Website


Spring 2017 CSCI 5890/8980
Multiview 3D Geometry in Computer Vision

Tue/Thr 4:00pm-5:15pm @ Rapson Hall 43

Description

(a) Head-mounted cameras
(b) 3D reconstruction of players
Paper Presentation

- Format
  - 20 min presentation with 15+ min questions
  - One presenter: this person defends the paper.
  - One committee: this person attacks the paper.

- You are free to choose a paper that
  - Is not written by you or your research group
  - Has a strong relevance to 3D vision (theory and application)
  - Is approved by me

- Or, you can choose a paper from my list.
Paper List

- Tomasi and Kanade, Shape and Motion from Image Streams under Orthography: a Factoriaztion Method, IJCV, 1992
Paper List

• Reid and Zisserman, Goal-directed Video Metrology, ECCV, 1996
Paper List

• Zhang, A Flexible New Technique for Camera Calibration, PAMI, 2000
Paper List

- Nister, An Efficient Solution to the Five-Point Relative Pose Problem, PAMI, 2004
Paper List

- Criminisi, Reid, and Zisserman, Single View Metrology, IJCV, 2000
Paper List

Paper List

- Izadi et al., KinectFusion, UIST, 2011
VERTIGO (1958)
3D Point Projection (Metric Space)

- $f_m$: Focal length in meter

3D point $(X, Y, Z)$
3D Point Projection (Metric Space)

\[ \tan \theta = \frac{X}{Z} \]

- \( f_m \): Focal length in meter

3D point \((X, Y, Z)\)
3D Point Projection (Metric Space)

\[ \tan \theta = \frac{X}{Z} \]

\[ u_{ccd} = -f_m \tan \theta = -f_m \frac{X}{Z} \]

\( f_m \): Focal length in meter

3D point \((X, Y, Z)\)
3D Point Projection (Metric Space)

\[ \tan \theta = \frac{X}{Z} \]

\[ u_{\text{ccd}} = -f_m \frac{X}{Z} \]

\[ v_{\text{ccd}} = -f_m \frac{Y}{Z} \]

\( f_m \): Focal length in meter

3D point \((X, Y, Z)\)
3D Point Projection (Metric Space)

\[ \tan \theta = \frac{X}{Z} \]

- \( f_m \) : Focal length in meter
- \( (u_{ccd}, v_{ccd}) \)

\[ u_{ccd} = -f_m \frac{X}{Z} \]
\[ v_{ccd} = -f_m \frac{Y}{Z} \]

\( (X, Y, Z) \rightarrow (u_{ccd}, v_{ccd}) = (-f_m \frac{X}{Z}, -f_m \frac{Y}{Z}) \)

2D projection onto CCD plane
3D Point Projection (Metric Space)

3D point \((X,Y,Z)\)

2D projection onto CCD plane

\[ u_{\text{ccd}} = f_m \frac{X}{Z} \]

\[ v_{\text{ccd}} = f_m \frac{Y}{Z} \]

\( (X,Y,Z) \rightarrow (u_{\text{ccd}}, v_{\text{ccd}}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z}) \)

- Pinhole
- Projection plane
- \( f_m \): Focal length in meter
3D Point Projection (Metric Space)

Projection plane

$(u_{ccd}, v_{ccd})$

Pinhole

$-f_m$

$u_{ccd} = f_m \frac{X}{Z}$

$X, Y, Z \rightarrow (u_{ccd}, v_{ccd}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z})$

2D projection onto CCD plane

$f_m :$ Focal length in meter

3D point $(X, Y, Z)$
3D Point Projection (Metric Space)

- Projection plane
- 3D point \((X, Y, Z)\)
- \(f_m\) : Focal length in meter

\[
u_{ccd} = f_m \frac{X}{Z} = f_m \frac{Y}{Z}
\]

\((X, Y, Z) \rightarrow (u_{ccd}, v_{ccd}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z})\)

2D projection onto CCD plane
3D Point Projection (Metric Space)

\[(u_{\text{img}}, v_{\text{img}}) \sim (u_{\text{ccd}}, v_{\text{ccd}}) \frac{?}{\text{Pixel}} \sim \frac{?}{\text{Metric}}\]

\[f_m : \text{Focal length in meter}\]

\[u_{\text{ccd}} = f_m \frac{X}{Z}\]

\[v_{\text{ccd}} = f_m \frac{Y}{Z}\]

\[(X, Y, Z) \rightarrow (u_{\text{ccd}}, v_{\text{ccd}}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z})\]

2D projection onto CCD plane.
3D Point Projection (Pixel Space)
3D Point Projection (Pixel Space)

\[(u_{\text{ccd}}, v_{\text{ccd}})\]

CCD sensor (mm)

\[(u_{\text{img}}, v_{\text{img}})\]

Image (pixel)
3D Point Projection (Pixel Space)

\[(u_{ccd}, v_{ccd})\] 

CCD sensor (mm)

\[h_{ccd}\]

\[w_{ccd}\]

Image (pixel)

\[(u_{img}, v_{img})\]
<table>
<thead>
<tr>
<th>Imager Sizes</th>
<th>Formats (Type)</th>
<th>~Diag.</th>
<th>Uses</th>
</tr>
</thead>
<tbody>
<tr>
<td>□ 1/7&quot; - 1.85 x 1.39mm</td>
<td>2.3</td>
<td>Cell phones, web cams, etc....</td>
<td></td>
</tr>
<tr>
<td>□ 1/6&quot; - 2.15 x 1.61mm</td>
<td>2.7</td>
<td>Cell phones, web cams, etc....</td>
<td></td>
</tr>
<tr>
<td>□ 1/5&quot; - 2.55 x 1.91mm</td>
<td>3.2</td>
<td>Cell phones, web cams, etc....</td>
<td></td>
</tr>
<tr>
<td>□ 1/4&quot; - 3.2 x 2.4mm</td>
<td>4.0</td>
<td>Cell phones, web cams, etc....</td>
<td></td>
</tr>
<tr>
<td>□ 1/3.6&quot; - 4.0 x 3.0mm</td>
<td>5.0</td>
<td>P&amp;S DSC</td>
<td></td>
</tr>
<tr>
<td>□ 1/3.2&quot; - 4.536 x 3.416mm</td>
<td>5.678</td>
<td>P&amp;S DSC</td>
<td></td>
</tr>
<tr>
<td>□ 1/3&quot; - 4.8 x 3.6mm</td>
<td>6.0</td>
<td>Casio QV-8000SX (1.2MP), Epson PhotoPC 700 (1.2MP)</td>
<td></td>
</tr>
<tr>
<td>□ 1/2.7&quot; - 5.27 x 3.96mm</td>
<td>6.592</td>
<td>Canon PowerShot A20 (1.92MP), HP PhotoSmart C818 (1.92)</td>
<td></td>
</tr>
<tr>
<td>□ 1/2&quot; - 6.4 x 4.8mm</td>
<td>8.0</td>
<td>Olympus C-2100Z (1.92MP), Epson PhotoPC 850Z (1.92)</td>
<td></td>
</tr>
<tr>
<td>□ 1/1.8&quot; - 7.176 x 5.319mm</td>
<td>8.932</td>
<td>Nikon Coolpix 995 (3.14MP), Olympus C-4040Z (3.9MP), Canon PowerShot G2 (3.8MP), Sony DSC-S85 (3.8MP)</td>
<td></td>
</tr>
<tr>
<td>□ 2/3&quot; - 8.8 x 6.6mm</td>
<td>11.0</td>
<td>Nikon Coolpix 5000 (4.92MP), Sony DSC-F707 (4.92MP), Olympus E-10 (3.7MP), Minolta DIMAGE 7 (4.92MP)</td>
<td></td>
</tr>
<tr>
<td>□ 1&quot; - 12.8 x 9.6mm</td>
<td>16.0</td>
<td>Not used in DSCs. Used in some high-end video cameras</td>
<td></td>
</tr>
</tbody>
</table>

Kodak KAF-5100CE CCD
17.8 x 13.4mm (4/3"")
22.28
Olympus announced development of a new camera and new lenses for this 4/3" size.
2614 x 1986 - 5.1MP - 6.8μm pixel

Foveon X3 F7-35X3-A25B
20.7 x 13.8mm
24.9
Sigma SD9 (X3)
2268 x 1512 = 3.43MP - 9.12μm pixel
1.74x Focal Length Multiplier (35mm film)
<table>
<thead>
<tr>
<th>Camera Model</th>
<th>Sensor Size</th>
<th>Resolution</th>
<th>Pixel Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Canon D30 CMOS</td>
<td>21.8 x 14.5 mm</td>
<td>2160 x 1440</td>
<td>3.11 MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.65x Focal Length Multiplier (35mm film)</td>
<td></td>
</tr>
<tr>
<td>Canon D60 CMOS</td>
<td>22.7 x 15.1 mm</td>
<td>3072 x 2048</td>
<td>6.3 MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.50x Focal Length Multiplier (35mm film)</td>
<td></td>
</tr>
<tr>
<td>Nikon D100 CCD</td>
<td>23.7 x 16.6 mm</td>
<td>3008 x 2000</td>
<td>6.1 MP</td>
</tr>
<tr>
<td>Nikon D1x CCD</td>
<td></td>
<td>4024 x 1324</td>
<td>5.24 MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.52x Focal Length Multiplier (35mm film)</td>
<td></td>
</tr>
<tr>
<td>APS Film</td>
<td>25.1 x 16.7 mm</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.44x Focal Length Multiplier (35mm film)</td>
<td></td>
</tr>
<tr>
<td>Canon EOS-1D CCD</td>
<td>27.0 x 17.8 mm</td>
<td>2464 x 1648</td>
<td>4.06 MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.34x Focal Length Multiplier (35mm Film)</td>
<td></td>
</tr>
<tr>
<td>Kodak KAF-6303CE CCD</td>
<td>27.8 x 18.5 mm</td>
<td>3088 x 2056</td>
<td>6.35 MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.30x Focal Length Multiplier (35mm film)</td>
<td></td>
</tr>
<tr>
<td>35mm Film</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Canon 1Ds</td>
<td>36.0 x 24.0 mm</td>
<td>4064 x 2704</td>
<td>10.99 MP</td>
</tr>
<tr>
<td>Kodak 14n</td>
<td></td>
<td>4536 x 3024</td>
<td>13.7 MP</td>
</tr>
<tr>
<td></td>
<td></td>
<td>8.85x Focal Length Multiplier (35mm film)</td>
<td></td>
</tr>
</tbody>
</table>
3D Point Projection (Pixel Space)

\[(u_{\text{ccd}}, v_{\text{ccd}})\]

CCD sensor (mm)

\[(u_{\text{img}}, v_{\text{img}})\]

Image (pixel)
3D Point Projection (Pixel Space)
3D Point Projection (Pixel Space)
3D Point Projection (Pixel Space)

Projection of pinhole

(0,0)

(w_img, h_img)

(p_x, p_y) : Image principal point

(uccd, vccd)

(u_img, v_img)

CCD sensor (mm)

Image (pixel)
3D Point Projection (Pixel Space)

\[
\frac{u_{\text{ccd}}}{w_{\text{ccd}}} = \frac{u_{\text{img}} - p_x}{w_{\text{img}}}
\]

- \((u_{\text{ccd}}, v_{\text{ccd}})\) represents the coordinates on the CCD sensor in millimeters.
- \((0,0)\) is the origin on the CCD sensor.
- \((u_{\text{img}}, v_{\text{img}})\) represents the coordinates on the image in pixels.
- \((p_x, p_y)\) is the image principal point.
- The projection of a pinhole maps points from the 3D space to the 2D image plane.
3D Point Projection (Pixel Space)

Projection of pinhole

\[ \frac{u_{\text{ccd}}}{w_{\text{ccd}}} = \frac{u_{\text{img}} - p_x}{w_{\text{img}}} \]

\[ \frac{v_{\text{ccd}}}{h_{\text{ccd}}} = \frac{v_{\text{img}} - p_y}{h_{\text{img}}} \]

\( (0,0) \)

\( (u_{\text{img}}, v_{\text{img}}) \):

Image (pixel)

Image principal point

\( (p_x, p_y) \):

\( (u_{\text{ccd}}, v_{\text{ccd}}) \):

CCD sensor (mm)

\[ (0,0) \]

\( w_{\text{img}} \)

\( h_{\text{img}} \)
3D Point Projection (Pixel Space)

Projection of pinhole

\[
\begin{align*}
    \frac{u_{\text{ccd}}}{w_{\text{ccd}}} &= \frac{u_{\text{img}} - p_x}{w_{\text{img}}} & \frac{v_{\text{ccd}}}{h_{\text{ccd}}} &= \frac{v_{\text{img}} - p_y}{h_{\text{img}}} \\
    u_{\text{img}} &= u_{\text{ccd}} \frac{w_{\text{img}}}{w_{\text{ccd}}} + p_x & v_{\text{img}} &= v_{\text{ccd}} \frac{h_{\text{img}}}{h_{\text{ccd}}} + p_y
\end{align*}
\]
3D Point Projection (Pixel Space)

\[(u_{\text{ccd}}, v_{\text{ccd}}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z}) : \text{Metric projection}\]
3D Point Projection (Pixel Space)

\[(u_{ccd}, v_{ccd}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z})\] : Metric projection

Pixel projection

\[u_{img} = u_{ccd} \frac{w_{img}}{w_{ccd}} + p_x\]
\[v_{img} = v_{ccd} \frac{h_{img}}{h_{ccd}} + p_y\]
3D Point Projection (Pixel Space)

Projection plane

$(X, Y, Z)$

$(u_{ccd}, v_{ccd})$

$O$

$f_m$

$Z$

$(u_{ccd}, v_{ccd}) = (f_m \cdot \frac{X}{Z}, f_m \cdot \frac{Y}{Z})$ : Metric projection

Pixel projection

$\longrightarrow u_{\text{img}} = u_{ccd} \cdot \frac{w_{\text{img}}}{w_{ccd}} + p_x = f_m \cdot \frac{w_{\text{img}}}{w_{ccd}} \cdot \frac{X}{Z} + p_x$

$v_{\text{img}} = v_{ccd} \cdot \frac{h_{\text{img}}}{h_{ccd}} + p_y = f_m \cdot \frac{h_{\text{img}}}{h_{ccd}} \cdot \frac{Y}{Z} + p_y$
3D Point Projection (Pixel Space)

Projection plane

\((X, Y, Z)\)

\((u_{\text{img}}, v_{\text{img}})\)

\(O\)

\(f\)

\(Z\)

\((u_{\text{ccd}}, v_{\text{ccd}})\)

\((0, 0)\)

\(w_{\text{img}}\)

\(h_{\text{img}}\)

\(w_{\text{ccd}}\)

\(h_{\text{ccd}}\)

\(u_{\text{cd}}\)

\(v_{\text{cd}}\)

\((p_x, p_y)\)

\(X\)

\(Y\)

\(Z\)

\(f_m\)

\(f_{m\text{\text{\text{\text{Z}}}}}\)

\((u_{\text{ccd}}, v_{\text{ccd}}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z}) : \text{Metric projection}\)

Pixel projection

\[ u_{\text{img}} = u_{\text{ccd}} \frac{w_{\text{img}}}{w_{\text{ccd}}} + p_x = f_m \frac{w_{\text{img}}}{w_{\text{ccd}}} \frac{X}{Z} + p_x \]

\[ v_{\text{img}} = v_{\text{ccd}} \frac{h_{\text{img}}}{h_{\text{ccd}}} + p_y = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y}{Z} + p_y \]

Focal length in pixel
3D Point Projection (Pixel Space)

\[ (u_{\text{ccd}}, v_{\text{ccd}}) = (f_m \frac{X}{Z}, f_m \frac{Y}{Z}) \] : Metric projection

Pixel projection

\[ \rightarrow u_{\text{img}} = u_{\text{ccd}} \frac{w_{\text{img}}}{w_{\text{ccd}}} + p_x = f_x \frac{X}{Z} + p_x \]

\[ v_{\text{img}} = v_{\text{ccd}} \frac{h_{\text{img}}}{h_{\text{ccd}}} + p_y = f_y \frac{Y}{Z} + p_y \]

Focal length in pixel

where \( f_x = f_m \frac{w_{\text{img}}}{w_{\text{ccd}}} \) and \( f_y = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \).
3D Point Projection (Pixel Space)

Projection plane

\[(X,Y,Z)\]

\[(u_{img}, v_{img})\]

\[O\]

\[f\]

\[Z\]

\[(u_{ccd}, v_{ccd})\]

\[(0,0)\]

\[w_{img}\]

\[h_{img}\]

\[w_{ccd}\]

\[h_{ccd}\]

CCD sensor (mm)

Image (pixel)

\[(u_{img}, v_{img})\]

\[(0,0)\]

\[w_{img}\]

\[h_{img}\]

\[w_{ccd}\]

\[h_{ccd}\]

Pixel projection

\[\rightarrow u_{img} = u_{ccd} \frac{w_{img}}{w_{ccd}} + p_x = f \frac{X}{Z} + p_x\]

\[v_{img} = v_{ccd} \frac{h_{img}}{h_{ccd}} + p_y = f \frac{Y}{Z} + p_y\]

Focal length in pixel

where

\[f = f_m \frac{w_{img}}{w_{ccd}} = f_m \frac{h_{img}}{h_{ccd}}\]

if

\[\frac{w_{img}}{w_{ccd}} = \frac{h_{img}}{h_{ccd}}\]
3D Point Projection (Pixel Space)

\[ u_{\text{img}} = f \frac{X}{Z} + p_x \]
\[ v_{\text{img}} = f \frac{Y}{Z} + p_y \]
3D Point Projection (Pixel Space)

\[ u_{\text{img}} = f \frac{X}{Z} + p_x \quad \rightarrow \quad Z u_{\text{img}} = f X + p_x Z \]

\[ v_{\text{img}} = f \frac{Y}{Z} + p_y \quad \rightarrow \quad Z v_{\text{img}} = f Y + p_y Z \]
3D Point Projection (Pixel Space)

\[ u_{\text{img}} = f \frac{X}{Z} + p_x \quad \rightarrow \quad Zu_{\text{img}} = fX + p_xZ \]

\[ v_{\text{img}} = f \frac{Y}{Z} + p_y \quad \rightarrow \quad Zv_{\text{img}} = fY + p_yZ \]

\[
\begin{bmatrix}
Z \\
Zv_{\text{img}}
\end{bmatrix}
= 
\begin{bmatrix}
f & p_x & X \\
& f & p_y & Y \\
& & 1 & Z
\end{bmatrix}
\]
3D Point Projection (Pixel Space)

\[ u_{\text{img}} = f \frac{X}{Z} + p_x \quad \longrightarrow \quad Zu_{\text{img}} = fX + p_xZ \]
\[ v_{\text{img}} = f \frac{Y}{Z} + p_y \quad \longrightarrow \quad Zv_{\text{img}} = fY + p_yZ \]

\[
\lambda \begin{bmatrix}
  u_{\text{img}} \\
  v_{\text{img}} \\
  1
\end{bmatrix} =
\begin{bmatrix}
  f & p_x \\
  f & p_y \\
  1 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
\]
Computer Graphics

\[ u_{\text{img}} = f \frac{X}{Z} + p_x \quad \rightarrow \quad Z u_{\text{img}} = fX + p_x Z \]

\[ v_{\text{img}} = f \frac{Y}{Z} + p_y \quad \rightarrow \quad Z v_{\text{img}} = fY + p_y Z \]

\[ \lambda \begin{bmatrix} u_{\text{img}} \\ v_{\text{img}} \\ 1 \end{bmatrix} = \begin{bmatrix} f & p_x \\ f & p_y \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]
Computer Vision = inv(Computer Graphics)

\[
\begin{align*}
    u_{\text{img}} &= f \frac{X}{Z} + p_x \\
    v_{\text{img}} &= f \frac{Y}{Z} + p_y
\end{align*}
\quad \rightarrow \quad
\begin{align*}
    Zu_{\text{img}} &= fX + p_x Z \\
    Zv_{\text{img}} &= fY + p_y Z
\end{align*}
\]

\[\lambda \begin{bmatrix} u_{\text{img}} \\ v_{\text{img}} \\ 1 \end{bmatrix} = \begin{bmatrix} f & p_x \\ f & p_y \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}\]
3D Point Projection (Pixel Space)

Projection plane

\( (X,Y,Z) \)

\( (u_{\text{img}}, v_{\text{img}}) \)

\( f \)

\( Z \)

Pixel space

\[
\begin{bmatrix}
  u_{\text{img}} \\
  v_{\text{img}} \\
  1
\end{bmatrix}
= \lambda
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
\]

\[
\lambda = f
\]

\[
\begin{bmatrix}
  u_{\text{img}} \\
  v_{\text{img}} \\
  w_{\text{img}}
\end{bmatrix}
= \begin{bmatrix}
  f \\
  p_x \\
  1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
\]

Metric space

\[
u_{\text{img}} = f \frac{X}{Z} + p_x \quad \rightarrow \quad Zu_{\text{img}} = fX + p_x Z
\]

\[
v_{\text{img}} = f \frac{Y}{Z} + p_y \quad \rightarrow \quad Zv_{\text{img}} = fY + p_y Z
\]

Pixel space

CCD sensor (mm)

(0,0)

\( h_{\text{ccd}} \)

\( w_{\text{ccd}} \)

Image (pixel)

(0,0)

\( h_{\text{img}} \)

\( w_{\text{img}} \)

(\( p_x \), \( p_y \))
3D Point Projection (Pixel Space)

Projection plane

\[(X, Y, Z)\]

\[(u_{\text{img}}, v_{\text{img}})\]

Pixel space

\[
\begin{bmatrix}
u_{\text{img}} \\
v_{\text{img}} \\
1
\end{bmatrix} =
\begin{bmatrix}
f & p_x & X \\
f & p_y & Y \\
0 & 0 & 1
\end{bmatrix}
\]

Metric space

\[
u_{\text{img}} = f \frac{X}{Z} + p_x \rightarrow Z u_{\text{img}} = fX + p_x Z
\]

\[
v_{\text{img}} = f \frac{Y}{Z} + p_y \rightarrow Z v_{\text{img}} = fY + p_y Z
\]

Camera intrinsic parameter

: metric space to pixel space
3D Point Projection (Pixel Space)

Projection plane

\[ (X, Y, Z) \]

\[ (u_{\text{img}}, v_{\text{img}}) \]

Pixel space

\[ \lambda \begin{bmatrix} u_{\text{img}} \\ v_{\text{img}} \\ 1 \end{bmatrix} = \begin{bmatrix} f & p_x \\ p_y & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]

Metric space

\[ u_{\text{img}} = f \frac{X}{Z} + p_x \quad \rightarrow \quad Zu_{\text{img}} = fX + p_xZ \]
\[ v_{\text{img}} = f \frac{Y}{Z} + p_y \quad \rightarrow \quad Zv_{\text{img}} = fY + p_yZ \]

Camera intrinsic parameter
: metric space to pixel space
Camera Intrinsic Parameter

Metric space

\[
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}
\]
Camera Intrinsic Parameter

\[
\begin{bmatrix}
f & p_x & X \\
p_y & Y \\
1 & Z
\end{bmatrix}
\]

Metric space

\[
(X, Y, Z)
\]

Projection plane
Camera Intrinsic Parameter

\[ \lambda \begin{bmatrix} u_{\text{img}} \\ v_{\text{img}} \\ 1 \end{bmatrix} = \begin{bmatrix} f \\ p_x \\ p_y \\ 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]
2D Inverse Projection

Pixel space:

\[
\begin{bmatrix}
u_{\text{img}} \\
v_{\text{img}} \\
1
\end{bmatrix}
= \begin{bmatrix}
f \\
p_x \\
p_y \\
1
\end{bmatrix}
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix}
\]

Metric space:
2D Inverse Projection

Pixel space
\[ \lambda \begin{bmatrix} u_{\text{img}} \\ v_{\text{img}} \\ 1 \end{bmatrix} = \begin{bmatrix} f \\ p_x \\ p_y \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]

Metric space

3D ray
\[ \lambda K^{-1} \begin{bmatrix} u_{\text{img}} \\ v_{\text{img}} \\ 1 \end{bmatrix} \]
2D Inverse Projection

Pixel space

\[
\lambda \begin{bmatrix}
u_{\text{img}} \\
v_{\text{img}} \\
1
\end{bmatrix} = K \begin{bmatrix}f \\ p_x \\ p_y \\ 1 \end{bmatrix} \begin{bmatrix}X \\ Y \\ Z \end{bmatrix}
\]

Metric space
2D Inverse Projection

The 3D point must lie in the 3D ray passing through the origin and 2D image point.
Exercise

What f to make the height of Eifel tower appear 960 pixel distance?

960 pix

1280 pix

324 m

21.8 mm

ccd size

1500 m

f_m?
Exercise

What f to make the height of Eifel tower appear 960 pixel distance?

\[ y_{\text{img}} = f \frac{Y}{Z} = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y}{Z} \]
Exercise

What f to make the height of Eifel tower appear 960 pixel distance?

\[ y_{img} = f \frac{Y}{Z} = f_m \frac{h_{img}}{h_{ccd}} \frac{Y}{Z} \]

\[ 960 = f_m \frac{1280}{0.0218} \frac{324}{1500} \rightarrow f_m = 0.0757 \text{ m} \]
Focal Length

Diagonal viewing angle for 35mm film

Ultra wide-angle  Wide-angle  Standard  Telephoto  Super telephoto

14mm  20mm  24mm  35mm  50mm  70mm  135mm  200mm  400mm  600mm

Normal view seen by the human eye
Exercise

What Z to make the height of Eifel tower appear 960 pixel distance?

\[ y_{\text{img}} = f \frac{Y}{Z} = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y}{Z} \]

\[ f_m = 50 \text{ mm} \]

324 m

960 pix

1280 pix

Z?
Exercise

What $Z$ to make the height of Eifel tower appear 960 pixel distance?

$$y_{\text{img}} = f \frac{Y}{Z} = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y}{Z}$$

$$960 = 0.05 \frac{1280}{0.0218} \frac{324}{Z} \rightarrow Z = 990.826 \text{m}$$
Exercise

What $Z_p$ to make the height of Eifel tower appear twice of the person?

$f_m = 50$ mm

$Z_p$

1500
Exercise

What $Z_p$ to make the height of Eifel tower appear twice of the person?

$$h_e = f \frac{Y}{Z} \quad h_p = f \frac{Y_p}{Z_p} \quad \text{s.t.} \quad h_p = \frac{h_e}{2}$$
Exercise

What $Z_p$ to make the height of Eifel tower appear twice of the person?

\[
h_e = f \frac{Y}{Z} \quad h_p = f \frac{Y_p}{Z_p} \quad \text{s.t.} \quad h_p = \frac{h_e}{2}
\]

\[
f \frac{Y_p}{Z_p} = f \frac{Y}{2Z} \quad \Rightarrow \quad Z_p = 2 \cdot 1500 \cdot \frac{1.7}{324} = 15.74 \text{m}
\]
Where Was I?

Sensor size

Sensor spec

\[ y = f \frac{Y}{Z} = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y}{Z} \]
Where Was I?

Sensor size

\[ y = f \frac{Y}{Z} = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y}{Z} \]
Where Was I?

Circa 1984
Where Was I?

Circa 1984
Where Was I?

Circa 1984

670 pix

250 pix

$ f_m = 50 \text{ mm} $

0.9 m

324 m
Where Was I?

Circa 1984

\[ y_1 = f \cdot \frac{Y_1}{Z_1} = f_m \cdot \frac{h_{\text{img}} \cdot Y_1}{h_{\text{ccd}} \cdot Z_1} \]
Where Was I?

Circa 1984

\[ y_1 = f \frac{Y_1}{Z_1} = f_m \frac{h_{\text{img}} Y_1}{h_{\text{ccd}} Z_1} \rightarrow Z_1 = f_m \frac{h_{\text{img}} Y_1}{h_{\text{ccd}} Y_1} \]
Where Was I?

Circa 1984

\[ y_1 = f \frac{Y_1}{Z_1} = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y_1}{Z_1} \rightarrow Z_1 = f_m \frac{h_{\text{img}}}{h_{\text{ccd}}} \frac{Y_1}{y_1} = 0.05 \frac{1280}{0.0218} \frac{0.9}{250} = 8.03\,\text{m} \]
Where Was I?

Circa 1984

\[
y_1 = f \frac{Y_1}{Z_1} = f_m \frac{h_{\text{img}} \ Y_1}{h_{\text{ccd}} \ Z_1} \rightarrow Z_1 = f_m \frac{h_{\text{img}} \ Y_1}{h_{\text{ccd}} \ Y_1} = 0.05 \frac{1280}{0.0218} \frac{0.9}{250} = 8.03 \text{m}
\]

\[
y_2 = f \frac{Y_2}{Z_2} = f_m \frac{h_{\text{img}} \ Y_2}{h_{\text{ccd}} \ Z_2} \rightarrow Z_2 = f_m \frac{h_{\text{img}} \ Y_2}{h_{\text{ccd}} \ Y_2} = 0.05 \frac{1280}{0.0218} \frac{324}{670} = 1079 \text{m}
\]
Where Was I?

\[
y_2 = f \frac{Y_2}{Z_2} = f \frac{h_{\text{img}}}{h_{\text{ccd}}} Y_2 \Rightarrow Z_2 = f \frac{h_{\text{img}}}{h_{\text{ccd}}} y_2 = 0.05 \frac{1280}{0.0218} = 670 = 1079 \text{m}
\]
Where Was I?

Photo Sphere - Jun 2014

[Map of Paris showing the Eiffel Tower and nearby areas]
Focal Length

\[ f \quad d \]

Strong perspective
Focal Length

\[ f \quad d \]

Weak perspective
Focal Length

\[ f \quad d \]
Focal Length

Diagonal viewing angle for 35mm film

Ultra wide-angle | Wide-angle | Standard | Telephoto | Super telephoto

14mm | 20mm | 24mm | 35mm | 50mm | 70mm | 135mm | 200mm | 400mm | 600mm

Normal view seen by the human eye
Dolly Zoom (Vertigo Effect)

VERTIGO (1958)
Dolly Zoom (Vertigo Effect)
Dolly Zoom

Given focal length ($f_m=100\text{mm}$), what $Z_{100}$ to make the height of the person remain the same as $f_m=50\text{mm}$?

$f_0 = 50\text{ mm}$

$Z_{50} = 157.41\text{ m}$
Given focal length \( f_m = 100\text{mm} \), what \( Z_{100} \) to make the height of the person remain the same as \( f_m = 50\text{mm} \)?

\[
h_{50} = f_{50} \frac{Y}{Z_{50}} \quad h_{100} = f_{100} \frac{Y}{Z_{100}} \quad \text{s.t.} \quad h_{100} = h_{50}
\]
Given focal length \( f_{100} = 100 \text{mm} \), what \( Z_{100} \) to make the height of the person remain the same as \( f_{m} = 50 \text{mm} \)?

\[
\begin{align*}
 h_{50} &= f_{50} \frac{Y}{Z_{50}} \\
 h_{100} &= f_{100} \frac{Y}{Z_{100}} \\
 \text{s.t.} \quad h_{100} &= h_{50} \\
 Z_{100} &= \frac{f_{100}}{f_{50}} Z_{50} \\
 Z_{100} &= \frac{100}{50} 157.41 = 314.8 \text{m}
\end{align*}
\]
Where am I with Dolly Zoom?

$h_1 = 400\text{pix}$

$h_2 = 120\text{pix}$
Where am I with Dolly Zoom?

How far I need to step back with zoom factor x2?
How will $h_2$ change?
Where am I with Dolly Zoom?

How far I need to step back with zoom factor x2?
How will $h_2$ change?
Where am I with Dolly Zoom?

Unknows: $f$, $d_1$

Top view

How far I need to step back with zoom factor x2?
How will $h_2$ change?
Where am I with Dolly Zoom?

Unknowns: $f$, $d_1$, $\Delta d$

How far I need to step back with zoom factor x2? How will $h_2$ change?
Where am I with Dolly Zoom?

Equations:
\[ h_1 = \frac{f \cdot H_1}{d_1} \]

How far I need to step back with zoom factor x2?
How will \( h_2 \) change?
Where am I with Dolly Zoom?

Equations:
\[ h_1 = f \frac{H_1}{d_1} \]
\[ h_1 = 2f \frac{H_1}{\Delta d + d_1} \]

How far I need to step back with zoom factor x2?

How will \( h_2 \) change?

Unknowns: \( f, d_1, \Delta d \)
Where am I with Dolly Zoom?

Equations:
\[ h_1 = \frac{f}{d_1} \quad h_t = 2f \frac{H_1}{\Delta d + d_1} \quad \Delta d = d_1 \]

How far I need to step back with zoom factor x2?
How will \( h_2 \) change?
Where am I with Dolly Zoom?

Equations:
\[ h_1 = f \frac{H_1}{d_1} \]
\[ h_2 = f \frac{H_2}{d_1 + d} \]
\[ h_1 = 2f \frac{H_1}{\Delta d + d_1} \rightarrow \Delta d = d_1 \]

How far I need to step back with zoom factor x2?
How will \( h_2 \) change?
Where am I with Dolly Zoom?

Equations:

\[ h_1 = 2f \frac{H_1}{\Delta d + d_1} \]
\[ h_2 = 2f \frac{H_2}{d_1 + d} \]

\[ d_1 = \frac{1}{1 - \frac{h_2 H_1}{h_1 H_2}} \cdot d = 2.5m \]

\[ \Delta d = 2.5m \]

Unknowns: \( f, d_1, \Delta d \)

\[ h_1 = 400\text{pix} \]
\[ h_2 = 120\text{pix} \]
\[ h_2' = 200\text{pix} \]
Where am I with Dolly Zoom?

**Equations:**

\[
\begin{align*}
    h_1 &= f \frac{H_1}{d_1} \\
    h_2 &= f \frac{H_2}{d_1 + d}
\end{align*}
\]

\[
\Delta d = d_1
\]

\[
\begin{align*}
    d_1 &= \frac{1}{1 - \frac{h_2 H_1}{h_1 H_2}} \\
        &= \frac{1}{1 - \frac{120 \times 6}{400 \times 4}} \\
        &= \frac{1}{1 - \frac{3}{5}} \\
        &= \frac{1}{\frac{2}{5}} \\
        &= \frac{5}{2} \\
        &= 2.5 \text{m}
\end{align*}
\]

\[
\Delta d = 2.5 \text{m} \
\]

\[
f = 250 \text{pix}
\]

Unknowns: \( f, d_1, \Delta d \)

\[
h_1 = 400 \text{pix} \
\]

\[
h_2 = 120 \text{pix} \
\]

\[
h_2' = 200 \text{pix}
\]
Where am I with Dolly Zoom?

Equations:

\[ h_1 = f \frac{H_1}{d_1} \]
\[ h_1 = 2f \frac{H_1}{\Delta d + d_1} \]
\[ h_2 = f \frac{H_2}{d_1 + d} \]
\[ h_2' = 2f \frac{H_2}{\Delta d + d_1 + d} \]

Unknowns: \( f, d_1, \Delta d \)

How far I need to step back with zoom factor x2?

How will \( h_2 \) change?

\( d_1 = \Delta d = 2.5 \text{m} \)
\( f = 250 \text{pix} \)

\( h_1 = 400 \text{pix} \)
\( h_2 = 120 \text{pix} \)
Where am I with Dolly Zoom?

Equations:

\[ h_1 = f \frac{H_1}{d_1} \]
\[ h_2 = f \frac{H_2}{d_1 + d} \]
\[ h_1' = 2f \frac{H_1}{\Delta d + d} \]
\[ h_2' = 2f \frac{H_2}{\Delta d + d_1 + d} = 429 \text{pix} \]

Unkowns: \( f, d_1, \Delta d \)

How far I need to step back with zoom factor x2?
How will \( h_2 \) change?
%% load data and set parameters
points = load('points.mat');

d_ref = 4;
f_ref = 400;
pos = 0 : -0.1 : -9.9;

H1 = points.points_A(1,2) - points.points_A(2,2);
H2 = points.points_C(1,2) - points.points_C(2,2);

%% Dolly Zoom: keep one object's height constant
f = compute_focal_length(d_ref, f_ref, pos);

for i = 1 : length(f)
    figure(1), hold on, axis equal;
    xlim([0,1920]), ylim([0,1080]);
    project_objects(f(i), pos(i), points, 1);
    pause(0.1);
end;

function [ f ] = compute_focal_length( d_ref, f_ref, pos )
f = (d_ref-pos) / d_ref * f_ref;
Where am I with Dolly Zoom?

1) Zoom in with factor x2
2) Step back to maintain $h_1$
3) Predict $h_2$
1) Where was I with respect to A?
2) How far B from A?