Computer Vision I - Projective Geometry and Geometry of a single Camera

Carsten Rother

24/11/2016
Roadmap for next four lectures

• Appearance-based Matching (sec. 4.1)

• Projective Geometry - Basics (sec. 2.1.1-2.1.4)

• Geometry of a Single Camera (sec 2.1.5, 2.1.6)
  • Camera versus Human Perception
  • The Pinhole Camera
  • Lens effects

• Geometry of two Views (sec. 7.2)
  • The Homography (e.g. rotating camera)
  • Camera Calibration (3D to 2D Mapping)
  • The Fundamental and Essential Matrix (two arbitrary images)

• Robust Geometry estimation for two cameras (sec. 6.1.4)

• Multi-View 3D reconstruction (sec. 7.3-7.4)
  • General scenario
  • From Projective to Metric Space
  • Special Cases
We look at these operations in: $R^2/R^3, P^2/P^3$

Primitives:
- Points in 2D/3D
- Lines in 2D/3D
- Planes in 3D
- Conic in 2D; Quadric in 3D

Operations with Primitives:
- Intersection
- Tangent

Transformations:
- Rotation
- Translation
- Projective
- ....
Reminder: Points at infinity

Points with coordinate \( \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \) are so-called “ideal points” or “points at infinity”

\( \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in P^2 \quad \text{Not defined in } R^2 \text{ since } w = 0 \)
Reminder: Line at infinity

• There is a “special” line, called line at infinity: (0,0,1)

• All points at infinity \((x, y, 0)\) lie on the line at infinity \((0,0,1)\):
  \[ x \cdot 0 + y \cdot 0 + 0 \cdot 1 = 0 \]
2D conic “Kegelschnitt”

• Conics are shapes that arise when a plane intersects a cone

• In compact form: \( x^t C x = 0 \) where \( C \) has the form:

\[
C = \begin{bmatrix}
a & b/2 & d/2 \\
b/2 & c & e/2 \\
d/2 & e/2 & f
\end{bmatrix}
\]

• This can be written in in-homogenous coordinates:

\( ax^2 + bxy + cy^2 + dx + ey + f = 0 \)

where \( \vec{x} = (x, y) \)

• \( C \) has 5DoF since unique up to scale:

\( x^t C x = k x^t C x = x^t kC x = 0 \)

• Properties: \( l \) is tangent to \( C \) at a point \( x \) if \( l = Cx \)
Example: 2D Conic

**A circle:**
\[ x^2 + y^2 - r^2 = 0 \]

**Parabola:**
\[ -x^2 + y = 0 \]
Given this square $S$ and a transformed shape $T$

What is the transformation that can brings $S$ into $T$?

1) Euclidian
2) Similarity
3) Affine
4) Projective
5) 3) and 4)
6) All
2D Transformations in $\mathbb{R}^2$

**Definition:**
- Euclidean: translation + rotation
- Similarity (rigid body transform): Euclidean + scaling
- Affine: Similarity + shearing
- Projective: arbitrary linear transform in homogenous coordinates
2D Transformations of points

• 2D Transformations in homogenous coordinates:

\[
\begin{pmatrix}
    x' \\
    y' \\
    w'
\end{pmatrix}
= \begin{bmatrix}
    a & b & d \\
    e & f & h \\
    i & j & l
\end{bmatrix}
\begin{pmatrix}
    x \\
    y \\
    1
\end{pmatrix}
\]

Transformation matrix

• Example: translation

\[
\begin{pmatrix}
    x' \\
    y' \\
    1
\end{pmatrix}
= \begin{bmatrix}
    1 & 0 & t_x \\
    0 & 1 & t_y \\
    0 & 0 & 1
\end{bmatrix}
\begin{pmatrix}
    x \\
    y \\
    1
\end{pmatrix}
\]

homogeneous coordinates

\[
\begin{pmatrix}
    x' \\
    y' \\
    1
\end{pmatrix}
= \begin{pmatrix}
    x \\
    y \\
    1
\end{pmatrix} + \begin{pmatrix}
    t_x \\
    t_y \\
    0
\end{pmatrix}
\]

inhomogeneous coordinates

Advantage of homogeneous coordinates (i.e. \( P^2 \))
# 2D Transformations of points

<table>
<thead>
<tr>
<th>Group</th>
<th>Matrix</th>
<th>Distortion</th>
<th>Invariant properties</th>
</tr>
</thead>
</table>
| Projective  | \[
| 8 dof      | \[
|            | \begin{bmatrix} h_{11} & h_{12} & h_{13} \\
|            | h_{21} & h_{22} & h_{23} \\
|            | h_{31} & h_{32} & h_{33} \end{bmatrix} \] | (of a square) | Concurrency, collinearity, order of contact: intersection (1 pt contact); tangency (2 pt contact); inflections (3 pt contact with line); tangent discontinuities and cusps. cross ratio (ratio of ratio of lengths). |
| Affine      | \[
| 6 dof      | \[
|            | \begin{bmatrix} a_{11} & a_{12} & t_x \\
|            | a_{21} & a_{22} & t_y \\
|            | 0 & 0 & 1 \end{bmatrix} \] | (two special points on the line at infinity) | Parallelism, ratio of areas, ratio of lengths on collinear or parallel lines (e.g. midpoints), linear combinations of vectors (e.g. centroids). The line at infinity, \( l_\infty \). |
| Similarity  | \[
| 4 dof      | \[
|            | \begin{bmatrix} sr_{11} & sr_{12} & t_x \\
|            | sr_{21} & sr_{22} & t_y \\
|            | 0 & 0 & 1 \end{bmatrix} \] | (see section 2.7.3.) | Ratio of lengths, angle. The circular points, \( I, J \) (two special points on the line at infinity) |
| Euclidean   | \[
| 3 dof      | \[
|            | \begin{bmatrix} r_{11} & r_{12} & t_x \\
|            | r_{21} & r_{22} & t_y \\
|            | 0 & 0 & 1 \end{bmatrix} \] | | Length, area |

Here \( r_{ij} \) is a 2 x 2 rotation matrix with 1 DoF, which can be written as: \[
\begin{bmatrix}
\cos \Theta & -\sin \Theta \\
\sin \Theta & \cos \Theta
\end{bmatrix}
\]
2D Transformation of lines

A point moves: \( x' = Hx \) then:

1) Line (defined by points) moves:
\[
l' = (H^{-1})^t \ l
\]

Proof:

1) Assume \( x_1 \) lies on \( l \), and \( l' = (H^{-1})^t \ l \).

Show that \( x'_1 \) lies on \( l' \).
\[
(x'_1)^t \ l' \overset{!}{=} 0 \Rightarrow (Hx_1)^t(H^{-1})^t \ l \overset{!}{=} 0 \Rightarrow \\
x_1^tH^t(H^{-1})^t \ l \overset{!}{=} 0 \Rightarrow x_1^t(H^{-1}H)^t \ l \overset{!}{=} 0 \Rightarrow x_1^t \ l = 0
\]
**Example: Projective Transformation**

1. Circles on the floor are circles in the image
2. Squares on the floor are squares in the image

**Affine transformation**

1. Circles on the floor are ellipse in the image
2. Squares on the floor are sheared in the image
3. Lines are still parallel

**Picture from the side (projective transformation)**

1. Lines converge to a vanishing point (not at infinity in the image)
In 3D: Points

- $x = (x, y, z) \in \mathbb{R}^3$ has 3 DoF

- With homogeneous coordinates: $(x, y, z, 1) \in \mathbb{P}^3$

- $\mathbb{P}^3$ is defined as the space $\mathbb{R}^4 \setminus (0,0,0,0)$ such that points $(x, y, z, w)$ and $(kx, ky, kz, kw)$ are the same for all $k \neq 0$

- Points: $(x, y, z, 0) \in \mathbb{P}^3$ are called points at infinity
In 3D: Planes

- Planes in $R^3$ are defined by 4 parameters (3 DoF):
  - Normal: $n = (n_x, n_y, n_z)$
  - Offset: $d$

- All points $(x, y, z)$ lie on the plane if:
  $$x n_x + y n_y + z n_z + d = 0$$

- With homogenous coordinates:
  $$x \pi = 0$$, where $x = (x, y, z, 1)$ and $\pi = (n_x, n_y, n_z, d)$

- Planes in $P^3$ are written as: $x \pi = 0$

- Points and planes are dual in $P^3$ (as points and lines have been in $P^2$)

- Plane at infinity is $\pi = (0,0,0,1)$ since all points at infinity $(x,y,z,0)$ lie on it.
Question

Statement:
“In $P^3$ there are many points at infinity and many lines at infinity. All of them lie on the plane at infinity."

Is the statement:
1) True
2) False
3) I don’t know
In 3D: Plane at infinity

All of these elements at infinity lie on the plane at infinity
Why is the plane at infinity important (see later)

Plane at infinity is important to visualize 3D reconstructions nicely.

Plane at infinity can be used to simplify 3D reconstruction.
What is the horizon?

- The ground plane is special (we stand on it)
- Horizon is a line at infinity where “plane at infinity” intersects ground plane

Ground plane: $(0,0,1,0)$
Plane at infinity: $(0,0,0,1)$

Many lines and planes in our real world meet at the horizon (since parallel to ground plane)
In 3D: Points at Infinity

- Points at infinity can be real points in a camera

\[
\begin{pmatrix} b \\ f \\ j \end{pmatrix} = \begin{bmatrix} a & b & c & d \\ e & f & g & h \\ i & j & k & l \end{bmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}
\]

Real point in the image
3x4 Camera Matrix
3D-\rightarrow\text{2D projection}
3D Point at infinity
In 3D: Lines

- Unfortunately not a compact form (as for points)

- A simple representation in $R^3$. Define a line via two points $p, q \in R^3$:
  \[ r = (1 - \lambda)p + \lambda q \]

- A line has 4 DoF (both points $p, q$ can move arbitrary on the line)

- A more compact, but more complex, way to define a 3D Line is to use Plücker coordinates:
  \[ L = pq^t - qp^t \text{ where } \det(L) = 0 \]
  here $L, p, q$ are in homogenous coordinates
In 3D: Quadrics

- Points $X$ on the quadric if: $X^T Q X = 0$
- A quadric $Q$ is a surface in $P^3$
- A quadric is a symmetric $4 \times 4$ matrix with 9 DoF
### In 3D: Transformation

<table>
<thead>
<tr>
<th>Group</th>
<th>Matrix</th>
<th>Distortion (of a cube)</th>
<th>Invariant properties</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projective</td>
<td>$\begin{bmatrix} A &amp; t \ v^T &amp; v \end{bmatrix}$</td>
<td>Intersection and tangency of surfaces in contact. Sign of Gaussian curvature.</td>
<td></td>
</tr>
<tr>
<td>15 dof</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affine</td>
<td>$\begin{bmatrix} A &amp; t \ 0^T &amp; 1 \end{bmatrix}$</td>
<td>Parallelism of planes, volume ratios, centroids. The plane at infinity, $\pi_\infty$, (see section 3.5).</td>
<td></td>
</tr>
<tr>
<td>12 dof</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Similarity</td>
<td>$\begin{bmatrix} sR &amp; t \ 0^T &amp; 1 \end{bmatrix}$</td>
<td>The absolute conic, $\Omega_\infty$, (see section 3.6).</td>
<td></td>
</tr>
<tr>
<td>7 dof</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Euclidean</td>
<td>$\begin{bmatrix} R &amp; t \ 0^T &amp; 1 \end{bmatrix}$</td>
<td>Volume.</td>
<td></td>
</tr>
<tr>
<td>6 dof</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Rotation $\mathbf{R}$ in 3D has 3 DoF. It is slightly more complex, and several options exist:

1) Euler angles: rotate around, $x$, $y$, $z$-axis in order (depends on order, not smooth in parameter space)

2) Axis/angle formulation:

$$\mathbf{R}(\mathbf{n}, \theta) = \mathbf{I} + \sin \theta \left[ \mathbf{n} \right]_\times + (1 - \cos \theta)\left[ \mathbf{n} \right]_\times^2$$

$n$ is the normal vector (2 DoF) and $\theta$ the angle (1 DoF)

3) Another option is unit quaternions (see book page 40)
Roadmap for next four lectures

- Appearance-based Matching (sec. 4.1)

- Projective Geometry - Basics (sec. 2.1.1-2.1.4)

- Geometry of a Single Camera (sec 2.1.5, 2.1.6)
  - Camera versus Human Perception
  - The Pinhole Camera
  - Lens effects

- Geometry of two Views (sec. 7.2)
  - The Homography (e.g. rotating camera)
  - Camera Calibration (3D to 2D Mapping)
  - The Fundamental and Essential Matrix (two arbitrary images)

- Robust Geometry estimation for two cameras (sec. 6.1.4)

- Multi-View 3D reconstruction (sec. 7.3-7.4)
  - General scenario
  - From Projective to Metric Space
  - Special Cases
The Human eye

- The retina contains different types of sensors: cones “Zapfen” (colors, 6 million) and rods “Stäbchen” (gray levels, 120 million)
- The resolution is much higher in fovea centralis
- Light first passes through the layer of neurons before it reaches photo sensors (smoothing). Only in the “fovea centralis” the light hits directly the photo sensors
- Signal goes out the other way: Retina → Ganglion cells (1 million) → Optic nerve → 1 MPixel Camera?
- Where the optical nerve “leaves” the retina is a “blind spot”
Spatial resolution
Spatial resolution

2MP Camera, far from the screen
Spatial resolution

5MP Camera, close to the screen
The resolution is much higher in fovea centralis.
- The Information is pre-processed by Ganglion cells
  (Compare: 3072×2304=7MPixel, 2.4 MB RGB JPEG loss-less)
- No still image, but a „Video“ (super-resolution)
- Scanning technique called Saccades
Eyes never move uniformly, but jump in **saccades** (approximately 15-100 ms duration between fixation points)

Saccades are driven by the “importance” of the scene parts (eyes, mouth etc).
What is light?
Spectrum, i.e. a function of the wavelength

Spectral resolution of the eye is relatively bad due to projection $\infty \rightarrow 3$
Compare to CCD Camera: Color - Filters

Bayer grid

Estimate missing components from neighboring values (demosaicing)

Why more green?

Human Luminance Sensitivity Function
There is a lot of image processing in a camera.
Roadmap for next four lectures

• Appearance-based Matching (sec. 4.1)

• Projective Geometry - Basics (sec. 2.1.1-2.1.4)

• Geometry of a Single Camera (sec 2.1.5, 2.1.6)
  • Camera versus Human Perception
  • The Pinhole Camera
  • Lens effects

• Geometry of two Views (sec. 7.2)
  • The Homography (e.g. rotating camera)
  • Camera Calibration (3D to 2D Mapping)
  • The Fundamental and Essential Matrix (two arbitrary images)

• Robust Geometry estimation for two cameras (sec. 6.1.4)

• Multi-View 3D reconstruction (sec. 7.3-7.4)
  • General scenario
  • From Projective to Metric Space
  • Special Cases
How can we capture the world

- Let’s design a camera
  - Idea 1: put a piece of film (or a CCD) in front of an object
  - Do we get a reasonable image?
• Add a barrier to block-off most of the rays
  • This reduces blurring
  • The opening is known as the *aperture* ("Blende")
  • For now we make the opening of aperture infinitely small (revisited later)
Pinhole camera model

Pinhole model:
- Captures pencil of rays – all rays through a single point
- Projected rays are straight lines
- The point where all rays meet is called center of projection (focal point)
- The image is formed on the image plane
Pinhole camera – Properties

• Many-to-one: any point along the same ray maps to the same point in the image

• Points map to points
  (But projection of points on focal plane is undefined)

• Lines map to lines (collinearity is preserved)
  (But line through focal point projects to a point)

• Planes map to planes (or half-plane)
  (But plane through focal point projects to line)
Dimensionality Reduction Machine (3D to 2D)

3D world

Point of observation

2D image
Question

Given an image of tiles. What properties of the 3D world are preserved in the image?

Questions:
1) Parallel lines are parallel
2) Angles
3) Length
4) Parallel lines meet in a point in $P^2$
5) None of the above
Pinhole camera model – in maths

- Similar triangles: \( \frac{y}{f} = \frac{Y}{Z} \)

- That gives: \( y = f \frac{Y}{Z} \) and \( x = f \frac{X}{Z} \)

- That gives: 
  \[
  \begin{bmatrix}
  x \\
  y \\
  1
  \end{bmatrix} = 
  \begin{bmatrix}
  f & 0 & 0 \\
  0 & f & 0 \\
  0 & 0 & 1
  \end{bmatrix}
  \begin{bmatrix}
  X \\
  Y \\
  Z
  \end{bmatrix}
  \] (remember “=” means equal up to scale)

\(2D\) homogenous coordinate \(3D\) inhomogenous coordinate
Pinhole camera model – in maths

That gives:

\[
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix} =
\begin{bmatrix}
  f & 0 & 0 \\
  0 & f & 0 \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
\]

Calibration matrix \( K \)

In short \( x = K \tilde{X} \) (here \( \tilde{X} \) means inhomogeneous coordinates)

Intrinsic Camera Calibration means we know \( K \)

We can also go from image points into the 3D world: \( \tilde{X} = K^{-1} x \)
Pinhole camera - Definitions

- **Principal axis**: line from the camera center perpendicular to the image plane
- **Normalized (camera) coordinate system**: camera center is at the origin of a world coordinate system, and the principal axis is the z-axis
- **Principal point (p)**: point where principal axis intersects the image plane (origin of normalized coordinate system)
Principal point \( p = (x_0, y_0) = (p_x, p_y) \)

- Camera coordinate system: origin is at the principal point
- Image coordinate system: origin is in the corner

*In practice: principal point in center of the image*
Adding principal point into $K$

That gives:

$$
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix}
= 
\begin{bmatrix}
  f & 0 & p_x \\
  0 & f & p_y \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
$$

Principal point $(p_x, p_y)$

Projection with principal point:

$$
y = f \frac{Y}{Z} + p_y = \frac{fY + Zp_y}{Z} \quad \text{and} \quad x = f \frac{X}{Z} + p_x = \frac{fX + Zp_x}{Z}
$$

"image coordinate system"  "camera coordinate system"
Pixel Size and Shape

- $f$ has a certain unit (m, mm, inch, ...)
- $m_x$ pixels per unit (m, mm, inch, ...) in horizontal direction
- $m_y$ pixels per unit (m, mm, inch, ...) in vertical direction
- $s'$ skew of a pixel

That gives:

$$
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix} =
\begin{bmatrix}
  f & 0 & px \\
  0 & f & py \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  m_x & s' & 0 \\
  0 & m_y & 0 \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix} =
\begin{bmatrix}
  f m_x & f s' & px \\
  0 & f m_y & py \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
$$

Simplify and choose a pixel as unit:

$$
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix} =
\begin{bmatrix}
  f & s & px \\
  0 & mf & py \\
  0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z
\end{bmatrix}
$$

$f$ now in units of pixels

Final calibration matrix $K$
Camera intrinsic parameters - Summary

• **Intrinsic parameters**
  • Principal point coordinates \((p_x, p_y)\)
  • Focal length \(f\)
  • Pixel y-direction magnification factors \(m\)
  • Skew (non-rectangular pixels) \(s\)

\[
K = \begin{bmatrix}
  f & s & p_x \\
  0 & mf & p_y \\
  0 & 0 & 1
\end{bmatrix}
\]

For later:
We sometimes have to only guess these values and then they are
optimized via e.g. bundle adjustment. A good guess is:
  • \(p\) in image center
  • \(s = 0, m = 1\)
  • \(f = \) EXIF tag (or guess, e.g. two times image dimension)
Camera intrinsic parameters - Summary

- **Intrinsic parameters**
  - Principal point coordinates \((p_x, p_y)\)
  - Focal length \(f\)
  - Pixel y-direction magnification factors \(m\)
  - Skew (non-rectangular pixels) \(s\)

The Mapping: \(x = K\bar{X}\) \(\text{(here } \bar{X} \text{ means inhomogeneous coordinates)}\)

\[
\begin{bmatrix}
    x \\
    y \\
    1
\end{bmatrix} =
\begin{bmatrix}
    f & s & p_x \\
    0 & mf & p_y \\
    0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
    X \\
    Y \\
    Z
\end{bmatrix}
\]
Given a 3D homogenous point $X^w$, and camera $C$ in world coordinate system

1) Translate from world to camera coordinate system:
\[ \tilde{X}^c = \tilde{X}^w - \tilde{C} \]
\[ \tilde{X}^c = (I_{3 \times 3} \mid - \tilde{C}) X^w \quad \text{where } I_{3 \times 3} \text{ is 3x3 identity matrix} \]

2) Rotate world coordinate system into camera coordinate system
\[ \tilde{X}^c = R (I_{3 \times 3} \mid - \tilde{C}) X^w \]

3) Apply camera matrix
\[ x = K R (I_{3 \times 3} \mid - \tilde{C}) X \]
Camera matrix

- Camera matrix \( P \) is defined as:

\[
x = K R \begin{pmatrix} I_{3\times3} & -\tilde{C} \end{pmatrix} X
\]

\( P \) (3 \times 4) camera matrix has 11 DoF

- In short we write: \( x = PX \)

- The camera center is the (right) nullspace of \( P \), since

\[
P C = K R (\tilde{C} - \tilde{C}) = 0
\]

Comments:
- The right nullspace is formed by the linear-independent vectors \( X \) that satisfies: \( PX = 0 \)
- \( P \) has full Rank 3
Camera parameters - Summary

• Camera matrix $P$ has 11 DoF

\[ x = PX \]
\[ x = KR (I_{3 \times 3} | -\tilde{C}) X \]

• Intrinsic parameters
  • Principal point coordinates $(p_x, p_y)$
  • Focal length $f$
  • Pixel y-direction magnification factors $m$
  • Skew (non-rectangular pixels) $s$

\[ K = \begin{bmatrix} f & s & p_x \\ 0 & mf & p_y \\ 0 & 0 & 1 \end{bmatrix} \]

• Extrinsic parameters
  • Rotation $R$ (3DoF) and translation $\tilde{C}$ (3DoF) relative to world coordinate system
Orthographic Projection

• Distance from center of projection to image plane is infinite (infinite focal length)

• Also called “parallel projection”

• Most simple form of orthographic projection

\[
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix} = \begin{bmatrix}
  1 & 0 & 0 & 0 \\
  0 & 1 & 0 & 0 \\
  0 & 0 & 0 & 1 \\
\end{bmatrix} \begin{bmatrix}
  X \\
  Y \\
  Z \\
  1
\end{bmatrix}
\]
Affine cameras

• Most general camera that does parallel projection is affine camera:

\[
\begin{bmatrix}
  x \\
  y \\
  1
\end{bmatrix} =
\begin{bmatrix}
  a & b & c & d \\
  e & f & g & h \\
  0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z \\
  w
\end{bmatrix}
\]

• Parallel lines in 3D map to parallel lines in 2D (since points at infinity stay)

\[
\begin{bmatrix}
  x \\
  y \\
  0
\end{bmatrix} =
\begin{bmatrix}
  a & b & c & d \\
  e & f & g & h \\
  0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix}
  X \\
  Y \\
  Z \\
  0
\end{bmatrix}
\]

• Affine cameras simplify the 3D reconstruction task (see later)

• See more details in: HZ (Hartley, Zissermann) chapter 6.3
Roadmap for next four lectures

• Appearance-based Matching (sec. 4.1)

• Projective Geometry - Basics (sec. 2.1.1-2.1.4)

• Geometry of a Single Camera (sec 2.1.5, 2.1.6)
  • Camera versus Human Perception
  • The Pinhole Camera
  • Lens effects

• Geometry of two Views (sec. 7.2)
  • The Homography (e.g. rotating camera)
  • Camera Calibration (3D to 2D Mapping)
  • The Fundamental and Essential Matrix (two arbitrary images)

• Robust Geometry estimation for two cameras (sec. 6.1.4)

• Multi-View 3D reconstruction (sec. 7.3-7.4)
  • General scenario
  • From Projective to Metric Space
  • Special Cases
So far: Pinhole Camera

- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the aperture ("Blende")
Home-made pinhole camera

Why so blurry?

http://www.debevec.org/Pinhole/
Why not make the aperture as small as possible?
- Less light gets through
- Diffraction effects...
Shrinking the aperture

Diffraction effects. Noise due to long exposure.
Adding a lens

- A lens focuses light onto the film
- Lets enough light through
- Rays passing through the center are not deviated
There is a specific distance at which objects are “in focus”.

Other points project to a “circle of confusion” in the image.

By moving the distance between film and lens we control what is in focus (this distance is related to the focal length).
Focal point in real camera

- All parallel rays converge to one point on a plane located at the focal length $f$
Depth of Field

http://www.cambridgeincolour.com/tutorials/depth-of-field.htm
Control the depth of field

Changing the aperture size affects depth of field

• A smaller aperture increases the range in which the object is approximately in focus (but longer exposure needed)
Field of View

FOV depends on focal length and size of the camera retina

Smaller FOV = larger Focal Length
Field of View / Focal Length

Close to affine camera (look at front light)

Large FOV, small f
Camera close to car

Small FOV, large f
Camera far from the car
Question:
1) Is \( A > B \) ?
2) Is \( B > A \) ?
3) I don’t know
Same effect for faces

wide-angle  
Small $f$

standard

telephoto  
Large $f$
Lens Flaws: Chromatic Aberration

Lens has different refractive indices for different wavelengths: causes color fringing.

High quality lens (top)
low quality lens (bottom)
blur + green edges

Purple fringing
Lens flaws: Vignetting
Lens distortion

- Caused by imperfect lenses
- Deviations are most noticeable for rays that pass through the edge of the lens

No distortion  Pin cushion  Barrel