1 \\
M & t \\
0^T & 1 \\
\end{bmatrix}
\]

- 3 x 4 : Perspective projection
- 4 x 4 : Translation
- 4 x 4 : Affine transformation (linear transformation + translation)
More about matrix transformations

\[
\begin{bmatrix}
M & t \\
0^T & 1
\end{bmatrix}
\]

\[M^T M = I\]

Euclidean
More about matrix transformations

\[
\begin{bmatrix}
M & t \\
0^T & 1
\end{bmatrix}
\]

\[M = sR\]

\[R^T R = I\]

Similarity transformation
More about matrix transformations

\[
\begin{bmatrix}
M & t \\
0^T & 1
\end{bmatrix}
\]

\[
M = \begin{bmatrix}
  s_x & 0 & 0 \\
  0 & s_y & 0 \\
  0 & 0 & s_z
\end{bmatrix}
\]

Anisotropic scaling and translation
More about matrix transformations

\[
\begin{bmatrix}
M & t \\
0^T & 1
\end{bmatrix}
\]

General affine transformation
Matrix transformations in 2D
Perspective projection in homogenous coordinates

\[
\vec{x}_{img} \equiv [I \ 0] \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \vec{x}_w
\]

\[
\vec{x}_{img} \equiv [R \ t] \vec{x}_w
\]
Matrix transformations in 2D

\[ \vec{x}_{img} \equiv K \begin{bmatrix} R & t \end{bmatrix} \vec{x}_w \]

\[ K = \begin{bmatrix} 1 & 0 & t_u \\ 0 & 1 & t_v \\ 0 & 0 & 1 \end{bmatrix} \]

Translation

\[ K = \begin{bmatrix} s_x & 0 & t_u \\ 0 & s_y & t_v \\ 0 & 0 & 1 \end{bmatrix} \]

Scaling of Image x and y (conversion from “meters” to “pixels”)

\[ K = \begin{bmatrix} s_x & \alpha & t_u \\ 0 & s_y & t_v \\ 0 & 0 & 1 \end{bmatrix} \]

Added skew if image x and y axes are not perpendicular
Final perspective projection

Camera extrinsics: where your camera is relative to the world. Changes if you move the camera

\[ \vec{x}_{img} \equiv K [R \ t] \vec{x}_w \]

Camera intrinsics: how your camera handles pixel. Changes if you change your camera

\[ \vec{x}_{img} \equiv P \vec{x}_w \]
Final perspective projection

\[ \vec{x}_{img} \equiv K \begin{bmatrix} R & t \end{bmatrix} \vec{x}_w \]

Camera parameters

\[ \vec{x}_{img} \equiv P \vec{x}_w \]