E1 216 COMPUTER VISION

LECTURE 02: CAMERA GEOMETRY

Venu Madhav Govindu
Department of Electrical Engineering
Indian Institute of Science, Bengaluru

2024

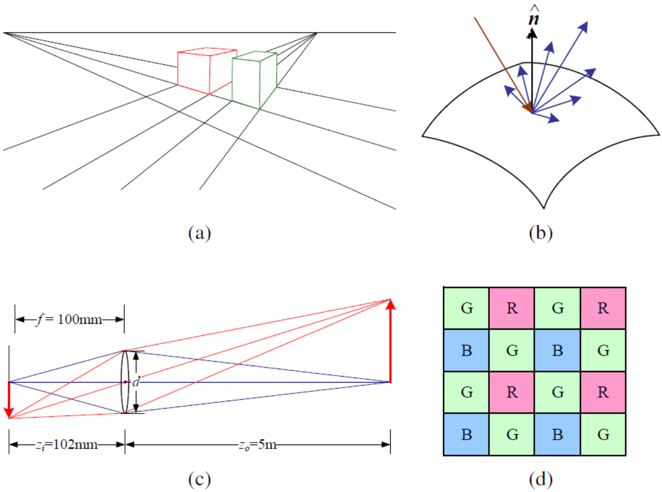
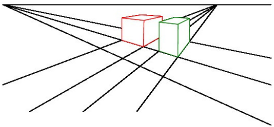
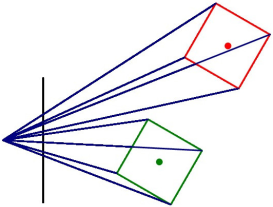


Figure 2.1 *A few components of the image formation process: (a) perspective projection; (b) light scattering when hitting a surface; (c) lens optics; (d) Bayer color filter array.*

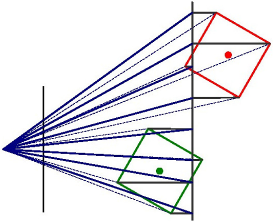
Projection Models



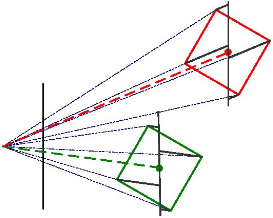
(a) 3D view



(e) perspective



(b) orthography



(d) para-perspective

Camera Geometry

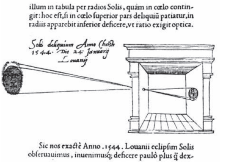
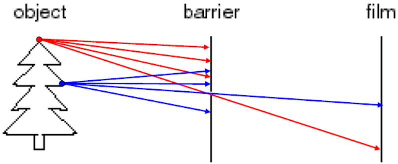


FIGURE 1.1. The camera obscura was used by Reinerus Gemma-Frisius in 1544 to observe an eclipse of the sun.



How do we capture light?

Science for the Curious Photographer, Steve Seitz

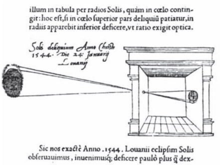
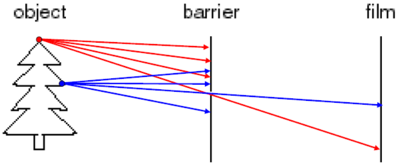


FIGURE 1.1. The camera obscura was used by Reinerus Gemma-Frisius in 1544 to observe an eclipse of the sun.



How do we capture light?

Pinhole Camera

Science for the Curious Photographer, Steve Seitz

Camera Geometry

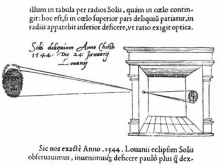
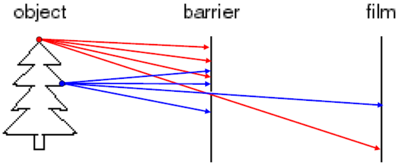


FIGURE 1.1. The camera obscura was used by Reinerus Gemma-Frisius in 1544 to observe an eclipse of the sun.



How do we capture light?

Pinhole Camera

Why?

Science for the Curious Photographer, Steve Seitz

What is a Camera?

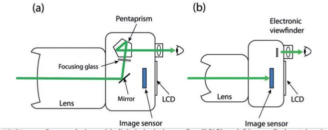
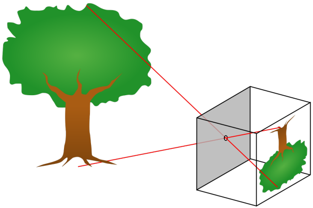


FIGURE 3.4. Camera designs: (a) digital single lens reflex (DSLR) and (b) rangefinder style mirrorless camera.



Camera = Pinhole

Science for the Curious Photographer, wikipedia

What is a Camera?

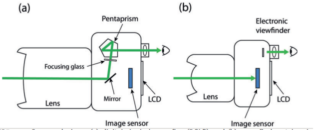
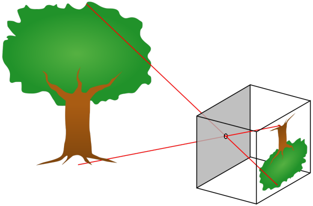


FIGURE 3.4. Camera designs: (a) digital single lens reflex (DSLR) and (b) rangefinder style mirrorless camera.

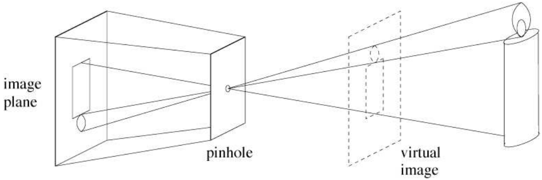


Camera = Pinhole

Powerful Mathematical Model

Science for the Curious Photographer, wikipedia

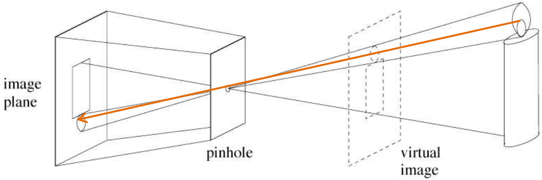
Camera Geometry



Pinhole Camera Model

- What are the consequences of this model?
- Imagine you project a 3D point onto the image plane
- Where did it come from?

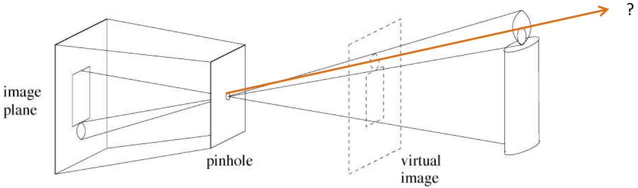
Camera Geometry



Pinhole Camera Model

- What are the consequences of this model?
- Imagine you project a 3D point onto the image plane
- Where did it come from?

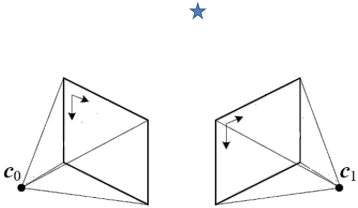
Camera Geometry



Pinhole Camera Model

- What are the consequences of this model?
- Imagine you project a 3D point onto the image plane
- Where did it come from?

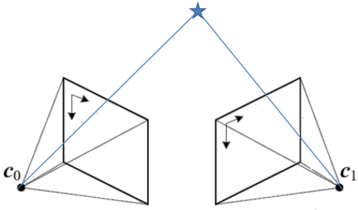
Camera Geometry



Recovering 3D Geometry

- Consider two cameras (one is never enough)
- Take pictures
- Maps to points on image planes
- Know linear constraint on 3D point from left camera
- Use right camera constraint to intersect

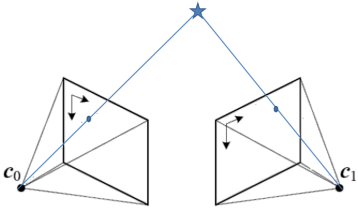
Camera Geometry



Recovering 3D Geometry

- Consider two cameras (one is never enough)
- Take pictures
- Maps to points on image planes
- Know linear constraint on 3D point from left camera
- Use right camera constraint to intersect

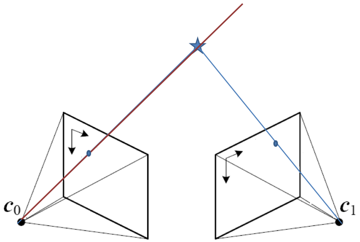
Camera Geometry



Recovering 3D Geometry

- Consider two cameras (one is never enough)
- Take pictures
- Maps to points on image planes
- Know linear constraint on 3D point from left camera
- Use right camera constraint to intersect

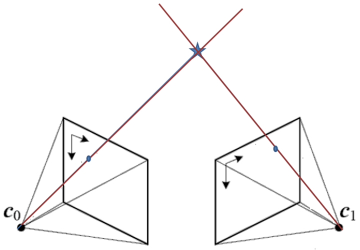
Camera Geometry



Recovering 3D Geometry

- Consider two cameras (one is never enough)
- Take pictures
- Maps to points on image planes
- Know linear constraint on 3D point from left camera
- Use right camera constraint to intersect

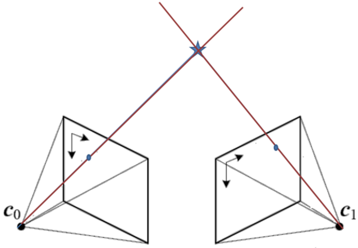
Camera Geometry



Recovering 3D Geometry

- Consider two cameras (one is never enough)
- Take pictures
- Maps to points on image planes
- Know linear constraint on 3D point from left camera
- Use right camera constraint to intersect

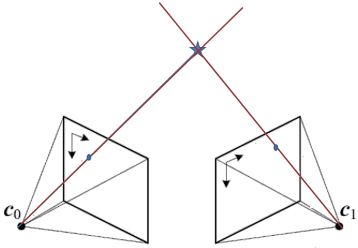
Camera Geometry



Many Considerations

- Do we know camera parameters? (intrinsic calibration)
- Do we know orientations of cameras? (extrinsic calibration)
- Match features (representation,matching,robustness)
- Do the backprojected rays intersect? (structure estimation)
- Extend this principle to multiple images
- Non-trivial, but many important advances
- State-of-the-art can handle large datasets ($> 10^4$ images)

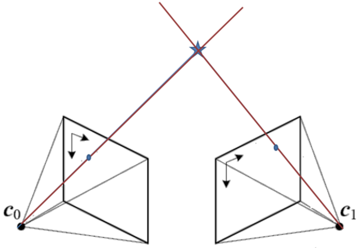
Camera Geometry



Many Considerations

- Do we know camera parameters? (intrinsic calibration)
- Do we know orientations of cameras? (extrinsic calibration)
- Match features (representation,matching,robustness)
- Do the backprojected rays intersect? (structure estimation)
- Extend this principle to multiple images
- Non-trivial, but many important advances
- State-of-the-art can handle large datasets ($> 10^4$ images)

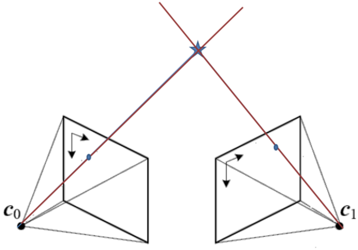
Camera Geometry



Many Considerations

- Do we know camera parameters? (intrinsic calibration)
- Do we know orientations of cameras? (extrinsic calibration)
- Match features (representation,matching,robustness)
- Do the backprojected rays intersect? (structure estimation)
- Extend this principle to multiple images
- Non-trivial, but many important advances
- State-of-the-art can handle large datasets ($> 10^4$ images)

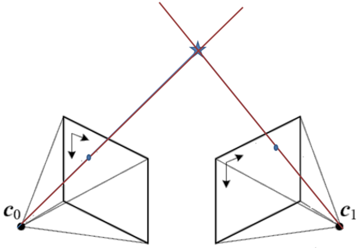
Camera Geometry



Many Considerations

- Do we know camera parameters? (intrinsic calibration)
- Do we know orientations of cameras? (extrinsic calibration)
- Match features (representation,matching,robustness)
- Do the backprojected rays intersect? (structure estimation)
- Extend this principle to multiple images
- Non-trivial, but many important advances
- State-of-the-art can handle large datasets ($> 10^4$ images)

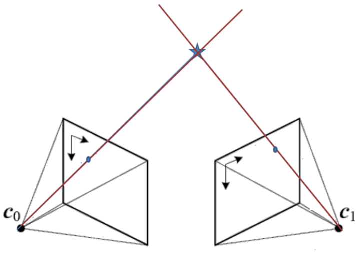
Camera Geometry



Many Considerations

- Do we know camera parameters? (intrinsic calibration)
- Do we know orientations of cameras? (extrinsic calibration)
- Match features (representation,matching,robustness)
- Do the backprojected rays intersect? (structure estimation)
- Extend this principle to multiple images
- Non-trivial, but many important advances
- State-of-the-art can handle large datasets ($> 10^4$ images)

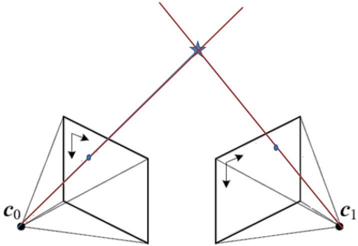
Camera Geometry



Many Considerations

- Do we know camera parameters? (intrinsic calibration)
- Do we know orientations of cameras? (extrinsic calibration)
- Match features (representation,matching,robustness)
- Do the backprojected rays intersect? (structure estimation)
- Extend this principle to multiple images
- Non-trivial, but many important advances
- State-of-the-art can handle large datasets ($> 10^4$ images)

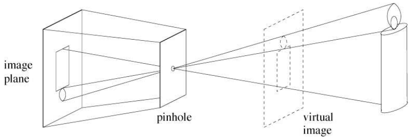
Camera Geometry



Many Considerations

- Do we know camera parameters? (intrinsic calibration)
- Do we know orientations of cameras? (extrinsic calibration)
- Match features (representation,matching,robustness)
- Do the backprojected rays intersect? (structure estimation)
- Extend this principle to multiple images
- Non-trivial, but many important advances
- State-of-the-art can handle large datasets ($> 10^4$ images)

What's a Good Camera Model?



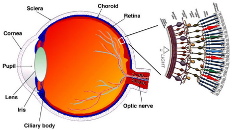
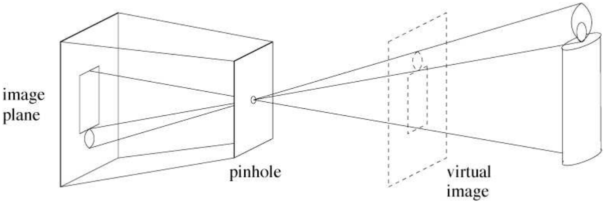


Fig. 1.1. A drawing of a section through the human eye with a schematic enlargement of the retina.

Camera Systems

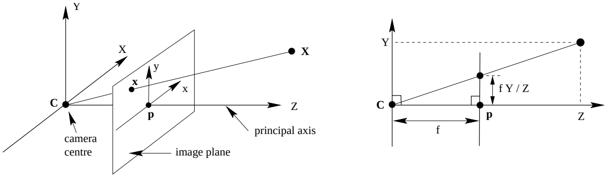
- Camera imaging surface - typically a rectangular plane
- Human retina is closer to a spherical surface
- Vastly different image plane geometries
- Fundamental 3D-2D imaging model is the same
- Spatial sampling is uniform for typical cameras
- Omnidirectional cameras

Camera Model : Perspective Projection



- Very simple geometry
- Sufficiently powerful representation
- *Virtual Image* considered in front of focus
- Real cameras *do* deviate from this model

Camera Model : Perspective Projection

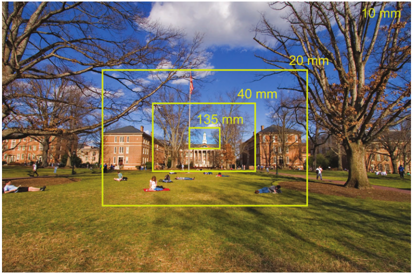


- Coordinate system with origin at camera centre
- World coordinates of point $\boldsymbol{P} = (\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z})$
- Image projection measured in *local* image coordinate system
- Image coordinates $\boldsymbol{p} = (\boldsymbol{x}, \boldsymbol{y})$

By simple similarity of triangles we have

$$\boldsymbol{x} = \frac{f\boldsymbol{X}}{\boldsymbol{Z}}$$
$$\boldsymbol{y} = \frac{f\boldsymbol{Y}}{\boldsymbol{Z}}$$

Camera Model : Perspective Projection



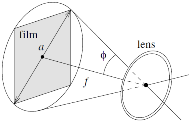


FIGURE 1.9: The field of view of a camera. It can be defined as 2ϕ, where $\phi \stackrel{\text{def}}{=} \arctan \frac{a}{2f}$, a is the diameter of the sensor (film, CCD, or CMOS chip), and f is the focal length of the camera.

Changing focal length

- Keep camera fixed, change focal length
- What happens to the volume imaged?

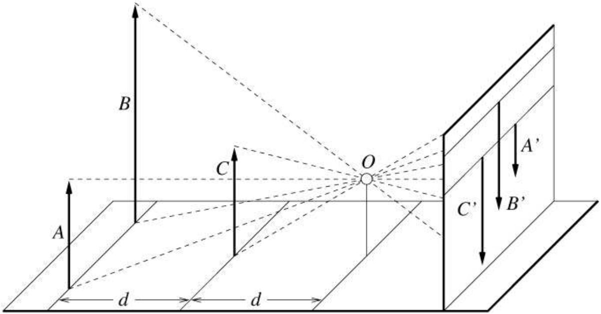
Science for the Curious Photographer; Forstyh and Ponce 2nd Edition.

Camera Model : Perspective Projection

$$x = \frac{fX}{Z}$$
$$y = \frac{fY}{Z}$$

Implications

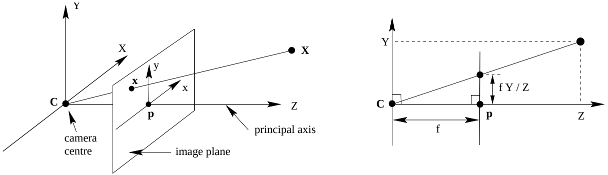
- Different points are scaled different according to depth
- Introduces non-linearities in the relationships
- Distant objects are smaller
- Cannot judge object size with a single image



Perspective projection

- Cannot judge object size with a single image
- Judgement of size can be wrong!

From twitter.com/rainmaker1973



Two co-ordinate systems!

- Remember that we have two measurements of interest
  - Measurements on the image plane
  - Measurements in the 3D world
- Our interest is to relate the two

Camera Model (contd.)

Consider perspective projection model

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Now let's translate the frame of reference (or camera), new co-ordinates

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{f}{Z + t_z} \begin{bmatrix} X + t_x \\ Y + t_y \end{bmatrix}$$

Camera Model (contd.)

Consider perspective projection model

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

Now let's translate the frame of reference (or camera), new co-ordinates

$$x^{'} = f\frac{(X + t_x)}{(Z + t_z)}$$

$$y^{'} = f\frac{(Y + t_y)}{(Z + t_z)}$$

Camera Model (contd.)

Or if we were to rotate the camera by rotation matrix $\boldsymbol{R}$

$$\boldsymbol{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

The new 3D coordinates would be

$$\begin{bmatrix} \boldsymbol{X}' \\ \boldsymbol{Y}' \\ \boldsymbol{Z}' \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \\ \boldsymbol{Z} \end{bmatrix}$$

Camera Model (contd.)

Therefore, the new image projections would look like

$$
\begin{aligned}
x &= f\frac{r_{11}X + r_{12}Y + r_{13}Z}{r_{31}X + r_{32}Y + r_{33}Z} \\
y &= f\frac{r_{21}X + r_{22}Y + r_{23}Z}{r_{31}X + r_{32}Y + r_{33}Z}
\end{aligned}
$$

- Now if we apply an additional transformation, the two rotations would get entangled
- End result of multiple transformations is very messy!
- Need a cleaner approach

Homogeneous Representations

To arrive at a solution, we take recourse to geometry

Geometric approaches

- "Purist" view - co-ordinate free approach to geometry
- Classical theorems due to Euclid
- Since Descartes, there's an algebraic view of geometric constructs
- Duality : Geometry $\leftrightarrow$ Algebra
- Circle : Centre + Radius $\leftrightarrow (\boldsymbol{p} - \boldsymbol{p}_0)^T (\boldsymbol{p} - \boldsymbol{p}_0) = r^2$

Homogeneous Representations

Consider a line $y = mx + c$
Rewrite as $mx - y + c = 0$
or generally as
ax+by+c = 0

Rewriting this we have

$$\begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$

Homogeneous Representation of a Line

$$\underbrace{\begin{bmatrix} a & b & c \end{bmatrix}}_{\boldsymbol{l}} \overbrace{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}^{\boldsymbol{p}} = 0$$

this results in a nice symmetric form

$$\boldsymbol{l}^T \boldsymbol{p} = 0$$

This form has many advantages over $y = mx + c$ form

Homogeneous Representation of a Line

Consider the intersection of two lines
To solve for the point of intersection

$$y = m_1 x + c_1$$
$$y = m_2 x + c_2$$

Solve simultaneous equations by substitution, $x = \frac{(y - c_1)}{m_1}$

$$\begin{aligned} y &= (y - c_1)\frac{m_2}{m_1} + c_2 \\ (1 - \frac{m_2}{m_1})y &= c_2 - \frac{c_1 m_2}{m_1} \\ y &= = \frac{(c_2 - \frac{c_1 m_2}{m_1})}{(1 - \frac{m_2}{m_1})} \end{aligned}$$

Quite a mess!!

Homogeneous Representation of a Line

In the homogeneous system of representation we have

$$\boldsymbol{l}_1{}^T \boldsymbol{p} = 0$$
$$\boldsymbol{l}_2{}^T \boldsymbol{p} = 0$$

Therefore, the co-ordinates of the intersection is given by

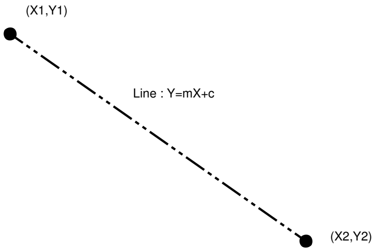
$$\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$$

Verify

- $\boldsymbol{l}_1{}^T(\boldsymbol{l}_1 \times \boldsymbol{l}_2) = 0$
- $\boldsymbol{l}_2{}^T(\boldsymbol{l}_1 \times \boldsymbol{l}_2) = 0$
- Much cleaner way of solving

Homogeneous Representation of a Line

Consider the line through two given points



Usual solution is messy
Instead, using homogeneous coordinates, we get the dual representation

$$\text{Line} : \boldsymbol{l} = \boldsymbol{p}_1 \times \boldsymbol{p}_2$$

Easily verified that this satisfies the requirements

- $(\boldsymbol{p}_1 \times \boldsymbol{p}_2)^T \boldsymbol{p}_1 = 0$
- $(\boldsymbol{p}_1 \times \boldsymbol{p}_2)^T \boldsymbol{p}_2 = 0$

Homogeneous Representation

The key relationship to note is that

$$\underbrace{\begin{bmatrix} a & b & c \end{bmatrix}}_{\boldsymbol{l}} \overbrace{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}}^{\boldsymbol{p}} = 0$$

results in a nice symmetric (and homogeneous) form

$$\boldsymbol{l}^T \boldsymbol{p} = 0$$

This form has many advantages over $y = mx + c$ form

Homogeneous Representation

In homogeneous form everything upto unknown scalar

Homogeneous

$\mathbb{R}^n \mapsto \mathbb{R}^{n+1}$

$$\begin{bmatrix} u \\ v \end{bmatrix} \mapsto \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

Inhomogeneous

$\mathbb{R}^n \mapsto \mathbb{R}^{n-1}$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \mapsto \begin{bmatrix} \frac{u}{w} \\ \frac{v}{w} \end{bmatrix}$$

Homogeneous Forms

- Embed in higher dimensions by appending a 1 (canonical)
- Homogeneous forms are equivalent upto scale
- Only ratios matter
- $[u, v, w] = \lambda [u, v, w], \forall \lambda \neq 0$
- Notice $[0, 0, 0]$ is not admissible

Homogeneous Representation

In homogeneous form everything upto unknown scalar

Homogeneous

$\mathbb{R}^n \mapsto \mathbb{R}^{n+1}$

$$\begin{bmatrix} u \\ v \end{bmatrix} \mapsto \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

Inhomogeneous

$\mathbb{R}^n \mapsto \mathbb{R}^{n-1}$

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} \mapsto \begin{bmatrix} \frac{u}{w} \\ \frac{v}{w} \end{bmatrix}$$

Homogeneous Forms

- Embed in higher dimensions by appending a 1 (canonical)
- Homogeneous forms are equivalent upto scale
- Only ratios matter
- $[u, v, w] = \lambda [u, v, w], \forall \lambda \neq 0$
- Notice $[0, 0, 0]$ is not admissible

7.7 Homogeneous Coordinates

Representing the points of the projective plane $\mathbb{RP}^2$ by lines through O gives *coordinates* to $\mathbb{RP}^2$ via the coordinates (x, y, z) of three-dimensional space. Such coordinates were invented by Möbius (1827) and Plücker (1830), and they are called *homogeneous* because each algebraic curve in $\mathbb{RP}^2$ is expressed by a homogeneous polynomial equation $p(x, y, z) = 0$. The simplest case is that of a projective line, which, as we saw in Section 7.5, is represented by a plane through O. Its equation therefore has the form

$$ax + by + cz = 0, \quad \text{for some constants } a, b, c, \text{ not all zero.}$$

Such an equation is called *homogeneous of degree* 1, because each nonzero term is of degree 1 in the variables x, y, z.

The *homogeneous coordinates of a point* P in $\mathbb{RP}^2$ are simply the coordinates of *all* points on the line through O that represents P. It follows that

Geometries in Computer Vision

- Geometry : Topological Space + Axioms
- Different set of axioms $\rightsquigarrow$ Different Geometries
  - Euclidean (Distances and Angles)
  - Affine (Parallelism)
  - Projective (Straight Line)
  - Non-linear (Riemannian Manifolds)

Stratification of transform space

Euclidean $\subset$ Affine $\subset$ Projective

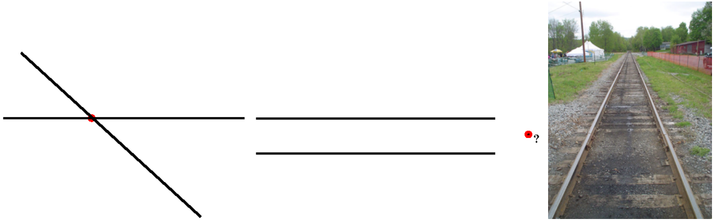
Euclidean Geometry

Axioms of incidence

- Familiar concepts from Euclidean geometry
- Length is a fundamental property of Euclidean Geometry
- Construction with straightedge and compass
- Axioms of Euclid

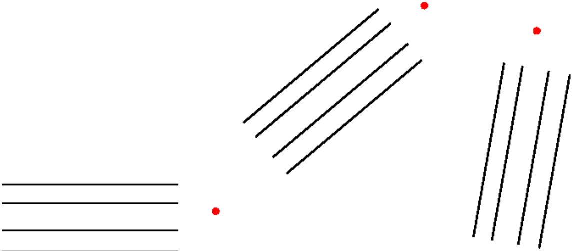
Following Hilbert state the axioms as

- For any two points A, B, a unique line passes through A, B
- Every line contains at least two points
- There exist three points not all on the same line
- **Parallel axiom** : For any line $\mathcal{L}$ and point $\mathcal{P}$ outside $\mathcal{L}$, there is exactly one line through $\mathcal{P}$ that does not meet $\mathcal{L}$



Wikipedia

- Two points have a unique line through them (join)
- Two lines have a unique intersection point (meet)
- What happens when the lines are parallel?
- What does it mean to say that they “intersect at ∞”?



- **Question** : Are all ∞ intersection points the same?
- The answer lies in the geometry of projective space
- Recall homogeneous representations

Homogeneous Forms

Parallel Lines

- Recall line equation: $\boldsymbol{l}^T\boldsymbol{p} = 0$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor $\boldsymbol{l}^T\boldsymbol{p} = (\lambda \boldsymbol{l})^T(\lambda' \boldsymbol{p}) = 0$
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms

Parallel Lines

- Recall line equation: $\boldsymbol{l}^T \boldsymbol{p} = 0$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor $\boldsymbol{l}^T \boldsymbol{p} = (\lambda \boldsymbol{l})^T (\lambda' \boldsymbol{p}) = 0$
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms

Parallel Lines

- Recall line equation: $\boldsymbol{l}^T\boldsymbol{p} = 0$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor $\boldsymbol{l}^T\boldsymbol{p} = (\lambda\boldsymbol{l})^T(\lambda'\boldsymbol{p}) = 0$
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms

Parallel Lines

- Recall line equation: $\boldsymbol{l}^T \boldsymbol{p} = 0$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor $\boldsymbol{l}^T \boldsymbol{p} = (\lambda \boldsymbol{l})^T (\lambda' \boldsymbol{p}) = 0$
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms

Parallel Lines

- Recall line equation: $\boldsymbol{l}^T \boldsymbol{p} = 0$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor $\boldsymbol{l}^T \boldsymbol{p} = (\lambda \boldsymbol{l})^T (\lambda' \boldsymbol{p}) = 0$
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms

Parallel Lines

- Recall line equation: $\boldsymbol{l}^T\boldsymbol{p} = 0$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor $\boldsymbol{l}^T\boldsymbol{p} = (\lambda\boldsymbol{l})^T(\lambda'\boldsymbol{p}) = 0$
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms

Parallel Lines

- Recall line equation: $\boldsymbol{l}^T \boldsymbol{p} = 0$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor $\boldsymbol{l}^T \boldsymbol{p} = (\lambda \boldsymbol{l})^T (\lambda' \boldsymbol{p}) = 0$
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms

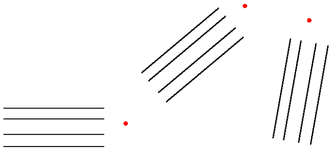
$$\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2 = \begin{bmatrix} (c^{'} - c)b & (c - c^{'})a & 0 \end{bmatrix} = [b, -a, 0]$$

What is the inhomogeneous form of $\boldsymbol{p}$?

Parallel Lines

- Recall line equation: $\boldsymbol{l} = \boldsymbol{p}_1 \times \boldsymbol{p}_2$
- $\boldsymbol{l}$ and $\boldsymbol{p}$ upto scale factor
- Intersection of two lines $\boldsymbol{p} = \boldsymbol{l}_1 \times \boldsymbol{l}_2$
- When are lines parallel?
- $\boldsymbol{l}_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $\boldsymbol{l}_2 = \begin{bmatrix} a & b & c^{'} \end{bmatrix}$
- Intersection point $\boldsymbol{p}$?

Homogeneous Forms



$$p = l_1 \times l_2 = \begin{bmatrix} (c' - c)b & (c - c')a & 0 \end{bmatrix}$$
$$= [b, -a, 0]$$

What is the inhomogeneous form of p?
Distinct "points at infinity"

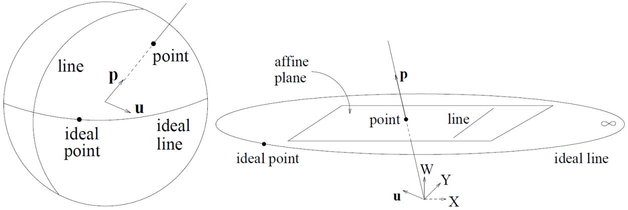
Parallel Lines

- Recall line equation: $l = p_1 \times p_2$
- l and p upto scale factor
- Intersection of two lines $p = l_1 \times l_2$
- When are lines parallel?
- $l_1 = \begin{bmatrix} a & b & c \end{bmatrix}$
- $l_2 = \begin{bmatrix} a & b & c' \end{bmatrix}$
- Intersection point p?

Projective Geometry

- Represent the projective plane as $\mathbb{P}^2$
- Obtained by adding all ∞ points
- ∞ points form a ‘line at infinity’. Why?
- Got rid of special case of parallel lines
- All lines have a unique intersection now
- So what is this space useful for?

Projective Geometry



(d) $\mathbb{P}^2 \equiv S^2$ (e) $\mathbb{P}^2 \equiv \mathbb{R}^3 \backslash \{\mathbf{0}\} / \simeq$

- Projective plane is topologically equivalent to unit sphere
- Associate with half-sphere to projective scale
- Where is the line at infinity on S^2?
- $\mathbb{P}^2$ is equivalent to $\mathbb{R}^3$ with origin removed, under equivalence relationship of scale

Homogeneous Form

Basic Definition

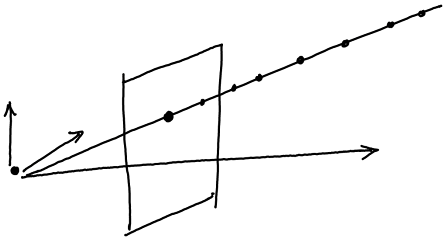
- n-dim real **affine space** is set of all points $(x_1, \cdots, x_n) \in \mathbb{R}^n$
- **Projective space** $\mathbb{P}^n$ given by
  - $(x_1, \cdots, x_n, x_{n+1}) \in \mathbb{R}^{n+1}$
  - at least one x_i is non-zero
  - for $\lambda \neq 0$, all $(\lambda x_1, \cdots, \lambda x_n, \lambda x_{n+1})$ are equivalent
- Homogeneous coordinates obtained by $(x_1, \cdots, x_n, 1)$

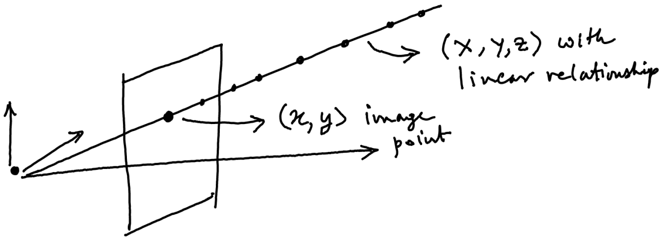
Homogeneous Form

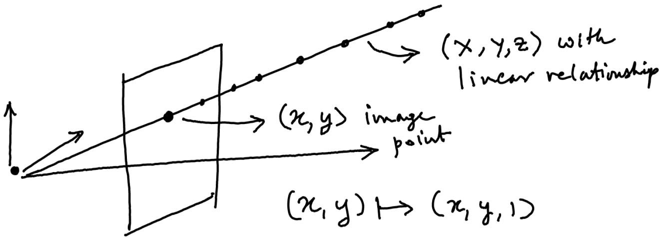
- Let the homogeneous form be $\boldsymbol{X} = (\boldsymbol{X}_1, \cdots, \boldsymbol{X}_{n+1})$
- Let the inhomogeneous form be $\boldsymbol{x} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n)$
- Equivalence relationship : $[\boldsymbol{x}, 1] = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n, 1) \simeq \boldsymbol{X}$
- $\boldsymbol{x}_i = \frac{\boldsymbol{X}_i}{\boldsymbol{X}_{n+1}}$

Line at Infinity

- **Question**: What is the homogeneous form for points at ∞?
- Is this homogeneous form $[\boldsymbol{x}, 1]$ always valid?
- $[\boldsymbol{x}, 0]$ is also in projective space
- $[\boldsymbol{x}, 0]$ does not have a finite inhomogenous form
- Projective Space: $[\boldsymbol{x}, 1]$ (affine space) $\cup$ $[\boldsymbol{x}, 0]$ (line at ∞)

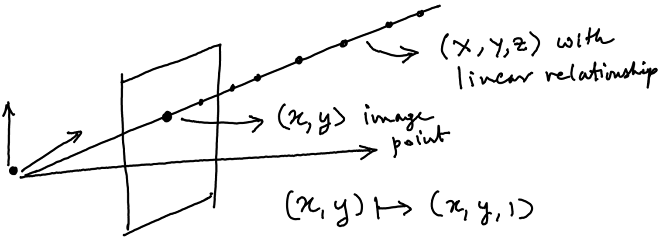



(X, Y, Z) with
linear relationship
(x, y) image
point



$(x, y) \mapsto (x, y, 1)$

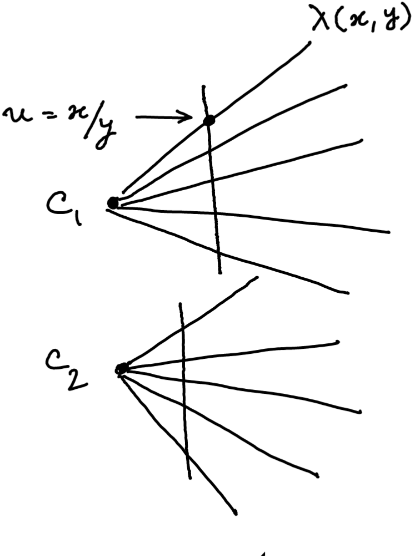
$(x, y, 1)$ is also point in (X, Y, Z)



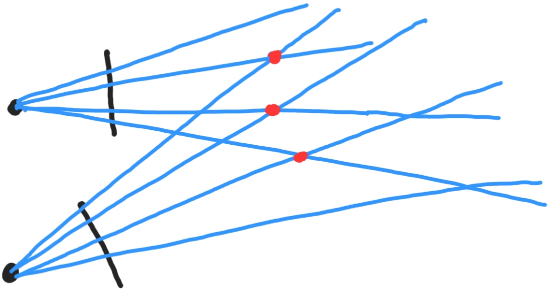
$$(x, y) \mapsto (x, y, 1)$$

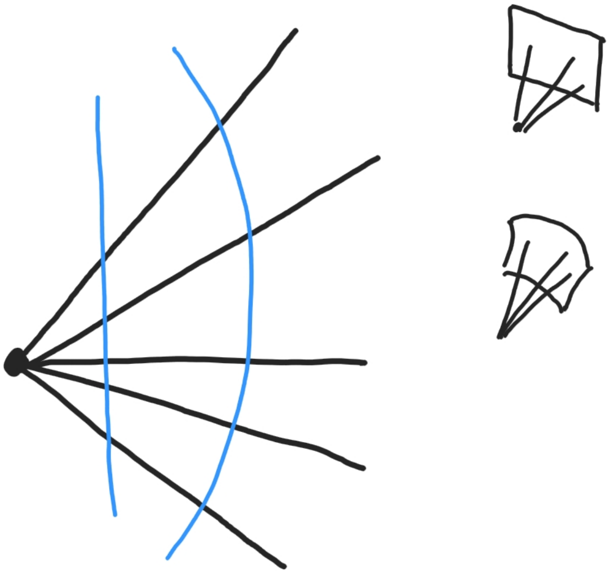
$(x, y, 1)$ is also point in (X, Y, Z)

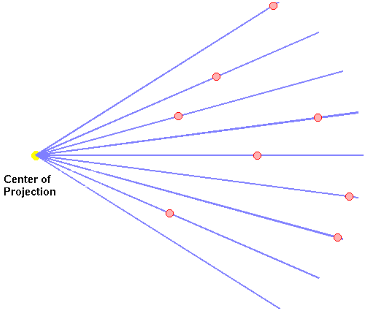
$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$



X(x, y)
u = x/y
C1
C2



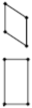
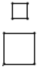




Camera = Camera Centre!

- Consider a centre of projection
- Establishes equivalence classes
- All points on ray are projectively equivalent (beads on wire)
- What happens when they line up?
- Camera model

Transformation Groups

Group	Matrix	Distortion	Invariant properties
Projective 8 dof	$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$		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	$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		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, $\mathbf{l}_\infty$.
Similarity 4 dof	$\begin{bmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Ratio of lengths, angle. The circular points, **I**, **J** (see section 2.7.3).
Euclidean 3 dof	$\begin{bmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$		Length, area

$$
\underset{\textit{Projective}}{\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}}
\quad
\underset{\textit{Affine}}{\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ 0 & 0 & 1 \end{bmatrix}}
\quad
\underset{\textit{Euclidean}}{\begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ 0 & 0 & 1 \end{bmatrix}}
$$

Euclidean

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Projective

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Two interpretations

- Euclidean vs. Projective transformations
- $\boldsymbol{H} : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ (9 dof)
- $\boldsymbol{H} : \mathbb{P}^2 \rightarrow \mathbb{P}^2$ (8 dof)

Since in projective space

$$\mathbb{P}^n, \text{ all } (\lambda x_1, \cdots, \lambda x_n, \lambda x_{n+1})$$

are equivalent, we can *linearise* our imaging model
Recall that

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} X \\ Y \end{bmatrix}$$

For now assume $f = 1$
then by **embedding** image and world points in projective spaces we have

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

We now have

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

Recall, that scaled points are projectively equivalent, i.e.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

$$\boldsymbol{p} = \boldsymbol{P}$$

We have now managed to linearise the relationship

Projective Geometry

Projective representations for both image and world points

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

Euclidean transformation of 3D points

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Projective Geometry

General process of taking a picture

- Apply Euclidean motion to 3D points
- Project onto image plane

Combining two steps we get

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

$$\Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \left[\, R \mid t \, \right] \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Projective Geometry

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Taking a Picture

- 3D Point in Homogeneous Form
- Rigid 3D Motion
- Ideal Pinhole Camera
- Image Projection
- $\mathbb{P}^3 \to \mathbb{P}^2$

Projective Geometry

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \boxed{\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}}$$

Taking a Picture

- 3D Point in Homogeneous Form
- Rigid 3D Motion
- Ideal Pinhole Camera
- Image Projection
- $\mathbb{P}^3 \longrightarrow \mathbb{P}^2$

Projective Geometry

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Taking a Picture

- 3D Point in Homogeneous Form
- Rigid 3D Motion
- Ideal Pinhole Camera
- Image Projection
- $\mathbb{P}^3 \rightarrow \mathbb{P}^2$

Projective Geometry

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Taking a Picture

- 3D Point in Homogeneous Form
- Rigid 3D Motion
- Ideal Pinhole Camera
- Image Projection
- $\mathbb{P}^3 \rightarrow \mathbb{P}^2$

Projective Geometry

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Taking a Picture

- 3D Point in Homogeneous Form
- Rigid 3D Motion
- Ideal Pinhole Camera
- Ideal Projection
- $\mathbb{P}^3 \longrightarrow \mathbb{P}^2$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Taking a Picture

- 3D Point in Homogeneous Form
- Rigid 3D Motion
- Ideal Pinhole Camera
- Image Projection
- $\mathbb{P}^3 \to \mathbb{P}^2$

Projective Geometry

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

Taking a Picture

- 3D Point in Homogeneous Form
- Rigid 3D Motion
- Ideal Pinhole Camera
- Image Projection
- $\mathbb{P}^3 \rightarrow \mathbb{P}^2$

Camera Model

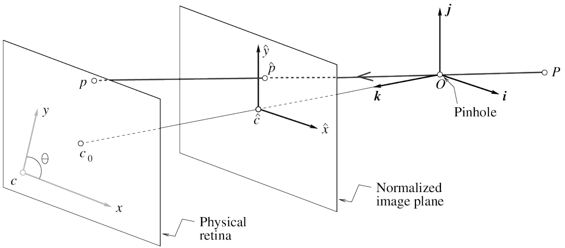


FIGURE 1.14: Physical and normalized image coordinate systems.

Intrinsic Parameters

- Focal length f
- Shift in origin or image center (u_0, v_0)
- Rectangular pixel dimensions (k_u, k_v)
- Imaging plane may be skewed by angle θ

Many deviations from an idealised model

Makes the entire imaging model very messy

Further, the effects of the camera parameters can be represented as a matrix form

$$\boldsymbol{K} = \begin{bmatrix} fk_u & -fk_u cot\theta & -u_0 \\ 0 & \frac{fk_v}{sin\theta} & -v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

General form for the transformation matrix is

$$\begin{bmatrix} fk_u & -fk_u cot\theta & -u_0 \\ 0 & \frac{fk_v}{sin\theta} & -v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \boldsymbol{t}_x \\ r_{21} & r_{22} & r_{23} & \boldsymbol{t}_y \\ r_{31} & r_{32} & r_{33} & \boldsymbol{t}_z \end{bmatrix}$$

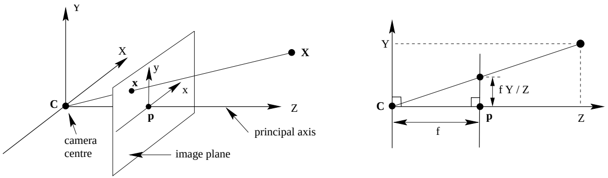
$$\underbrace{\begin{bmatrix} fk_u & -fk_u cot\theta & -u_0 \\ 0 & \frac{fk_v}{sin\theta} & -v_0 \\ 0 & 0 & 1 \end{bmatrix}}_{intrinsic} \overbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix}}^{extrinsic}$$

Put simply the general form of the projective transformation is

$$\boldsymbol{P} = \begin{bmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \boldsymbol{P}_{13} & \boldsymbol{P}_{14} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \boldsymbol{P}_{23} & \boldsymbol{P}_{24} \\ \boldsymbol{P}_{31} & \boldsymbol{P}_{32} & \boldsymbol{P}_{33} & \boldsymbol{P}_{34} \end{bmatrix}$$

Has $3 \times 4 - 1 = 11$ degrees of freedom

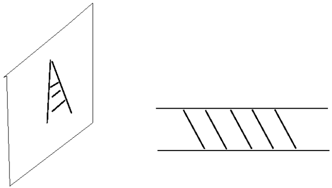
REMINDER!



Two co-ordinate systems!

- Remember that we have two measurements of interest
  - Measurements on the image plane
  - Measurements in the 3D world
- Our interest is to relate the two

Why Projective Geometry?



- A camera is a projective engine
- Simpler representation than affine or Euclidean forms
- Can handle points-at-∞ naturally (fewer special cases)
- Most general representation for our problems

Projective Representations

Reminder

- We are dealing with three types of Projective transformations or mappings
  - Transformations of image plane ($\boldsymbol{H}_{3\times 3} : \mathbb{P}^2 \rightarrow \mathbb{P}^2$)
  - Imaging by a pinhole camera ($\boldsymbol{P}_{3\times 4} : \mathbb{P}^3 \rightarrow \mathbb{P}^2$)
  - Projective change of basis for 3D space ($\boldsymbol{H}_{4\times 4} : \mathbb{P}^3 \rightarrow \mathbb{P}^3$)

Representation of the Projective Camera

$$\boldsymbol{P} = \boldsymbol{K} \left[\begin{array}{ccc|c} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} & \boldsymbol{R}_{13} & \boldsymbol{T}_1 \\ \boldsymbol{R}_{21} & \boldsymbol{R}_{22} & \boldsymbol{R}_{23} & \boldsymbol{T}_2 \\ \boldsymbol{R}_{31} & \boldsymbol{R}_{32} & \boldsymbol{R}_{33} & \boldsymbol{T}_3 \end{array} \right] \quad vs. \quad \left[\begin{array}{ccc|c} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} & \boldsymbol{P}_{13} & \boldsymbol{P}_{14} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} & \boldsymbol{P}_{23} & \boldsymbol{P}_{24} \\ \boldsymbol{P}_{31} & \boldsymbol{P}_{32} & \boldsymbol{P}_{33} & \boldsymbol{P}_{34} \end{array} \right]$$

- Distinction between perspective and projective cameras
- Perspective is a model for a true Euclidean (rigid) transformation
- Perspective camera is a special case of projective camera
- Projective camera is a *purely* mathematical engine
- Projective camera is not necessarily physically realisable
- What about degrees of freedom?

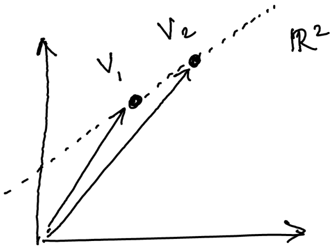
Euclidean Transformation in $\mathbb{R}^3$

$$\begin{aligned} \boldsymbol{P}^{'} &= \boldsymbol{R}\boldsymbol{P} + \boldsymbol{T} \\ \boldsymbol{P}^{'} &= \boldsymbol{R}(\boldsymbol{P} + \boldsymbol{T}) \end{aligned}$$

- First rotate then translate
- Second translate then rotate
- Both are valid representations
- We will prefer the first form over the second
- **Warning : Always understand which one is used!**

EXTRA MATERIAL
NOT PART OF SYLLABUS

Affine Geometry



Consider $v_1, v_2 \in \mathbb{R}^2$
Linear combination: Span $\{v_1, v_2\}$
Affine combination: Line in $\mathbb{R}^2$

Affine Combinations

Consider vectors $v_1, \cdots, v_k \in \mathbb{R}^n$
Affine Combination: $\lambda_1 v_1 + \cdots \lambda_k v_k$
$\lambda_1, \cdots, \lambda_k \in \mathbb{R}$
Restriction: $\lambda_1 + \cdots + \lambda_k = 1$

Linear Combinations

Consider vectors $v_1, \cdots, v_k \in \mathbb{R}^n$
Linear Combination:
$\lambda_1 v_1 + \cdots + \lambda_k v_k \in \mathbb{R}^n$
$\lambda_1, \cdots, \lambda_k \in \mathbb{R}$

Convex Combinations

Restriction: $\lambda_1 + \cdots + \lambda_k = 1$
Further restriction $\lambda_i \geq 0$

Affine Geometry

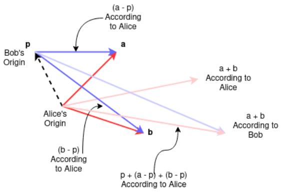
Vector Subspace

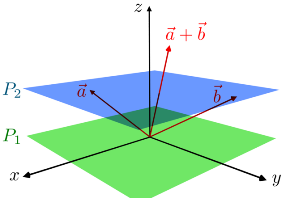
- $A \subseteq \mathbb{R}^n$
- $\mathbf{0} \in A$
- $\boldsymbol{a} \in A \Rightarrow \lambda \boldsymbol{a} \in A$
- $\boldsymbol{a}, \boldsymbol{b} \in A \Rightarrow \boldsymbol{a} + \boldsymbol{b} \in A$
- Points and vectors coincide
- Equipped with inner product
- Distances and angles preserved

Affine Subspace

- $A \subseteq \mathbb{R}^n$
- No origin
- $\boldsymbol{a} \in A \Rightarrow \lambda \boldsymbol{a} \in A$
- $\boldsymbol{a}, \boldsymbol{b} \in A \Rightarrow \lambda \boldsymbol{a} + (1 - \lambda) \boldsymbol{b} \in A$
- $A - \boldsymbol{a}$ is a vector space for any $\boldsymbol{a} \in A$
- Vectors only as differences (translations)
- Only parallelism is preserved

Affine Geometry





Affine Subspaces

- Consider the 2D plane, but forget origin
- What can two independent observers agree upon?
- Second observer assumes that $\boldsymbol{p}$ is the origin
- Adding two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ results in $\boldsymbol{p} + (\boldsymbol{a} - \boldsymbol{p}) + (\boldsymbol{b} - \boldsymbol{p})$
- When linear combination is $\lambda \boldsymbol{a} + (1 - \lambda)\boldsymbol{b}$, observers agree
- Observers know the “affine structure” but not the “linear structure”
- Direction is a fundamental property here, not length

wikipedia