Stereo 1

CSC420
David Lindell
University of Toronto
cs.toronto.edu/~lindell/teaching/420
Slide credit: Babak Taati ← Ahmed Ashraf ← Sanja Fidler, Yannis Gkioulekas
Logistics

- A4 is out. Due date is March 29

- Final exam April 18\textsuperscript{th} BA1130 9AM – 12 PM
  - multiple choice, short answer, long answer
Overview

- Recap camera matrix and perspective projection
- Two-view geometry
Overview

• Recap camera matrix and perspective projection

• Two-view geometry
The camera as a coordinate transformation

A camera is a mapping from:

the 3D world
to:

a 2D image

3D object

3D to 2D transform
(camera)

2D image

2D to 2D transform
(image warping)
The camera as a coordinate transformation

A camera is a mapping from:

the 3D world

to:

a 2D image

What are the dimensions of each variable?

\[ x = PX \]

homogeneous coordinates

2D image point

camera matrix

3D world point
The camera as a coordinate transformation

\[ x = \mathbf{P}X \]

\[
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}
= 
\begin{bmatrix}
p_1 & p_2 & p_3 & p_4 \\
p_5 & p_6 & p_7 & p_8 \\
p_9 & p_{10} & p_{11} & p_{12}
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]

homogeneous image coordinates 3 x 1  
camera matrix 3 x 4  
homogeneous world coordinates 4 x 1
World-to-camera coordinate system transformation

Coordinate of the camera center in the world coordinate frame

World coordinate system

Camera coordinate system

\( \mathbf{X}_w \)
World-to-camera coordinate system transformation

Why aren't the points aligned?

\[(\tilde{X}_w - \tilde{C})\]

translate
World-to-camera coordinate system transformation

\[
R \cdot (\tilde{X}_w - \tilde{C})
\]

points now coincide

rotate translate
Modeling the coordinate system transformation

In heterogeneous coordinates, we have:

\[ \tilde{X}_c = R \cdot (\tilde{X}_w - \tilde{C}) \]
Modeling the coordinate system transformation

In heterogeneous coordinates, we have:

\[
\mathbf{X}_c = \mathbf{R} \cdot (\mathbf{X}_w - \tilde{\mathbf{C}})
\]

In homogeneous coordinates, we have:

\[
\begin{bmatrix}
X_c \\
Y_c \\
Z_c \\
1
\end{bmatrix} = \begin{bmatrix}
\mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\
0 & 1
\end{bmatrix} \begin{bmatrix}
X_w \\
Y_w \\
Z_w \\
1
\end{bmatrix}
\text{ or } \mathbf{X}_c = \begin{bmatrix}
\mathbf{R} & -\mathbf{R}\tilde{\mathbf{C}} \\
0 & 1
\end{bmatrix} \mathbf{X}_w
\]
Putting it all together

We can write everything into a single projection:

\[ x = PX_w \]
Putting it all together

We can write everything into a single projection:

\[ \mathbf{x} = \mathbf{PX}_w \]

The camera matrix now looks like:

\[
\mathbf{P} = \begin{bmatrix}
    f & 0 & p_x \\
    0 & f & p_y \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    \mathbf{I} & 0
\end{bmatrix}
\begin{bmatrix}
    \mathbf{R} & -\mathbf{R}\mathbf{C} \\
    0 & 1
\end{bmatrix}
\]
Putting it all together

We can write everything into a single projection:

\[ \mathbf{x} = \mathbf{PX}_w \]

The camera matrix now looks like:

\[
\mathbf{P} = \begin{bmatrix}
  f & 0 & p_x \\
  0 & f & p_y \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  \mathbf{I} & 0 \\
  \mathbf{R} & -\mathbf{R}\mathbf{\tilde{C}}
\end{bmatrix}
\]

intrinsic parameters (3 x 3): correspond to camera internals (image-to-image transformation)
Putting it all together

We can write everything into a single projection:

\[ \mathbf{x} = \mathbf{PX}_w \]

The camera matrix now looks like:

\[
\mathbf{P} = \begin{bmatrix}
    f & 0 & p_x \\
    0 & f & p_y \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    1 & 0 \\
    0 & 1 \\
    0 & 0
\end{bmatrix}
\begin{bmatrix}
    \mathbf{R} & -\mathbf{R}\mathbf{C} \\
    0 & 1
\end{bmatrix}
\]

intrinsic parameters (3 x 3): correspond to camera internals (image-to-image transformation)

perspective projection (3 x 4): maps 3D to 2D points (camera-to-image transformation)
Putting it all together

We can write everything into a single projection:

\[
x = PX_w
\]

The camera matrix now looks like:

\[
P = \begin{bmatrix} f & 0 & px \\ 0 & f & py \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} I & 0 \end{bmatrix} \begin{bmatrix} R & -R\hat{C} \\ 0 & 1 \end{bmatrix}
\]

- **Intrinsic parameters** (3 x 3): correspond to camera internals (image-to-image transformation)
- **Perspective projection** (3 x 4): maps 3D to 2D points (camera-to-image transformation)
- **Extrinsic parameters** (4 x 4): correspond to camera externals (world-to-camera transformation)
Putting it all together

We can write everything into a single projection:

\[ \mathbf{x} = \mathbf{PX}_w \]

The camera matrix now looks like:

\[
\mathbf{P} = \begin{bmatrix}
    f & 0 & \mathbf{p}_x \\
    0 & f & \mathbf{p}_y \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    \mathbf{R} & -\mathbf{RC}
\end{bmatrix}
\]

**Intrinsic parameters (3 x 3):** correspond to camera internals

**Extrinsic parameters (3 x 4):** correspond to camera externals (world-to-image transformation)
We can decompose the camera matrix like this:

\[ P = KR[I] - C \]

(translate first then rotate)
General pinhole camera matrix

We can decompose the camera matrix like this:

\[ P = KR[I - C] \]

(translate first then rotate)

Another way to write the mapping:

\[ P = K[R|t] \]

where  \( t = -RC \)

(rotate first then translate)
General pinhole camera matrix

\[ P = K[R|t] \]
General pinhole camera matrix

\[ P = K[R|t] \]

\[
P = \begin{bmatrix}
f & 0 & p_x \\
0 & f & p_y \\
0 & 0 & 1 \\
\end{bmatrix}
\begin{bmatrix}
r_1 & r_2 & r_3 & t_1 \\
r_4 & r_5 & r_6 & t_2 \\
r_7 & r_8 & r_9 & t_3 \\
\end{bmatrix}
\]

intrinsic parameters  extrinsic parameters
General pinhole camera matrix

\[ P = K[R|t] \]

\[
P = \begin{bmatrix}
f & 0 & p_x \\
0 & f & p_y \\
0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
r_1 & r_2 & r_3 & t_1 \\
r_4 & r_5 & r_6 & t_2 \\
r_7 & r_8 & r_9 & t_3
\end{bmatrix}
\]

- **intrinsic parameters**
- **extrinsic parameters**

\[
R = \begin{bmatrix}
r_1 & r_2 & r_3 \\
r_4 & r_5 & r_6 \\
r_7 & r_8 & r_9
\end{bmatrix}
\]

\[
t = \begin{bmatrix}
t_1 \\
t_2 \\
t_3
\end{bmatrix}
\]

- **3D rotation**
- **3D translation**
Recap

What is the size and meaning of each term in the camera matrix?

\[ P = KR[I] - C \]

Recap

What is the size and meaning of each term in the camera matrix?

\[ P = KR[I| - C] \]

3x3 intrinsics
Recap

What is the size and meaning of each term in the camera matrix?

\[ P = KR[I - C] \]

3x3 intrinsics 3x3 3D rotation
Recap

What is the size and meaning of each term in the camera matrix?

\[ P = K R [I] - C \]

- 3x3 intrinsics
- 3x3 3D rotation
- 3x3 identity
- ?
Recap

What is the size and meaning of each term in the camera matrix?

\[ P = KR[I] - C \]

- 3x3 intrinsics
- 3x3 3D rotation
- 3x3 identity
- 3x1 3D translation
Quiz

The camera matrix relates what two quantities?

\[ x = PX \]
Quiz

The camera matrix relates what two quantities?

$$ x = PX $$

homogeneous 3D points to 2D image points
Quiz

The camera matrix relates what two quantities?

\[ x = PX \]

homogeneous 3D points to 2D image points

The camera matrix can be decomposed into?
Quiz

The camera matrix relates what two quantities?

\[ x = PX \]

homogeneous 3D points to 2D image points

The camera matrix can be decomposed into?

\[ P = K[R|t] \]

intrinsic and extrinsic parameters
Perspective distortion
Forced perspective
The Ames room illusion
The Ames room illusion
Magnification depends on depth

real-world object

depth 2 Z
Magnification depends on depth

real-world object

depth Z

depth 2 Z
Magnification depends on depth

What happens as we change the focal length?

real-world object

depth $Z$

depth $2Z$
Magnification depends on focal length
What if…

real-world object

1. Set focal length to half

depth 2 Z

focal length f

focal length 2 f
What if...

1. Set focal length to half
2. Set depth to half

Is this the same image as the one I had at focal length 2f and distance 2Z?
Perspective distortion

long focal length  
mid focal length  
short focal length
Perspective distortion
Vertigo effect

Named after Alfred Hitchcock’s movie
• also known as “dolly zoom”
Vertigo effect

How would you create this effect?
Other camera models
What if…

... we continue increasing $Z$ and $f$ while maintaining the same magnification?

$f \to \infty$ and $\frac{f}{Z} = \text{constant}$
camera is close to object and has small focal length

camera is far from object and has large focal length
Different cameras

Perspective camera

Weak perspective camera
Weak perspective vs perspective camera

\[
[X \ Y \ Z]^\top \mapsto [fX/Z_0 \ fY/Z_0]^\top
\]

- magnification does not change with depth
- constant magnification depending on \( f \) and \( Z_0 \)

[magnification changes with depth] \([X \ Y \ Z]^\top \mapsto [fX/Z \ fY/Z]^\top\)
When can we assume a weak perspective camera?

1. When the scene (or parts of it) is very far away.

Weak perspective projection applies to the mountains.
Orthographic camera

Special case of weak perspective camera where:
• constant magnification is equal to 1.
Orthographic camera

Special case of weak perspective camera where:
• constant magnification is equal to 1.
• there is no shift between camera and image origins.
Orthographic camera

Special case of weak perspective camera where:
• constant magnification is equal to 1.
• there is no shift between camera and image origins.
• the world and camera coordinate systems are the same.
Orthographic camera

Special case of weak perspective camera where:
• constant magnification is equal to 1.
• there is no shift between camera and image origins.
• the world and camera coordinate systems are the same.

What is the camera matrix in this case?
Orthographic camera

Special case of weak perspective camera where:
• constant magnification is equal to 1.
• there is no shift between camera and image origins.
• the world and camera coordinate systems are the same.

\[
P = \begin{bmatrix}
1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 1
\end{bmatrix}
\]
Overview

• Recap camera matrix and perspective projection

• Two-view geometry
Homography

• In Lecture 8 we said that a homography is a transformation that maps a projective plane to another projective plane.

• Defined by the following:

\[
\begin{bmatrix}
    x' \\
    y' \\
    1
\end{bmatrix} =
\begin{bmatrix}
    a & b & c \\
    d & e & f \\
    g & h & i
\end{bmatrix}
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix}
\]
Homography

• Let’s revisit our transformation in the (new) light of perspective projection.
Homography

- Let’s revisit our transformation in the (new) light of perspective projection.

We have our object in two different worlds, in two different poses relative to camera, two different photographers, and two different cameras.
Homography

- Let’s revisit our transformation in the (new) light of perspective projection.

Our object is a plane. Each plane is characterized by one point \( d \) on the plane and two independent vectors \( a \) and \( b \) on the plane.
Let’s revisit our transformation in the (new) light of perspective projection.

Then any other point $X$ on the plane can be written as: $X = d + \alpha a + \beta b$; where $\alpha$ and $\beta$ are in the DVD’s coordinate system defined by its basis vectors and origin.
Homography

- Let’s revisit our transformation in the (new) light of perspective projection.

Any two Chicken Run DVDs on our planet are related by some transformation $T$. We’ll compute it, don’t worry.
Homography

• Let’s revisit our transformation in the (new) light of perspective projection.

Each object is seen by a different camera and thus projects to the corresponding image plane with different camera intrinsics.
Homography

• Let’s revisit our transformation in the (new) light of perspective projection.

Given this, the question is what’s the transformation that maps the DVD on the first image to the DVD in the second image?
Homography

- Each point on a plane can be written as: $X = d + \alpha \cdot a + \beta \cdot b$, where $d$ is a point, and $a$ and $b$ are two independent directions on the plane.
Homography

• Each point on a plane can be written as: $X = d + \alpha \cdot a + \beta \cdot b$, where $d$ is a point, and $a$ and $b$ are two independent directions on the plane.

• Let’s have two different planes in 3D:

  First plane: $X_1 = d_1 + \alpha \cdot a_1 + \beta \cdot b_1$

  Second plane: $X_2 = d_2 + \alpha \cdot a_2 + \beta \cdot b_2$
Homography

- Each point on a plane can be written as: $X = d + \alpha \cdot a + \beta \cdot b$, where $d$ is a point, and $a$ and $b$ are two independent directions on the plane.

- Let’s have two different planes in 3D:

  First plane : $X_1 = d_1 + \alpha \cdot a_1 + \beta \cdot b_1$

  Second plane : $X_2 = d_2 + \alpha \cdot a_2 + \beta \cdot b_2$
Homography

• Each point on a plane can be written as: \( X = d + \alpha \cdot a + \beta \cdot b \), where \( d \) is a point, and \( a \) and \( b \) are two independent directions on the plane.

• Let’s have two different planes in 3D:

  First plane : \( X_1 = d_1 + \alpha \cdot a_1 + \beta \cdot b_1 \)
  Second plane : \( X_2 = d_2 + \alpha \cdot a_2 + \beta \cdot b_2 \)

• Via \( \alpha \) and \( \beta \), the two points \( X_1 \) and \( X_2 \) are in the same location relative to each plane (correspondences!)
Homography

- Each point on a plane can be written as: \( X = d + \alpha \cdot a + \beta \cdot b \), where \( d \) is a point, and \( a \) and \( b \) are two independent directions on the plane.

- Let's have two different planes in 3D:
  \[
  \text{First plane : } X_1 = d_1 + \alpha \cdot a_1 + \beta \cdot b_1 \\
  \text{Second plane : } X_2 = d_2 + \alpha \cdot a_2 + \beta \cdot b_2
  \]

- Via \( \alpha \) and \( \beta \), the two points \( X_1 \) and \( X_2 \) are in the same location relative to each plane.

- We can rewrite this using homogeneous coordinates:
  \[
  \text{First plane : } X_1 = \begin{bmatrix} a_1 & b_1 & d_1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix} = A_1 \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix}
  \]
Homography

- Each point on a plane can be written as: $X = d + \alpha \cdot a + \beta \cdot b$, where $d$ is a point, and $a$ and $b$ are two independent directions on the plane.

- Let's have two different planes in 3D:

  First plane: $X_1 = d_1 + \alpha \cdot a_1 + \beta \cdot b_1$

  Second plane: $X_2 = d_2 + \alpha \cdot a_2 + \beta \cdot b_2$

- Via $\alpha$ and $\beta$, the two points $X_1$ and $X_2$ are in the same location relative to each plane.

- We can rewrite this using homogeneous coordinates:

  First plane: $X_1 = \begin{bmatrix} a_1 & b_1 & d_1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix} = A_1 \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix}$

  Second plane: $X_2 = \begin{bmatrix} a_2 & b_2 & d_2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix} = A_2 \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix}$
Homography

• Each point on a plane can be written as: \( X = d + \alpha \cdot a + \beta \cdot b \), where \( d \) is a point, and \( a \) and \( b \) are two independent directions on the plane.

• Let’s have two different planes in 3D:

  First plane: \( X_1 = d_1 + \alpha \cdot a_1 + \beta \cdot b_1 \)

  Second plane: \( X_2 = d_2 + \alpha \cdot a_2 + \beta \cdot b_2 \)

• Via \( \alpha \) and \( \beta \), the two points \( X_1 \) and \( X_2 \) are in the same location relative to each plane.

• We can rewrite this using homogeneous coordinates:

\[
\begin{align*}
\text{First plane: } X_1 &= \begin{bmatrix} a_1 & b_1 & d_1 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix} = A_1 \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix} \\
\text{Second plane: } X_2 &= \begin{bmatrix} a_2 & b_2 & d_2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix} = A_2 \begin{bmatrix} \alpha \\ \beta \\ 1 \end{bmatrix}
\end{align*}
\]

• \( A_1 = \begin{bmatrix} a_1 & b_1 & d_1 \end{bmatrix} \) and \( A_2 = \begin{bmatrix} a_2 & b_2 & d_2 \end{bmatrix} \) are 3 x 3 matrices.
Homography

• In 3D, a transformation between the planes is given by:

\[ X_2 = T X_1 \]

There is one transformation \( T \) between every pair of points \( X_1 \) and \( X_2 \).
Homography

• In 3D, a transformation between the planes is given by:

\[ x_2 = T x_1 \]

There is one transformation \( T \) between every pair of points \( x_1 \) and \( x_2 \).

• Expand it:

\[
\begin{bmatrix}
\alpha \\
\beta \\
1
\end{bmatrix}
= T
\begin{bmatrix}
\alpha \\
\beta \\
1
\end{bmatrix}
\text{ for every } \alpha, \beta
Homography

• In 3D, a transformation between the planes is given by:

\[ X_2 = T X_1 \]

There is one transformation \( T \) between every pair of points \( X_1 \) and \( X_2 \).

• Expand it:

\[
\begin{bmatrix}
\alpha \\
\beta \\
1
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta \\
1
\end{bmatrix} =
\begin{bmatrix}
A_2 \\
A_1
\end{bmatrix}
\begin{bmatrix}
\alpha \\
\beta \\
1
\end{bmatrix}
\]

for every \( \alpha, \beta \)

• Then it follows: \( T = A_2 A_1^{-1} \), with \( T \) a \( 3 \times 3 \) matrix.
Homography

• Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

\[
\begin{align*}
    w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} &= K_1 X_1 \\
    w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} &= K_2 X_2
\end{align*}
\]
Homography

- Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

  \[
  w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K_1 X_1 \quad \text{and} \quad w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 X_2
  \]

- Insert $X_2 = T X_1$ into the equality on the right

  \[
  w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 T X_1
  \]
Homography

• Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

$$ \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K_1 X_1 \quad \text{and} \quad \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 X_2 $$

• Insert $X_2 = T X$ into the equality on the right

$$ \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 T X_1 = K_2 T (K_1^{-1} K_1) X_1 $$
Homography

• Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

\[
w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K_1 X_1 \quad \text{and} \quad w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 X_2
\]

• Insert $X_2 = T X$ into the equality on the right

\[
w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 T X_1 = K_2 T (K_1^{-1} K_1) X_1
\]

\[
w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}
\]
Homography

- Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

\[
\begin{align*}
\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} &= K_1 X_1 \\
\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} &= K_2 X_2
\end{align*}
\]

- Insert $X_2 = T X$ into the equality on the right

\[
\begin{align*}
\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} &= K_2 \ T \ X_1 \\
&= K_2 \ T \ (K_1^{-1} K_1) X_1 \\
&= w_1 \ K_2 \ T \ K_1^{-1} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}
\end{align*}
\]
Homography

- Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

$$w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K_1 X_1 \quad \text{and} \quad w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 X_2$$

- Insert $X_2 = T X$ into the equality on the right

$$w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 T X_1 = K_2 T (K_1^{-1} K_1) X_1 = w_1 \underbrace{K_2 T K_1^{-1}}_{3 \times 3 \text{ matrix}} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$
Homography

- Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

$$w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K_1 X_1 \quad \text{and} \quad w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 X_2$$

- Insert $X_2 = T X$ into the equality on the right

$$w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 T X_1 = K_2 T (K^{-1}_1 K_1) X_1 = w_1 \underbrace{K_2 T K^{-1}_1}_{3 \times 3 \text{ matrix}} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

- Finally, divide through by $w_1$

$$w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$
Homography

- Let’s look at what happens in projective (image) plane. Note that we have each plane in a separate image and the two images may not have the same camera intrinsic parameters. Denote them with $K_1$ and $K_2$.

$$w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K_1 X_1 \quad \text{and} \quad w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 X_2$$

- Insert $X_2 = TX$ into the equality on the right

$$w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K_2 T X_1 = K_2 T (K_1^{-1} K_1) X_1 = w_1 \underbrace{K_2 T K_1^{-1}}_{3 \times 3 \text{ matrix}} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

- Finally, divide through by $w_1$

$$w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

what is this?
Homography

• If we want to compute correspondences between images and we have the homography, what else do we need?
  • 3D positions?
  • Camera intrinsics?
Homography

- If we want to compute correspondences between images and we have the homography, what else do we need?
  - 3D positions?
  - Camera intrinsics?

- Still one more loose end from lecture 8 to recap…
Remember Panorama Stitching from Lecture 9?

Take a tripod, rotate camera and take pictures

[Source: Fernando Flores-Mangas]
Remember Panorama Stitching from Lecture 9?

• Each pair of images is related by homography. Why?

[Source: Fernando Flores-Mangas]
Rotating the Camera

• Rotating my camera with $R$ is the same as rotating the 3D points with $R^T$ (inverse of $R$):

$$X_2 = R^T X_1$$

• where $X_1$ is a 3D point in the coordinate system of the first camera and $X_2$ the 3D point in the coordinate system of the rotated camera.
Rotating the Camera

• Rotating my camera with $R$ is the same as rotating the 3D points with $R^T$ (inverse of $R$):

$$X_2 = R^T X_1$$

• where $X_1$ is a 3D point in the coordinate system of the first camera and $X_2$ the 3D point in the coordinate system of the rotated camera.

• We can use the same trick as before, where we have $T = R^T$:

$$w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = KX_1 \quad \text{and} \quad w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = KX_2$$

$$w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = w_1 \underbrace{KR^TK^{-1}}_{3 \times 3 \text{ matrix}} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$
Rotating the Camera

• Rotating my camera with $R$ is the same as rotating the 3D points with $R^T$ (inverse of $R$):

$$X_2 = R^T X_1$$

• where $X_1$ is a 3D point in the coordinate system of the first camera and $X_2$ the 3D point in the coordinate system of the rotated camera.

• We can use the same trick as before, where we have $T = R^T$:

$$w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K X_1 \quad \text{and} \quad w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K X_2$$

$$w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = w_1 \underbrace{K R^T K^{-1}}_{3 \times 3 \text{ matrix}} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

what is this?
What If I Move The Camera?

• So if I take a picture, rotate the camera, and take a second picture...

• How are the first and second images related?
What If I Move The Camera?

- So if I take a picture, rotate the camera, and take a second picture…

- How are the first and second images related?

- by a Homography (assuming the scene didn’t change)
What If I Move The Camera?

• So if I take a picture, rotate the camera, and take a second picture…

• How are the first and second images related?

• by a Homography (assuming the scene didn’t change)

• What if I move the camera?
What If I Move The Camera?

- If I move the camera by \( t \), then: \( X_2 = X_1 - t \). Let’s try the same trick again:

\[
\begin{bmatrix}
  x_2 \\
  y_2 \\
  1
\end{bmatrix}
\begin{bmatrix}
w_2 \\
  0
\end{bmatrix}
= K X_2
\]
What If I Move The Camera?

• If I move the camera by t, then: $X_2 = X_1 - t$. Let’s try the same trick again:

\[
\begin{bmatrix}
  x_2 \\
  y_2 \\
  1
\end{bmatrix}
= K X_2 = K (X_1 - t)
\]
What If I Move The Camera?

- If I move the camera by $t$, then: $X_2 = X_1 - t$. Let’s try the same trick again:

\[
\begin{bmatrix}
    x_2 \\
    y_2 \\
    1
\end{bmatrix}
= K \begin{bmatrix}
    X_2 \\
    1
\end{bmatrix}
= K \left( X_1 - t \right)
\]

\[
\begin{bmatrix}
    x_1 \\
    y_1 \\
    1
\end{bmatrix}
\]
What If I Move The Camera?

• If I move the camera by \( t \), then: \( X_2 = X_1 - t \). Let’s try the same trick again:

\[
\begin{bmatrix}
  x_2 \\
  y_2 \\
  1
\end{bmatrix}
= K X_2 = K (X_1 - t) =
\begin{bmatrix}
  x_1 \\
  y_1 \\
  1
\end{bmatrix}
- K t
\]

• What’s the problem here?
What If I Move The Camera?

• If I move the camera by \( t \), then: \( X_2 = X_1 - t \). Let’s try the same trick again:

\[
\begin{bmatrix}
  x_2 \\
  y_2 \\
  1
\end{bmatrix}
= K X_2 = K (X_1 - t) = w_1
\begin{bmatrix}
  x_1 \\
  y_1 \\
  1
\end{bmatrix}
- K t
\]

• Now, different values of \( w_1 \) give different points in the second image!
What If I Move The Camera?

• If I move the camera by \( t \), then: \( X_2 = X_1 - t \). Let’s try the same trick again:

\[
\begin{bmatrix}
  x_2 \\
  y_2 \\
  1
\end{bmatrix} = K X_2 = K (X_1 - t) = \begin{bmatrix}
  x_1 \\
  y_1 \\
  1
\end{bmatrix} - K t
\]

• Now, different values of \( w_1 \) give different points in the second image!

• So, even if I have \( K \) and \( t \) I can’t compute where a point from the first image projects to in the second image.
What If I Move The Camera?

• If I move the camera by \( t \), then: \( X_2 = X_1 - t \). Let’s try the same trick again:

\[
\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = K X_2 = K (X_1 - t) = w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - K t
\]

• Now, different values of \( w_1 \) give different points in the second image!

• So, even if I have \( K \) and \( t \) I can’t compute where a point from the first image projects to in the second image.

• From

\[
\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = K X_1
\]

we know that different \( w_1 \) map to different points \( X_1 \) on the projective line

• Where \((x_1, y_1)\) maps to in the 2\(^{nd}\) image depends on the 3D location of \( X_1 \)
What If I Move The Camera?

- **Summary**: if I move the camera, I can’t easily map one image to the other. The mapping depends on the 3D scene behind the image.
What If I Move The Camera?

- **Summary**: if I move the camera, I can’t easily map one image to the other. The mapping depends on the 3D scene behind the image.

- What about the opposite, what if I know that points \((x_1, y_1)\) in the first image and \((x_2, y_2)\) in the second belong to the same 3D point?
What If I Move The Camera?

- **Summary**: if I move the camera, I can’t easily map one image to the other. The mapping depends on the 3D scene behind the image.

- What about the opposite, what if I know that points \((x_1, y_1)\) in the first image and \((x_2, y_2)\) in the second belong to the same 3D point?

\[
w_2 \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = w_1 \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} - Kt
\]

We know this

- We can compute \(w_1\) and \(w_2\)
- We can compute point in 3D!
What If I Move The Camera?

- **Summary**: if I move the camera, I can’t easily map one image to the other. The mapping depends on the 3D scene behind the image.

- What about the opposite, what if I know that points \((x_1, y_1)\) in the first image and \((x_2, y_2)\) in the second belong to the same 3D point?

- This allows triangulating 3D points, leads to **stereo** vision and **two-view** geometry
Summary – Stuff You Need To Know

Perspective Projection

- If point \( Q \) is in camera’s coordinate system:

\[
Q = (X, Y, Z)^T \quad \rightarrow \quad q = \left( f \frac{X}{Z} + p_x, f \frac{Y}{Z} + p_y \right)^T
\]
Summary – Stuff You Need To Know

Perspective Projection

- If point \( Q \) is in camera’s coordinate system:
  \[
  Q = (X, Y, Z)^T \rightarrow q = \left( \frac{f \cdot X}{Z} + p_x, \frac{f \cdot Y}{Z} + p_y \right)^T
  \]

- Same as: \( Q = (X, Y, Z)^T \rightarrow \begin{bmatrix} w \cdot x \\ w \cdot y \\ w \end{bmatrix} = K \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow q = \begin{bmatrix} x \\ y \end{bmatrix} \]
Summary – Stuff You Need To Know

Perspective Projection

• If point $Q$ is in camera’s coordinate system:

$$Q = (X, Y, Z)^T \rightarrow q = \left( \frac{f \cdot X}{Z} + p_x, \frac{f \cdot Y}{Z} + p_y \right)^T$$

• Same as: $Q = (X, Y, Z)^T \rightarrow \begin{bmatrix} w \cdot x \\ w \cdot y \\ w \end{bmatrix} = K \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \rightarrow q = \begin{bmatrix} x \\ y \end{bmatrix}$

where $K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix}$ is camera intrinsic matrix
Summary – Stuff You Need To Know

Perspective Projection

• If point \( Q \) is in camera’s coordinate system:

\[
Q = (X, Y, Z)^T \quad \rightarrow \quad q = \left(\frac{f \cdot X}{Z} + p_x, \frac{f \cdot Y}{Z} + p_y\right)^T
\]

• Same as: \( Q = (X, Y, Z)^T \quad \rightarrow \quad \begin{bmatrix} w \cdot x \\ w \cdot y \\ w \end{bmatrix} = K \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \quad \rightarrow \quad q = \begin{bmatrix} x \\ y \end{bmatrix}
\]

where \( K = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \) is camera intrinsic matrix

• If \( Q \) is in world coordinate system, then the full projection is characterized by a 3x4 matrix \( P \):

\[
\begin{bmatrix} w \cdot x \\ w \cdot y \\ w \end{bmatrix} = K[R | t] P \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}
\]
Summary – Stuff You Need To Know

Perspective Projection

• All parallel lines in 3D with the same direction meet in one, so-called vanishing point in the image
• All lines that lie on a plane have vanishing points that lie on a line, so-called vanishing lines.
• All parallel planes in 3D have the same vanishing line in the image
Summary – Stuff You Need To Know

Perspective Projection

• All parallel lines in 3D with the same direction meet in one, so-called vanishing point in the image
• All lines that lie on a plane have vanishing points that lie on a line, so-called vanishing lines.
• All parallel planes in 3D have the same vanishing line in the image

Orthographic Projection

• Projections simply drops the Z coordinate:

\[
Q = \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}
\]

• Parallel lines in 3D are parallel in the image
Two-view Geometry
Triangulation

Given camera 1 with matrix \( P \) and camera 2 with matrix \( P' \).
Triangulation

Which 3D points map to $x$?

image 1

image 2

camera 1 with matrix $P$
camera 2 with matrix $P'$
Triangulation

How can you compute this ray?

camera 1 with matrix $P$

camera 2 with matrix $P'$
Triangulation

Create two points on the ray:
1) find the camera center; and
2) apply the pseudo-inverse of $P$ on $x$.

Then connect the two points.
This procedure is called backprojection.
How do we find the exact point on the ray?

\[ \mathbf{P}'x' \]

image 1

image 2

camera 1 with matrix \( \mathbf{P} \)

camera 2 with matrix \( \mathbf{P}' \)
Triangulation

Find 3D object point

Will the lines intersect?

camera 1 with matrix \( P \)

image 1

image 2

camera 2 with matrix \( P' \)

\( x \)

\( x' \)
Triangulation

Find 3D object point (no single solution due to noise)

image 1

image 2

camera 1 with matrix $P$
camera 2 with matrix $P'$
Triangulation

Given a set of (noisy) matched points

\[ \{ x_i, x'_i \} \]

and camera matrices

\[ P, P' \]

Estimate the 3D point

\[ X \]
Can we compute $\mathbf{X}$ from a single correspondence $\mathbf{x}$?
This is a similarity relation because it involves homogeneous coordinates

$$x = P X$$

(homogeneous coordinate)

Same ray direction but differs by a scale factor

$$x = \alpha PX$$

(homogeneous coordinate)

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \alpha \begin{bmatrix} p_1 & p_2 & p_3 & p_4 \\ p_5 & p_6 & p_7 & p_8 \\ p_9 & p_{10} & p_{11} & p_{12} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$
Linear algebra reminder: cross product

Vector (cross) product takes two vectors and returns a vector \textit{perpendicular} to both

\[ c = a \times b \]

\[ a \times b = \begin{bmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{bmatrix} \]

cross product of two vectors in the same direction is zero vector
\[ a \times a = 0 \]

remember this!!!

\[ c \cdot a = 0 \quad \text{and} \quad c \cdot b = 0 \]
Linear algebra reminder: cross product

Cross product

\[ \mathbf{a} \times \mathbf{b} = \begin{bmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{bmatrix} \]

Can also be written as a matrix multiplication

\[ \mathbf{a} \times \mathbf{b} = [\mathbf{a}] \times \mathbf{b} = \begin{bmatrix} 0 & -a_3 & a_2 \\ a_3 & 0 & -a_1 \\ -a_2 & a_1 & 0 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} \]

Skew symmetric
Compare with: dot product

\[ c = a \times b \]

\[ c \cdot a = 0 \quad c \cdot b = 0 \]

dot product of two orthogonal vectors is (scalar) zero
Back to triangulation

\[ x = \alpha PX \]

Same direction but differs by a scale factor

*How can we rewrite this using vector products?*
\[ x = \alpha PX \]

Same direction but differs by a scale factor

\[ x \times PX = 0 \]

Cross product of two vectors of same direction is zero
(this equality removes the scale factor)
\[
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix}
= \alpha
\begin{bmatrix}
p_1 & p_2 & p_3 & p_4 \\
p_5 & p_6 & p_7 & p_8 \\
p_9 & p_{10} & p_{11} & p_{12}
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]
\[
\begin{bmatrix}
    x \\
    y \\
    z
\end{bmatrix} = \alpha
\begin{bmatrix}
    p_1 & p_2 & p_3 & p_4 \\
    p_5 & p_6 & p_7 & p_8 \\
    p_9 & p_{10} & p_{11} & p_{12}
\end{bmatrix}
\begin{bmatrix}
    X \\
    Y \\
    Z \\
    1
\end{bmatrix}
\]

\[
\begin{bmatrix}
    x \\
    y \\
    z
\end{bmatrix} = \alpha
\begin{bmatrix}
    \vdots & p_1 \vdots \\
    \vdots & p_2 \vdots \\
    \vdots & p_3 \vdots
\end{bmatrix}
\begin{bmatrix}
    X
\end{bmatrix}
\]
\[
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix} = \alpha \begin{bmatrix}
p_1 & p_2 & p_3 & p_4 \\
p_5 & p_6 & p_7 & p_8 \\
p_9 & p_{10} & p_{11} & p_{12} \\
\end{bmatrix} \begin{bmatrix}
X \\
Y \\
Z \\
1 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix} = \alpha \begin{bmatrix}
\ldots & \overset{1}{p_1} & \ldots \\
\ldots & \overset{2}{p_2} & \ldots \\
\ldots & \overset{3}{p_3} & \ldots \\
\end{bmatrix} \begin{bmatrix}
X \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
x \\
y \\
z \\
\end{bmatrix} = \alpha \begin{bmatrix}
\overset{1}{p_1}X \\
\overset{2}{p_2}X \\
\overset{3}{p_3}X \\
\end{bmatrix}
\]
\[
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = \alpha \begin{bmatrix}
p_1 & p_2 & p_3 & p_4 \\
p_5 & p_6 & p_7 & p_8 \\
p_9 & p_{10} & p_{11} & p_{12}
\end{bmatrix} \begin{bmatrix}
X \\
Y \\
Z \\
1
\end{bmatrix}
\]

\[
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = \alpha \begin{bmatrix}
\overrightarrow{p_1} & \overrightarrow{p_2} & \overrightarrow{p_3}
\end{bmatrix} \begin{bmatrix}
X
\end{bmatrix}
\]

\[
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} = \alpha \begin{bmatrix}
p_{1}^\top X \\
p_{2}^\top X \\
p_{3}^\top X
\end{bmatrix}
\]

\[
\begin{bmatrix}
x \\
y \\
z
\end{bmatrix} \times \begin{bmatrix}
p_{1}^\top X \\
p_{2}^\top X \\
p_{3}^\top X
\end{bmatrix} = \begin{bmatrix}
yp_{3}^\top X - p_{2}^\top X \\
p_{1}^\top X - xp_{3}^\top X \\
 xp_{2}^\top X - yp_{1}^\top X
\end{bmatrix} = \begin{bmatrix}
0 \\
0 \\
0
\end{bmatrix}
\]
Using the fact that the cross product should be zero

\[
\mathbf{x} \times \mathbf{PX} = 0
\]

\[
\begin{bmatrix}
y p_3^\top \mathbf{X} - p_2^\top \mathbf{X} \\
p_1^\top \mathbf{X} - x p_3^\top \mathbf{X} \\
x p_2^\top \mathbf{X} - y p_1^\top \mathbf{X}
\end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\]

Third line is a linear combination of the first and second lines.
(x times the first line plus y times the second line)

One 2D to 3D point correspondence give you 3 equations
Using the fact that the cross product should be zero

\[ \mathbf{x} \times \mathbf{PX} = 0 \]

\[
\begin{bmatrix}
yp_3^T \mathbf{X} - p_2^T \mathbf{X} \\
p_1^T \mathbf{X} - xp_3^T \mathbf{X} \\
xp_2^T \mathbf{X} - yp_1^T \mathbf{X}
\end{bmatrix}
= \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\]

Third line is a linear combination of the first and second lines. (x times the first line plus y times the second line)

One 2D to 3D point correspondence give you 2 equations
Now we can make a system of linear equations
(two lines for each 2D point correspondence)

\[
\begin{bmatrix}
yp_3^T X - p_2^T X \\
p_1^T X - xp_3^T X
\end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

Remove third row, and rearrange as system of unknowns

\[
\begin{bmatrix}
yp_3^T \\
p_1^T - xp_3^T
\end{bmatrix} X = \begin{bmatrix} 0 \\ 0 \end{bmatrix}
\]

\[
A_i X = 0
\]
How do we solve homogeneous linear system?

Concatenate the 2D points from both images

\[
\begin{bmatrix}
yp_3^\top - p_2^\top \\
q_1^\top - xp_3^\top \\
y'p_3'^\top - p_2'^\top \\
p_1'^\top - x'p_3'^\top
\end{bmatrix} \begin{bmatrix}
X
\end{bmatrix} = \begin{bmatrix}
\begin{pmatrix}0 \\
0 \\
0 \\
0
\end{pmatrix}
\end{bmatrix}
\]

\text{sanity check! dimensions?}

\[
AX = 0
\]
How do we solve homogeneous linear system?

Concatenate the 2D points from both images

\[
\begin{bmatrix}
yp_3^\top - p_2^\top \\
p_1^\top - xp_3^\top \\
y'p_3^\top - p_2^\top \\
p_1'^\top - x'p_3'^\top
\end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}
\]

\[\mathbf{A}\mathbf{X} = \mathbf{0}\]

*How do we solve homogeneous linear system?*

\[\mathbf{S} \quad \mathbf{V} \quad \mathbf{D} \quad \mathbf{!}\]
How do we solve homogeneous linear system?

\[ \begin{bmatrix} y p_3^T - p_2^T \\ p_1^T - x p_3^T \\ y' p_3' - p_2' \\ p_1' - x' p_3' \end{bmatrix} X = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \]

\[ AX = 0 \]

How do we solve homogeneous linear system?

S V D !

This is triangulation!
Triangulation recap

Given a set of (noisy) matched points
\[ \{x_i, x'_i\} \]
and camera matrices
\[ P, P' \]

Estimate the 3D point
\[ X \]

- use relationship  \( x \times PX = 0 \)
Triangulation recap

Given a set of (noisy) matched points \( \{x_i, x'_i\} \) and camera matrices \( P, P' \)

Estimate the 3D point \( X \)

- use relationship \( x \times PX = 0 \)
- formulate system of equations (2 for each correspondence)
Triangulation recap

Given a set of (noisy) matched points $\{x_i, x'_i\}$ and camera matrices $P, P'$

Estimate the 3D point $X$

- use relationship $x \times PX = 0$
- formulate system of equations (2 for each correspondence)
- Solve with SVD
Epipolar geometry
Epipolar geometry

Image plane
Epipolar geometry
Epipolar geometry
Epipolar geometry

- Epipole: projection of $o'$ on the image plane
- Baseline
- Epipolar plane
- Image plane

Epipole (projection of $o'$ on the image plane)
Epipolar geometry

- **Epipole** (projection of $o'$ on the image plane)
- **Epipolar line** (intersection of Epipolar plane and image plane)
- **Epipolar plane**
- **Baseline**
- **Image plane** (projection of $o'$ on the image plane)
What is this?
What is this?

Epipolar plane
Quiz

Epipolar plane
(Intersection of Epipolar plane and image plane)

Epipolar line

Epipolar plane

o o'

l e l'
e e'
What is this?
Quiz

Epipolar line
(intersection of Epipolar plane and image plane)

Epipole
(projection of o' on the image plane)

Epipolar plane

\(p\)

\(o\) \(o'\)

\(l\) \(l'\)

\(e\)
What is this?

Epipolar plane

Epipolar line
(intersection of Epipolar plane and image plane)

Epipole
(projection of o’ on the image plane)
Quiz

Epipolar plane

Epipolar line
(intersection of Epipolar plane and image plane)

Epipole
(projection of o' on the image plane)

Baseline

p

o

l

l'

e

e'

o'
Another way to construct the epipolar plane, this time given $\mathbf{x}$,
Epipolar Constraint

Potential matches for $x$ lie on the epipolar line $l'$.
The point $x$ (left image) maps to a ___________ in the right image.

The baseline connects the ___________ and ____________.

An epipolar line (left image) maps to a ___________ in the right image.

An epipole $e$ is a projection of the ____________ on the image plane.

All epipolar lines in an image intersect at the ____________.
Converging cameras

Where is the epipole in this image?
Converging cameras

Where is the epipole in this image? It's not always in the image.
Parallel cameras

Where is the epipole?
Parallel cameras

epipole at infinity
The epipolar constraint is an important concept for stereo vision.

**Task:** Match point in left image to point in right image

*How would you do it?*
Potential matches for $x$ lie on the epipolar line $l'$.
The epipolar constraint is an important concept for stereo vision.

**Task:** Match point in left image to point in right image

Want to avoid search over entire image

Epipolar constraint reduces search to a single line
The epipolar constraint is an important concept for stereo vision.

**Task:** Match point in left image to point in right image

![Left image](image1.png) ![Right image](image2.png)

Want to avoid search over entire image.
Epipolar constraint reduces search to a single line.

*How do you compute the epipolar line?*
The essential matrix
Recall: Epipolar Constraint

Potential matches for $x$ lie on the epipolar line $l'$. 

$\text{Potential matches for } x \text{ lie on the epipolar line } l'$
Given a point in one image, multiplying by the **essential matrix** will tell us the **epipolar line** in the second view.

\[ E x = l' \]
The Essential Matrix is a 3 x 3 matrix that encodes **epipolar geometry**

Given a point in one image, multiplying by the **essential matrix** will tell us the **epipolar line** in the second image.
Representing the epipolar line

\[ ax + by + c = 0 \]

in vector form \( \mathbf{l} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \)

If the point \( \mathbf{x} \) is on the epipolar line \( \mathbf{l} \) then

\[ \mathbf{x}^\top \mathbf{l} = ? \]
Representing the epipolar line

\[ ax + by + c = 0 \]  \[ l = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \]

If the point \( x \) is on the epipolar line \( l \) then

\[ x^\top l = 0 \]
So if $x'^\top l' = 0$ and $Ex = l'$ then

$$x'^\top Ex = ?$$
So if $x' \top l' = 0$ and $Ex = l'$ then

$$x' \top Ex = 0$$
What’s the difference between the essential matrix and a homography?

Essential Matrix vs Homography
What’s the difference between the essential matrix and a homography?

They are both 3 x 3 matrices but …

**Essential Matrix vs Homography**

\[ l' = Ex \quad x' = Hx \]

Essential matrix maps a **point** to a **line**

Homography maps a **point** to a **point**
Where does the essential matrix come from?
\[ x' = R(x - t) \]
\[ x' = R(x - t) \]

*Does this look familiar?*
$x' = R(x - t)$

**Camera-camera** transform just like **world-camera** transform
These three vectors are coplanar
\( x, t, x' \)
If these three vectors are coplanar $\mathbf{x}, t, \mathbf{x}'$ then

$$\mathbf{x}^\top (t \times \mathbf{x}) = ?$$
If these three vectors are coplanar $\mathbf{x}, t, \mathbf{x}'$ then

$$\mathbf{x}^\top (t \times \mathbf{x}) = 0$$

dot product of orthogonal vectors

cross-product: vector orthogonal to plane
If these three vectors are coplanar $\mathbf{x}, t, \mathbf{x}'$ then

$$(\mathbf{x} - t)^\top (t \times \mathbf{x}) = ?$$
If these three vectors are coplanar $\mathbf{x}, t, \mathbf{x}'$ then

$$(\mathbf{x} - t)^\top (t \times \mathbf{x}) = 0$$
Putting it together

rigid motion

\[ x' = R(x - t) \]

coplanarity

\[ (x - t)^\top (t \times x) = 0 \]

\[ (x'\top R)(t \times x) = 0 \]
Putting it together

rigid motion

\[ x' = R(x - t) \]

cooplanarity

\[ (x - t)^\top (t \times x) = 0 \]

use skew-symmetric matrix to represent cross product

\[ (x'^\top R)(t \times x) = 0 \]

\[ (x'^\top R)([t \times] x) = 0 \]
Putting it together

rigid motion

\[ x' = R(x - t) \]

coplanarity

\[ (x - t)^T (t \times x) = 0 \]

\[ (x'^T R)(t \times x) = 0 \]

\[ (x'^T R)([t \times] x) = 0 \]

\[ x'^T (R[t \times]) x = 0 \]
Putting it together

\[ x' = R(x - t) \quad (x - t)^\top (t \times x) = 0 \]

\[
(x'\top R)(t \times x) = 0 \\
(x'\top R)([t \times] x) = 0 \\
x'\top (R[t \times]) x = 0 \\
x'\top E x = 0
\]
Putting it together

rigid motion

\[ x' = R(x - t) \]

coplanarity

\[ (x - t)^\top (t \times x) = 0 \]

\[ (x'{}^\top \mathbf{R})(t \times x) = 0 \]

\[ (x'{}^\top \mathbf{R})([t \times] \mathbf{x}) = 0 \]

\[ x'{}^\top (\mathbf{R}[t \times]) \mathbf{x} = 0 \]

\[ x'{}^\top \mathbf{E} \mathbf{x} = 0 \]

Essential Matrix

[Longuet-Higgins 1981]
properties of the E matrix

Longuet-Higgins equation

$$x'^\top E x = 0$$

(2D points expressed in camera coordinate system)
properties of the E matrix

Longuet-Higgins equation

\[ x'^\top E x = 0 \]

Epipolar lines

\[ x^\top l = 0 \quad \quad x'^\top l' = 0 \]
\[ l' = Ex \quad \quad l = E^T x' \]

(2D points expressed in camera coordinate system)
properties of the E matrix

**Longuet-Higgins equation**
\[ x'\mathbf{E}x = 0 \]

**Epipolar lines**
\[ x^\top l = 0 \quad \quad x'^\top l' = 0 \]
\[ l' = \mathbf{E}x \quad \quad l = \mathbf{E}^\top x' \]

**Epipoles**
\[ e'^\top \mathbf{E} = 0 \quad \quad \mathbf{E}e = 0 \]

(2D points expressed in camera coordinate system)
Given a point in one image, multiplying by the **essential matrix** will tell us the **epipolar line** in the second view.

\[ \mathbf{Ex} = l' \]

**Assumption:**

2D points expressed in camera coordinate system (i.e., intrinsic matrices are identities)
How do you generalize to non-identity intrinsic matrices?
The fundamental matrix
The fundamental matrix is a generalization of the essential matrix, where the assumption of identity matrices is removed.
\hat{x}'^\top E \hat{x} = 0

The essential matrix operates on image points expressed in 2D coordinates expressed in the camera coordinate system.

\[
\hat{x}' = K'^{-1} \hat{x}' \\
\hat{x} = K^{-1} x
\]

- \hat{x}' = \text{camera point}'
- \hat{x} = \text{image point}
\[ \hat{x}' \top E \hat{x} = 0 \]

The essential matrix operates on image points expressed in 2D coordinates expressed in the camera coordinate system.

\[ \hat{x}' = K'^{-1} x' \quad \hat{x} = K^{-1} x \]

Writing out the epipolar constraint in terms of image coordinates:

\[ K'^{-\top} E K^{-1} x = 0 \]
\[ x'^\top (K'^{-\top} E K^{-1}) x = 0 \]
\[ x'^\top F x = 0 \]
Same equation works in image coordinates!

\[ x'^\top F x = 0 \]

it maps pixels to epipolar lines
properties of the $E$ matrix

**Longuet-Higgins equation**

$$x'^\top E x = 0$$

**Epipolar lines**

$$x^\top l = 0 \quad x'^\top l' = 0$$

$$l' = E x \quad l = E^\top x'$$

**Epipoles**

$$e'^\top E = 0 \quad E e = 0$$

(points in image coordinates)
Breaking down the fundamental matrix

\[
F = K'^{-T} E K^{-1}
\]

\[
F = K'^{-T} [t_x] R K^{-1}
\]

Depends on both intrinsic and extrinsic parameters
Breaking down the fundamental matrix

\[ F = K'^{-T} E K^{-1} \]

\[ F = K'^{-T} [t_x] R K^{-1} \]

Depends on both intrinsic and extrinsic parameters

*How would you solve for F?*

\[ x'_m F x_m = 0 \]
The 8-point algorithm
Assume you have $M$ matched *image* points

$$\{ x_m, x'_m \} \quad m = 1, \ldots, M$$

Each correspondence should satisfy

$$x'_m^T F x_m = 0$$

*How would you solve for the 3 x 3 $F$ matrix?*
Assume you have $M$ matched image points

$$\{ \mathbf{x}_m, \mathbf{x'}_m \} \quad m = 1, \ldots, M$$

Each correspondence should satisfy

$$\mathbf{x'}_m \mathbf{F} \mathbf{x}_m = 0$$

How would you solve for the 3 x 3 $\mathbf{F}$ matrix?

$$ \mathbf{S} \quad \mathbf{V} \quad \mathbf{D}$$
Assume you have $M$ matched *image* points

\[
\{ \mathbf{x}_m, \mathbf{x}'_m \} \quad m = 1, \ldots, M
\]

Each correspondence should satisfy

\[
\mathbf{x}'_m^T \mathbf{F} \mathbf{x}_m = 0
\]

*How would you solve for the 3 x 3 $\mathbf{F}$ matrix?*

Set up a homogeneous linear system with 9 unknowns
\[ x'_m \mathbf{F} x_m = 0 \]

\[
\begin{bmatrix}
  x'_m & y'_m & 1
\end{bmatrix}
\begin{bmatrix}
f_1 & f_2 & f_3 \\
f_4 & f_5 & f_6 \\
f_7 & f_8 & f_9
\end{bmatrix}
\begin{bmatrix}
x_m \\
y_m \\
1
\end{bmatrix} = 0
\]

How many equation do you get from one correspondence?
\[
\begin{bmatrix}
  x'_m & y'_m & 1
\end{bmatrix}
\begin{bmatrix}
  f_1 & f_2 & f_3 \\
  f_4 & f_5 & f_6 \\
  f_7 & f_8 & f_9
\end{bmatrix}
\begin{bmatrix}
  x_m \\
  y_m \\
  1
\end{bmatrix} = 0
\]

ONE correspondence gives you ONE equation

\[
x_m x'_m f_1 + x_m y'_m f_2 + x_m f_3 + y_m x'_m f_4 + y_m y'_m f_5 + y_m f_6 + x'_m f_7 + y'_m f_8 + f_9 = 0
\]
Set up a homogeneous linear system with 9 unknowns

\[
\begin{bmatrix}
  f_1 & f_2 & f_3 \\
  f_4 & f_5 & f_6 \\
  f_7 & f_8 & f_9 \\
\end{bmatrix}
\begin{bmatrix}
  x_m \\
  y_m \\
  1 \\
\end{bmatrix} = 0
\]

Set up a homogeneous linear system with 9 unknowns

\[
\begin{bmatrix}
  x_1 & x_1' & x_1 & y_1 & y_1' & x_1' & y_1' & 1 \\
  \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
  x_M & x_M' & x_M & y_M & y_M' & x_M' & y_M' & 1 \\
\end{bmatrix}
\begin{bmatrix}
  f_1 \\
  f_2 \\
  f_3 \\
  f_4 \\
  f_5 \\
  f_6 \\
  f_7 \\
  f_8 \\
  f_9 \\
\end{bmatrix} = 0
\]

How many equations do you need?
Each point pair (according to epipolar constraint) contributes only one scalar equation

$$x'_m \mathbf{F} x_m = 0$$

**Note:** This is different from the Homography estimation where each point pair contributes 2 equations.

We need at least 8 points

Hence, the 8 point algorithm!
How do you solve a homogeneous linear system?

\[ AX = 0 \]
How do you solve a homogeneous linear system?

\[ \mathbf{AX} = 0 \]

Total Least Squares

minimize \( \| \mathbf{Ax} \|^2 \)

subject to \( \| \mathbf{x} \|^2 = 1 \)
How do you solve a homogeneous linear system?

\[ AX = 0 \]

Total Least Squares

minimize \[ \|Ax\|^2 \]

subject to \[ \|x\|^2 = 1 \]
Eight-Point Algorithm

0. (Normalize points)
1. Construct the M x 9 matrix $A$
2. Find the SVD of $A$
3. Entries of $F$ are the elements of column of $V$ corresponding to the least singular value
4. (Enforce rank 2 constraint on $F$)
5. (Un-normalize $F$)
Eight-Point Algorithm

0. (Normalize points)
1. Construct the M x 9 matrix \( \mathbf{A} \)
2. Find the SVD of \( \mathbf{A} \)
3. Entries of \( \mathbf{F} \) are the elements of column of \( \mathbf{V} \) corresponding to the least singular value
4. (Enforce rank 2 constraint on \( \mathbf{F} \))
5. (Un-normalize \( \mathbf{F} \))

See Hartley-Zisserman for why we do this
Eight-Point Algorithm

0. (Normalize points)
1. Construct the M x 9 matrix $\mathbf{A}$
2. Find the SVD of $\mathbf{A}$
3. Entries of $\mathbf{F}$ are the elements of column of $\mathbf{V}$ corresponding to the least singular value
4. (Enforce rank 2 constraint on $\mathbf{F}$)
5. (Un-normalize $\mathbf{F}$)

How do we do this?
Eight-Point Algorithm

0. (Normalize points)
1. Construct the M x 9 matrix $A$
2. Find the SVD of $A$
3. Entries of $F$ are the elements of column of $V$ corresponding to the least singular value
4. (Enforce rank 2 constraint on $F$)
5. (Un-normalize $F$)

How do we do this?

S V D !
Enforcing rank constraints

Problem: Given a matrix $F$, find the matrix $F'$ of rank $k$ that is closest to $F$,

$$\min_{F'} \| F - F' \|^2$$

$$\text{rank}(F') = k$$

Solution: Compute the singular value decomposition of $F$,

$$F = U\Sigma V^T$$

Form a matrix $\Sigma'$ by replacing all but the $k$ largest singular values in $\Sigma$ with 0.

Then the problem solution is the matrix $F'$ formed as,

$$F' = U\Sigma' V^T$$
Eight-Point Algorithm

0. (Normalize points)
1. Construct the M x 9 matrix $\mathbf{A}$
2. Find the SVD of $\mathbf{A}$
3. Entries of $\mathbf{F}$ are the elements of column of $\mathbf{V}$ corresponding to the least singular value
4. (Enforce rank 2 constraint on $\mathbf{F}$)
5. (Un-normalize $\mathbf{F}$)
Example
epipolar lines
\[ F = \begin{bmatrix}
-0.00310695 & -0.0025646 & 2.96584 \\
-0.028094 & -0.00771621 & 56.3813 \\
13.1905 & -29.2007 & -9999.79
\end{bmatrix} \]

\[ x = \begin{bmatrix}
343.53 \\
221.70 \\
1.0
\end{bmatrix} \]

\[ l' = Fx = \begin{bmatrix}
0.0295 \\
0.9996 \\
-265.1531
\end{bmatrix} \]
\[ l' = Fx \]

\[
\begin{bmatrix}
0.0295 \\
0.9996 \\
-265.1531
\end{bmatrix}
\]
Where is the epipole?

How would you compute it?
The epipole is in the right null space of $\mathbf{F}$

$\mathbf{F} \mathbf{e} = 0$

*How would you solve for the epipole?*
The epipole is in the right null space of $F$

$Fe = 0$

How would you solve for the epipole?

S V D!
Next Time:
Stereo depth estimation